# Markovian Framework for Foreground-Background-Shadow Separation of Real World Video Scenes

Csaba Benedek, Tamás Szirányi
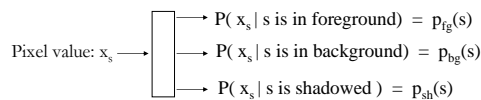benedek@digitus.itk.ppke.hu, sziranyi@sztaki.hu

## 1. Introduction, research goals

- **Foreground detection under complex illumination conditions**
- **Assumptions**
  - Static cameras
  - Background is stationary within a short observation period
- **Properties of the scenes**
  - Dynamic changes of the background objects and the lightning conditions
  - 'Crowded' and 'empty' scenarios alternate
  - Presence of background colored/textured objects
  - Shadow effects
- **Considered features**
  - Pixel level information
  - Neighborhood connectivity

## 2. Segmentation model

- A likelihood classification model for pixel 's'

Pixel value: $x_s$ →

$P(x_s \mid s \text{ is in foreground}) = p_{fg}(s)$

$P(x_s \mid s \text{ is in background}) = p_{bg}(s)$

$P(x_s \mid s \text{ is shadowed}) = p_{sh}(s)$

- Markov random field (Pott model)
  - Video image: 2 dimensional grid of pixels
  - $w_s$: label of pixel s, $w_s \in \{bg, fg, sh\}$
  - Field energy for each segmentation (global labeling):

$$E = -\sum_s \log P(x_s \mid w_s) + \sum_{\substack{\{s,r\} \\ neighbours}} V_{s,r}(w_s, w_r) \qquad V_{s,r}(w_s, w_r) = \begin{cases} -\beta & w_s = w_r \\ +\beta & w_s \neq w_r \end{cases}$$

## 3. Background model

- Stauffer-Grimson algorithm



Time histogram for 's': GMM approximation
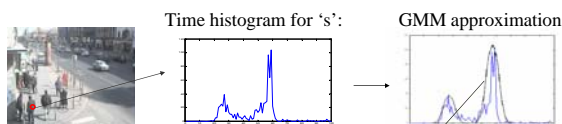
$p_{bg}(s) = \eta(x(s), \mu_{bg}(s), \Sigma_{bg}(s))$ ← Greatest peak: background

## 4. Shadow model

- Classical model: shadow is approximated as a linear transform of background pixel values:

$$\begin{pmatrix} r_{shadow} \\ g_{shadow} \\ b_{shadow} \end{pmatrix} = \begin{bmatrix} a_r & 0 & 0 \\ 0 & a_g & 0 \\ 0 & 0 & a_b \end{bmatrix} \cdot \begin{pmatrix} r_{background} \\ g_{backgrouns} \\ b_{background} \end{pmatrix}$$

  - This model assumes the background surfaces to be homogenous!
- Improved model: darkening ratio is a random variable

$$r_r = \frac{r_{actual}}{r_{background}}, r_g = \frac{g_{actual}}{g_{background}}, r_b = \frac{b_{actual}}{b_{background}} \longrightarrow R_1 = \frac{r_r + r_g + r_b}{3}, R_2 = \frac{r_b}{r_r}, R_3 = \frac{r_b}{r_g}$$
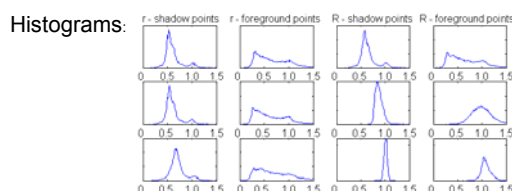
Histograms:



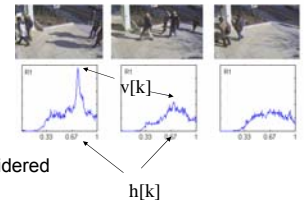R vector: lower correlation between the components

## 5. Shadow parameter adaptation

- Shadow parameters:

$p_{sh}(s) = \eta(R(s), \mu_{sh}, \Sigma_{sh})$

$\Sigma_{sh} = diag\{\sigma_{sh}^2(R_1), \sigma_{sh}^2(R_2), \sigma_{sh}^2(R_3)\}$

$\mu_{sh} = [\mu_{sh}(R_1), \mu_{sh}(R_2), \mu_{sh}(R_3)]$

- $\mu_{sh}(R_1)$ is the mean darkening ratio in gray scale
  - Other parameters can be considered constant in the scene
- Adaptation rule:

$$\mu_{sh}(R_1)^{[k+1]} = \rho \cdot h[k] + (1 - \rho) \cdot \mu_{sh}(R_1)^{[k]} \qquad \rho = \alpha \cdot v[k] \cdot \frac{v[k]}{\overline{v}[k]}$$

Histogram of the $R_1$ values for non-background points in the images:



## 6. Foreground probabilities

- Temporal statistics is not available
  - In the literature *uniform* distribution is used $p_{fg}(s)=u$, which produces low performance in several cases
- Preprocessing step:
  - s is foreground ↔ $p_{sh}(s)<u$ AND $p_{bg}(s)<u$
- Assumption for a given foreground pixel:
  - In the neighborhood there are some correctly classified foreground pixels
  - The color of the pixel matches to the color distribution of the set of the neighboring foreground pixels.
- Approximation the color statistics of the probably foreground pixels in the neighborhood
  - Close-in-color pixels to s are grouped in one weighted Gaussian term
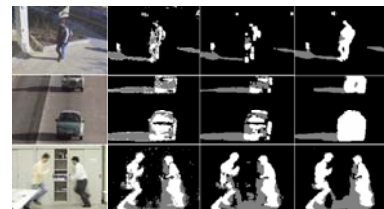


Uniform fg. model

Proposed fg. model

## 7. Results

- Different illumination conditions (surveillance videos)
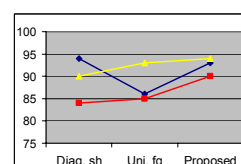


- Segmentation results by the proposed MRF model, compared to the preliminary step, and a morphology-based approach



- Segmentation results compared to the previous MRF models

**Foreground detection rate:** TP/(FN+TP)

**Accuracy rate:** TP/(TP+FP)



University
Highway
Laboratory

Diag. sh — Uni. fg — Proposed