Class 3: Advanced Moving Object Detection and Alert Detection

Feb. 18, 2008

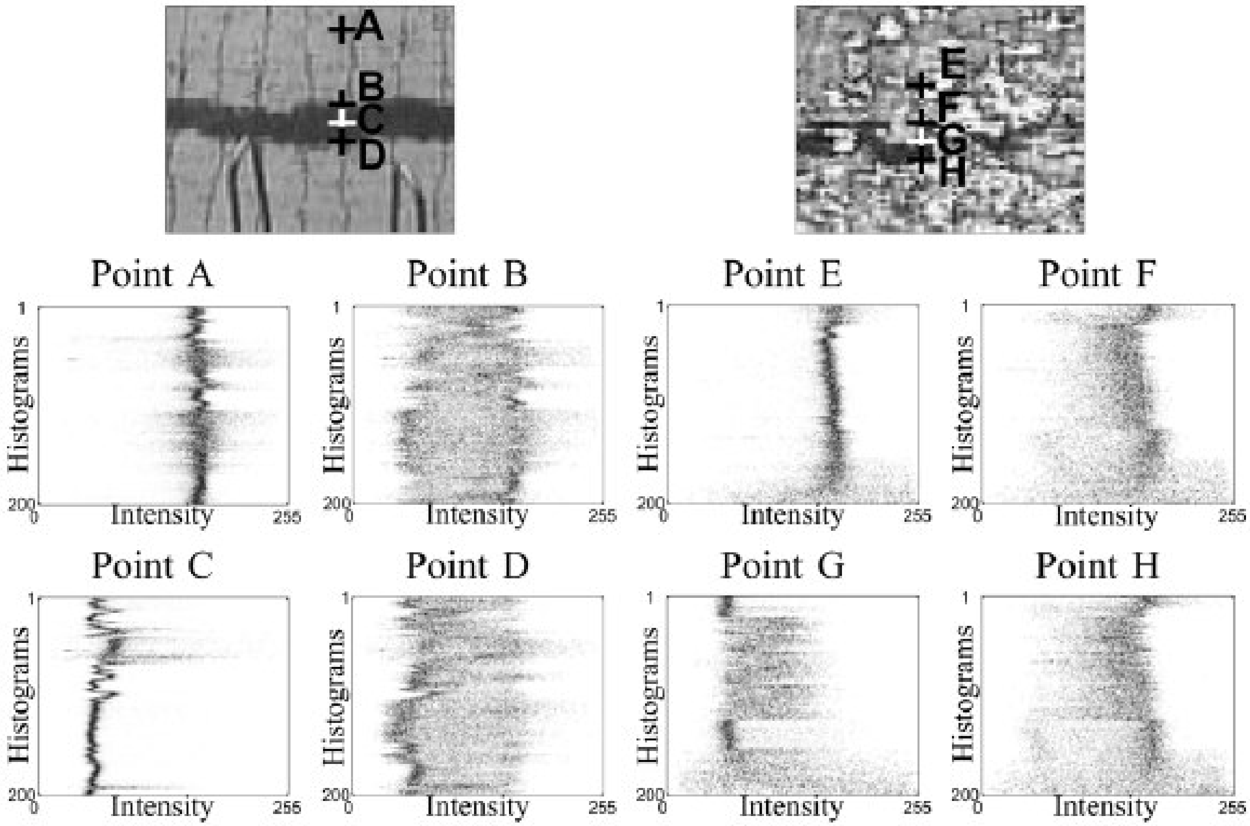
Instructor: YingLi Tian

Video Surveillance E6998-007 Senior/Feris/Tian

Outlines

- Moving Object Detection with Distraction Motions
  - Region-based mixture of Gaussians
  - Statistical framework for BGS
  - Motion-based moving object detection
- Interaction between BGS and Tracking
- Moving Object from Moving Cameras
- Real-time alerts of video surveillance

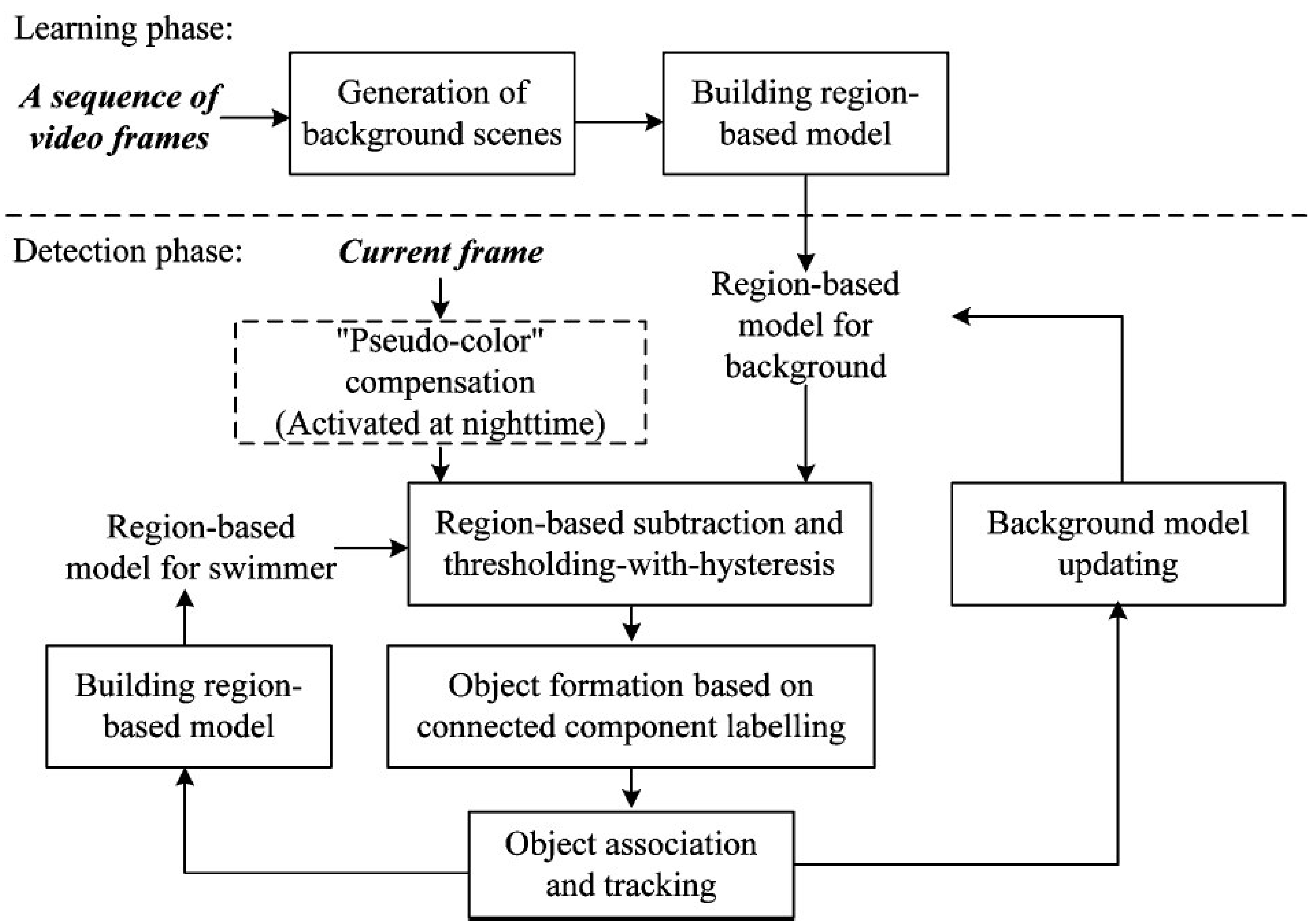
Region-based BGS (Eng et al. 2006) – (1)



Intensity histogram for different points of a typical pool

From Eng et al. 2006

Region-based BGS – (2)



From Eng et al. 2006

Region-based BGS – (3)

- A sequence of N1* N2 background frames,
- Each frame is divided into n1*n2 non-overlapping blocks (s*s)
- Each block, homogeneous background is generated.
- Computer the mean and covariance matrix of a region $\boldsymbol{R}_{a,b}^{k}$

Region-based BGS – (4)

$$\boldsymbol{\mu}_{\boldsymbol{R}^k_{a,b}} = \{\mu^1_{\boldsymbol{R}^k_{a,b}}, \ldots, \mu^d_{\boldsymbol{R}^k_{a,b}}\}\Big|$$

d=3 (dimension of the color space)

$$P(\boldsymbol{x}_{i,j}|\boldsymbol{\mu}_{\boldsymbol{R}^k_{a,b}}, \Sigma_{\boldsymbol{R}^k_{a,b}}) = \frac{1}{(2\pi)^{d/2}|\Sigma_{\boldsymbol{R}^k_{a,b}}|^{1/2}} \times \exp\left\{-\frac{1}{2}(\boldsymbol{x}_{i,j} - \boldsymbol{\mu}_{\boldsymbol{R}^k_{a,b}})\Sigma^{-1}_{\boldsymbol{R}^k_{a,b}}(\boldsymbol{x}_{i,j} - \boldsymbol{\mu}_{\boldsymbol{R}^k_{a,b}})^T\right\}$$

From Eng et al. 2006

Region-based BGS – (5)

