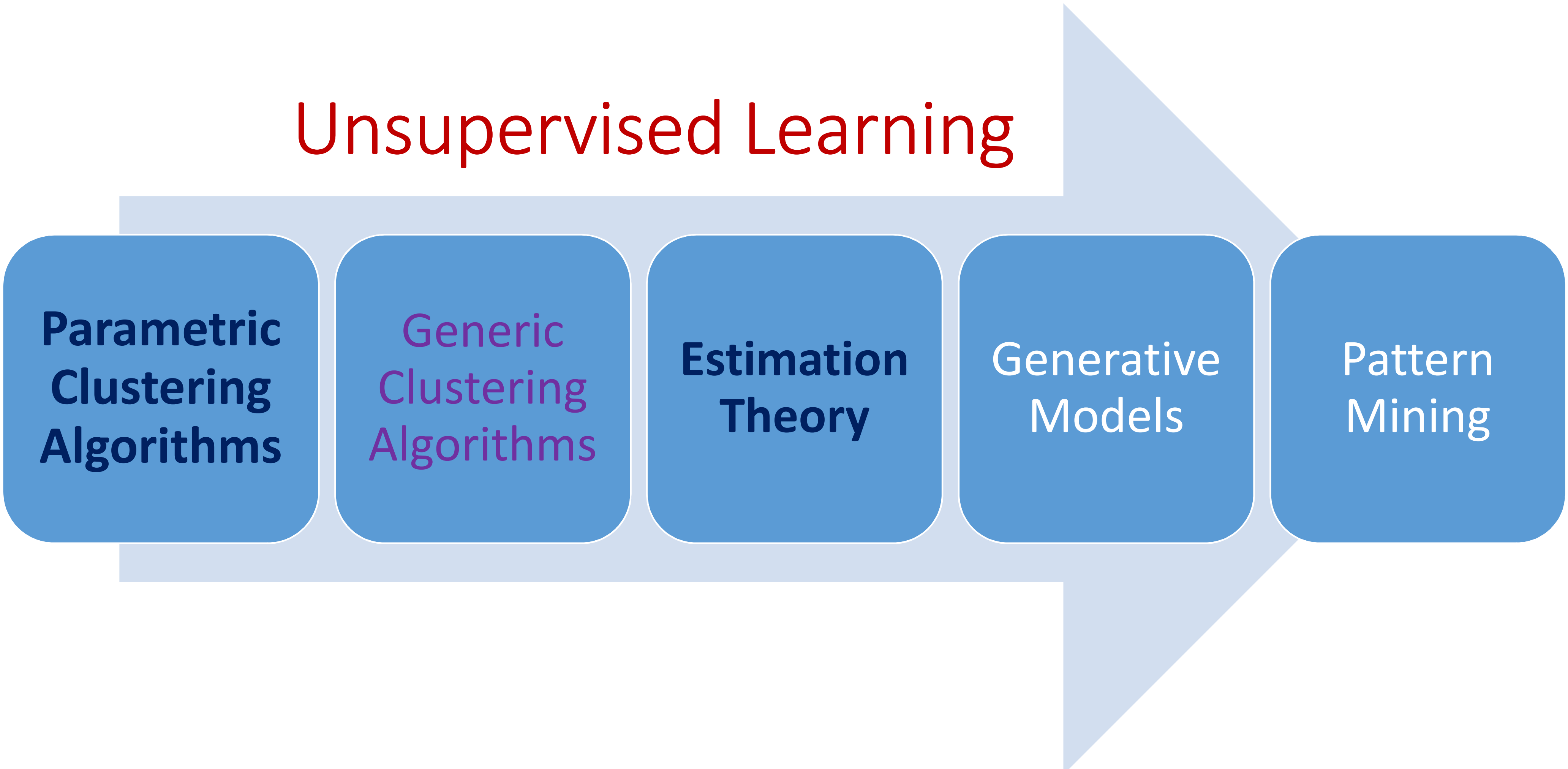
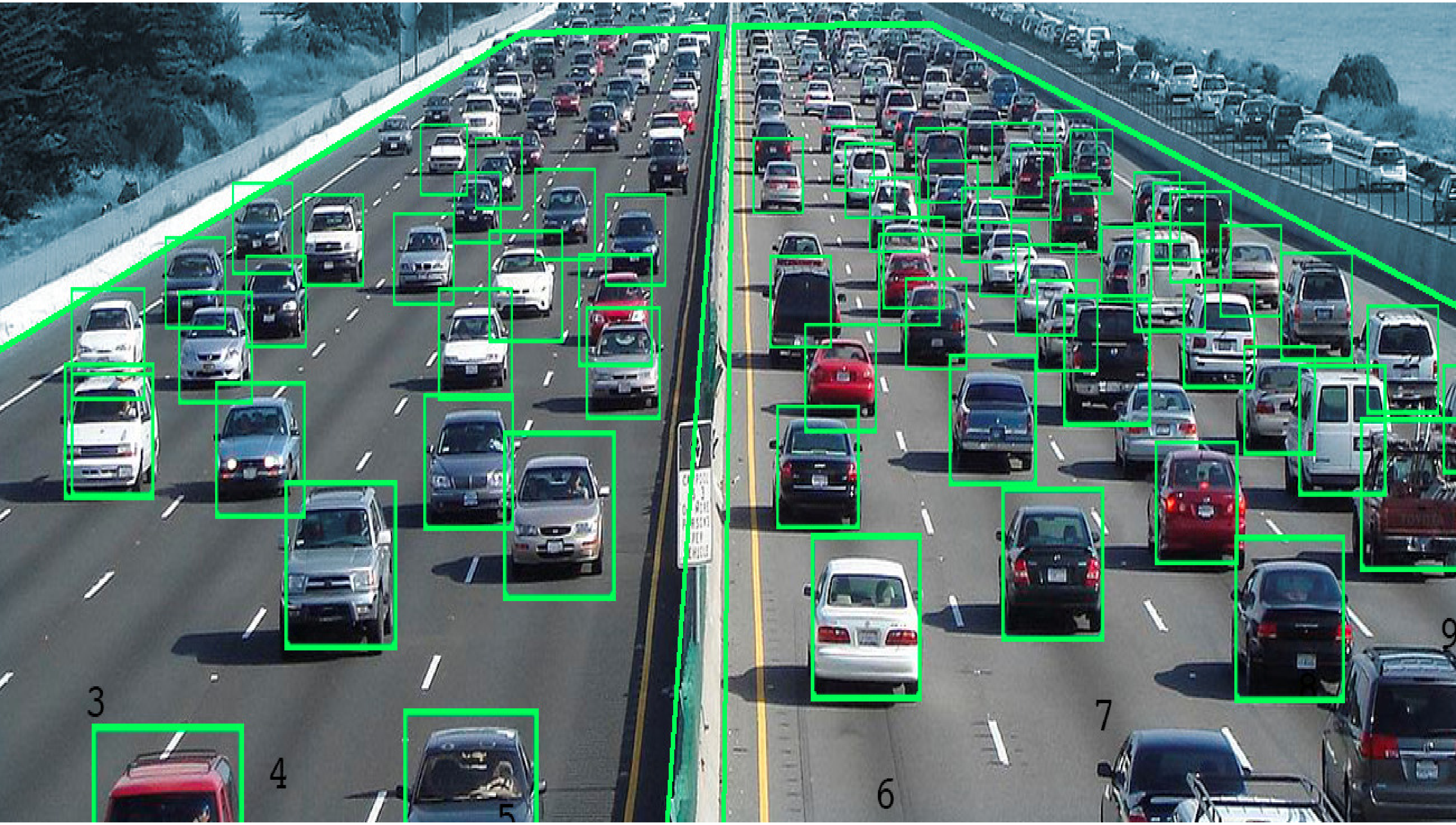
Video Surveillance: Intrusion Detection & Tracking

Prithwijit Guha
Dept. of EEE, IIT Guwahati


Unsupervised Learning
Parametric Clustering Algorithms
Generic Clustering Algorithms
Estimation Theory
Generative Models
Pattern Mining


3
4
6
7

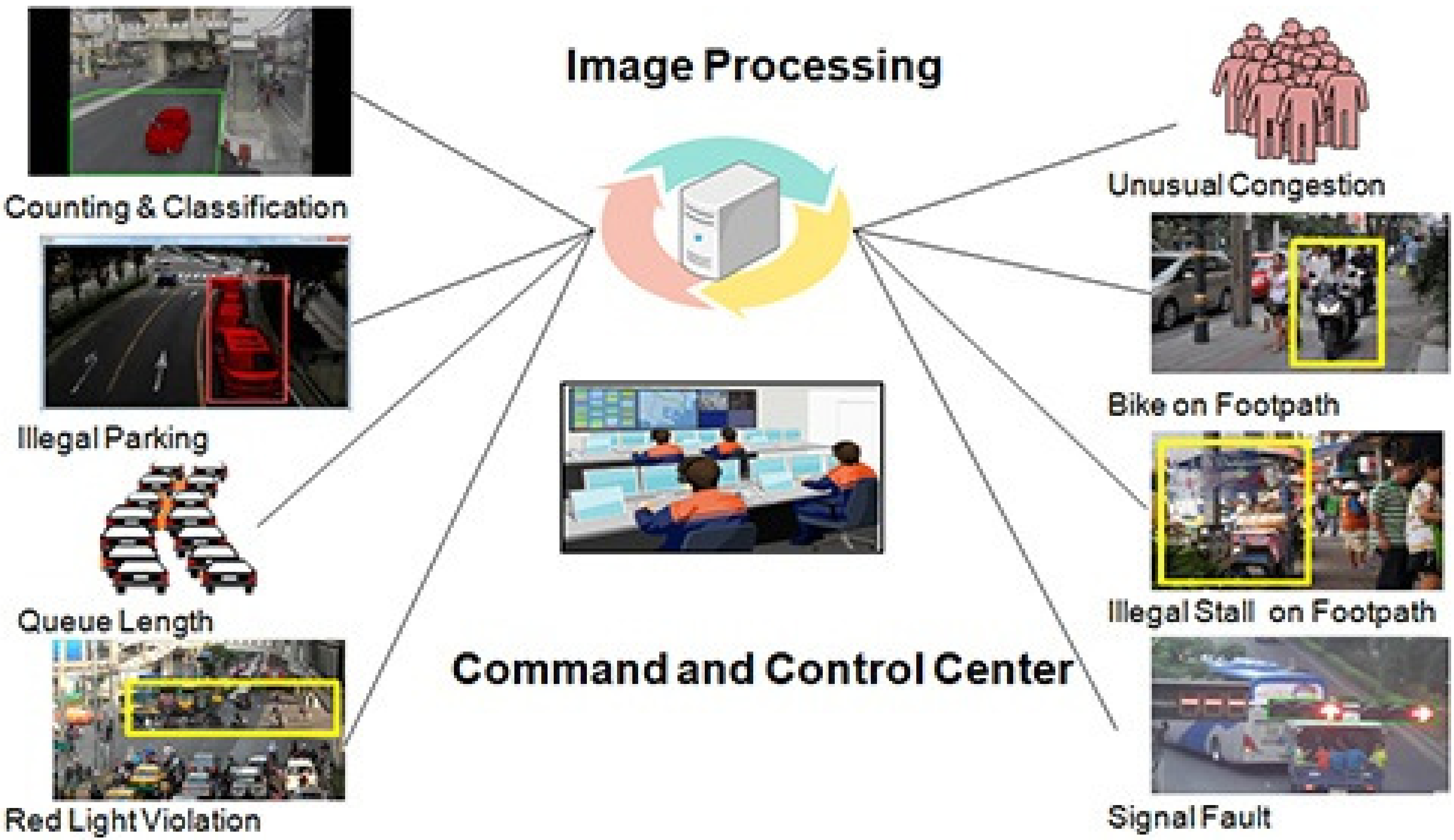


Image Processing
Counting & Classification
Illegal Parking
Queue Length
Red Light Violation
Command and Control Center
Unusual Congestion
Bike on Footpath
Illegal Stall on Footpath
Signal Fault

