Playing Robot Soccer under Natural Light: A Case Study

Gerd Mayer, Hans Utz, and Gerhard K. Kraetzschmar

University of Ulm, James-Franck-Ring, 89069 Ulm, Germany

Abstract. The recent debate in the RoboCup middle-size community about natural light conditions shows that a more in-depth analysis of the problems incurred by this is necessary in order to draft out a focused and realistic roadmap for research. Based on real-world images taken under varying lighting conditions, we performed descriptive and statistical analysis of the effects on color-based vision routines. The results show that pure color-based image processing is not likely to perform well under varying lighting conditions, even if the vision system is calibrated on a per-game base. We conclude that color-based vision has to be combined with other methods and algorithms in order to work robustly in more difficult environments with varying illumination.

1 Introduction

One of the long-term goals of the RoboCup initiative is to construct soccer robots capable of playing in typical soccer playgrounds, including outdoor soccer fields. As a consequence, soccer robots must be able to play under varying lighting conditions, ranging from temporally stable artificial illumination (fluorescent or floodlight light) to natural illumination with varying temporal dynamics (slow changes during the day, fast changes by clouds). An open challenge for the community is to draft a roadmap for research, which ultimately yields soccer robots playing well in all these environments.

Currently, the middle-size league plays robot soccer in an environment with strictly controlled artificial lighting conditions. Providing this environment is laborious and expensive for tournament organizers. For this reason, the RoboCup middle-size community discussed on its mailing list, whether competitions can be held under daylight conditions in the near future or not. Some teams have already tried to play under natural lighting conditions, e.g. at fairs or exhibition games, and problems seemed to be not so bad. Nevertheless, permitting daylight constitutes a significant change in the environment and is expected to severely affect game play and robot performance. Although everyone seems to intuitively agree on that, the roadmap discussion would greatly benefit from more detailed scientific investigations on this topic. This paper contributes to this discussion by providing an analysis of the effects of various lighting conditions on color-based image processing for soccer robots.

The paper is organized as follows: Section 2 discusses various lighting conditions, and their effect on image processing and vision routines. The image data

D. Polani et al. (Eds.): RoboCup 2003, LNAI 3020, pp. 238–249, 2004.

used and the methods applied for further analysis are described in sections 3 and 4. Experiments and results are presented in Section 5. Section 6 draws some conclusions and discusses potential solutions.

2 Survey of Lighting Conditions

In order to gain a better understanding of the causes and effects of varying lighting conditions, we analyze several different scenarios for playing (robot) soccer.

The first scenario is the one currently in use for RoboCup tournaments: An indoor soccer field is illuminated by *artificial light*. The current rules constrain light intensity to lie between 800 and 1200 lux, but do not impose additional constraints, e.g. on color temperature. Special care is often taken to provide diffuse lighting. Once calibrated, the lighting system remains unchanged during the whole tournament. Although there is no a-priori specification of lighting conditions, teams can perform an on-site, one-time calibration of their vision systems and optimize their performance for the lighting situation on a particular field. The calibration should account for the remaining small variations of lighting across the field. This process needs some care, but can be done with reasonable effort. It results in robots with vision systems well tuned to the lighting situation on site, but which cannot deal with unexpected disturbances like those caused by flashlight.

Fig. 1. The computer science faculty building at TU Munich, where most test images were taken.

The next scenario is characterized by *small changes over time*. An example would be the slow movement of the sun during the day. In addition, the shadows thrown by objects in the environment, e.g. from trees, pillars, or even the robot itself (see the rightmost image[1] in Figure 2), will wander over time, and dramatically change local perception of other objects like the goal or the ball. This effect can be even stronger in indoor environments illuminated by natural light through windows, which are often located on one side of the room. Assuming that lighting changes remain small within the relevant match period, robot teams can handle this scenario with suitable technology for fast on-site, per-game vision calibration. Due to the presence of shadows, the remaining variations in lighting conditions within the field are harder to deal with.

A third scenario is characterized by *sudden, discrete variations* of illumination. This happens especially on days with partially cloudy skies, when a cloud temporarily covers the sun. Very sudden lighting changes also occur, if something shadows the sun for a short moment only, like people bending over a robot's camera. As these changes are hardly predictable, pre-game calibration to cope with such effects should be very difficult, if not impossible.

