Moving Cast Shadow Detection using Physics-based Features

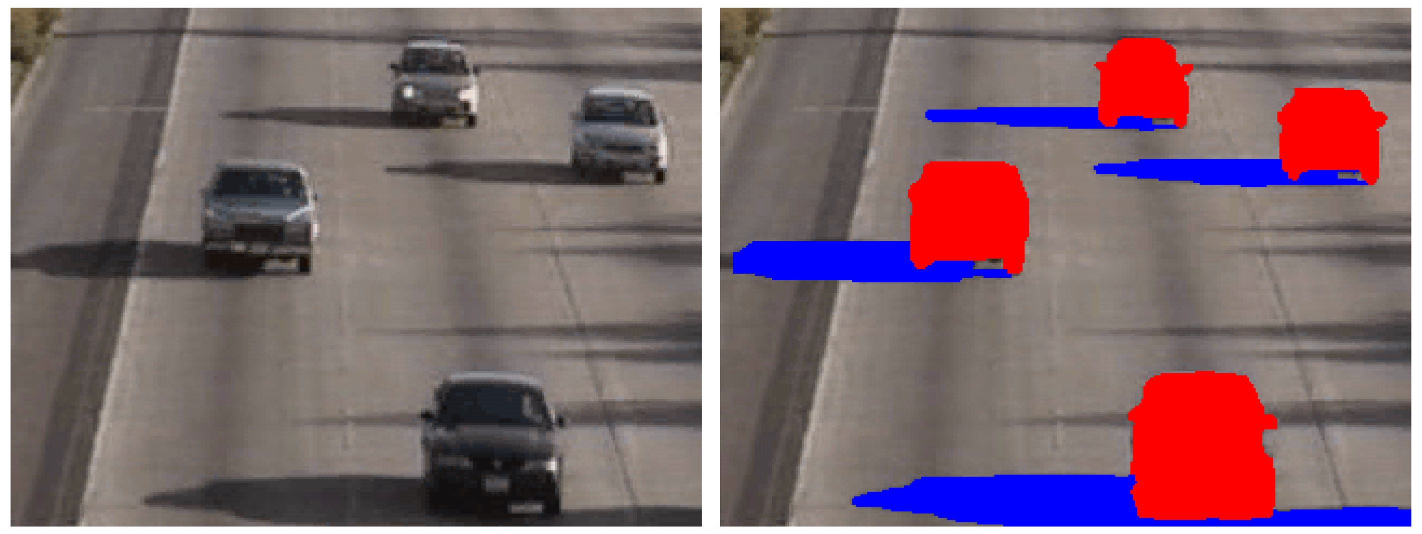
Jia-Bin Huang and Chu-Song Chen

Institute of Information Science, Academia Sinica, Taipei, Taiwan

E-Mail: jbhuang@ieee.org, song@iis.sinica.edu.tw

Goal

- Detecting moving objects in a scene



Input: Video sequence Output: Label fields

Challenges

- Shadows may be misclassified as foreground because:
- 1) they are typically significantly different from background
- 2) they have the same motion as foreground objects
- 3) they are usually attached to foreground objects

Related Works

- **Shadow detection with static parameter settings**
  - Require significant parameter tuning
  - Cannot adapt to environment changes
- **Shadow detection using statistical learning methods**
  - **[1]** Shadow flow (Porikli et al., ICCV 2005)
  - **[2]** Gaussian mixture shadow model (Martel-Brisson et al., CVPR 2005)
  - **[3]** Local and global features (Liu et al., CVPR 2007)
  - **[4]** Physical model of cast shadows (Martel-Brisson et al., CVPR 2008)
- **Major Drawback**:
  - All are pixel-based methods: need numerous foreground activities to learn the parameters.

Contributions

- A global shadow model learned from *physics-based features*.
- Does not require numerous foreground activities or high frame rates to learn the shadow model parameters
- Can be used for fast learning of local features in pixel-based models

Physics-Based Shadow Model

- **Bi-illuminant Dichromatic Reflection Model**
  - Decompose reflected light into four types: body and surface reflection for direct light sources and ambient illumination (Maxwell et al. CVPR 2008)

$$I(\lambda) = c_b(\lambda)[m_b l_d(\lambda) + M_{ab}(\lambda)] + c_s(\lambda)[m_s l_d(\lambda) + M_{as}(\lambda)].$$

  - Consider only matte surfaces, the camera sensor response

$$g_i = \alpha F_i m_b c_b^i l_d^i + F_i c_b^i M_{ab}^i, i \in \{R, G, B\}$$

  defines a line segment in the RGB color space: shadowed pixel to fully lit pixel.