Sensors: Modality

CAMERA

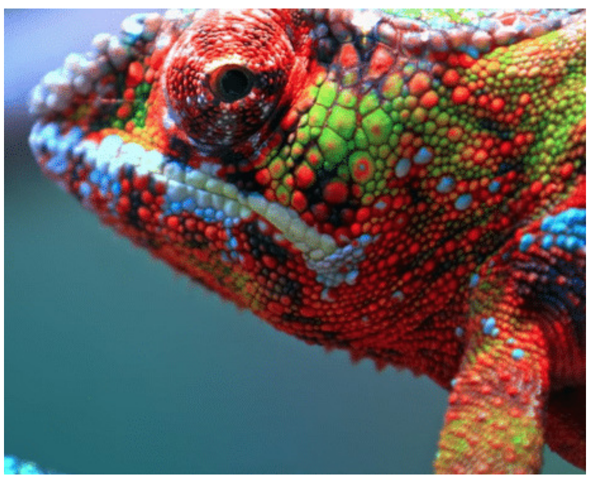
LIDAR

RADAR

Sensors: Multiplicity

Sensors: Mobility

Image & Feature Distribution



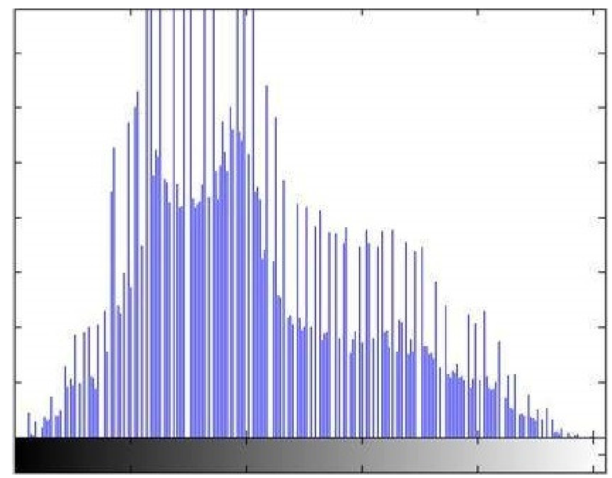


Image: R-G-B Channels

Tri-chrome

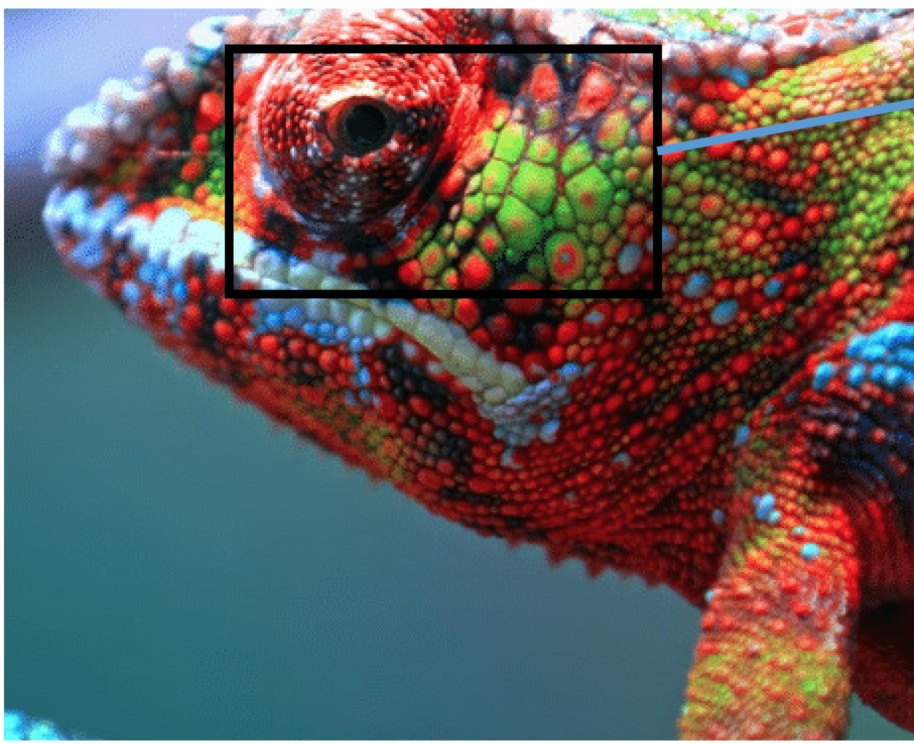
$I(x, y) \in \{0, \dots 255\}^3$

Monochrome

$I(x, y) \in \{0, 1, \dots 255\}$

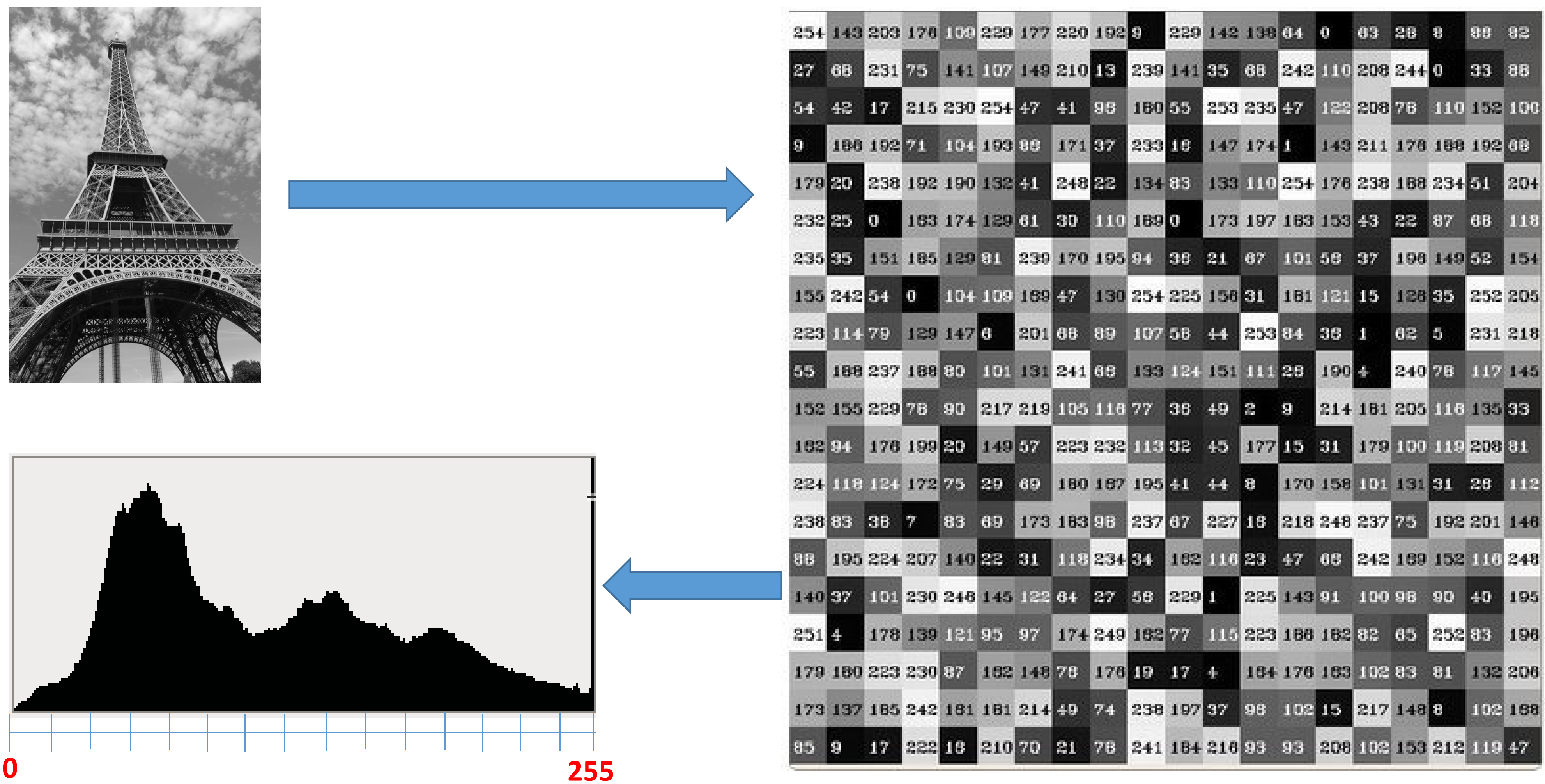
R-G-B RED GREEN BLUE

Image Region of Interest (RoI)

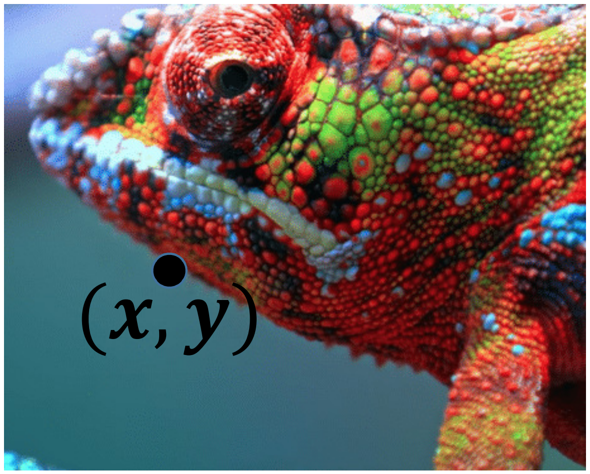


Region of
Interest (RoI)