[1] All color images of this paper can be found at
`http://smart.informatik.uni-ulm.de/DataSets/RoboCup/NatLight/`

3 Image Data

A set of soccer robot camera images covering all previously described scenarios was collected in two different locations in order to allow for further analysis. One location was the new computer science faculty building of Technical University of Munich, where we were invited for a few friendly games against the Agilo robot soccer team. The soccer field was set up in the large central hall (see Figure 1), the outer hull of which consists almost completely of windows. The images were made during a period of about 5 hours around midday and early afternoon, and at the late evening during setup. The other location was a lab room at the University of Ulm with large windows on one side of the room. In both locations, the same carpet, goals, and balls were used.

Image set I (Munich, evening) The first set of images were taken during setup in the late evening. So, no natural light was present. The field was illuminated only by the fluorescent lights. Thus, illumination consisted of constant, very diffuse light, so neither hard shadows nor strong reflections posed a major problem. Although light intensity was probably at the lower end of the range, this setting was certainly closest to the RoboCup setting currently specified.

Image set II (Munich, daytime) The second set of images were taken over a period of 5 hours around midday and early afternoon of the next day. The weather was sunny with occasional clouds. Illumination of the field was both with the indoor lighting structure and heavy natural light influence through the dyed windows. Because the sun was very low and surrounding buildings shaded her from throwing hard shadows on the field, the overall light was still rather diffuse, although the dyed windows influenced the color spectrum significantly (see the center image in Figure 2). In addition, illumination changed slightly over time; see Section 5 for examples on this.

Image set III (Ulm, daytime) A third series of images was made in our laboratory with windows on only one side of the room. Because they were collected at a sunny day, illumination consisted of very directed light, which led to hard shadows and non-uniform lighting.

Fig. 2. Examples of the three different lighting conditions used for the experiments. Neon light in Munich in the leftmost picture, pure daylight in Munich in the middle one, and again natural but directed light in our lab in the right image.

Some image examples are shown in Figure 2. Note the differences in color and illumination uniformity, although exactly the same carpet and the same goals were used.

4 Method

In order to investigate the effects of varying lighting conditions, it would be best to take a number of test images for precisely the same situation on the playground, but under different, carefully controlled lighting conditions. A direct comparison of these images would then be possible *without* assuming or implying any particular method for image processing. Unfortunately, such a data set is virtually impossible to obtain, because neither can the weather be controlled nor would it be possible to guarantee comparability of images taken under motion, so need to use other investigative means for evaluation.

As color-based image processing still is the dominating approach in RoboCup middle-size league, we evaluate the effects on the color information in sets of real-world images taken under different lighting conditions. Furthermore, we investigate how these variations affect common color-based image processing methods. In our team, the early visual processing consists of three steps: *Transformation* (map image from RGB to HSV color space), *Classification* (Assign a color class to each pixel) and *Blob detection* (find regions of contiguous pixels with the same color class). Most further processing, like object recognition or field line detection for self-localization, make somehow use of the color-segmented regions image delivered by this process. Many teams use a similar approach, so that the results presented here should apply to them as well.

The second step is the one of interest here, because the quality of the classification process has a strong influence on any further processing. In our case, colors are mapped from the HSV color space to a set of target color classes, containing all colors relevant to RoboCup, and a catch-all color class *gray*, to which all colors not of interest are mapped. The mapping is simply defined by manually specifying for each target color class a lower and upper bound in each color space dimension. If the target color classes can be sufficiently narrow defined and are far away from each other in the color space (high separability), then the calibration process trades off misclassifications between a target color class and the catch-all color class. Often, however, target classes are bordering each other or would even have to overlap (i.e. there exist particular pixel values belonging to different classes, sometimes even in the same image), so calibration must directly trade off misclassifications between two target color classes.

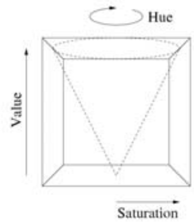

Fig. 3. Illustration of the HSV color space structure.

