A color-coded real-time vision system for a NAO humanoid robot

Alina Liliana Trifan
alina.trifan@ua.pt
António J.R. Neves
an@ua.pt
Manuel B. Cunha
mbc@det.ua.pt
Nuno Lau
nunolau@ua.pt

Transverse Activity on Intelligent Robotics
IEETA/DETI, University of Aveiro
3810-193 Aveiro

Abstract

For both humans and robots vision is a very important sense that has the task of interpreting spatial data, indexed by more than one dimension. For a humanoid robot, a robust vision system should be able to provide accurate information about the environment and a precise description of the objects of interest. This paper presents an implementation of a vision system for a humanoid robot designed to perform in color-coded environments. From acquiring images, processing them and detecting the objects of interest based on color classification, all the algorithms have been tested on the NAO soccer playing humanoid robot. For this robot the world is simplified to a number of colors that are meaningful in the mentioned context. The vision system that we propose can perform in real time and it has proven its practical efficiency. Moreover, we present an innovative algorithm for self-calibration of the most important intrinsic parameters of the camera, as well as two external applications developed for debugging and color calibration. One of the most important features of the vision system that we propose is its modularity, which allows it to be used on a wide range of robotic platforms, with a minimum number of changes.

1 Introduction

Humanoid robotics is the branch of robotics that focuses on developing robots that have an overall appearance similar to the human body and can perform tasks that until now were strictly designated for humans.

Probably the most important sense for a humanoid robot is vision. Just like in the case of humans, the main way for a robot to understand the world with all the objects that are surrounding him is by means of vision. The performance of a humanoid vision system is strongly influenced not just by the hardware architecture of the robot but mostly by its body movements.

In this paper we provide a concise description of a real-time modular vision system based on color classification for a humanoid robot. We start by presenting an overview about the system, outlining its modularity which makes it intuitively easy for being exported to various humanoid platforms. Then we propose an algorithm for self-calibration of the intrinsic parameters of the camera based on the histogram of intensities of the acquired images and a white area, known in advance. For the color segmentation algorithm a lookup table and horizontal or vertical scan lines are used. Finally, we present some validation approaches for a good recognition of the objects of interest by the robot.

The vision system that we present in this paper has been tested and used with the NAO robot, currently used as the standard robotic platform in the RoboCup Standard Platform League [2], [5]. In this environment, the robots play soccer on a green field with white lines, with an orange ball and they defend yellow and blue goals. Thus, the representation of the surrounding world of such robot is slightly simplified with base on color codes.

2 System overview

The architecture of the vision system that we propose can be divided into three main parts: access to the device and image acquisition, calibration of the camera parameters and object detection and classification. Moreover, apart from these modules, two applications have also been developed for calibrating the colors of interest (NaoCalib) and for debugging purposes (NaoViewer). These two applications run on an external computer and communicate with the robot by means of a TCP module of the type client-server that we have developed. Figure 1 shows the proposed modular architecture.

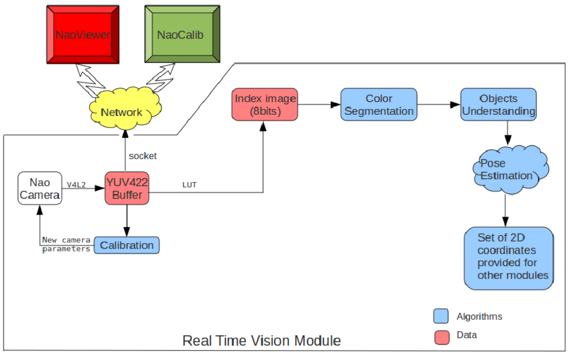

Figure 1: Block diagram of the proposed vision system.

The camera is accessed using V4L2 API, a kernel interface for analog radio and video capture and output drivers. The calibration module is executed whenever the environment or the lighting conditions change, having the purpose of setting the parameters of the camera so that the images acquired give the best possible representation of the surrounding world. Details of the algorithm for self-calibration of the camera are presented in Section 3.