Monochrome Image Histogram



Color Image Histogram



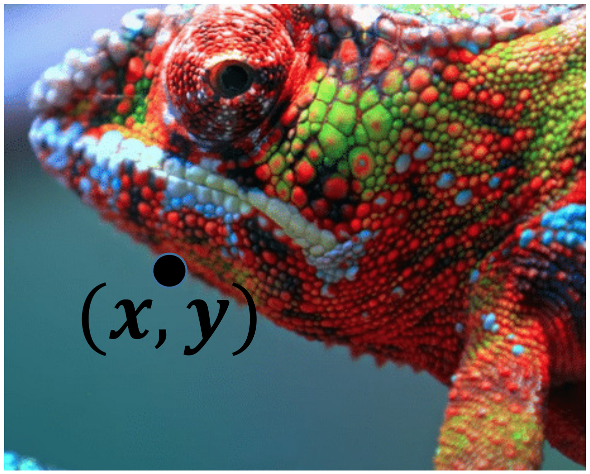
$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \boldsymbol{I}(x, y) \longrightarrow (RGB)_{256}$$

$$(324)_{10} = 3 \times 10^2 + 2 \times 10^1 + 4 \times 10^0$$

Three Hundred Twenty Four in Decimal Number System (Base 10)

Digits in $\{0,1 \ldots (10-1)\}$

Color Image Histogram



$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \boldsymbol{I}(x, y) \Rightarrow (RGB)_{256}$$

$$(RGB)_{256} \equiv R \times 256^2 + G \times 256^1 + B \times 256^0$$

R-G-B in a Number System of Base 256

Digits in $\{0,1 \ldots (256-1)\}$

Color Image Histogram

$(RGB)_{256} \rightarrow 256^3$ Colors

$$R \times 256^2 + G \times 256^1 + B \times 256^0$$

b_{pc}: Bins Per Color Channel

$$\text{Bin Size} = b_s = \frac{256}{b_{pc}}$$

$$r = \left\lfloor \frac{R}{b_s} \right\rfloor \quad g = \left\lfloor \frac{G}{b_s} \right\rfloor \quad b = \left\lfloor \frac{B}{b_s} \right\rfloor$$

$$bI = r \times b_{pc}^2 + g \times b_{pc}^1 + b \times b_{pc}^0$$

Image: Color Distribution

$\{\boldsymbol{x}_i; i = 1, \ldots n\}$: Image Pixels

$\boldsymbol{I}(\boldsymbol{x}_i)$: RGB Color Vector at Pixel $\boldsymbol{x}_i$

$bI(\boldsymbol{x}_i)$: Color Bin Index Value at Pixel $\boldsymbol{x}_i$

b_{pc} : Bins Per Color Channel

$$\boldsymbol{H}[j] = \frac{1}{n}\sum_{i=1}^{n} \delta[bI(\boldsymbol{x}_i) - \mathrm{j}\,] \qquad j = 0, \ldots (b_{pc}^3 - 1)$$

Image: Weighted Color Distribution

$\{\boldsymbol{x}_i; i = 1, \dots n\}$: Image Pixels

$bI(\boldsymbol{x}_i)$: Color Bin Index Value at Pixel $\boldsymbol{x}_i$

$\omega(\boldsymbol{x}_i)$: Weight Value of Pixel $\boldsymbol{x}_i$

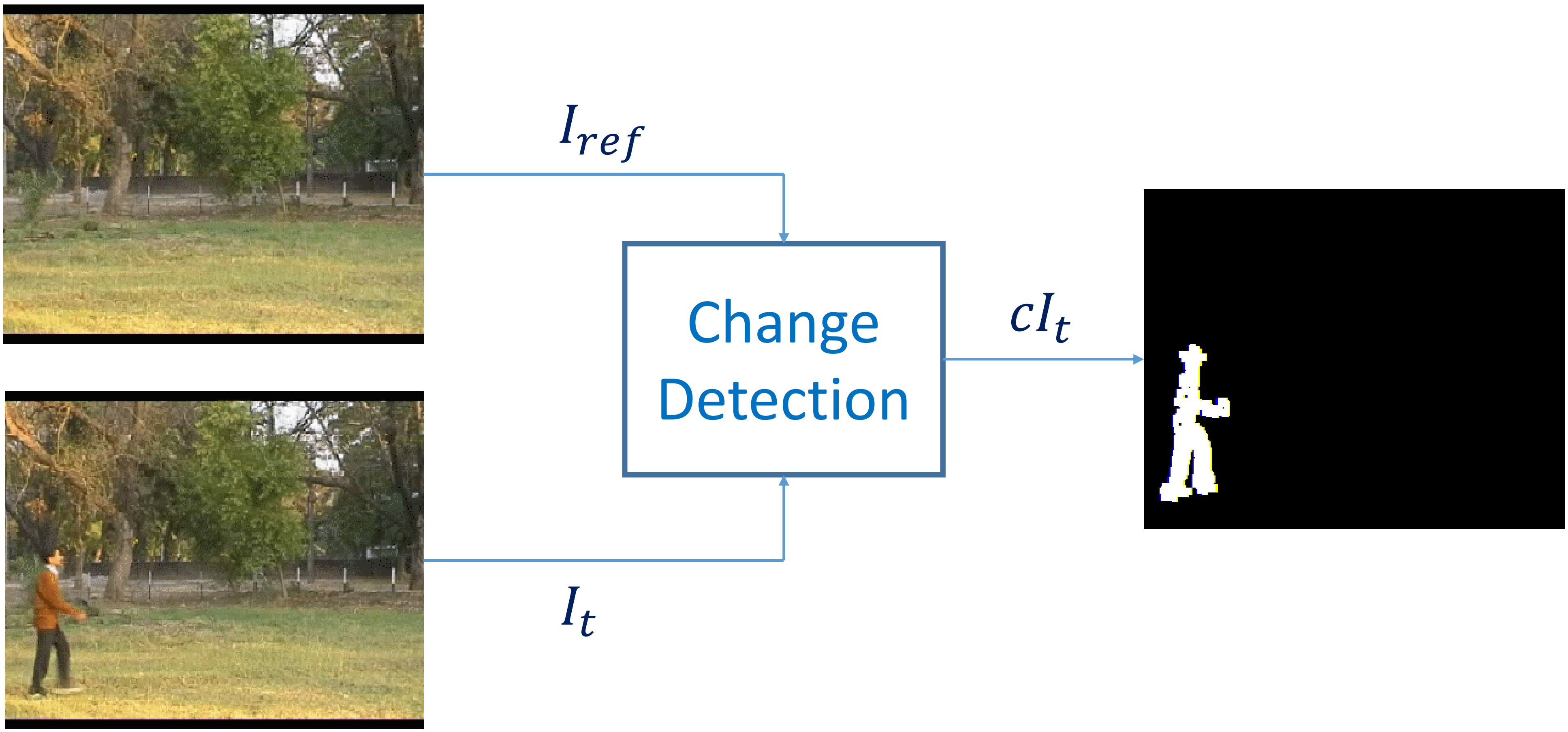
$$j = 0, \dots (b_{pc}^3 - 1)$$

$$\boldsymbol{V}[j] = \sum_{i=1}^{n} \omega(\boldsymbol{x}_i)\delta[bI(\boldsymbol{x}_i) - \mathrm{j}\,]$$

$$\boldsymbol{H}[j] = \frac{\boldsymbol{V}[j]}{\sum_{r=1}^{m} \boldsymbol{V}[r]}$$

Intrusion Detection

Change Detection



Change Detection

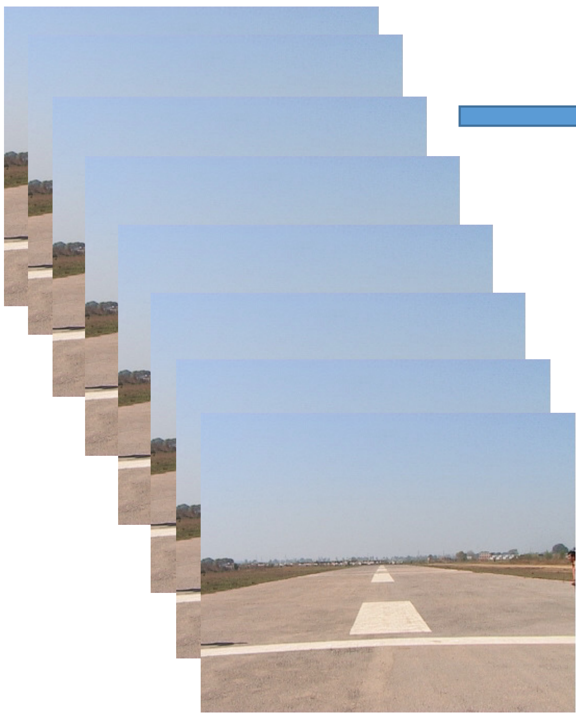
I_{ref}

$$\neg C(x, y, t) \Rightarrow \wedge_{k=0}^{2} \left[\left| I_t(x, y, k) - I_{ref}(x, y, k) \right| \leq \eta_c \right]$$

