Robust Autonomous Structure-based Color Learning on a Mobile Robot

Mohan Sridharan
Electrical and Computer Engineering
University of Texas at Austin

Motivation

The scientific community is working towards the creation of fully autonomous mobile robots capable of interacting with the proverbial *real world*. Autonomous robots are predominantly dependent on sensory information but the ability to accurately sense the complex world is still missing. Visual input, in the form of color images from a camera, should be an excellent and rich source of such information, considering the significant amount of progress made in machine vision. But color, and images in general, have been used sparingly on mobile robots, where people have mostly focussed their attention on other sensors such as sonar and laser. There are three main reasons for this reliance on other relatively low-fidelity sensors. First, most sophisticated vision algorithms require substantial amount of computational and/or memory resources making them infeasible to use on mobile robotic systems that typically have constrained memory and computational resources but demand real-time processing. Second, most vision algorithms assume a stationary or slowly moving camera and hence cannot account for the rapid non-linear camera motion that is characteristic of mobile robot domains. Third, the variation of illumination over the operating environment causes a nonlinear shift in color distributions that is difficult to model; mobile robots, while moving around the world, often go into places with changing illumination.

Figure 1: An Image of the Aibo and the field.

Even in the face of these formidable challenges, we aim to use the visual information in a robust manner. One factor that can be leveraged is that many mobile robot applications involve a *structured* environment with objects of unique shape(s) and color(s) – information that can be exploited to help overcome the problems mentioned above. In this thesis, the focus is on designing algorithms that can work within the robot's computational and environmental constraints while making the best use of the available information. We aim to provide a mobile robot vision system that autonomously *learns* colors from *structure* and adapts to illumination changes.

I shall first describe the work done to date towards achieving that goal, followed by the tasks I hope to complete as part of the thesis. All algorithms are fully implemented and tested on SONY Aibo robots (Sony 2004). One test domain that these robots are used in is the RoboCup initiative where teams of robots play a competitive game of soccer on an indoor field (see Figure 1), though the ultimate goal is to play a game of soccer against a human team by the year 2050. The camera is the primary sensor and all processing for vision, localization, motion planning and action-selection has to be performed on-board the robot.

Summary - Completed Work

Here, I provide a brief summary of the work completed to date. Please refer to (Sridharan & Stone 2004; 2005b; 2005a) for complete details and the associated references.

Baseline Implementation

First, I developed a prototype vision system that works on-board an autonomous robot and addresses two main vision challenges — color segmentation and object recognition, within the robot's real-time constraints. The system is robust to the presence of jerky non-linear camera motion and noisy images.

The baseline vision system begins with color segmentation: assigning color labels to image pixels. We represent this mapping as a *color map/color cube* which is created via an off-board training process by hand-labeling a set of images captured using the robot's camera, such that the robot learns the range of $\{Y, Cb, Cr\}$ values that map to each desired color. During the on-board phase, this color map is used to map each pixel in the raw YCbCr input image into a color class label (m_i), one of nine different colors ($i \in [0, 8]$) in our domain.[1] As the pixels in the image are segmented, they are organized into run-lengths represented as the start point and length in pixels of a contiguous color strip. The run-lengths of the same color that are within a threshold distance are merged together to result in regions, each associated with a rectangular boundary (*bounding box*) that stores the region's properties. These regions are then used to deter-

[1]pink, yellow, blue, orange, red, darkblue, white, green, black.

mine the useful color-coded objects in the robot's environment using heuristic constraints. An efficient algorithm was incorporated to find the lines in the field, an important input to localization. Complete details on the innovations in the color segmentation phase using different color spaces, and the heuristic constraints used for forming regions and recognizing objects can be found in (Sridharan & Stone 2005b).

However, the baseline system relies on *manually* labeled training data and operates in *constant* (and reasonably uniform) illumination conditions.

Using Structure to learn colors

To eliminate the time-consuming and brittle manual calibration process, I developed an algorithm that exploits the *structure* inherent in the environment: the robot uses the known model of its world to autonomously learn the desired colors. This drastically reduced the time required for the color calibration process (from an hour or more to ≈ 5 minutes).

