Background Estimation Using Shared Gaussians

Henry Atkins, Chris Stauffer & Eric Grimson

Artificial Intelligence Laboratory
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

http://www.ai.mit.edu

The Problem: This work addresses the problem of maintaining an accurate and robust representation of background information for the purposes of foreground segmentation. This is an augmentation of the previous work of Stauffer et al. [2], which utilizes a single mixture of gaussians to represent the background model.

Motivation: Background subtraction remains a necessary step for activity monitoring. An accurate representation of the background is important to clearly detect foreground objects of interest in a scene. We believe that by utilizing a single mixture of gaussians for all the pixels in an scene we can achieve a more robust model of the background. Figure 1 shows that there is significant structure in the pixel values across the scene. It has also been shown that there are regularities in the behavior of similar colors[1]. Every gaussian will be estimated from multiple pixel processes. This should enable each gaussian to adapt better to changes in the lighting and background. Also, gaussians can be reused for moving objects.

Previous Work: Most applications utilize a pixel level background representation [2, 3]. Every pixel is represented independently. None of the systems attempt to make use of the redundancy of information between pixels in the image. For example, pixels on a road may tend to have the same intensity values and distributions. In addition [3] utilizes a "region level" model which occurs after the pixel level processing to avoid problems where the interior of a moving object is considered background. Further, [3] adds a "frame level" to react to global scene changes as when turning on a light. Pfinder[4] utilizes both color and spacial information in a gaussian representation of the foreground human body for motion analysis and interaction but only uses a full-covariance unimodal background model.

Approach: We represent individual pixels with a shared mixture of gaussians allowing a single gaussian to represent multiple pixels. [2, 3] represent each pixel's history independently. With our representation (Figure 2), pixels with similar intensities and histories will be represented by the same gaussians. In Figure 2 the pixels on the side of the building have the same histories, as do the pixels on the grass. However, the pixels in the parking lot potentially have different histories because different cars were parked in the two spaces in the past. From one perspective, this representation is simply a color based segmentation over time. However, this segmentation is not dependent on pixel location. Instead, the segmentation is a side-effect of our representation.

Future Work: Our background model contains an implicit segmentation of the scene. This segmentation may enable more robust background and foreground estimation in the presence of camera motion, sleeping/waking objects[3], and shadows. We will also investigate applications of this model to handling global scene changes (e.g., flipping a light switch).

Research Support: This work has been funded by DARPA.

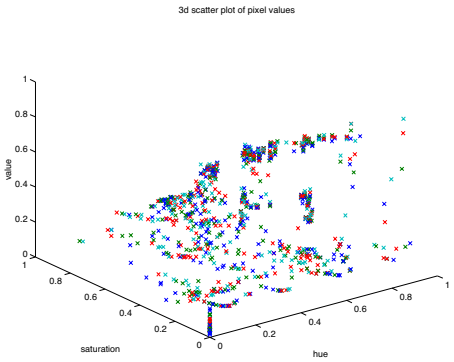

Figure 1: This figure shows a 3d scatterplot of the {h,s,v} pixel values of a portion of the image in Figure 2. It is evident that there is significant structure in the pixel values.

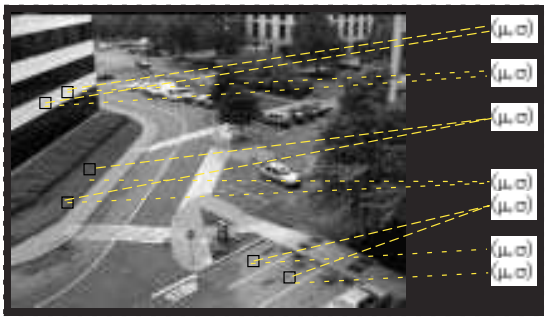

Figure 2: This is a visualization of our representation. Assuming that each of the six cubes is a pixel, each pixel is represented by a set of gaussians as shown by the dashed lines connecting them. Lines with higher densities represent relationships to gaussians with higher weights (more recent support).

References:

[1] E. Miller and K. Tieu. Color eigenflows: Statistical modeling of joint color changes. In *Int. Conf. Comp. Vis.*, 2001.

[2] C. Stauffer and W.E.L. Grimson. Adaptive background mixture models for real-time tracking. In *Proc. Computer Vision and Pattern Recognition*, pages 246–252, 1999.

[3] K. Toyama, J. Krumm, B. Brummit, and B. Meyers. Wallflower: Principles of background maintenance. In *Proc. International Conference on Computer Vision*, Corfu, Greece, 1999.

[4] C. Wren, A. Azerbayejani, T. Darrell, and A. Pentland. Pfinder: Real-time tracking of the human body. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):780–785, July 1997.