I_t

$$cI_t(x, y) = \begin{cases} 0, & \neg C(x, y) \\ 255, & Otherwise \end{cases}$$

Unimodal Background Model

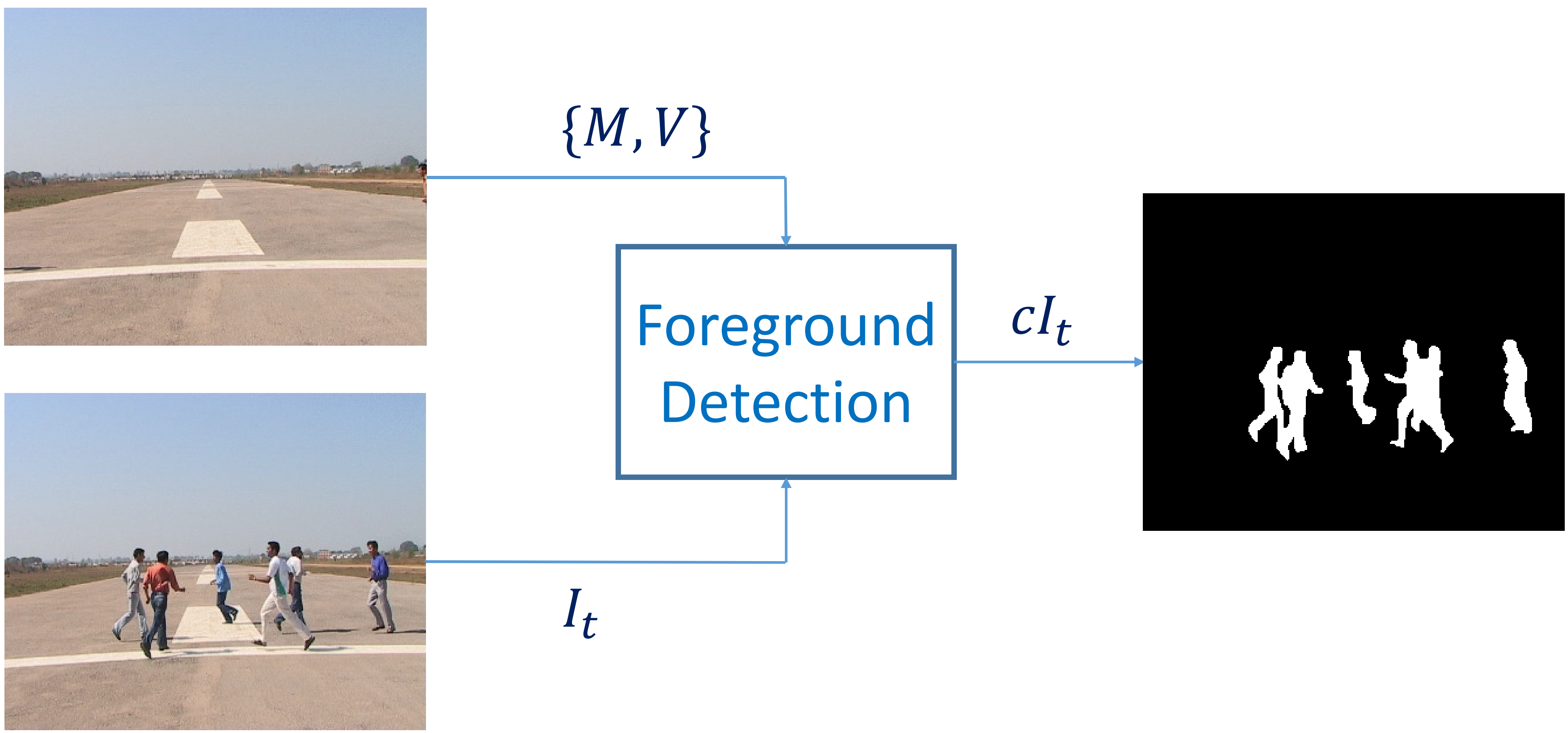


Background Modeling

$M_t(x, y, k)$

$V_t(x, y, k)$

Foreground Detection



Foreground Detection

$\{M_{t-1}, V_{t-1}\}$

$$d_t(x, y, k) = I_t(x, y, k) - M_{t-1}(x, y, k)$$

$$\neg C(x, y, t) \Rightarrow \wedge_{k=0}^{2} \left[d_t^2(x, y, k) \leq \lambda^2 V_{t-1}(x, y, k)\right]$$

I_t

$$cI_t(x, y) = \begin{cases} 0, & \neg C(x, y, t) \\ 255, & Otherwise \end{cases}$$

Background Model Update

cI_t

$\{M_t, V_t\}$

$$d_t(x,y,k) = I_t(x,y,k) - M_{t-1}(x,y,k)$$

IF $cI(x,y,t) = 0$

$$M_t(x,y,k) = (1-\alpha)M_{t-1}(x,y,k) + \alpha I_t(x,y,k)$$

$$M_t(x,y,k) = M_{t-1}(x,y,k) + \alpha d_t(x,y,k)$$

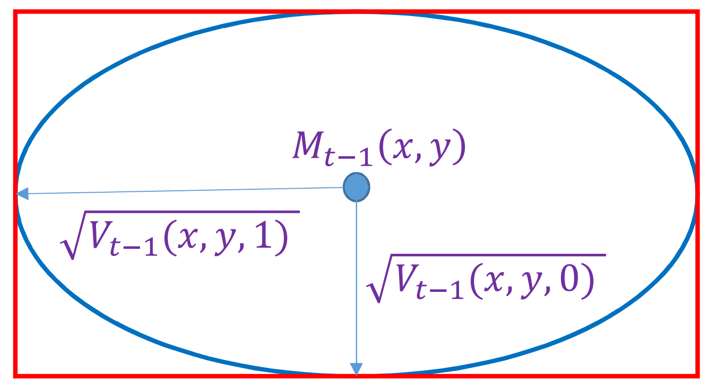
$$V_t(x,y,k) = (1-\alpha)[V_{t-1}(x,y,k) + \alpha d_t^2(x,y,k)]$$

$k = 0(R), 1(G), 2(B)$

An Implementation Issue

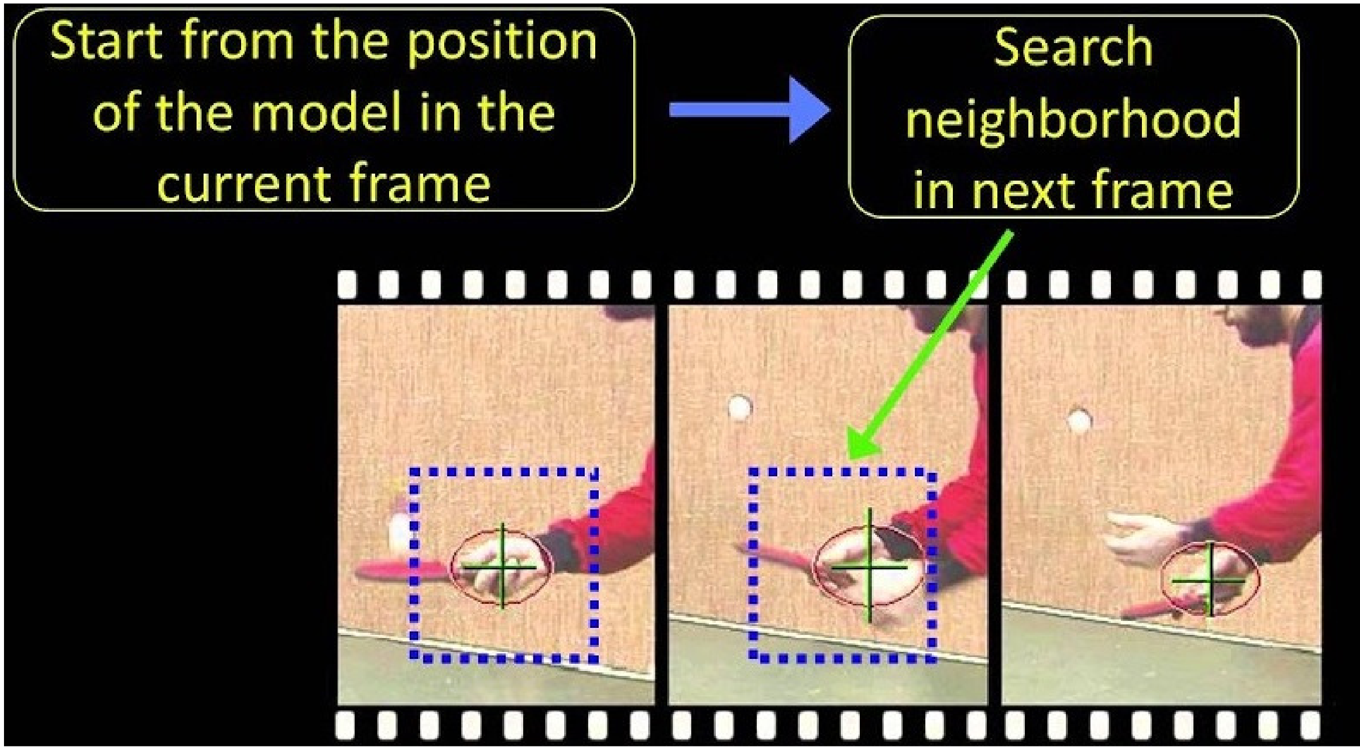
$$d_t(x,y,k) = I_t(x,y,k) - M_{t-1}(x,y,k)$$

$$\sum_{k=0}^{2} \frac{\{I_t(x,y,k) - M_{t-1}(x,y,k)\}^2}{V_{t-1}(x,y,k)} \leq \lambda^2$$



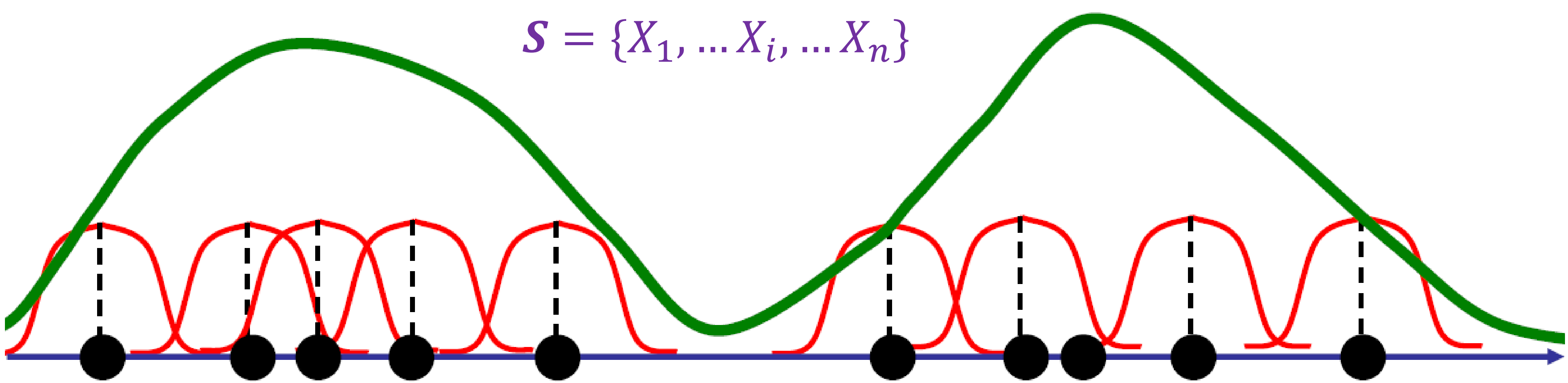
$$\wedge_{k=0}^{2}\,[d_t^2(x,y,k) \leq \lambda^2 V_{t-1}(x,y,k)]$$