The HSV color space is a circular, cone-like representation of the dimensions hue, saturation and brightness. Black is at the cone peak on the bottom, white at the innermost part of the upper plate and the colors arranged orbitally around the white color (illustrated in Figure 3).

For small variations, like those even present on the current RoboCup field and lighting setup, the calibration process can make cuboids of target colors large enough to compensate for these variations and still obtain good classification results. The open question is whether this can also be done for larger variations, like those present in scenarios involving daylight. In order to answer this question, we perform the following experiments:

1. We first compare the target class boundaries for several scenarios. For each scenario, the boundaries are the result of a manual calibration process based on a representative set of images.
2. Using the manually calibrated color classificator for each scenario, we color-classify a larger set of images. The spatial distribution of pixel values in each color class is characterized for each scenario by computing and comparing their mean and standard deviation.
3. In order to show that the results are not significantly influenced by the chosen color classification procedure and its calibration process, the previous results are compared with equivalent distributions of pixel values in HSV space obtained from idealized color classifications, which were generated by manually masking in some images areas with objects of interest and mapping their pixel values to HSV space.
4. Finally, a few special examples were selected to illustrate some situations, whose effects are insufficiently described by the above statistics. Even though these effects may be marginal problems or side-issues, they cannot be ignored and, under different circumstances, they may occur more frequently.

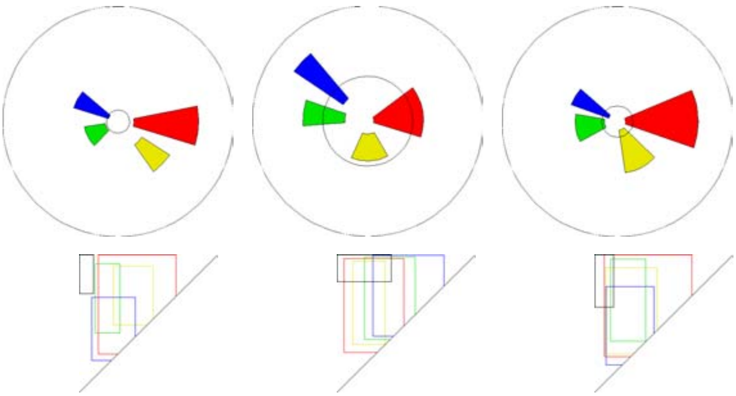

Fig. 4. Results of manually calibrating color classificators for different scenarios. The cuboids for the target color classes are projected onto the hue/saturation plane in the upper row, while the lower row contains projections onto the saturation/brightness half-plane. Figures on the left relate to image set I, figures in the middle to image set II and figures on the right relate to image set III.

5 Experiments and Results

5.1 Comparing Manual Calibration Results for Different Scenarios

Figure 4 illustrates the resulting calibrations for the different scenarios. The conflict resulting from the overlap between the white cuboid and other color classes is resolved with a priority rule, which gives any other target color class priority over white.

The figures show that the color cubes for the midday and lab settings are larger and closer to each other (e.g. the blue and green cone on the rightmost calibration for the lab scenario). This means an increased risk to misclassify pixels from different color classes. The nearer the color descriptions come together, the smaller gets the catch-all class *gray* in between them.

5.2 Variations in Spatial Distribution of Color Classes

For this step, over one hundred images were processed with the appropriately calibrated color classificator for a particular image set. Then, for each color class the frequency and spatial distribution of HSV color values were collected from the original images. Figure 5 illustrates the blobs within the HSV color space resulting from this processing step for the fluorescent light scenario.

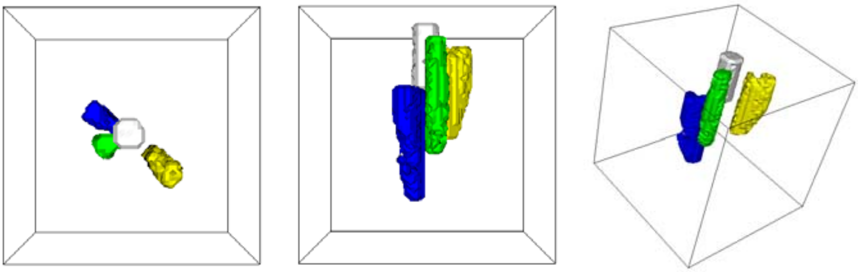

Fig. 5. Three-dimensional plots of the distribution of pixel values in HSV color space for several target color classes, as obtained by color-classifying a single test image taken from image set I. Colors in the figures represent the target color classes (e.g. yellow for the yellow goal). The left image show a top-down view, the middle one the view from one side and the right an isometric version of the same plot. The leftmost plot corresponds with the top-left image in Figure 4, the middle with the bottom-left image in Figure 4. The plot on the right gives a better view of the three-dimensional structure.

