CMSC 317: The Computational Image

Assignment 8: Foreground/Background Segmentation

Due April 24th, 2019

In this assignment we will compare techniques for image segmentation, in particular, modeling the background in a video sequence. Each method models the background with increasing sophistication. Once the background model is computed a pixel can be classified as background or foreground using a simple threshold decision rule. For this lab, implement and compare the following methods using grayscale or RGB/HSV images.

- **Baseline** - Take a static image and subtract that from subsequent frames. Then decide a pixel x is background only if the intensity difference is less than T:

$$|x - x_b| < T$$

- **Average** - Compute an average value for each pixel's intensity. A running average can be computed using the following efficient update rule.

$$\hat{\mu} = \frac{1}{t}\sum_{i=1}^{t} x_i, \qquad \hat{\mu}_t = \frac{t-1}{t}\hat{\mu}_{t-1} + \frac{1}{t}x = \hat{\mu}_{t-1} + \frac{1}{t}(x - \hat{\mu}_{t-1}) \qquad |x - \hat{\mu}| < T$$

- **Moving Average** - Compute a moving-average value for each pixel's intensity to model the background (where $0 \leq \alpha \leq 1$).

$$\hat{\mu}_t = (1-\alpha)\hat{\mu}_{t-1} + \alpha x = \hat{\mu}_{t-1} + \alpha(x - \hat{\mu}_{t-1}) \qquad |x - \hat{\mu}| < T$$

- **Gaussian** - Assume each pixel's intensity in the background image can be modeled probabilistically using a Gaussian probability distribution. Once we have this model, we can compute the likelihood of a particular pixel value.[1] You can estimate μ and σ^2 from a set of data x_i using the following equations:

$$\hat{\mu} = \frac{1}{t}\sum_{i=1}^{t} x_i, \qquad \hat{\sigma}^2 = \frac{1}{t-1}\sum_{i=1}^{t}(x_i - \hat{\mu})^2 \qquad p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2}$$

$$\hat{\mu}_t = \hat{\mu}_{t-1} + \frac{x - \hat{\mu}_{t-1}}{t} \qquad S_t = S_{t-1} + (x_t - \hat{\mu}_{t-1})(x_t - \hat{\mu}_t) \qquad \hat{\sigma}_t^2 = \frac{1}{t-1}S_t$$

  This method can used in the same fashion as the previous four thresholding functions, but using the Mahalanobis distance rather than the absolute difference. The Mahalanobis is also known as the z-score in one dimension.

$$r^2 = (\mathbf{x} - \boldsymbol{\mu})^\mathsf{T}\boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}) \qquad r = \frac{|x - \mu|}{\sigma} \qquad \frac{|x - \mu|}{\sigma} < T$$

Deliverable

Via Moodle, your Processing sketch and a PDF file documenting the comparison between the methods. The sketch should be inside a directory named your last name.

[1] Feel free to use a different pixel model: a) multiple independent univariate Gaussian distributions to describe the RGB or YUV of each pixel b) a multivariate Gaussian distribution c) a Gaussian mixture model.