Mean Shift Tracking

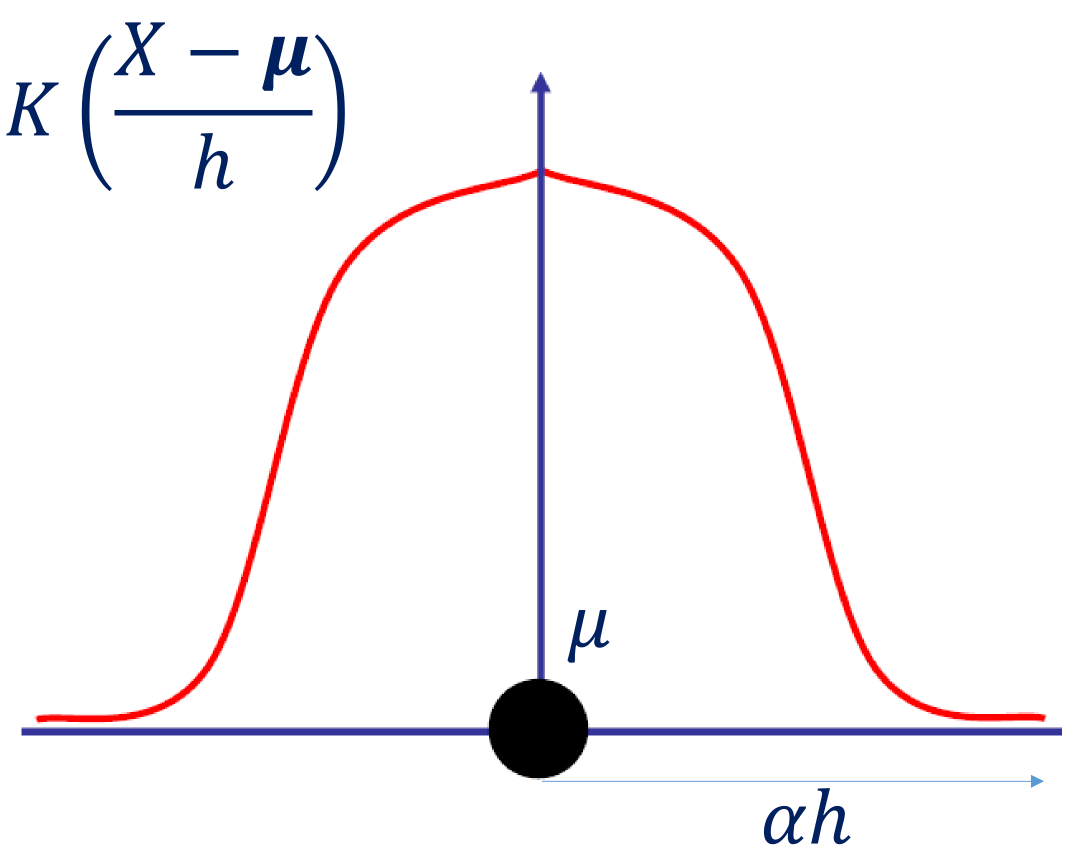


Distribution of Multivariate Data

$$P(X) = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{K}\left(\frac{X - X_i}{h}\right) = \frac{1}{n}\sum_{i=1}^{n} c\boldsymbol{k}\left(\left\|\frac{X - X_i}{h}\right\|^2\right)$$



Kernel Functions



Decays to ZERO As Moves Away From $\boldsymbol{\mu}$

Rate of Decay Controlled by Bandwidth h

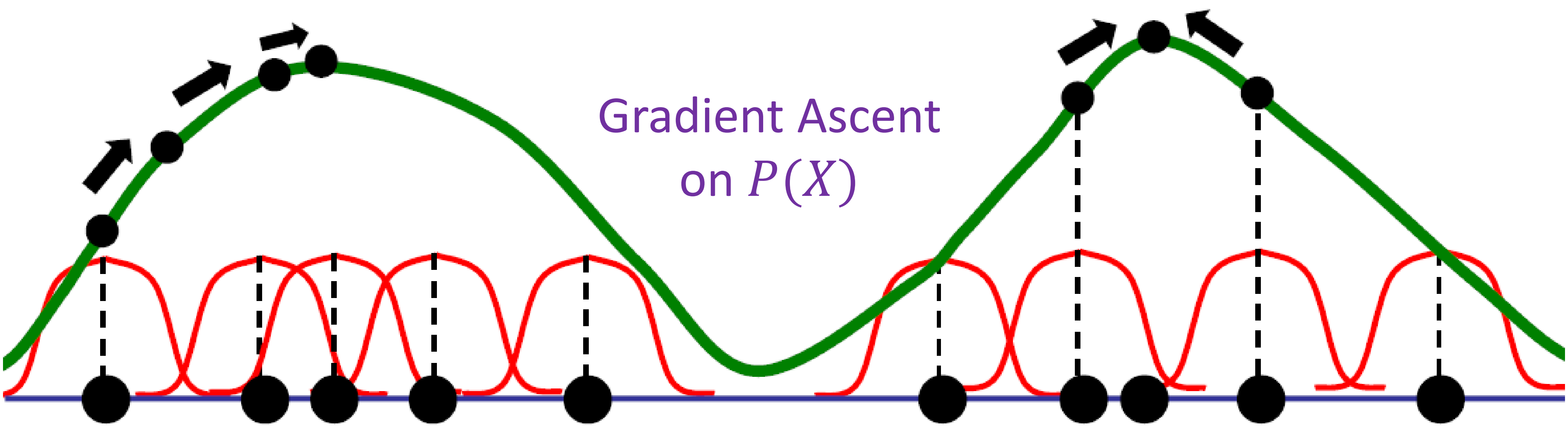
Maximum at $\boldsymbol{\mu}$

Symmetric around $\boldsymbol{\mu}$

$$\int_{\mathbb{R}^N} K\left(\frac{X-\mu}{h}\right) dX = 1$$

Seeking the Modes of $P(X)$

$$X^{(t+1)} = X^{(t)} + \eta_t \nabla_X P\left(X^{(t)}\right)$$



Mean-Shift Iterations

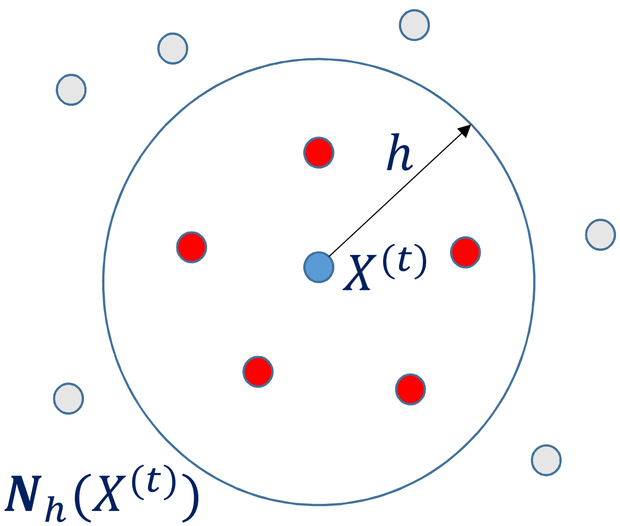
$$\nabla_X P(X) = \frac{2c}{nh^2} \sum_{i=1}^{n} (X_i - X)\,\boldsymbol{g}\left(\left\|\frac{X - X_i}{h}\right\|^2\right)$$

$$X^{(t+1)} = \frac{\sum_{i=1}^{n} X_i\,\boldsymbol{g}\left(\left\|\frac{X^{(t)} - X_i}{h}\right\|^2\right)}{\sum_{i=1}^{n} \boldsymbol{g}\left(\left\|\frac{X^{(t)} - X_i}{h}\right\|^2\right)}$$

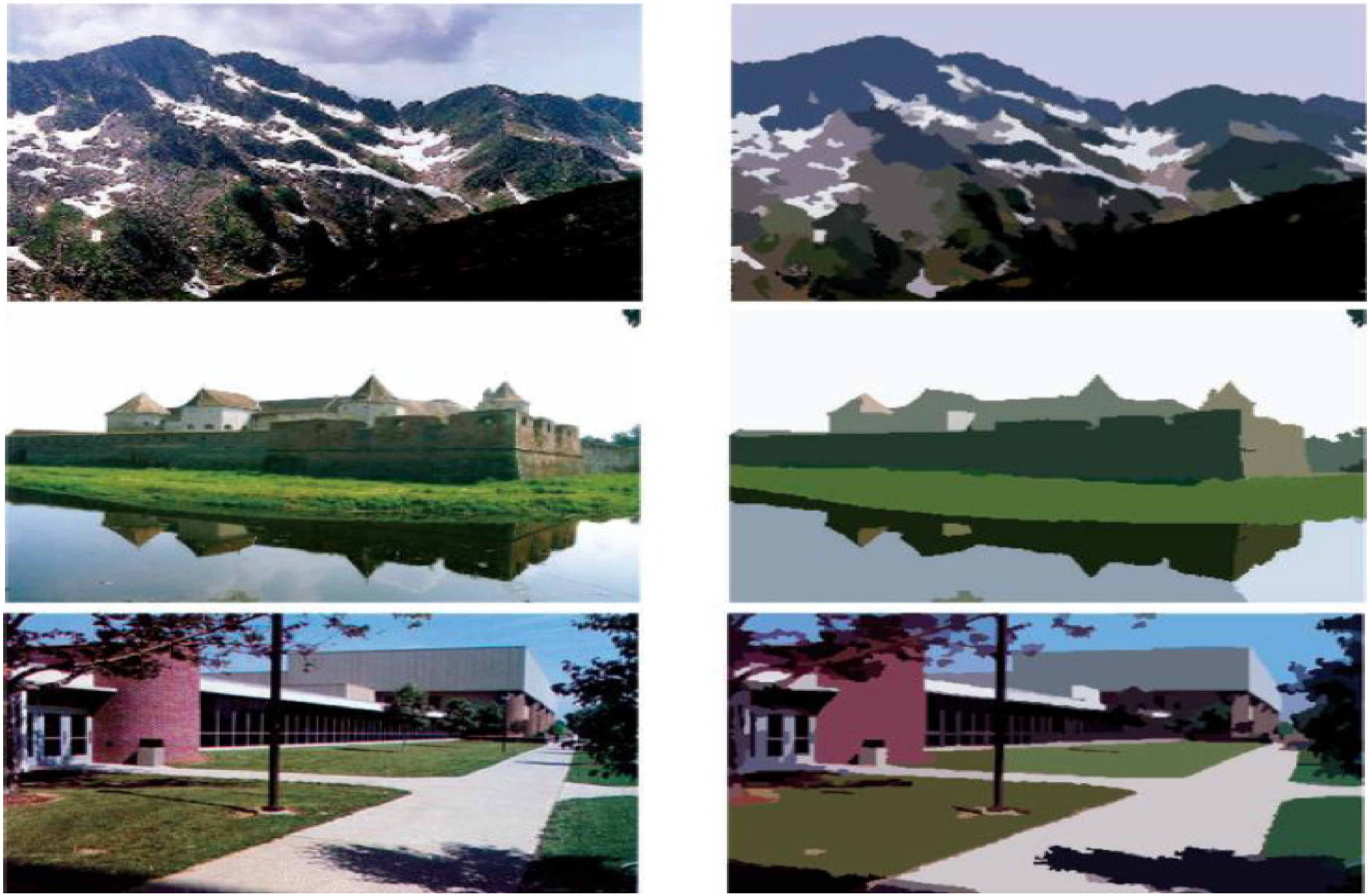
Input: $\boldsymbol{S} = \{X_i; i = 1, \dots n\}$ $\quad$ $X^{(0)} \in \boldsymbol{S}$