For the set of images of each scenario, the standard deviation and the weighted arithmetical mean are calculated for the distributions of pixel values for each color class. Table 1 show these values for the three different scenarios and three different colors.

The standard deviation for the different colors increases the more difficult lighting conditions become. The standard deviation for the hue values remains relatively constant for all illuminations, but the standard deviation values for saturation and even more for brightness increase significantly. Note f.i. the values for the green floor: During the night in Munich under constant fluorescent light, the whole playing field was illuminated quite uniformly. As a result, standard deviation for saturation was very low. The same location during midday shows an increased standard deviation for saturation. The increased value is possibly due to slowly varying illumination over time rather than a non-uniform lighting. Standard deviations increase also for the laboratory setting. Because these images are recorded during a short period, the increased standard deviation is the result of the non-uniform, very directed light, that throws hard shadows and leaves parts of the playing field darker. The weighted mean also strengthens the previous insight. Finally, an effect of the colored windows is that while the hue value for the blue goal remains almost unaffected by daylight influence during midday, the centers of the other colors change significantly.

Table 1. The standard deviation and the arithmetical weighted mean over the three dimensions in the HSV color space (hue, saturation and brightness). Each value is specified for four different (G=green, B=blue, Y=yellow and O=orange) color blobs.

	Hue				Saturation				Brightness			
	G	B	Y	O	G	B	Y	O	G	B	Y	O
Standard Deviation												
Evening	6.3	4.6	3.2	6.4	7.1	10.0	12.4	47.2	22.8	28.4	15.2	59.1
Midday	3.3	2.8	9.8	11.4	17.6	34.9	13.3	25.7	30.6	37.4	26.4	45.1
Laboratory	6.9	6.6	8.3	8.3	16.0	13.7	12.1	46.2	31.7	49.7	22.8	64.2
Weighted Mean												
Evening	151	210	44	4	49	62	112	96	157	116	181	167
Midday	193	218	78	352	94	119	62	64	169	175	160	155
Laboratory	176	206	64	5	67	37	37	99	160	116	85	174

Spatial Distributions of Idealized Color Classifications

To exclude the possibility that the results from the previous section are too strongly influenced by the quality of the manual calibration procedure, we compute spatial distributions of idealized color classifications and compare them with the previous results. For each scenario and each target color class, several images with different views of the object are selected. In each image, the object represented by the target color class is manually masked and assigned to the color class. Then, for each color class the spatial distribution of HSV color values were collected from the original images, just like in the previous section.

Figure 6 illustrates the spatial distributions in the three different setups for the target color classes green (on top-left, representing the carpet on the floor), blue (top-right, representing the blue goal), and yellow (bottom, representing the yellow goal). The three-dimensional structure is visualized using an isometric perspective. Note, that the colors used in these images do not represent target color classes, but different lighting situations. The green blob describes the object under fluorescent light, the blue cluster the same object under midday natural light conditions, and the red blob the same object under directed daylight in our laboratory. The different level of transparency in the plots describe different iso-levels of the density distribution, with a convex hull drawn over all points with the same histogram value (i.e. the opaque part has a higher density then the transparent part, but are the same cluster from the same color).