For the detection process, with the use of a look-up table, and by means of the OpenCV library, the raw buffer can be converted into an 8-bit grayscale image in which only the colors of interest are mapped using a one color to one bit relationship (orange, green, white, yellow, blue, pink and blue sky, while gray stands for no color). The next step is the search for the colors of interest in the grayscale image, which we call an index image, using vertical or horizontal scan lines, followed by the formation of blobs from pixels that have the same color. The blobs are then marked as objects if they pass the validation criteria which are constructed based on different measurements of the blobs (bounding box, area, center of mass of the blob). The color segmentation and object detections are detailed in Section 4.

Having the possibility of running the vision module as a server, the two applications that we have developed, NaoCalib and NaoViewer can act as clients that can receive, display and manipulate the data coming from the robot. Thus, NaoViewer is a graphical application that allows the display both of the original image as well as the corresponding index image containing the validation marks for each object of interest that was found. This application was essential in terms of understanding what the robot “sees” since most humanoid robots have small processing capabilities that do not support the use of any graphical interface that might be used for the display and manipulation of images. NaoCalib is a very helpful application that we developed for the calibration of the colors of interest and it is presented in more details in Subsection 3.2.

3 Calibration of the vision system

The accuracy of the representation of the colors in an image captured by the camera of the robot is related to the intrinsic parameters of the camera such as: brightness, saturation, gain, contrast or white balance. By controlling these parameters relatively to the illumination of the environment we can acquire images that accurately represent the surrounding world [4].

3.1 Self-calibration of the camera intrinsic parameters

The algorithm uses the histogram of intensities of the acquired images for calculating some statistic measurements of the images which are then used for compensating the values of the gain and exposure by means of a PI controller. Moreover, a white area, whose location in the image is known in advance, is used for calibrating the white balance. The human intervention is only needed for positioning a white object in the predefined area. The algorithm only needs an average number of 20 frames to converge and the processing time of each frame is approximately 300ms.

From the gray level histogram the Mean Sample Value (MSV) can be computed based on the following formula: $MSV = \frac{\Sigma_{j=0}^{4}(j+1)x_j}{\Sigma_{j=0}^{4}x_j}$, where x_j is the sum of the gray values in region j of the histogram. The MSV represents a useful measure of the balance of the tonal distribution in the image. . The histogram is divided into five regions. The image is considered to have the best quality when the MSV ≈ 2.5. The values for the gain and exposure are compensated with the help of the PI controller until the value of the MSV for the images acquired is ≈ 2.5.

For the calibration of the white balance, the algorithm that we are proposing assumes that the white area should appear white in the acquired image. In the YUV color space, this means that the average value of U and V should be close to 127 when both components are coded with 8 bits. If the white-balance is not correctly configured, these values are different from 127 and the image does not have the correct colors. The white-balance parameter is composed by two values, blue chroma and red chroma, directly related to the values of U and V.

3.2 Calibration of the colors of interest

Along with the calibration of the parameters of the camera, a calibration of the color range associated to each color class has to be performed whenever the environment or the illumination conditions change. These two processes are co-dependent and crucial for image segmentation and object detection.

NaoCalib is an application used for the manual calibration of the colors of interest and it allows the creation of a configuration file that contains the Hue, Saturation and Value minimum and maximum values of the colors of interest [1]. The configuration file is a binary file that apart from the H, S and V maximum and minimum value also contains the current values of the intrinsic parameters of the camera. It is then exported to the robot and loaded when the vision module starts.

4 Object detection

In this application color is a strong hint for the detection and validation of an object of interest. However, it is not sufficient. This section presents our approach for the detection and validation of the objects of interest, based on color segmentation followed by blob formation and measurements computations for the validation of the blobs.

The results obtained both in terms of processing times and accuracy percentages prove the reliability of our system. The vision system that we propose can be used in real time, with the camera working at 30fps since the total processing time spent by the image processing algorithms is 28ms. This processing time was obtained using the hardware architecture of the NAO robot, which is based on a single-core 500MhZ CPU and 256 RAM. Out of this, the most time-consuming task is the one of converting the YUV image to an index one. Moreover, the percentage of accurate detections of all the objects of interest when having a static robot is of 99% and 33% when the robot is performing some kind of locomotion.