Seeking Modes of $P(X)$: Choice of Kernel

$$X^{(t+1)} = \frac{\sum_{i=1}^{n} X_i \boldsymbol{g_E}\left(\left\|\frac{X^{(t)} - X_i}{h}\right\|^2\right)}{\sum_{i=1}^{n} \boldsymbol{g_E}\left(\left\|\frac{X^{(t)} - X_i}{h}\right\|^2\right)}$$



$$X^{(t+1)} = \frac{\sum_{X \in \boldsymbol{N}_h(X^{(t)})} X \cdot 1}{\sum_{X \in \boldsymbol{N}_h(X^{(t)})} 1}$$

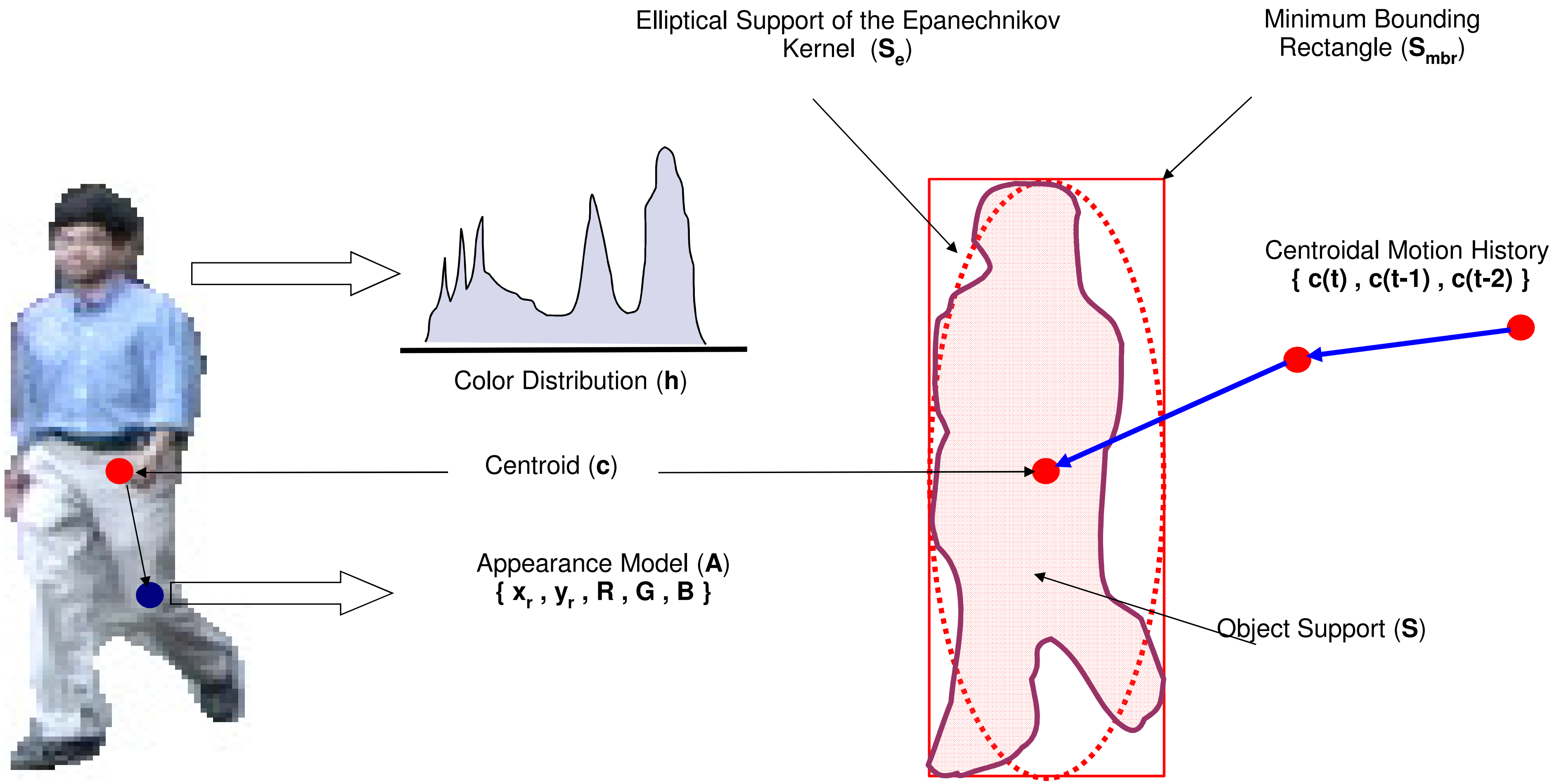



MONMOUTH 7
RACE
Stupeflix

Challenges in Visual Tracking

- Illumination Changes
- Object Appearance Changes
- Background Clutter
- Object Deformations & Articulations
- Motion Discontinuities
- Occlusions – Partial & Complete
- Camera Motions

Object Feature: Color Distribution



Tracking with Motion Model

$$s = s_0 + u\Delta t + \frac{1}{2} f \Delta t^2$$

Equation of Motion

$$x(t + \Delta t) = x(t) + (\Delta t)x'(t) \ + \frac{1}{2}(\Delta t)^2 x''(t)$$

Taylor's Series

Tracking with Motion Model

δt : Time Domain Sampling Interval

$$\frac{dp}{dt} \equiv \frac{p_n - p_{n-1}}{\delta t} \quad \Rightarrow \quad x'(t) \equiv \frac{x_n - x_{n-1}}{\delta t}$$

$$x''(t) \equiv \frac{\frac{x_n - x_{n-1}}{\delta t} - \frac{x_{n-1} - x_{n-2}}{\delta t}}{\delta t} = \frac{x_n - 2x_{n-1} + x_{n-2}}{(\delta t)^2}$$