- **Extracting Useful Features**
  - Spectral ratio

$$S_i = \frac{SD^i}{BG^i - SD^i} = \frac{\alpha}{1-\alpha} + \frac{M_{ab}^i}{(1-\alpha) m_b l_d^i}$$

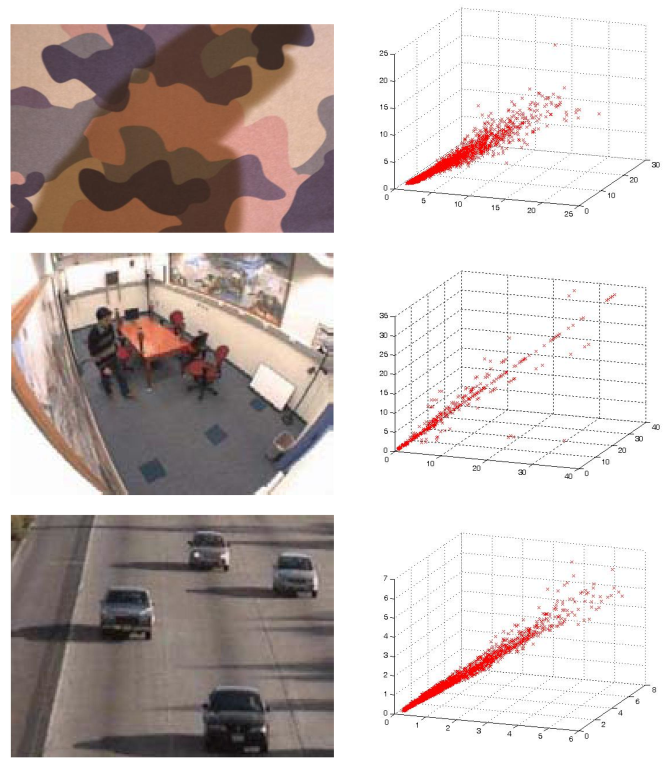
  - Subtract the bias term

$$\gamma_i = S_i - \frac{\alpha}{1-\alpha} = \frac{M_{ab}^i}{(1-\alpha) m_b l_d^i}$$

  - Normalized spectral ratio

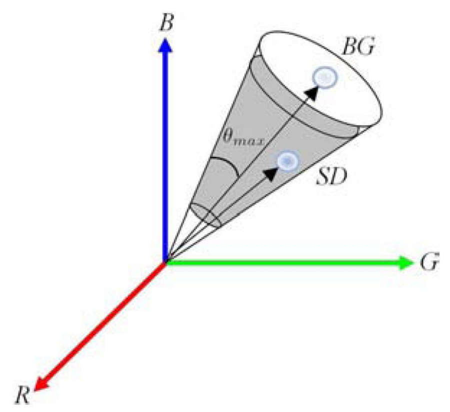
$$\hat{\gamma}_i = \frac{M_{ab}^i}{(1-\alpha) m_b l_d^i}\left(\frac{1}{|\gamma|}\right),$$

  where $|\gamma| = \frac{1}{(1-\alpha)m_b}\sqrt{(\frac{M_{ab}^R}{l_d^R})^2 + (\frac{M_{ab}^G}{l_d^G})^2 + (\frac{M_{ab}^B}{l_d^B})^2}$



Learning Cast Shadows

- **Weak Shadow Detector**
  - Evaluate every moving pixels detected by the background model to filter out impossible ones



- **Global and Local Shadow Model**
  - Learn a global shadow model for the whole scene using Gaussian Mixture Model (GMM) with the Expectation-Maximization (EM) algorithm:

$$p(\hat{r}|\mu, \Sigma) = \sum_{k=1}^{K} \pi_k G_k(\hat{r}, \mu_k, \Sigma_k)$$

  - Cast shadows are expected to form a compact cluster (with large mixing weight and small variance)
  - Build pixel-based GMM to learn local features (gradient density distortion)
- **Confidence-Rated Gaussian Mixture Learning**
  - Adaptive learning rate for mixing weights and Gaussian parameters:

$$\rho_\pi = C(\hat{\gamma}) * \left(\frac{1-\rho_{default}}{\sum_{j=1}^{K} c_j}\right) + \rho_{default}$$

$$\rho_G = C(\hat{\gamma}) * \left(\frac{1-\rho_{default}}{c_k}\right) + \rho_{default}$$