4.1 Look-up table and the image of labels

Color classes are defined with the use of a look-up table (LUT) for fast color classification. A LUT represents a data structure, in this case an array used for replacing a runtime computation with a basic array indexing operation. This approach has been chosen in order to save significant processing time. The image acquired in the YUV format is converted to an index image (image of labels) using an appropriate LUT.

The table consists of 16,777,216 entries (2^{24}, 8 bits for Y, 8 bits for U and 8 bits for V). Each bit expresses whether one of the colors of interest (white, green, blue, yellow, orange, red, blue sky, gray - no color) is within the corresponding class or not. A given color can be assigned to multiple classes at the same time. For classifying a pixel, first the value of the color of the pixel is read and then used as an index into the table. The 8-bit value read from the table is called the "color mask" of the pixel.

4.2 Color segmentation and blob formation

Having the colors of interest labeled, scan lines are used for detecting transitions between colors [3]. For the detection of goals and field lines vertical search lines are used, while for the detection of the ball both vertical and horizontal scan lines can be used. Both for vertical and horizontal scan lines, it is possible to configure the number of scan lines to be used in order to reduce the processing time. Both types of scan lines start in the upper left corner of the image and go along the width and the height, respectively, of the image. For each scan line, the run length information of each color of interest is saved and further used for blob formation. The scan lines are then validated and can be further used for the blob formation.

Blobs of certain colors of interest are formed based on the run-length information of parallel scan lines. Having the blobs formed, several validation criteria are applied in the case of the orange ball and of the blue or yellow goals, respectively. In order to be considered a yellow goal, a yellow blob has to have the size larger than a predefined number of pixels. In the situation in which the robot sees both posts of the goals, the middle point of the distance between the two posts is marked as the point of interest for the robot. In the case when just one of the posts is seen, its mass center is marked. For the validation of the ball, the areas of the orange blobs are calculated and the blob validated as being the ball will be the one that has the area over a predefined minimum value and is closest to the robot.

5 Conclusions and Future Work

This paper presents a real-time reliable vision system for a humanoid robot. From calibrating the intrinsic parameters of the camera, to color classification and object detection the results presented prove the efficiency of our vision system. The main advantages of our approach is its modularity, which allows it to be used with a large number of different humanoid robots and the real-time capabilities allow us to use the camera at 30fps even with a low processor as the one used in the NAO robot.

Future developments of our work include more validation criteria for the ball detection based on circular histograms and classifiers training which are more generic and are not color dependent. Also the algorithm for finding transitions from one color of interest to another will be further applied for the detection of the team markers.

References

[1] P. M. R. Caleiro, A. J. R. Neves, and A. J. Pinho. Color-spaces and color segmentation for real-time object recognition in robotic applications. *Revista do DETUA*, 4(8):940–945, June 2007.

[2] D. Gouaillier, V. Hugel, P. Blazevic, C. Kilner, J. Monceaux, P. Lafourcade, B. Marnier, J. Serre, and B. Maisonnier. The NAO humanoid: a combination of performance and affordability. *ArXiv e-prints*, July 2008.

[3] A. J. R. Neves, D. A. Martins, and A. J. Pinho. A hybrid vision system for soccer robots using radial search lines. In *Proc. of the 8th Conference on Autonomous Robot Systems and Competitions, Portuguese Robotics Open - ROBOTICA'2008*, pages 51–55, Aveiro, Portugal, April 2008.

[4] A. J. R. Neves, A. J. Pinho B. Cunha, and I. Pinheiro. Autonomous configuration of parameters in robotic digital cameras. In *Pattern Recognition and Image Analysis*, volume 5524 of *Lecture Notes in Computer Science*, pages 80–87. Springer, June 2009.

[5] A.J.R Neves, N. Lau, L.P. Reis, and A. Moreira. Portuguese Team Team Description, 2011.