Optimization:

$$P(\boldsymbol{x}_{i,j}|\boldsymbol{\mu}_{\boldsymbol{R}^k_{a,b}}, \boldsymbol{\sigma}_{\boldsymbol{R}^k_{a,b}}) = \prod_{m=1}^{d} \frac{1}{\sqrt{2\pi}\sigma^m_{\boldsymbol{R}^k_{a,b}}} \exp\left\{-\frac{(x^m_{i,j} - \mu^m_{\boldsymbol{R}^k_{a,b}})^2}{2(\sigma^m_{\boldsymbol{R}^k_{a,b}})^2}\right\}$$

$$\boldsymbol{\sigma}_{\boldsymbol{R}^k_{a,b}} = \{\sigma^1_{\boldsymbol{R}^k_{a,b}}, \ldots, \sigma^d_{\boldsymbol{R}^k_{a,b}}\}$$

$$D(\boldsymbol{x}_{i,j}|\boldsymbol{\mu}_{\boldsymbol{R}^k_{a,b}}, \boldsymbol{\sigma}_{\boldsymbol{R}^k_{a,b}}) = \sqrt{\frac{\sum_{m=1}^{d} (x^m_{i,j} - \mu^m_{\boldsymbol{R}^k_{a,b}})^2}{(\sigma^m_{\boldsymbol{R}^k_{a,b}})^2}}.$$

From Eng et al. 2006

Region-based BGS – (6)

1. Generating background frames (pixel-based)
   a) temporal vector filter
   b) swimmer skin model
2. Generating initial background model – region-based (S x S)
3. Updating the background models

$$\boldsymbol{\mu}^{t}_{\boldsymbol{R}^{k}_{a,b}} \leftarrow (1-\rho)\boldsymbol{\mu}^{t-1}_{\boldsymbol{R}^{k}_{a,b}} + \rho\boldsymbol{\mu}^{t}_{\boldsymbol{R}^{k}_{a,b}}$$

$$\boldsymbol{\sigma}^{t}_{\boldsymbol{R}^{k}_{a,b}} \leftarrow (1-\rho)\boldsymbol{\sigma}^{t-1}_{\boldsymbol{R}^{k}_{a,b}} + \rho\boldsymbol{\sigma}^{t}_{\boldsymbol{R}^{k}_{a,b}}$$

From Eng et al. 2006

Region-based BGS – (7)

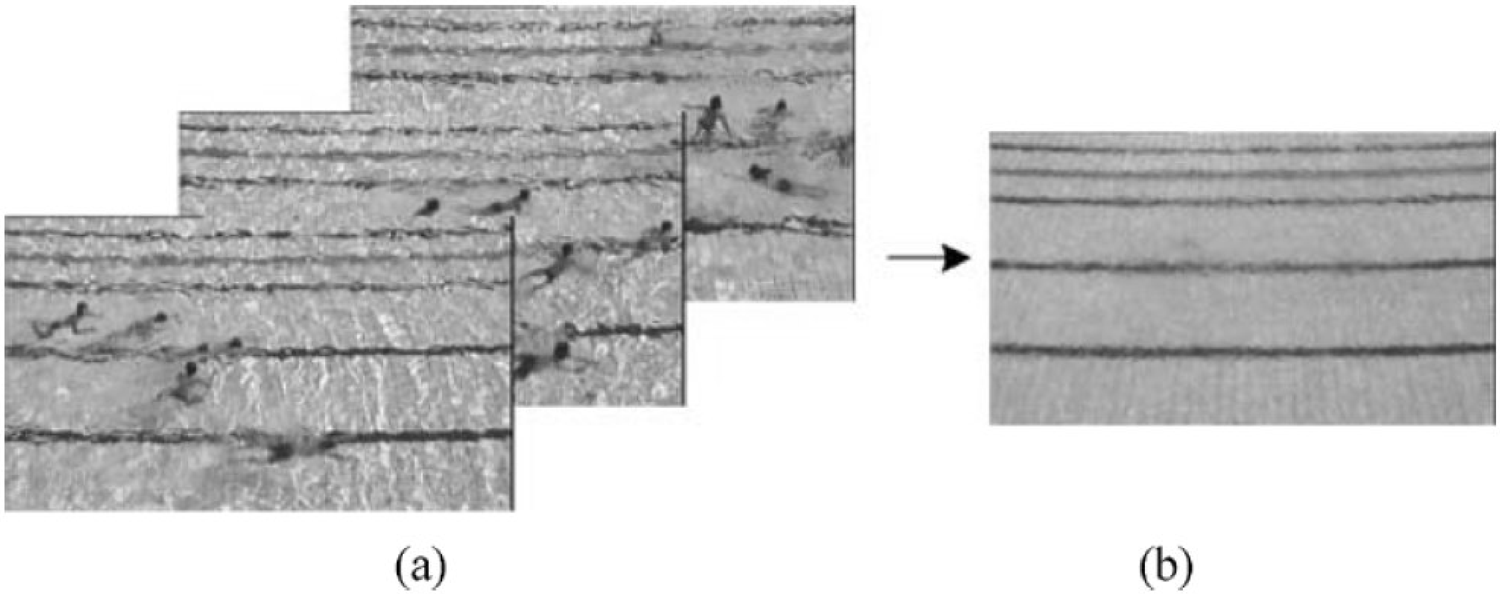


Fig. 4. Generation of a background scene using a temporal vector median filter. (a) A sequence of frames contains foreground swimmers. (b) Generated background scene.

Region-based BGS – (8)

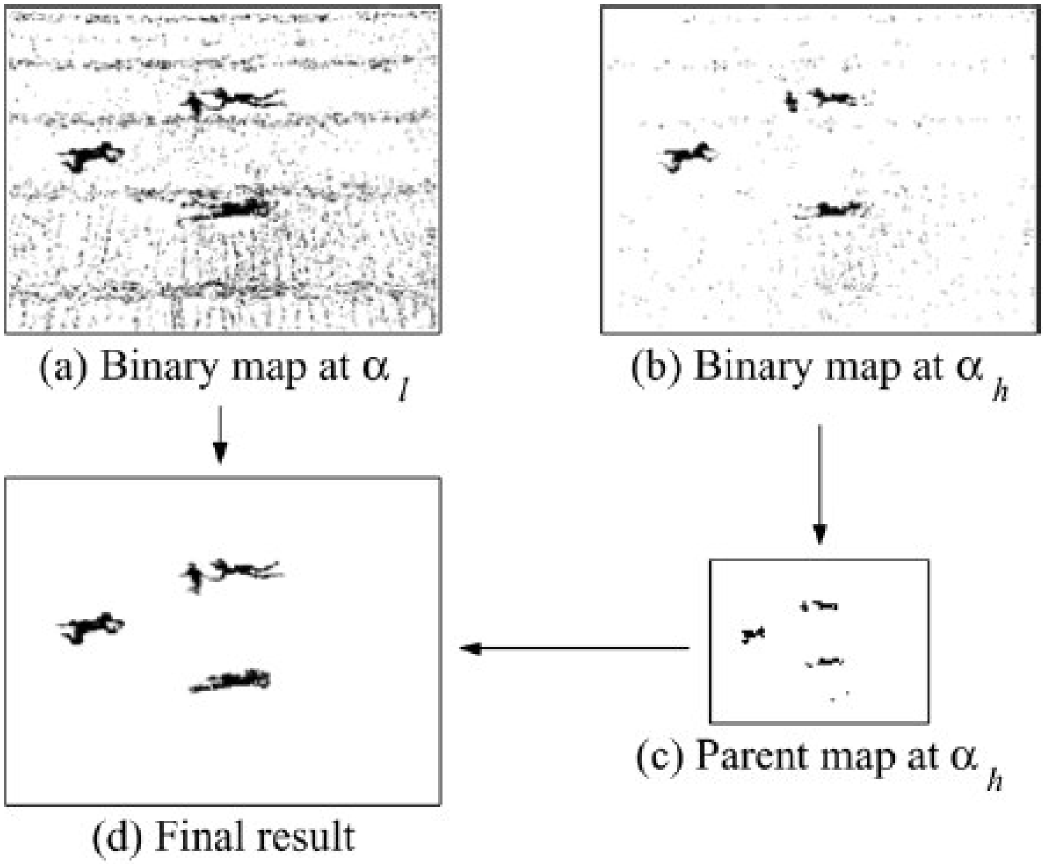
1. Foreground Detection

$$D_{i,j}^{\min} = \min\{D(\boldsymbol{x}_{i,j}|\boldsymbol{\mu}_{\boldsymbol{R}^k_{a+q,b+r}}, \boldsymbol{\sigma}_{\boldsymbol{R}^k_{a+q,b+r}})\}$$