The robot learns the desired colors using the environmental structure — known locations and shape(s) of color-coded objects. The color map is learnt on-board the robot with no human intervention. The robot starts from a known location in the field and it is provided with three lists: the list of colors to be learnt (*Colors[]*), a list of corresponding positions that are appropriate to learn those colors (*Pose[]*), and a list of corresponding objects, defined as shapes, that can be used to learn the colors — *Descriptors[]*. The robot moves through the positions, automatically selecting suitable candidate regions at each position and learning the colors' parameters, modeled as 3D Gaussians. The segmentation performance of the learnt color map is comparable to that of the color map generated by manual calibration. This method works at different illumination conditions and is independent of the color labels. If such a scheme were, for example, employed in a robot working in an office environment with similar amount of structure, repainting walls or objects to different colors would not cause a problem. Complete details on the algorithm and experiments can be found in (Sridharan & Stone 2005a).

Robustness to Illumination changes

Next, I attempted to remove the other main limitation of the baseline system, its sensitivity to illumination changes. To provide a certain degree of robustness to illumination changes, I addressed a subset of the illumination invariance problem.

For a discrete set of illumination conditions (three in our case, named *Bright, Dark, Intermediate*), the robot is provided with a color map and image statistics (color space distributions generated offline). In the online phase, the robot periodically samples image statistics and determines the *closest* illumination condition (using a robust comparison measure: KL-divergence) among the training samples and transitions to using the corresponding color map. The training of multiple color maps is greatly simplified using our color learning scheme. The results show that the robot is able to perform its tasks efficiently while periodically checking for and adapting to changing illumination conditions. The results also indicate that with this representation, the robot can perform efficiently even in illumination conditions in between the ones that it has been explicitly trained for. Complete details can be found in (Sridharan & Stone 2004).

Though our approach does seem to require a fair amount of human input, in situations where the environment changes less frequently than the illumination conditions or object colors, it would require significantly less human effort than manually labeling pixel values. In many applications, it is a lot easier (and faster) to generate the world model with the positions of the various objects of interest than hand-labeling several images. The robot could autonomously learn the colors and automatically adjust for variations in operating conditions (illumination) and between operating environments (objects/walls of different colors).

Proposed Work

Here, I propose extensions to the developed algorithms so as to provide an efficient vision system that *learns autonomously* from the environment and performs well in the presence of illumination changes. These extensions are motivated by the drawbacks of the existing system.

First, I shall reduce the human input required for color learning. The robot shall, in addition to learning the colors, plan its motion sequence autonomously based on the knowledge of the locations (and shapes) of the useful objects in the environment. The robot can then learn the colors from any given position on the field. This could be extended such that the robot learns the colors starting from an unknown position in its world.

Second, modeling colors as 3D Gaussians may not hold outside the controlled lab setting. I shall develop a more general representation that enables the robot to learn the desired colors in uncontrolled settings outside the lab.

Third, I shall combine the color learning algorithm with the method that provides robustness to illumination changes. The robot should be able to learn colors based on the environmental structure and should be able to detect changes in illumination based on the drifts in the color distributions. It should be able to adapt for minor illumination changes by re-learning specific colors and tackle major illumination changes by re-learning the entire color map.

Vision on mobile robots is a highly daunting problem that is far from being solved. Ultimately, we aim to develop efficient algorithms that enable a mobile robot to function under completely uncontrolled natural lighting conditions.

References

Sony. 2004. The Sony Aibo robots. `http://www.sonystyle.com`.

Sridharan, M., and Stone, P. 2004. Towards illumination invariance in the legged league. In *The Eighth International RoboCup Symposium*.

Sridharan, M., and Stone, P. 2005a. Autonomous color learning on a mobile robot. In *The Twentieth National Conference on Artificial Intelligence (AAAI)*.

Sridharan, M., and Stone, P. 2005b. Real-time vision on a mobile robot platform. In *The IEEE International Conference on Intelligent Robots and Systems (IROS)*.