Results

- **Qualitative Results**

(a) (b) (c) (d) (e)

Figure: (a) Input frame. (b) Background posterior probability $P(BG|x_p)$. (c) Confidence map by the global shadow model. (d) The shadow posterior probability $P(SD|x_p)$. (e) The foreground posterior probability $P(FG|x_p)$.

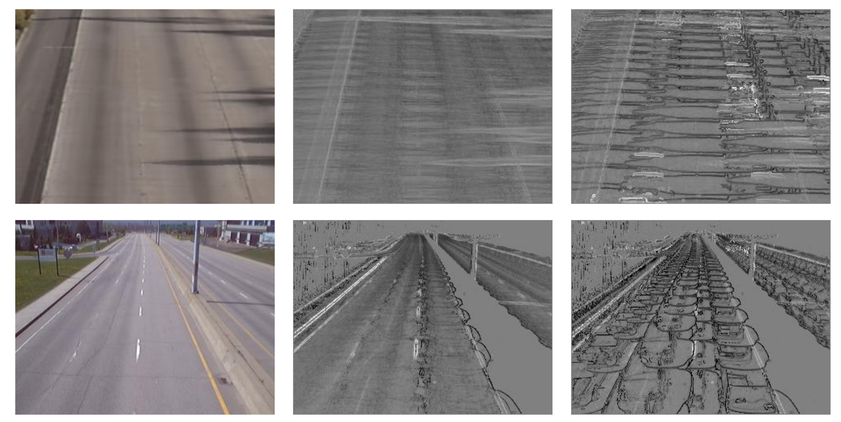
- **Quantitative Results**
  - Evaluation metrics: shadow detection and discrimination rate:

$$\eta = \frac{TP_S}{TP_S + FN_S}; \xi = \frac{TP_F}{TP_F + FN_F}$$

Table: Quantitative results on surveillance sequences

Sequence	Highway I		Highway II		Hallway	
Methods	η%	ξ%	η%	ξ%	η%	ξ%
Proposed	**70.83**	82.37	**76.50**	**74.51**	**82.05**	**90.47**
Kernel [4]	70.50	**84.40**	68.40	71.20	72.40	86.70
LGf [3]	72.10	79.70	-	-	-	-
GMSM [2]	63.30	71.30	58.51	44.40	60.50	87.00

- **Effect of Confidence-Rated Gaussian Mixture Learning**



(a) (b) (c)

Figure: (a)Background image. (b) The mean map with confidence-rated learning. (c) The mean map w/o using confidence-rated learning.

- **Handling Scene with Few Foreground Activities**

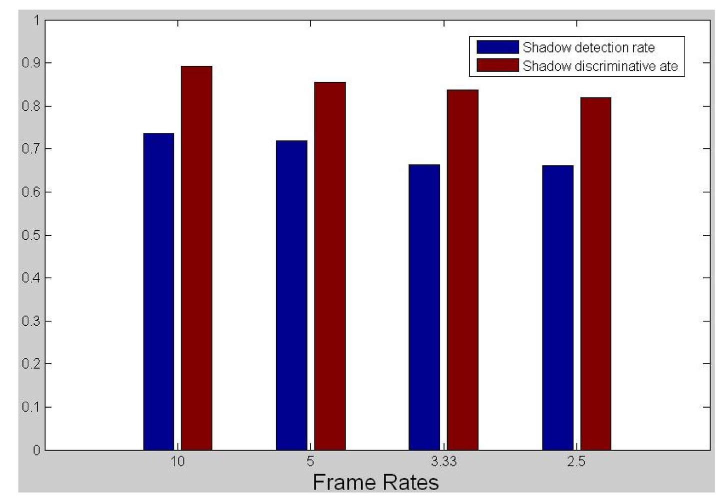


Figure: Quantitative results of the Intelligent Room sequence under different frame rates settings (10, 5, 3.33, and 2.5 frame/sec)