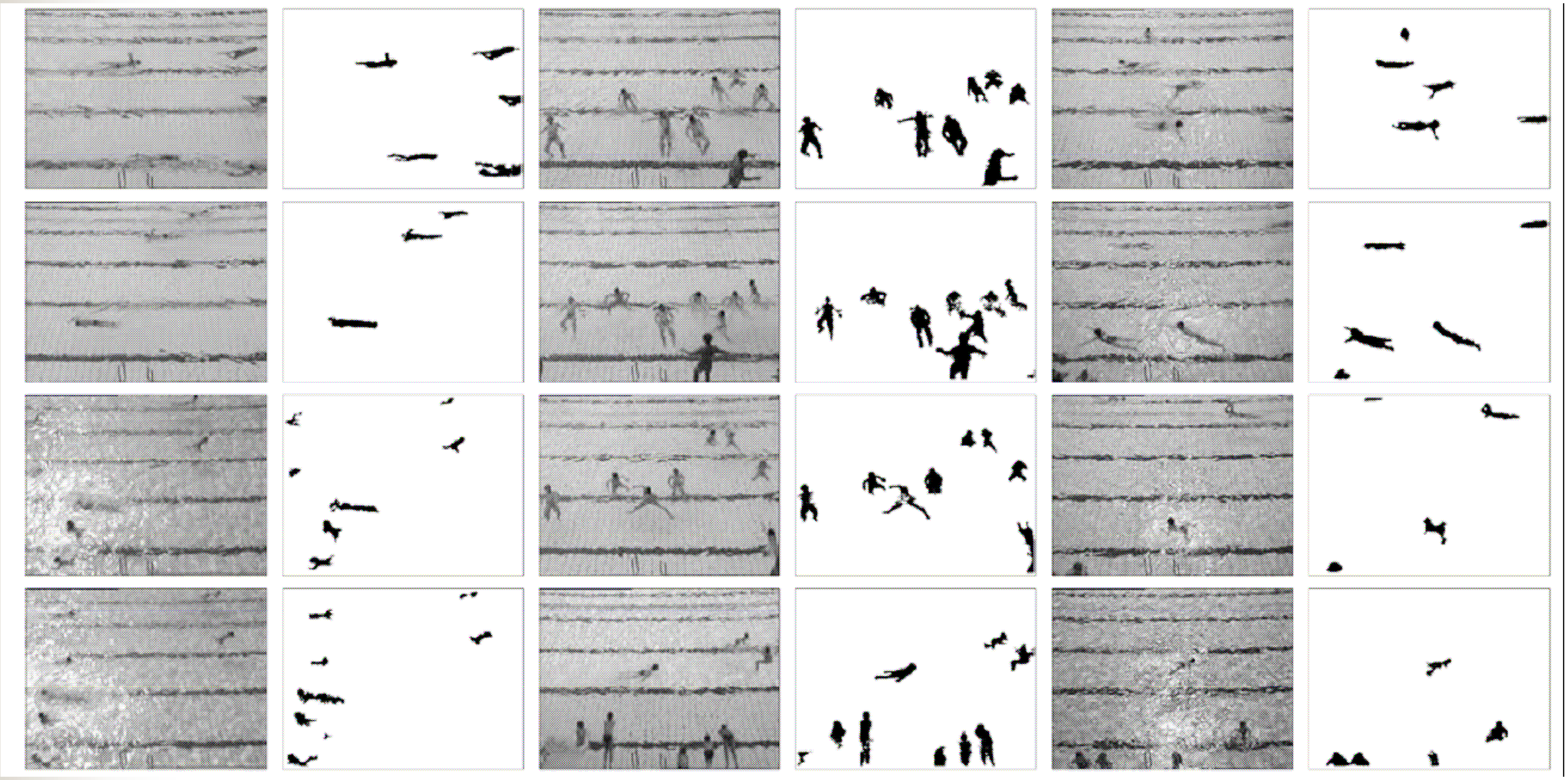
$$M_{i,j} = \begin{cases} 0 \text{ background}, & D_{i,j}^{\min} < \alpha \\ 1 \text{ foreground}, & \text{otherwise.} \end{cases}$$

From Eng et al. 2006

Region-based BGS – (9)

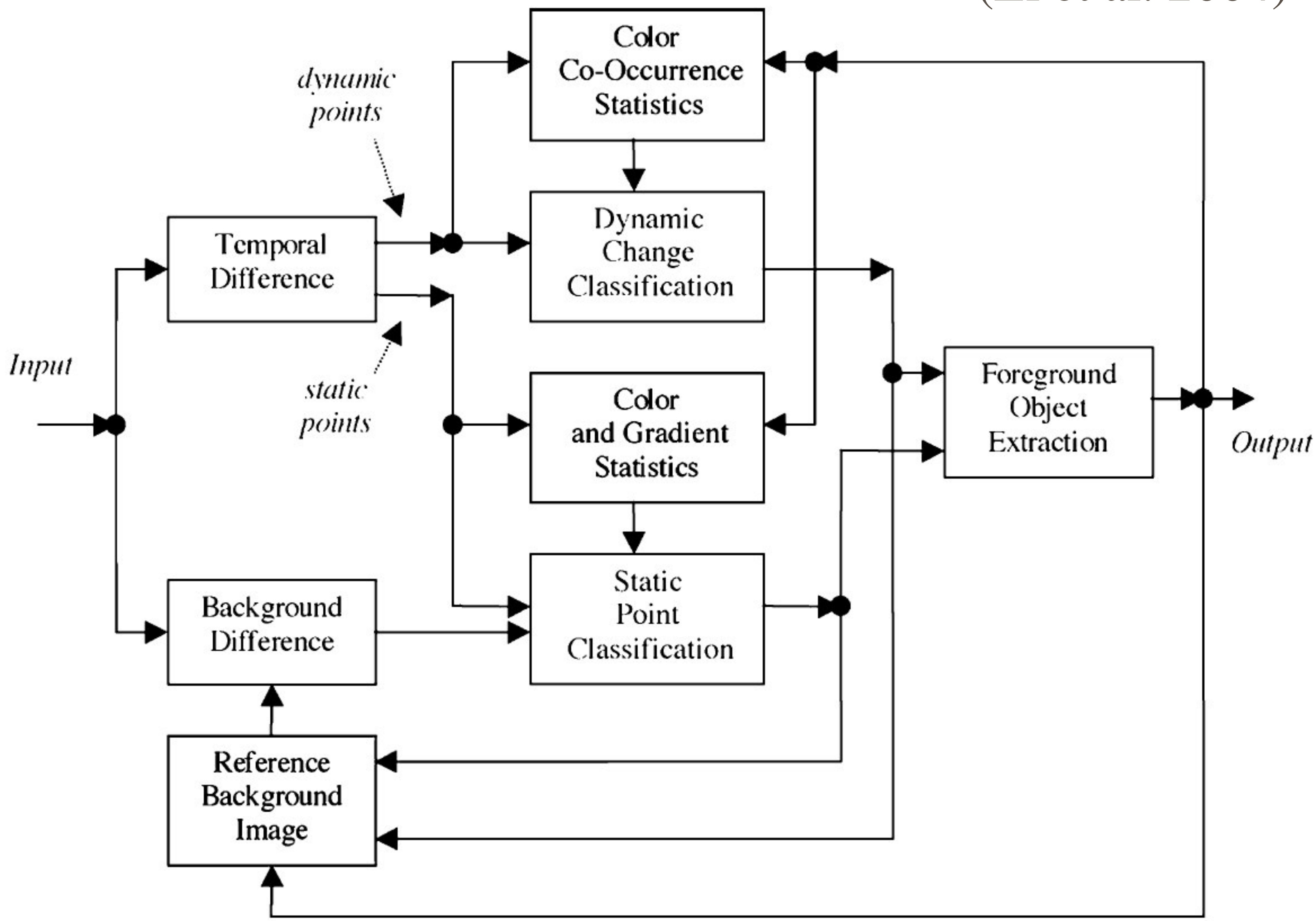


Region-based BGS – (10)



Statistical Modeling for BGS – (1)

(Li et al. 2004)



Statistical Modeling for BGS – (2)

- Bayesian framework by using spatial, temporal and spectral information
- Posterior probability for BG and FG:

$$P_s(b|\mathbf{v}) = \frac{P_s(\mathbf{v}|b)P_s(b)}{P_s(\mathbf{v})} \qquad P_s(f|\mathbf{v}) = \frac{P_s(\mathbf{v}|f)P_s(f)}{P_s(\mathbf{v})}$$

If $P_s(b|\mathbf{v}) > P_s(f|\mathbf{v})$, the pixel belongs to BG

V is the feature vector.

Statistical Modeling for BGS – (3)

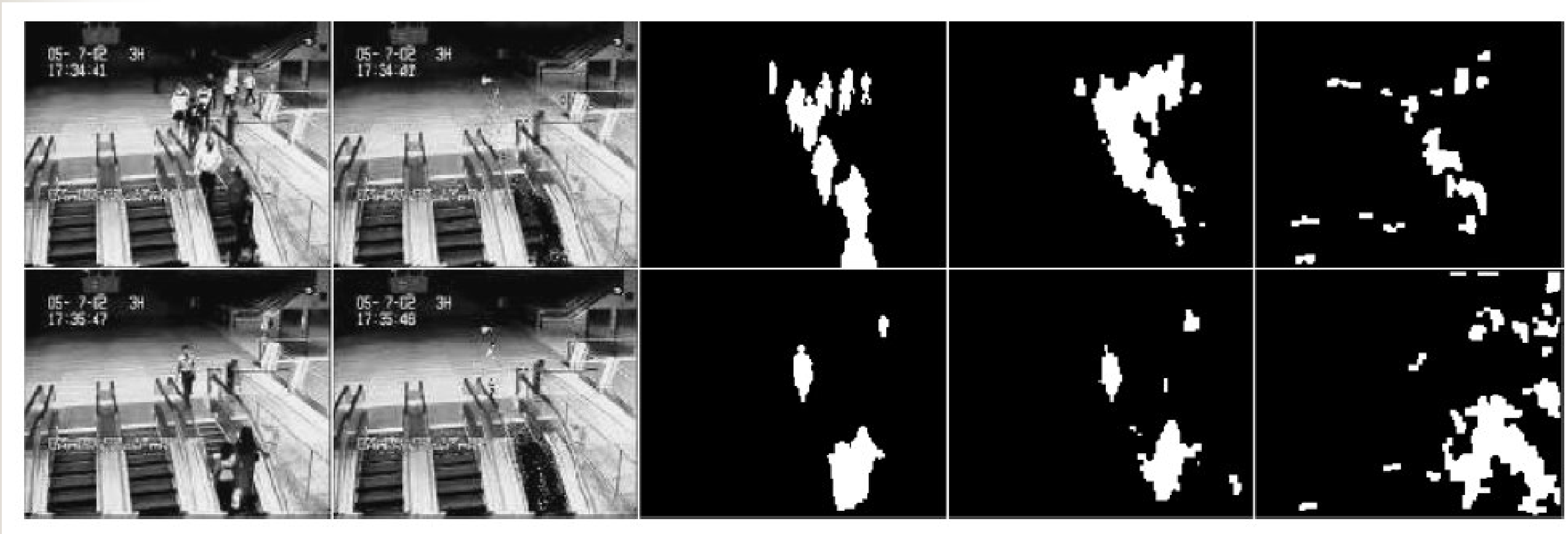
- Features
  - Color and gradient (static BG)
  - Color co-occurrence between consecutive frame (dynamic BG)
- Principal features: histogram of features

Statistical Modeling for BGS – (4)