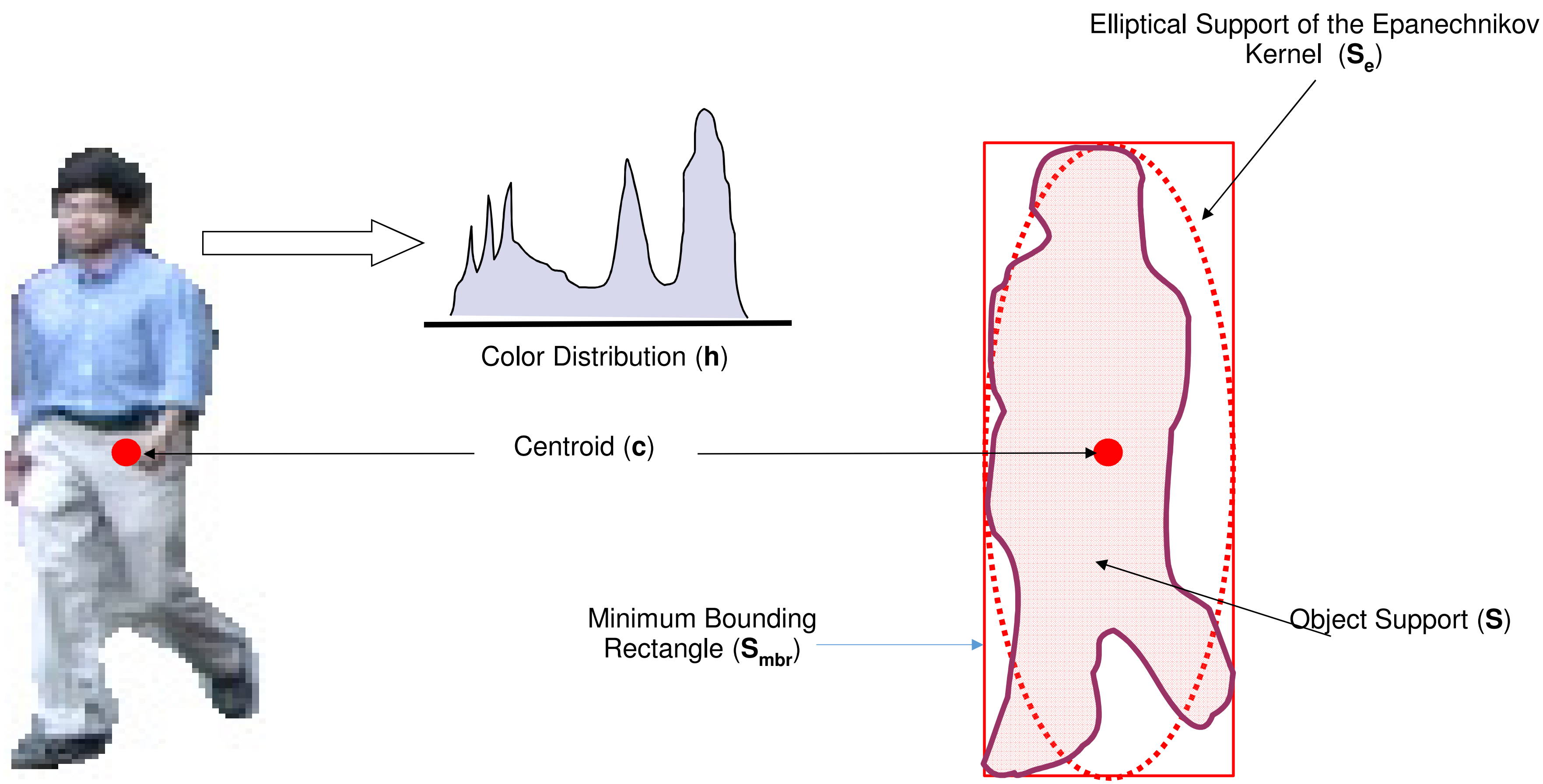
Tracking with Motion Model

$$x(t+\Delta t) = x(t) + (\Delta t)x'(t) + \frac{1}{2}(\Delta t)^2 x''(t)$$

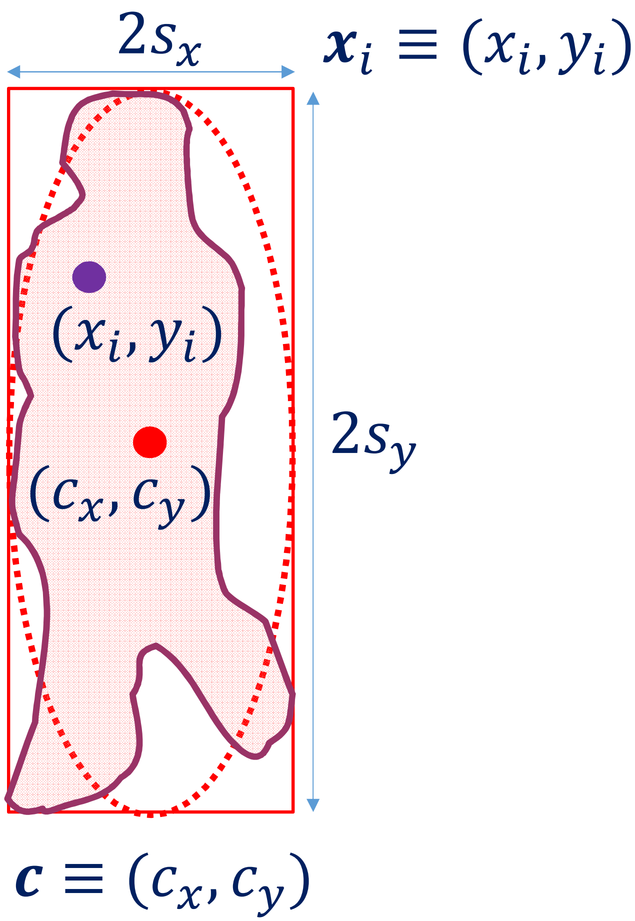
$$\hat{x}_n = x_{n-1} + (\delta t)\frac{x_{n-1} - x_{n-2}}{\delta t} + \frac{1}{2}(\delta t)^2 \frac{x_{n-1} - 2x_{n-2} + x_{n-3}}{(\delta t)^2}$$

$$\boxed{\hat{x}_n = 2.5x_{n-1} - 2x_{n-2} + 0.5x_{n-3}}$$

Tracking with Appearance



Epanechnikov Kernel on Bounding Box

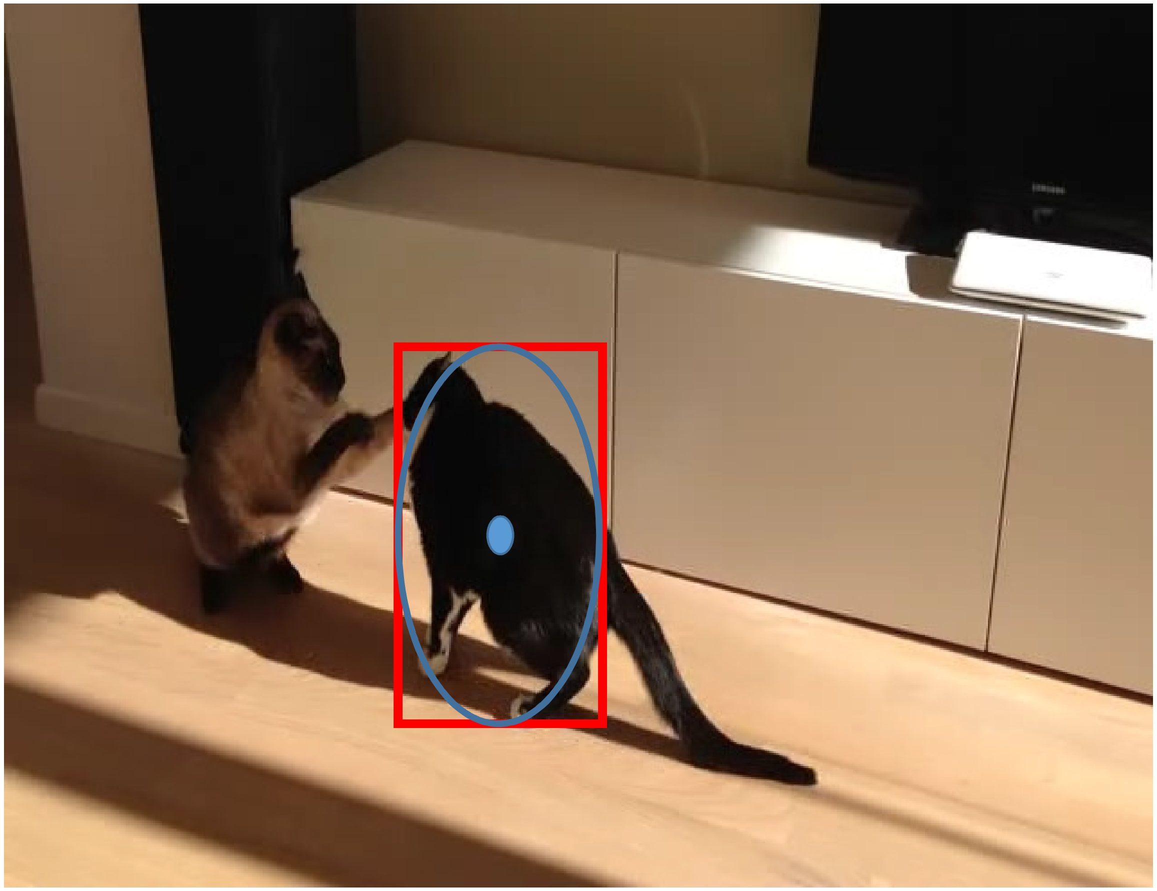


Epanechnikov Kernel Weight

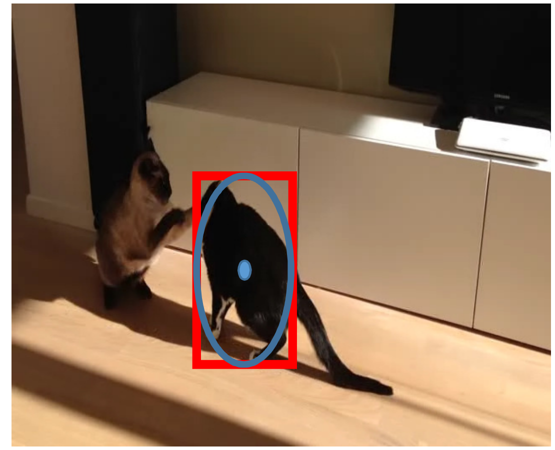
$$\omega(\boldsymbol{x}_i) = 1 - \left\{\left(\frac{x_i - c_x}{s_x}\right)^2 + \left(\frac{y_i - c_y}{s_y}\right)^2\right\}$$

$$\omega(\boldsymbol{c}) = \begin{cases} \omega(\boldsymbol{x}_i), \omega(\boldsymbol{x}_i) \geq 0 \\ 0, \qquad Otherwise \end{cases}$$

