Adaptive background mixture models for real-time tracking

Mohammad Shirazi
Brendan Morris
Improving Background Subtraction

• Adaptive Background Mixture Models for Real Time Tracking
  ▫ Chris stauffer and W.E.L. Grimson
  ▫ Over 4000 citations since 1999
Motivation

• Robust background subtraction should handle lighting changes, repetitive motion from clutter and long term scene changes

• **Standard method of adaptive background (Adv)**
  ▫ It is effective where object moves continuously
  ▫ Background is visible a significant portion of the time
Motivation

- **Standard method of adaptive background (Disadv)**
  - It is not robust to scenes with many moving objects particularly if they move very slowly
    - [video link](http://www.youtube.com/watch?v=YA_lWWhePW8)
  - Recovers slowly when background is uncovered
  - It has single predetermined threshold for the entire scene
Motivation

- Standard method of adaptive background (Disadv)
  - It can not handle bimodal backgrounds
  - Assume rainy situation
  - Pixel intensity values of background for 16 frames:
    - 139, 140, 141, 141, 138, 140, 140, 139, 240, 241, 243, 244, 180, 141, 140, 142
  - Modeling background for each 8 frame with Gaussian distribution
    - $\mu_1 = 139.75$, $\delta_1 = 1.22$
    - $\mu_2 = 196.37$, $\delta_2 = 50.43$
Motivation

- \[139, 140, 141, 141, 138, 140, 140, 139, 240, 241, 243, 244, 180, 141, 140, 142\]

\[\mu_1 = 139.75, \delta_1 = 1.22\]

\[\mu_2 = 196.37, \delta_2 = 50.43\]
Motivation

RG plots of a single pixel

Differing threshold over time

Bimodal distribution over time
Approach

• Modeling the values of a particular pixel as a mixture of Gaussians
• Based on the persistence and the variance of each Gaussians of the mixture, we determine which Gaussians may correspond to background color
• Pixel values that do not fit the background distributions are considered foreground
Gaussian Distribution

- **Univariate**

\[ \mathcal{N}(x | \mu, \sigma^2) = \frac{1}{\sqrt{2\pi \sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \]

- **Multivariate**

\[ \mathcal{N}(x | \mu, \Sigma) = \frac{1}{(2\pi)^{D/2} |\Sigma|^{1/2}} e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)} \]
Online Mixture Model

- History of a pixel is known up to current time $t$
  \[
  \{X_1, ..., X_t\} = \{I(x_o, y_o, i): 1 \leq i \leq t\}
  \]

- Model the history as mixture of $K$ Gaussian Distributions
  \[
  P(X_t) = \sum_{i=1}^{K} w_{i,t} \mathcal{N}(X_t | u_{i,t}, \Sigma_{i,t})
  \]
  $w_{i,t}$ - prior probability (weight) of Gaussians $i$
Mixture Model Example

- For grayscale image with $K=5$
Update the Mixture Model (Stage 1)

- Every new pixel value, $X_{t+1}$, is checked against the existing $K$ Gaussian distributions until a match is found.

- A match is defined as a pixel value within 2.5 standard deviations of a distribution.
Stage 2

- Match

\[
\begin{align*}
\mu_{i,t+1} &= (1 - \rho) \mu_{i,t} + \rho X_{t+1} \\
\sigma^2_{i,t+1} &= (1 - \rho) \sigma^2_{i,t} + \rho (X_{t+1} - \mu_{i,t})^2 \\
\rho &= \alpha \mathcal{N}(X_{t+1} | \mu_{i,t}, \sigma^2_{i,t}) \\
\alpha &- \text{is a learning rate}
\end{align*}
\]

- Prior weights of Gaussians are updated

\[
\begin{align*}
w_{i,t+1} &= (1 - \alpha) w_{i,t} + \alpha (M_{i,t+1}) \\
M_{i,t+1} &= 1 \text{ for matching Gaussian or } M_{i,t+1} = 0 \text{ for all others}
\end{align*}
\]
Stage 3

- **No Match**
  - If none of the $K$ distributions match the current pixel value, the least probable distribution is go out.
  - A new distribution with the current value as its mean value, an initially high variance, and low prior weight, is enter.
  - The mean and variance remain unchanged.
Background Model Estimation

- Gaussians are ordered by the value \( \omega / \sigma \)
- We are interested in the Gaussian distributions which have the most **supporting evidence** = and the **least variance**. Why??

- For “background” distributions when a static, persistent object is visible, leading to high weight and relatively low variance.

- New object occludes the background object creation of a distribution or the increase in the variance of an existing distribution, so the variance of the moving object is expected to remain larger than a background pixel until the moving object stops
Background Estimation Example

- First B distributions are selected as the background

\[ B = \text{argmin}_b(\sum_{i=1}^{b} w_i > T) \]

- As we remember, \( W_i \) is portion of data that is accounted for by \( i \) Gaussian
- \( T \) minimum portion of image expected to be background
Background Estimation Example

- After background estimation red are backgrounds and blacks are foreground
Discussion

• Advantageous
  ▫ Different threshold for each pixel
  ▫ Pixel-wise thresholds adapt over time
  ▫ Objects are allowed to become part of the background without destroying the existing background model
  ▫ Provides fast recovery

• Disadvantages
  ▫ Can not handle sudden, drastic lighting changes
  ▫ Must have good Gaussian initialization (median filtering)
  ▫ There are number of parameters to tune
Summary

• Simple background subtraction approaches such as fame diff, mean, and median filtering are fast
• Constant thresholds make them ill-suited for challenging real-world problems
• Adaptive background mixture model approach can handle challenging situations
• Bimodal backgrounds, long-term scene changes, and repetitive motion
• Improvements include upgrade the approach with temporal information or using region-based techniques
Thank You

• Questions?

My Background subtraction implementation using GMM at OpenCV