- ⑤ Principal feature update

$$
\begin{aligned}
p_{\mathbf{v}}^{t+1}(b) &= (1-\alpha)p_{\mathbf{v}}^{t}(b) + \alpha L_b^t \\
p_{\mathbf{v}_i}^{t+1} &= (1-\alpha)p_{\mathbf{v}_i}^{t} + \alpha L_{\mathbf{v}_i}^t \\
p_{\mathbf{v}_i|b}^{t+1} &= (1-\alpha)p_{\mathbf{v}_i|b}^{t} + \alpha \left(L_b^t L_{\mathbf{v}_i}^t\right)
\end{aligned}
$$

Statistical Modeling for BGS – (5)



Input Image　BG Image　GT　Proposed Method　MoG

Salient Motion Detection – (1)

BGS can handle:

✓ **Cluttered background**

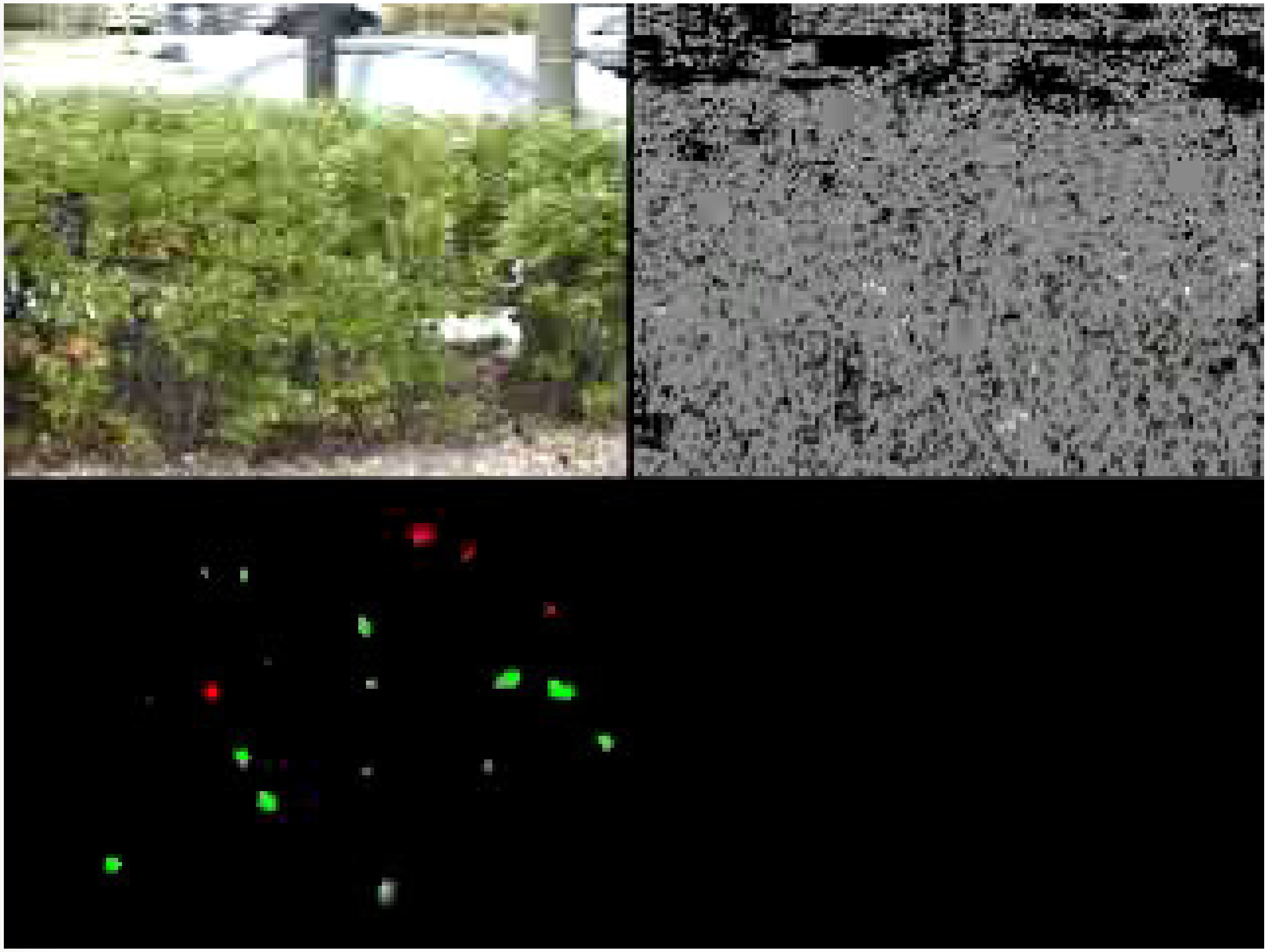
BGS cannot handle:

✗ **Large Distracting Motion**

(a) Original Image

(b) Background Subtraction

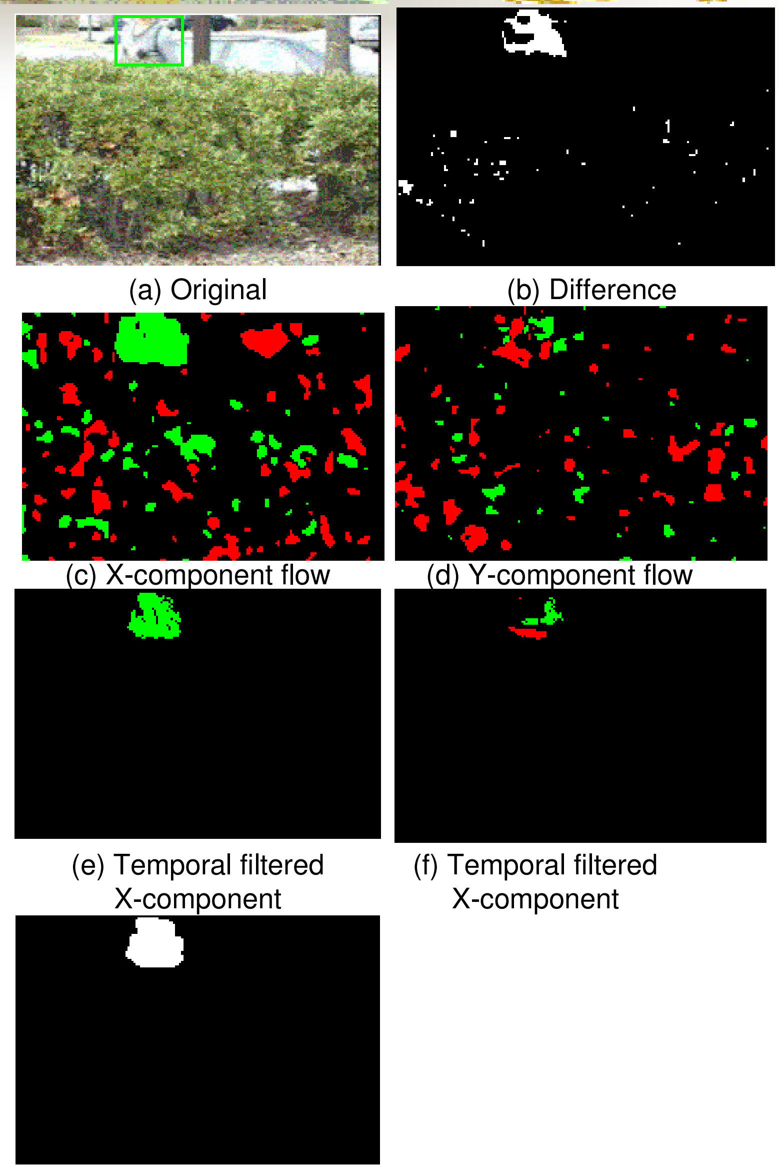
Handling Distracting Motion / Lighting Changes – (2)



Salient Motion Detection – (3)

Salient Motion: motion that is likely to result from a typical surveillance target, e.g. a person or vehicle traveling with a sense of direction through a scene.

- Accumulated Temporal Difference
- Motion – Optical Flow
- Temporal Filter
- **Multi-sources Fusion**
- Region Growing



(a) Original
(b) Difference
(c) X-component flow
(d) Y-component flow
(e) Temporal filtered X-component
(f) Temporal filtered X-component
(g) Salient object

Salient Motion Detection – (4)

⑤ Accumulated Temporal Difference:

$$I_{difference}(x, y, t+1) = \begin{cases} 1, if\ (I_{accum}(x, y, t+1) > T) \\ 0, otherwise \end{cases},$$

(1)

And

$$I_{accum}(x, y, t+1) = (1 - W_{accum}) I_{accum}(x, y, t) + W_{accum} \left| I(x, y, t+1) - I(x, y, t) \right|.$$

Salient Motion Detection – (5)

- Motion Extraction – Optical Flow:

$$I_{t+1}(x+d) - I_t(x) = 0$$

(1)
And

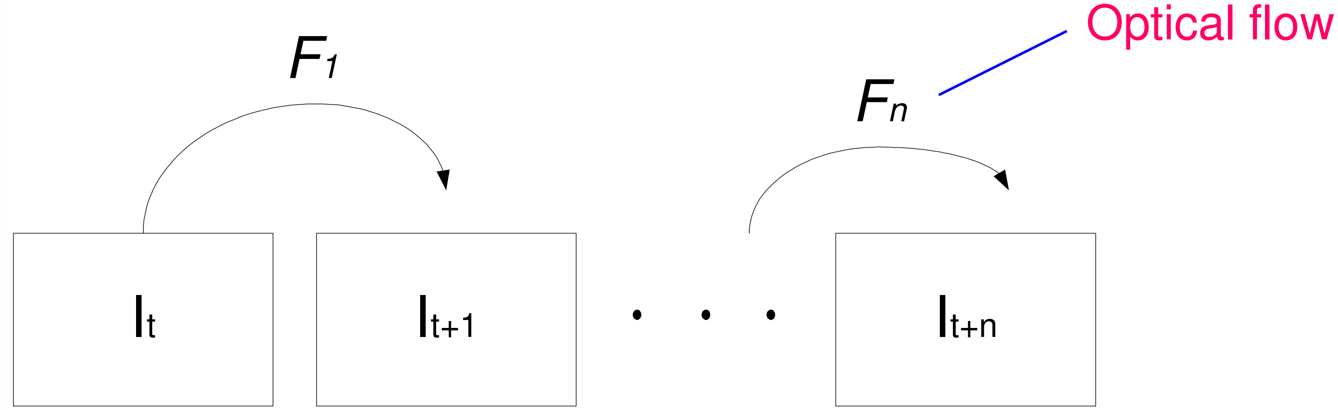
$$E = \sum_{x \in R} [I_{t+1}(x+d) - I_t(x)]^2$$

$$d_{n+1} = d_n + \left\{ \sum_{x \in R} \left(\frac{\partial I}{\partial x}\right)^T \Big|_{x+d_n} [I_t(x) - I_{t+1}(x)] \right\}$$

$$\left[\sum_{x \in R} \left(\frac{\partial I}{\partial x}\right)\left(\frac{\partial I}{\partial x}\right)^T \Big|_{x+d_n} \right]^{-1}$$