Target Appearance Model



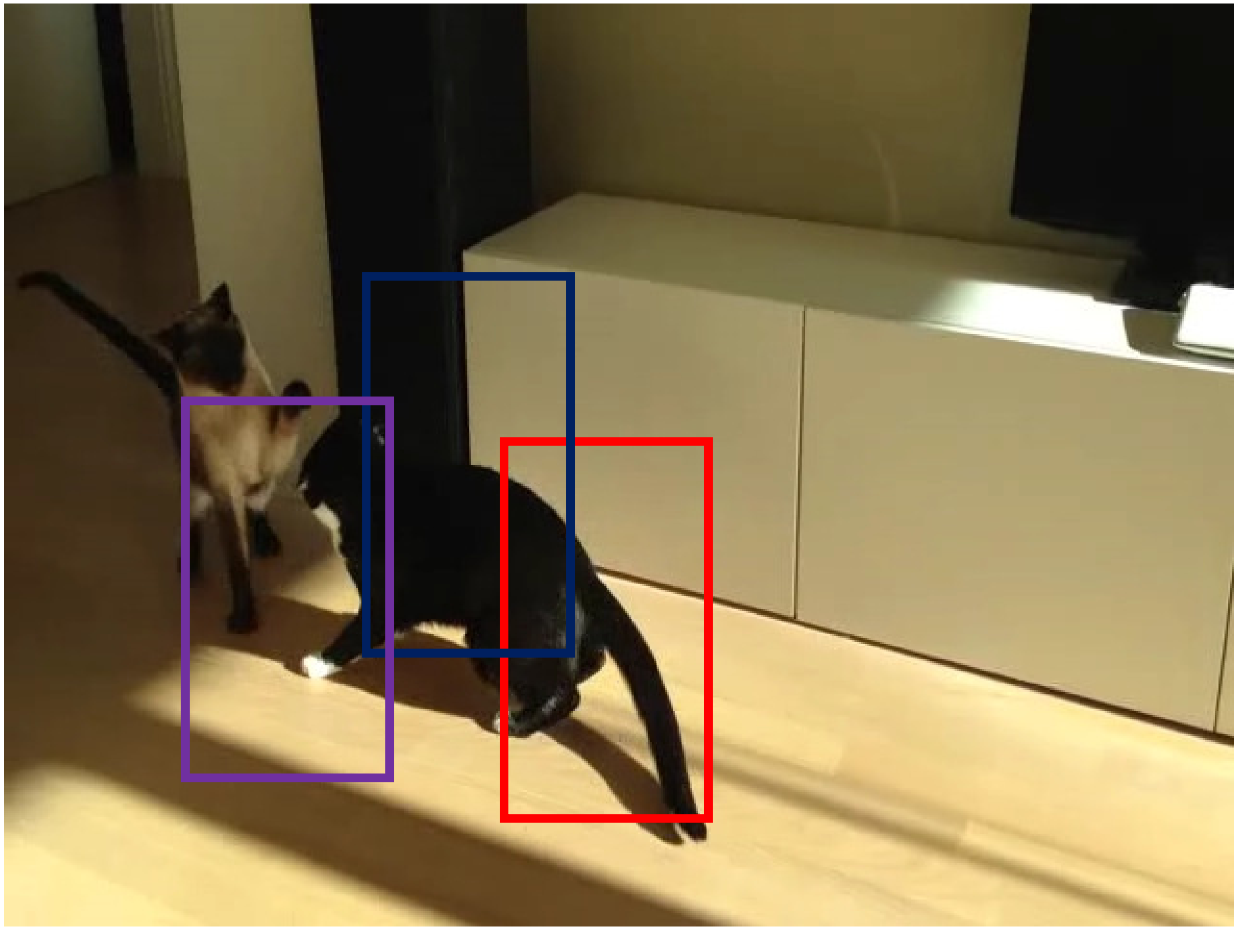
Target Appearance Model



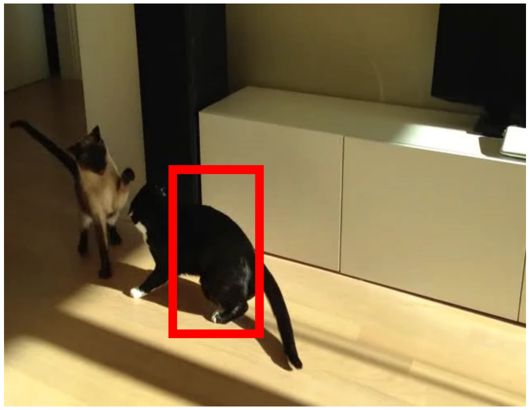
$$\boldsymbol{T}[j] = C_T \sum_{i=1}^{n} k\left(\left\|\frac{\boldsymbol{x}_i - \boldsymbol{c}}{h}\right\|^2\right) \delta[bI(\boldsymbol{x}_i) - \mathrm{j}\,]$$

$j = 0,1, \ldots (b_{pc}^3 - 1)$

Candidate Appearance Model



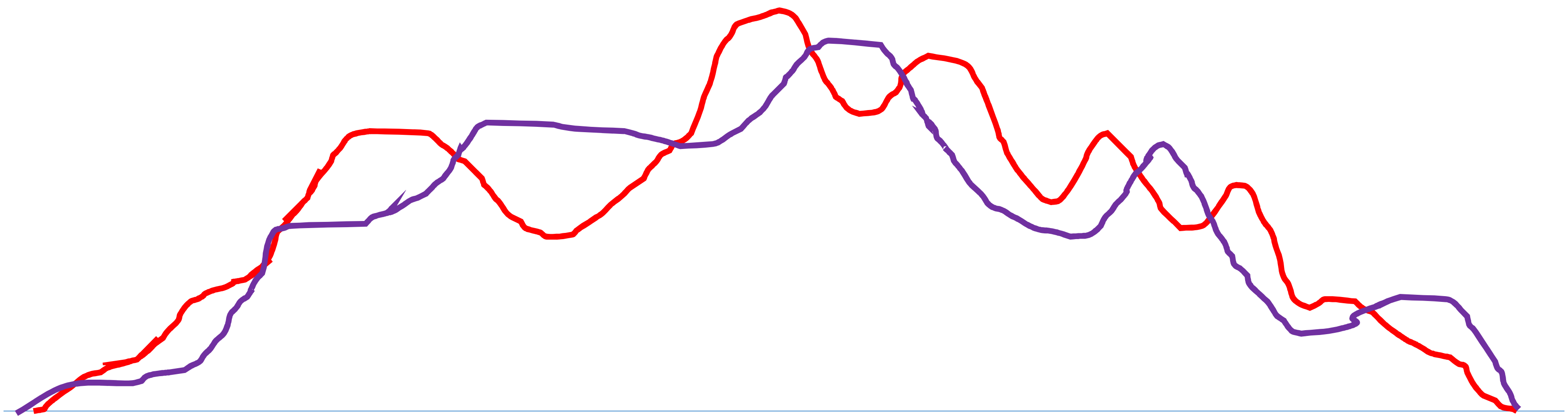
Candidate Appearance Model



$$\boldsymbol{Q}_{\boldsymbol{y}}[j] = C_Q \sum_{i=1}^{n} k\left(\left\|\frac{\boldsymbol{z}_i - \boldsymbol{y}}{h}\right\|^2\right)\delta[bI(\boldsymbol{z}_i) - \mathrm{j}\,]$$

$j = 0,1, \ldots (b_{pc}^3 - 1)$

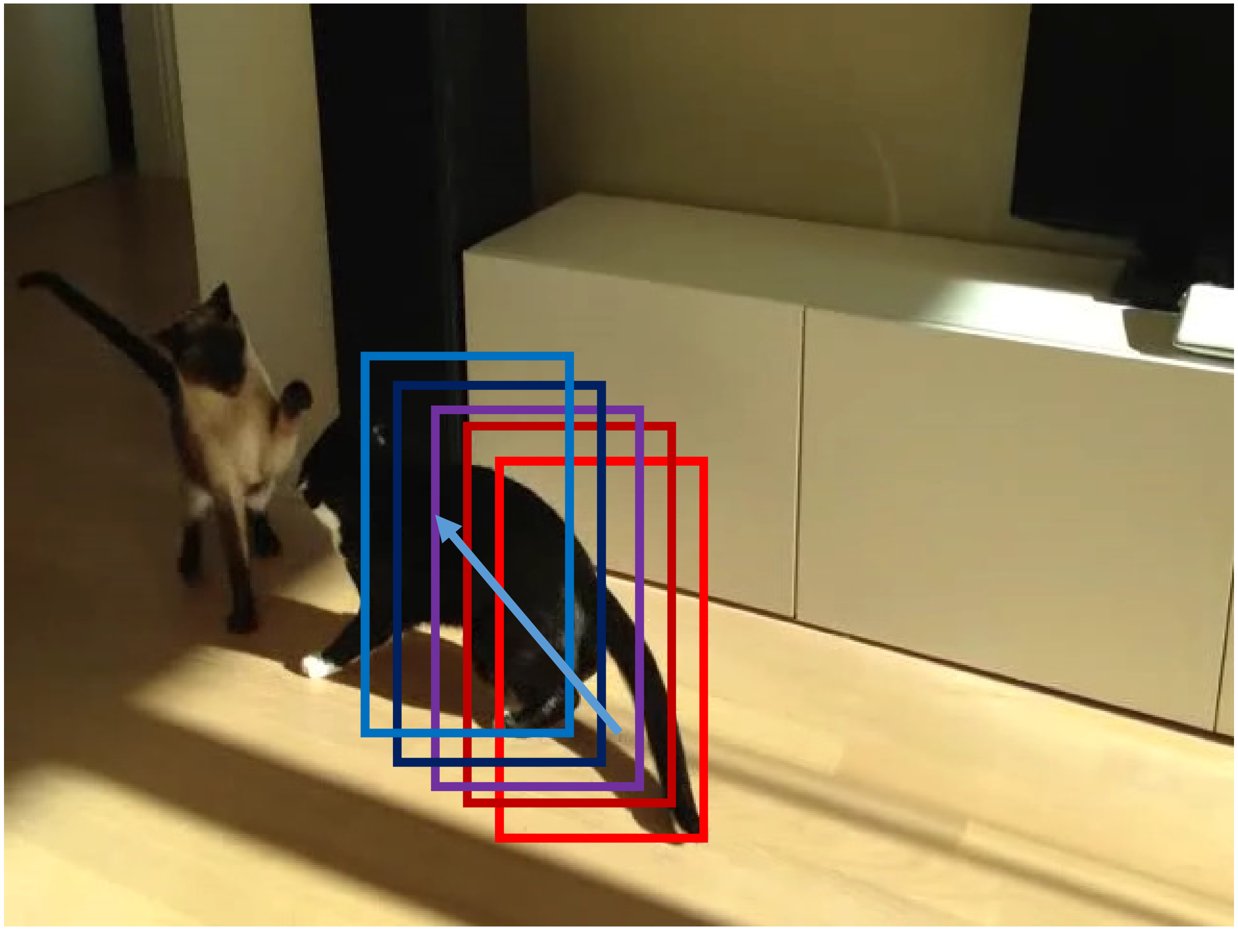
Comparing Distributions: Bhattacharya Coefficient



$$BC(P_1, P_2) = \sum_{j=0}^{m-1} \sqrt{P_1[j]P_2[j]}$$

$$0 \leq BD(P_1, P_2) \leq 1 \qquad BD(P_1, P_2) = 1 - BC(P_1, P_2)$$

Mean-Shift Tracking



Mean-Shift Tracking

$$\boldsymbol{BC}(\boldsymbol{T}, \boldsymbol{Q_y}) = \sum_{j=0}^{m-1} \sqrt{\boldsymbol{T}[\boldsymbol{j}]\boldsymbol{Q_y}[\boldsymbol{j}]}$$

Maximize Similarity Between $\boldsymbol{T}$ and $\boldsymbol{Q_y}$ to Localize the Target

$$\boldsymbol{y}^{(t+1)} = \boldsymbol{y}^{(t)} + \eta_t \nabla_{\boldsymbol{y}} \boldsymbol{BC}(\boldsymbol{T}, \boldsymbol{Q_y})$$

$m = b_{pc}^3$

Mean-Shift Tracking

$$w_i = \sum_{j=0}^{m-1} \delta[bI(\boldsymbol{x}_i) - j]\sqrt{\frac{\boldsymbol{Q_y}[\boldsymbol{j}]}{\boldsymbol{T}[\boldsymbol{j}]}}$$

$$\boldsymbol{y}^{(\boldsymbol{t}+\boldsymbol{1})} = \frac{\sum_{i=1}^{n} \boldsymbol{x}_i w_i \boldsymbol{g}\left(\left\|\frac{\boldsymbol{x}_i - \boldsymbol{y}^{(t)}}{h}\right\|^2\right)}{\sum_{i=1}^{n} w_i\, \boldsymbol{g}\left(\left\|\frac{\boldsymbol{x}_i - \boldsymbol{y}^{(t)}}{h}\right\|^2\right)}$$

Discussions

- Ruled Visual Tracking for a Decade!!!
- Very Fast If Used With Only Color
- Multiple Modifications are Proposed
- Works Well under Deformations & Partial Occlusions
- A Precomputed Mask Makes Tracking Faster
- Motion Model Initialization Improves Performance

Summary

- Video Surveillance System
- Image Pixels and Color Distribution
- Change Detection
- Unimodal Background Model
- Mean-Shift Tracking

Thank You