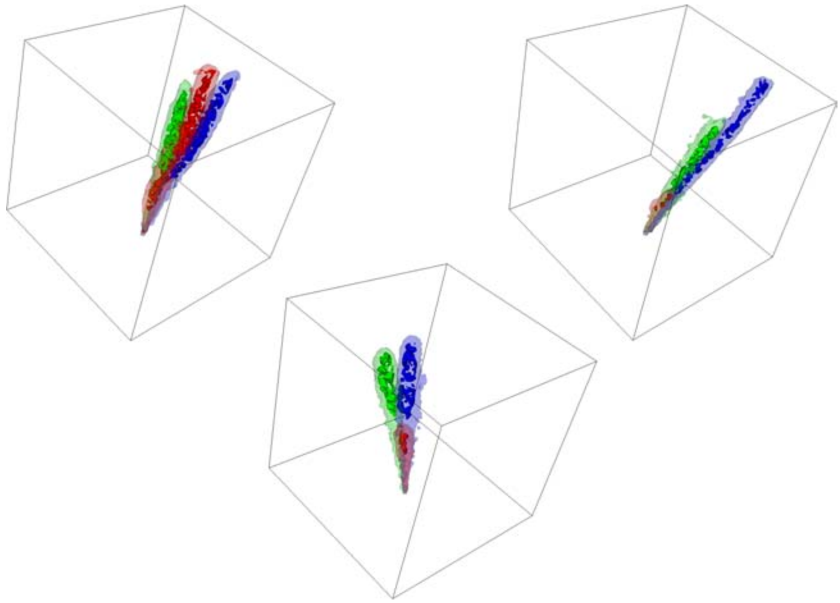

Fig. 6. Three-dimensional plots of the arrangement of the same object respectively under different illuminations within the HSV color space. The first image describe the green color, the top-right plot the blue goal and the bottom image the colors of the yellow goal.

The plots in Figure 6 seem to confirm the values from Table 1. For example, in the first image in Figure 6, displaying the spatial distribution of the green floor color, the opaque part of the red distribution (describing image set III) is spatially much larger in the brightness dimension then the others. This illustrates nicely the extremely varying illumination in this case. Note also, that for the yellow target color displayed in the bottom image, the red blob remains constrained on very low brightness values. This is explained in more detail in Section 5.2.

The same statistics – standard deviations of the different color channels within the HSV color space distribution and weighted mean – are calculated for the idealized color classifications and displayed in Table 2. Because the yellow and blue corner posts were also masked out of the images for the idealized color classification, the values for these two colors decrease slightly. The brightness value for the yellow goal in the lab setting remains on a very low level. The hue value for the blue goal is virtually the same, whereas for the other colors it changes significantly. The top-right plot in Figure 6 illustrates the large stan-

Table 2. The standard deviation and the arithmetical weighted mean over the three dimensions in the HSV color space (hue, saturation and brightness). Each value is specified for four different (G=green, B=blue, Y=yellow and O=orange) color blobs.

	Hue				Saturation				Brightness			
	G	B	Y	O	G	B	Y	O	G	B	Y	O
Standard Deviation												
Evening	6.5	3.9	2.9	3.6	8.3	18.0	12.0	29.5	24.5	26.4	16.4	42.3
Midday	3.6	3.3	6.7	28.6	19.1	40.0	12.1	26.4	32.0	49.7	28.7	50.0
Lab	10.7	13.1	8.4	25.4	18.9	7.9	7.4	59.3	41.9	10.9	9.8	85.1
Weighted Mean												
Evening	151	211	44	7	49	72	115	148	152	124	183	221
Midday	194	218	74	348	92	135	68	63	162	182	166	151
Laboratory	174	208	65	2	60	25	36	92	149	58	82	173

dard deviation for the blue goals during the games under natural light (the blue cluster).

A remarkable congruence of the data obtained from idealized color classifications (Table 2) with those obtained from a manually calibrated color classificator (Table 1) can be observed. This is a strong indication that the effects described in Section 5.2 are not significantly biased by our choice of color classification method and the calibration procedure.

Some Specific Examples

Finally, a few specific examples are presented which may serve as worst-case scenarios for certain aspects. The first example illustrates the problem of distance-dependent variations of color recognition. In Figure 7 two views on the yellow goal (within our laboratory) and the appropriately segmented images are shown. The yellow goal can only be detected in the second image, where the robot is closer to the goal. However, this problem is not due to a badly adjusted calibration of the color classificator. In the right image, parts of the background are already getting classified as blue, which indicates that the calibration is at the limit for the darker boundaries. Note that the yellow part of the corner post, which is lightened directly by the daylight through the window, can be detected in both images, while the yellow goal, whose side panel shadows the back part of the goal, cannot.