Salient Motion Detection – (6)

- Temporal Filter:

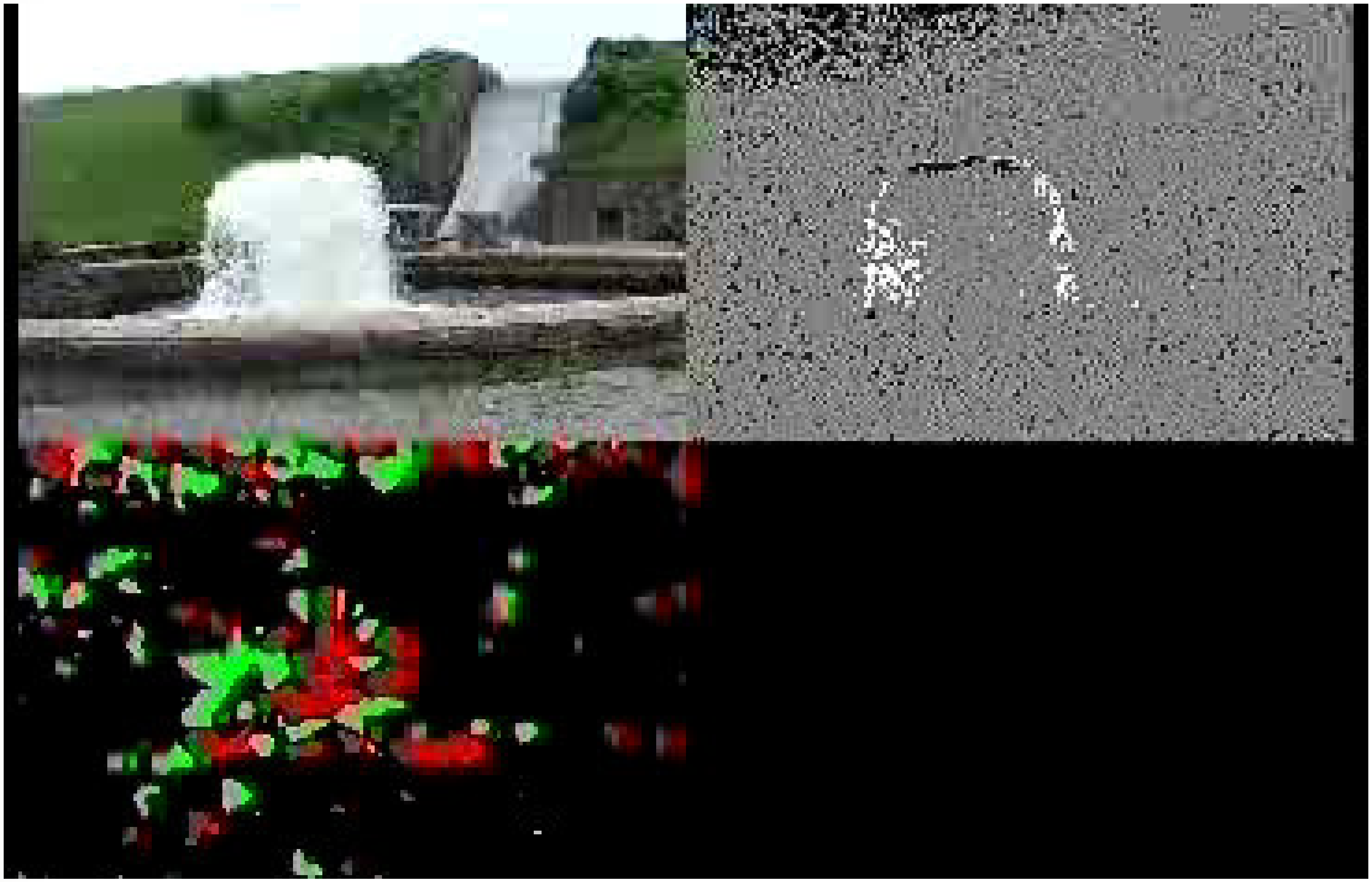


(1)
And

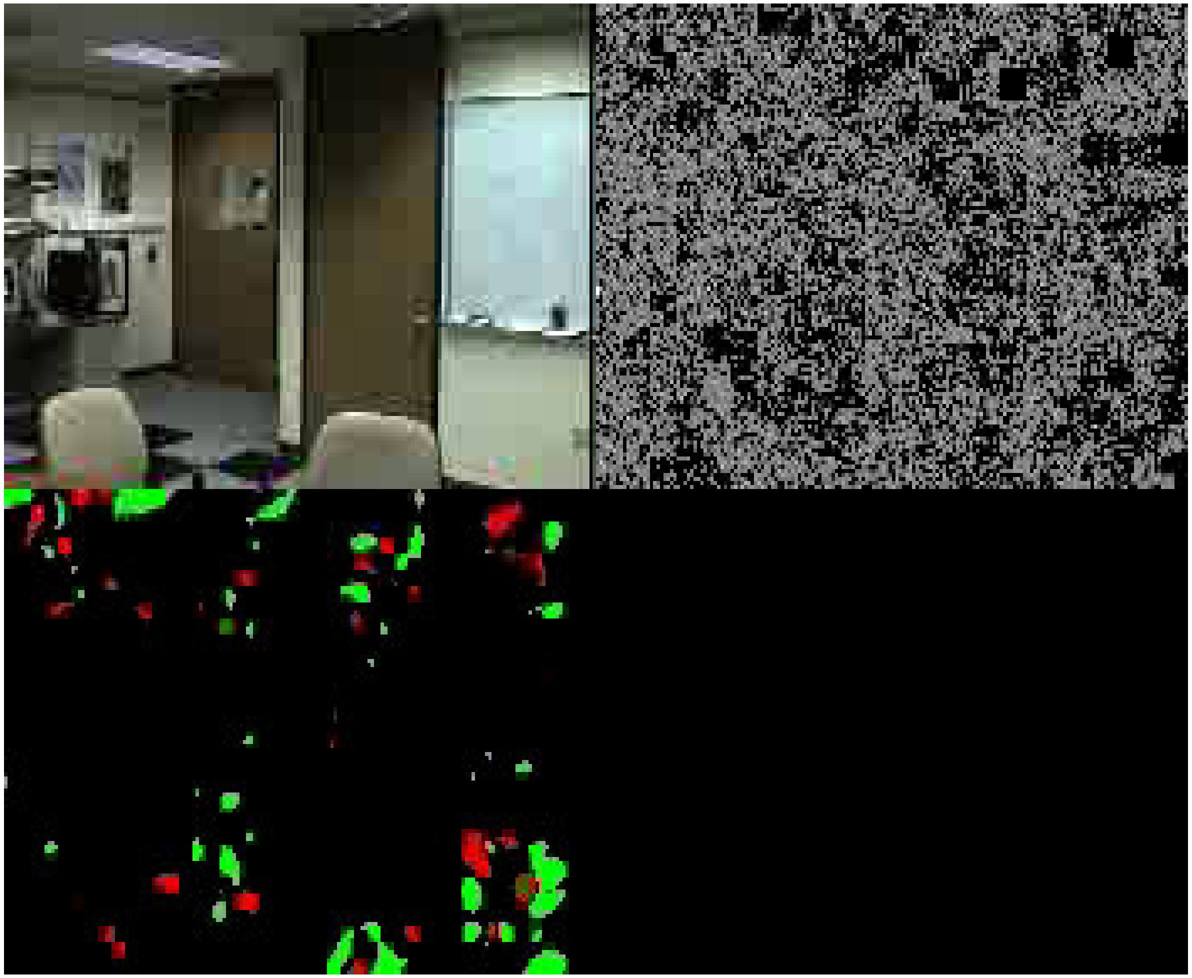
- **Multi-sources Fusion**

$$I_{salient}(x, y, t) = I_{difference}(x, y, t) \bigcap (I_{X-temporal}(x, y, t) \bigcup I_{Y-temporal}(x, y, t))$$

Salient Motion Detection – (7)



Salient Motion Detection – (8)



Salient Motion Detection – (9)

- Salient Motion Detection
  - Deal with large distracting motion
  - Assumptions of object motion
  - Cannot detect the object when it is stop
- Need interaction between higher level processing -- tracking

BGS with higher level feedback

- Frame level
  - Reset BGM
- Tracking
  - Hold an object
  - Heal an object
- Time
  - Different BGM for different time

BGS and Tracker Interaction

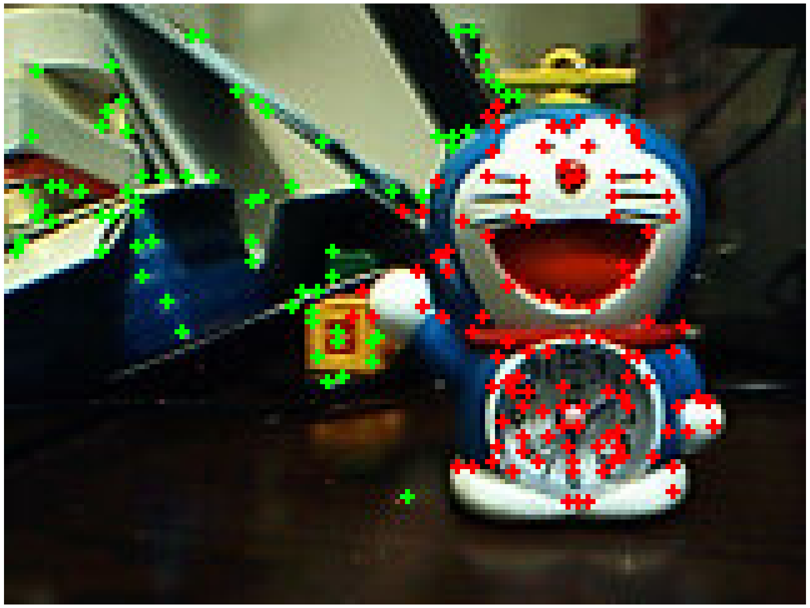
- BGS get feedback from Tracker
  - Slow moving object tracking
  - Stopped object healing
- Different situations
  - Tracker sends “Heal request”, BGS will push the region to BG model
  - Tracker sends out “Unheal request” and provide the image which BGS can use it for BG model, BGS update the BG Model.
  - Tracker sends out “Hold a region”, BGS will not update that region.
  - BGS sends out “Heal request” (auto heal process), tracker decides if do it.

Moving Object Detection from moving camera – (1)

1. Find good feature to track
2. Track features
3. Classify foreground and background features
4. Decide region of foreground object

Moving Object Detection from moving camera – (2)

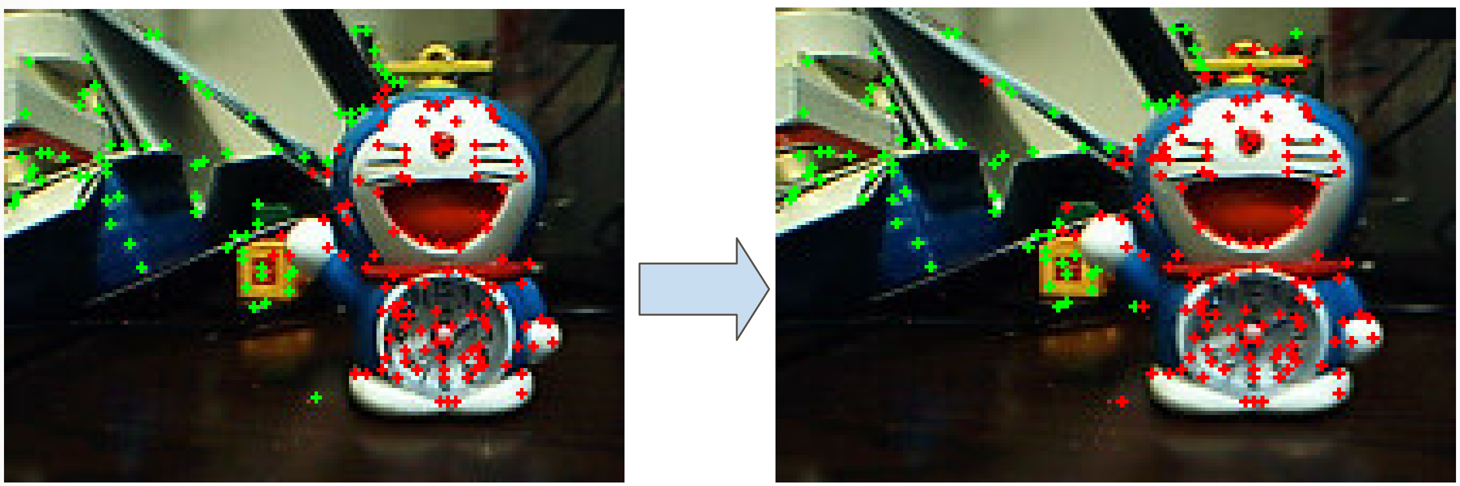
- Finding good feature to track
  - Shi and Tomasi 's method



Images from Martin Chang

Moving Object Detection from moving camera – (3)

⑤ Track features – Optical follow



Images from Martin Chang

Moving Object Detection from moving camera – (4)

- Classify foreground and background feature points
  - Optical flow
    - Moving direction of feature
    - Length of moving direction

Affine Motion Model for Background Registration

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} a_0 \\ a_3 \end{pmatrix} + \begin{pmatrix} a_1 & a_2 \\ a_4 & a_5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

- The affine model describes the vector at each point in the image
- Need to find values for the parameters that best fit the motion present
- Point feature tracker for correspondence between frame pairs
- Iterative reweighted least squares to avoid the features in moving objects (P. W. Holland et al, Robust regression using iteratively reweighted least sqares, Communications in Statistics, A6(9):813-827, 1977)

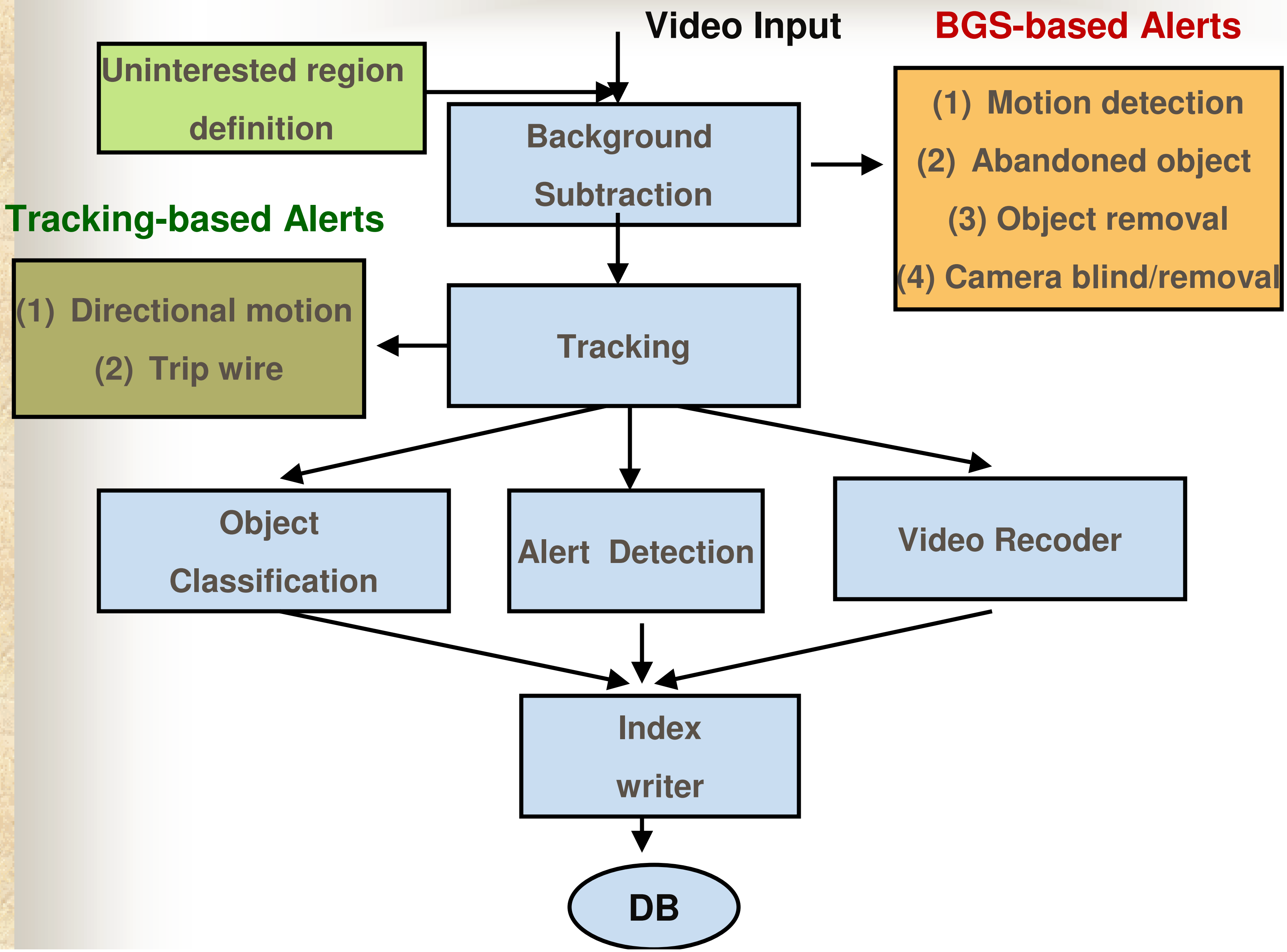
Alerts for Video Surveillance – (1)

- **User defined alerts**
  - Generating real-time alerts from video analytics
  - Generating alerts based on the index – speeding, big car, …
- **Learning-based alerts**
  - loitering, …
  - Recalculate alerts

Alerts for Video Surveillance – (2)

- Motion detection (Trigger alarm when motion detected)
- Directional motion detection (Trigger alarm when motion in the direction detected)
- Trip wire (Trigger alarm when cross boundary)
- Abandoned object (Trigger alarm when abandoned object detected)
- Object removal (Trigger alarm when monitored object removed)
- Camera blind/removal (Trigger alarm when camera being blocked/moved)
- Compound Alarms – (sequential or temporal)
- Region alert
- Camera move stopped
- Slip/fall
- Running
- Gathering (become crowded)
- Speeding

Alerts for Video Surveillance – (3)