Another example illustrates the influence of the colored windows during our friendly game. Figure 8 shows two images with exactly the same robot with identical cyan color markers in front of the same white border. The left image is taken from image set I, the right image from image set II. The magnified details illustrate, how dramatically the color values are changing and especially the distance between the cyan and white values decreases. Under fluorescent light, the (spatial) arithmetical distance between the two magnified RGB values (255,255,255) and (199,255,255) is 56 (assuming all values are within a range of [0..255]). Under natural daylight conditions with dyed windows, the distance

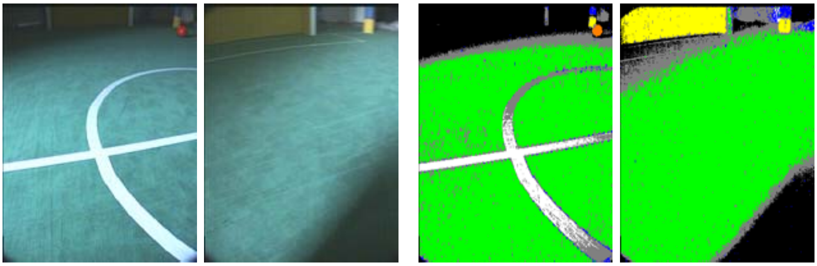

Fig. 7. Examples on how the same color can be detected differently dependent on the distance. The left two images shows the original recording, the right the color-segmented image.

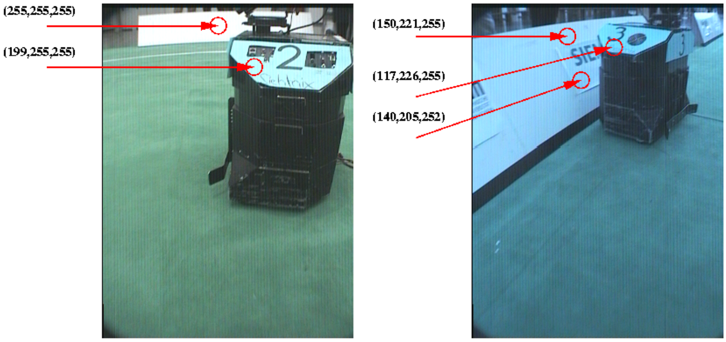

Fig. 8. Examples on how colored panes can influence the color. The left image shows the robot with color marker in front of a white border under fluorescent light, the right image the same robot in front of the same border, this time under daylight conditions.

between the upper detail (150,221,255) and the color marker detail (117,226,255) decreases to 33 or even to 31 between the color marker and the lower wall-detail. As an immediate consequence, we were not able to create a manual calibration, that allowed to separate the color markers reliably from other objects.

The last example illustrates the effects of quick changes in natural daylight illumination within a short time period. The images in Figure 9 were recorded within 14 minutes. The leftmost image looks much like the example in Figure 2 for the natural daylight scenario. Note the bluish coloring of the floor, caused by dyed windows. The second image shows almost the same view on the field a short time later on, when the sun was briefly covered by clouds, and the field was more or less illuminated by the additional fluorescent light. The RGB values for the green color range within each image illustrates this effect nicely: In the first image the weighted mean for all green values within the RGB color space is (81,190,168), whereas the weighted mean for the green color in the second image is (88,149,155). Note the strong shift for the green (middle) value. During such situations, a fixed calibration of the color classificator causes classification quality to deter significantly, as the right images of Figure 9 demonstrates. The second image from right segments the floor almost perfectly, assigning green color to most parts of the carpet, whereas the rightmost image shows a significant increase of the catch-all color class gray within the floor. Remember, that this effect happens in our case during only 14 minutes – and could happen in even shorter intervals, so that it may occur within a single half of a match, with no opportunity of team coaches to perform any kind of manual or manually initiated re-calibration.

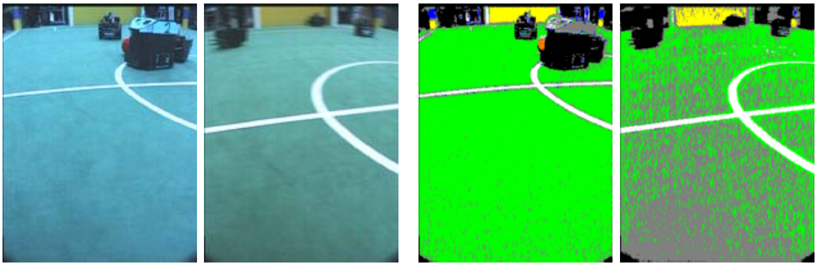

Fig. 9. Example on how the illumination may change within a short time. The left two images shows the original recording, the right the segmented images.

6 Conclusions

From the experiments and statistical analysis performed we can draw the following conclusions:

- Strong variations in field illumination have a strong effect on the spatial distribution of target color classes. Between certain situations differences are so strong that the spatial distribution of certain color classes do not even overlap any more (see Figure 4).
- Due to these strong variations of the target color classes in color space it will not be possible to obtain (neither manually nor by some automated procedure) a color classificator performing well under strongly varying lighting conditions. Making color class cuboids large enough to account for lighting variations causes too much noise and too many misclassifications for any particular, locally stable lighting setup.
- Because of most teams still being largely dependent on working color-based vision methods, introducing natural light to the field setup is likely to have dramatic effects on game performance.

We suggest to undertake the following steps in order to obtain robot vision systems which are more robust against lighting variations:

- As lighting will always be different on different playgrounds, teams should develop and use automated on-site vision calibration routines (see e.g. [1]), which the robot can perform autonomously when put into a new field.
- Another possible path to improve color-based image processing in RoboCup is to enhance and apply algorithms for improving color constancy. These

algorithms try to "stabilize" color recognition, so that the same colors look almost the same under different illuminations. Exemplary work in this area was performed by Forsyth [2, 3], or by Jobson and Rahman [4, 5].
- A third thread of work necessary for RoboCup is on using additional visual features aside of color for object detection and tracking, and possibly integrating them with color-based methods. The realtime processing requirements implied by the RoboCup setup pose a particular challenge, especially for many already existing methods in this area. A notable example for work in this area which has already been proven in several RoboCup competitions is by Hanek et. al. [6, 7]. They use deformable models (snakes), which are fitted to known objects within the images by an iterative refining process based on local image statistics.

New results and methods in any of these areas would be of interest to everyone working on sophisticated applications in robotics, e.g. service robotics in normal, naturally lighted habitations or any variant of outdoor robotics. The RoboCup middle-size league should foster research in this area by defining appropriate technical challenges involving variations in lighting conditions.

Acknowledgment

The work described in this paper was done in project *Adaptivity and Learning in Teams of Cooperating Mobile Robots* supported by DFG SPP-1125. We thank Günther Palm and the University of Ulm for supporting our robot team, and the Agilo team from TU Munich (Michael Beetz, Sebastian Buck, Robert Hanek, Thorsten Schmitt and Freek Stulp) for inviting us to the adventure of holding a friendly game under daylight conditions and the pleasant teamwork we enjoyed with them.

References

1. Mayer, G., Utz, H., Kraetzschmar, G.K.: Towards autonomous vision self-calibration for soccer robots. Proceeding of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS-2002) **1** (2002) 214–219
2. Forsyth, D.: A novel approach for color constancy. International Journal of Computer Vision **5** (1990) 5–36
3. Forsyth, D.: Sampling, resampling and colour constancy. In: Proceedings Computer Vision and Pattern Recognition. (1999) 300–305
4. Rahman, Z., Jobson, D.J., Woodell, G.A.: Multiscale retinex for color image enhancement. In: International Conference on Image Processing. (1996)
5. Jobson, D.J., Rahman, Z., Woodell, G.A.: Properties and performance of a center/surround retinex. (1997)
6. Hanek, R.: The contracting curve density algorithm and its application to model-based image segmentation. IEEE Conference Computer Vision and Pattern Recognition **1** (2001) 797–804
7. Hanek, R., Schmitt, T., Buck, S., Beetz, M.: Towards robocup without color labeling. RoboCup International Symposium 2002 (2002)