Motion Detection Alert

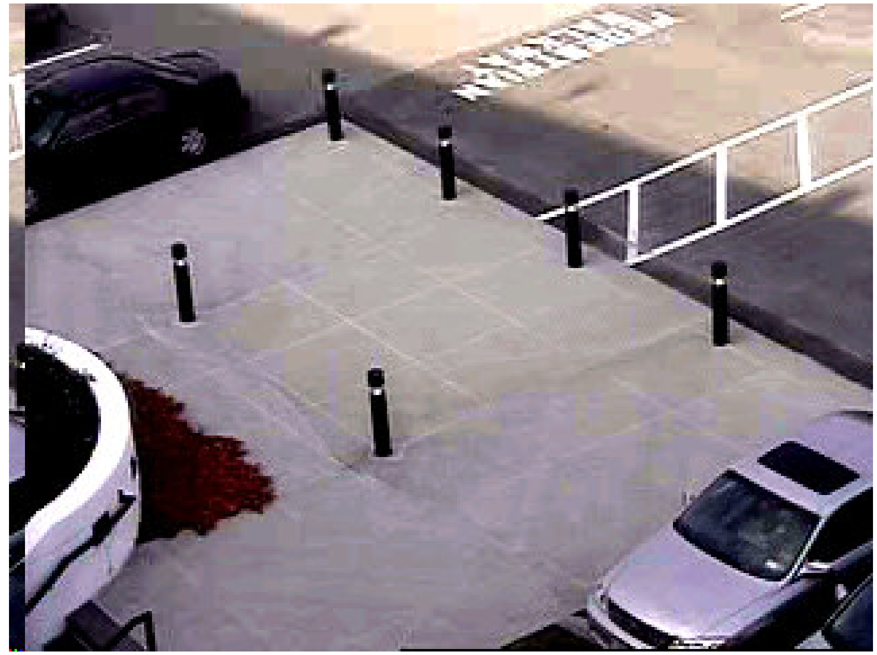
- Can be tracking based or only BGS based
- Region of interest
- Min detected object size:
- Max detected object size
- Number of frames with motion
  - Alarm will be triggered after detecting number of frames with motion
- Min number of moving objects
  - Input at the parameter window (1, 2..)

Motion Detection

Camera Blind/moved Alert: – BGS-based

- Time for pre-event video recording (in seconds):
- Sensitivity to camera movement
  - high
  - Medium high
  - Medium
  - Medium low
  - low

Camera Move/blind



Directional Motion Alert

- Tracking based
- Motion-based – crowded environment

Directional Motion Alert – Tracking-based

- Region of interest
- Define Direction of Motion
- Accuracy degrees (how many degrees can be tolerated)
- Object type
- Object Color
- Object Speed

Directional Motion Detection

Trip Wire Alert – Tracking-based

- Define Trip Wire:
- Min detected object size:
- Max detected object size
- Object type (person, car)
- Object Speed
- Object color

Trip Wire Alert

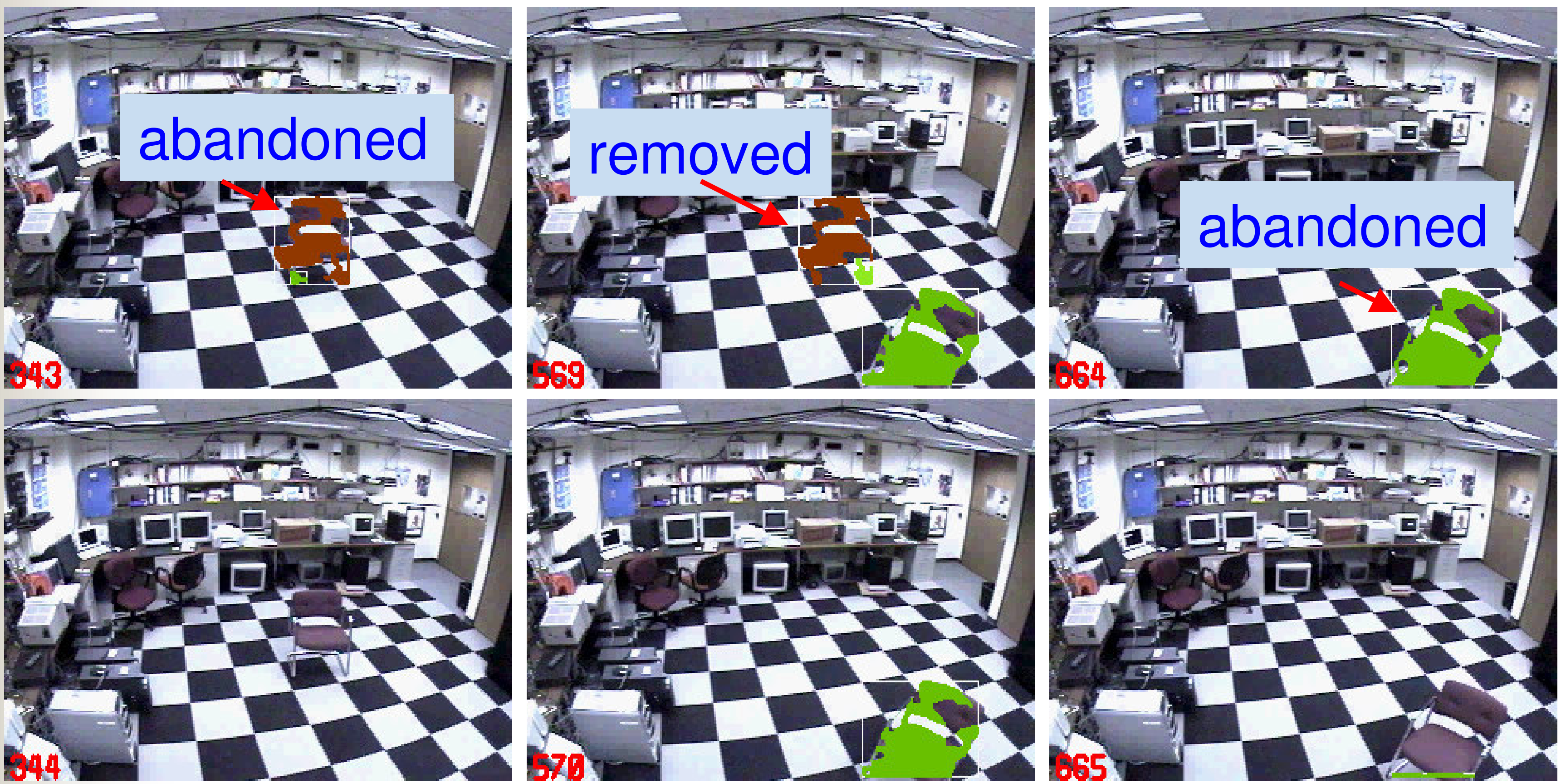
Abandoned/removed Object Detection – (1)

- Detect Static Object
  - Using 2nd Gaussian Model
- When to heal the static region
  - When the static region start to shrink
- Detect heal type
  - Region growing for BG image and input image by using the heal region as seeds (abandoned, removed, unclear)
- Match the region of the input image and the heal region
- Trigger the alert if it meet all the requirements of the alert definition

Abandoned Object Alert – (2)

- Region of interest
- Min detected object size (in pixels):
- Max detected object size (in pixels)
- Waiting time before trigger the alarm (in seconds):
  - Input at the parameter window (1, 2..)

Abandoned/removed Object Detection – (3)



(a) Frame 343 & 344

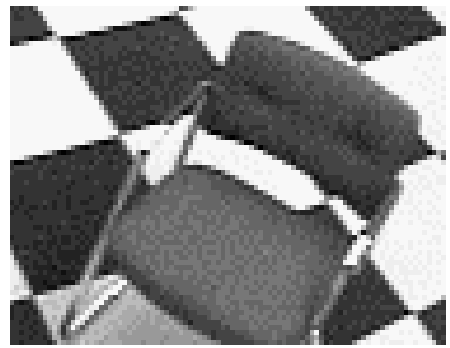
(b) Frame 569 & 570

(c) Frame 664 & 665

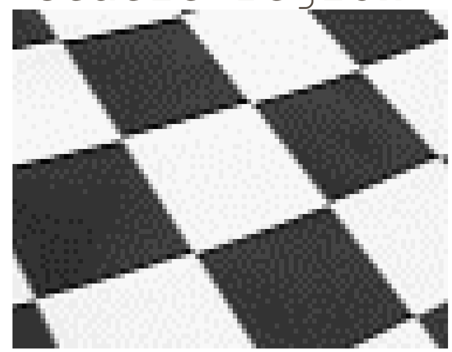
Abandoned or removed Object Detection – (4)

(a) Static Region

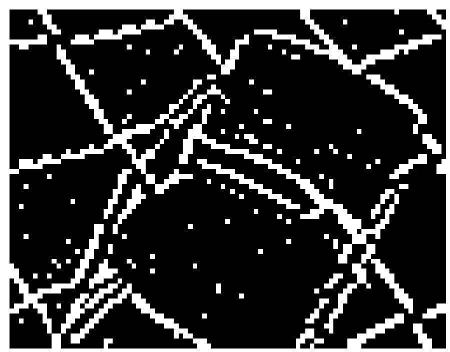
(b) Boundary of static region



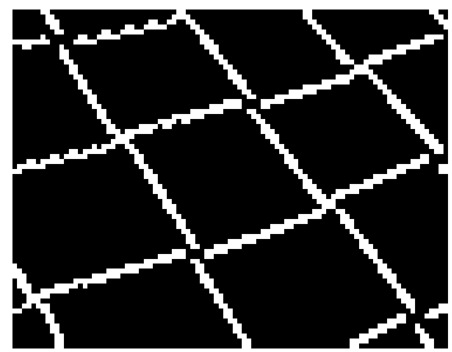
(c) Original image



(d) BG image

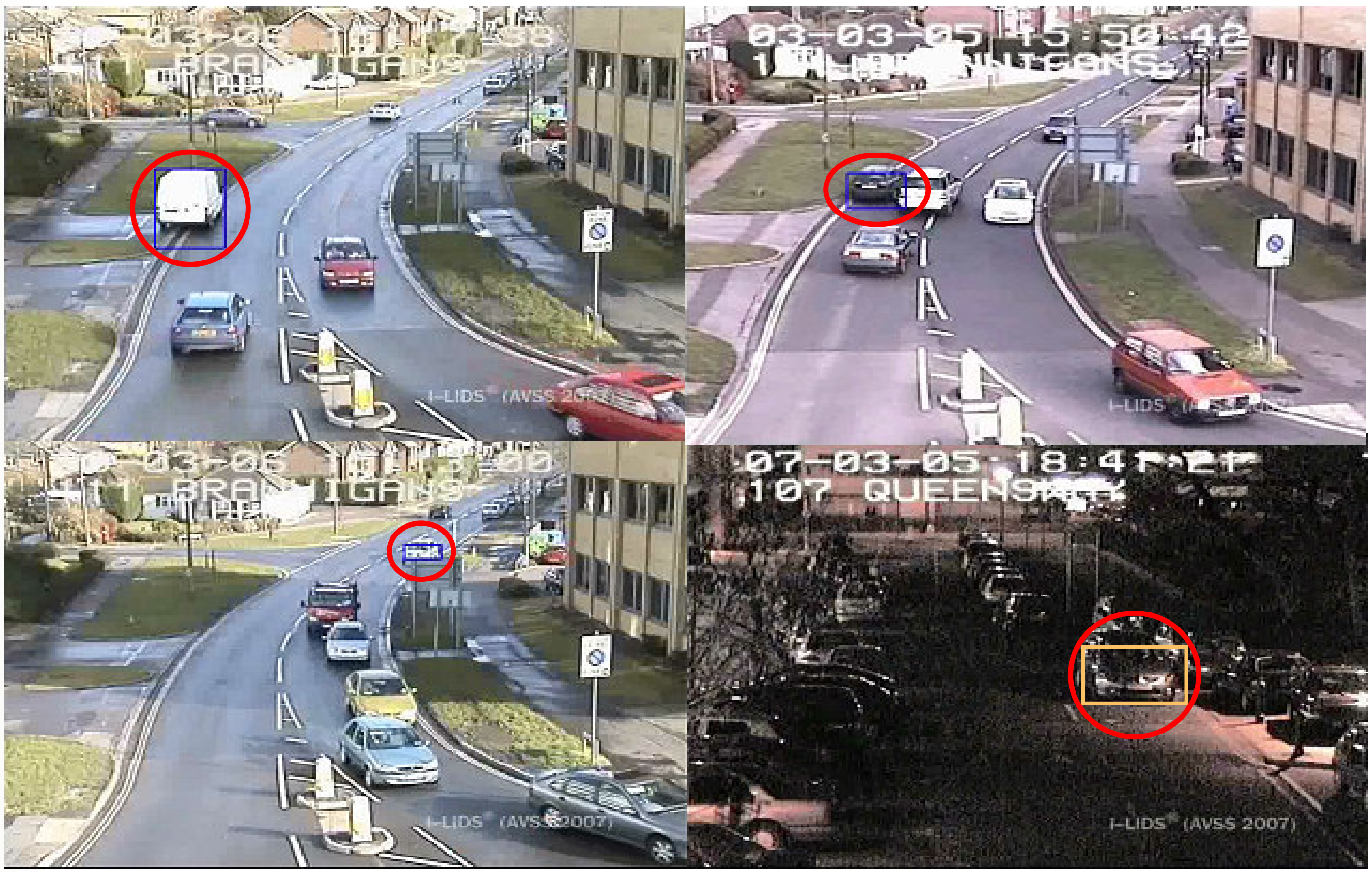


(e) Edge image of Static region

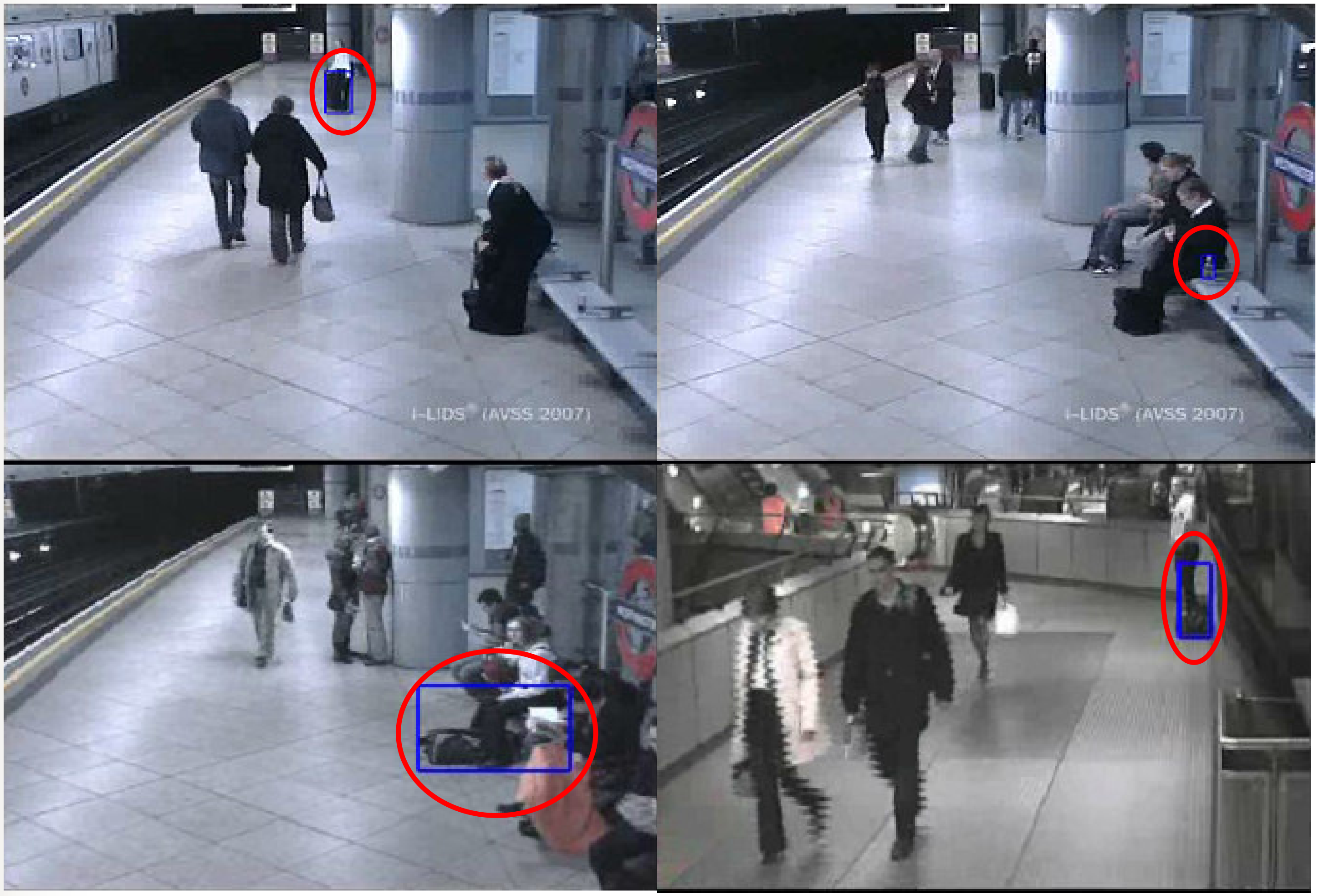


(f) BG edge image

Abandoned/removed Object Detection – (5)



Abandoned/removed Object Detection – (6)



Abandoned/removed Object Detection – (7)

Abandoned Object Detection – (8)

Object Removal Alert – BGS-based

- Region of monitoring
- Sensitivity to changes in the monitoring region:
  - high
  - Medium high
  - Medium
  - Medium low
  - low

Object Removal – museum mode

Summary

* Moving Object Detection with Distraction Motions
  * Region-based mixture of Gaussians
  * Statistical framework for BGS
  * Motion-based moving object detection
    * Salience Motion
* Interaction between BGS and Tracking
* Moving Object from Moving Cameras
* Real-time alerts of video surveillance

References – (1)

- H. Eng, J. Wang, A. Kam, A. Siew, and W. Yau, "Robust Human Detection within a Highly Dynamic Aquatic Environment in Real Time," IEEE Transaction on Image Processing, Vol. 15, No. 6, 2006.
- M. Harville, "A Framework for High-level Feedback to adaptive, per-pixel, Mixture-of-Gaussian Background Models", Proceedings on ECCV, 2002.
- L. Li, W. Huang, I.Y.H. Gu, and Q. Tian, "Statistical Modeling of Complex Backgrounds for Foreground Object Detection", IEEE Transaction on Image Processing, Vol. 13, No. 11, 2004.
- A. Monnet, A. Mittal, N. Paragios, and V. Ramesh, "Background Modeling and Subtraction of Dynamic Scenes", Proc. on ICCV, Pages 1305–1312, 2003.

References – (2)

- A. Mittal and N. Paragios, "Motion-based Background Subtraction using Adaptive Kernel Density Estimation," Proceedings on CVPR, 2004.
- Ying-Li Tian and Arun Hampapur, "Robust Salient Motion Detection with Complex Background for Real-time Video Surveillance," IEEE Workshop on Motion and Video Computing, Jan, 2005.
- Zhang, et al, "Segmentation of Moving Objects in Image Sequence: A Review ", Circuits, Systems, and Signal Processing, Vol. 20(2), 2001, pp. 143-183.
- J. Shi and Carlo Tomasi, "Good Features to Track" CVPR'94, pp. 593-600.
- P. W. Holland et al, Robust regression using iteratively reweighted least squares, Communications in Statistics, A6(9):813-827, 1977