NonStationary Shape Activities: Tracking & Abnormality Detection

Namrata Vaswani (namrata@iastate.edu)

Dept. of ECE, Iowa State University

http://www.ece.iastate.edu/∼namrata

and

Rama Chellappa (rama@cfar.umd.edu)

Dept. of ECE, Univ. of Maryland, College Park

http://www.cfar.umd.edu/∼rama

Outline

- **Main Idea**
- **Landmark Shape Dynamical Model for an "Activity"**
  - **Tracking**
- **Abnormal Activity Detection: Change Detection**
- **Types of Shape Activity Models & Applications**

Problem Formulation

- Modeling activity performed by a group of moving and interacting point "objects" ("landmarks").
- "Landmarks": People, Vehicles, Robots, Human body parts,...
- **Changing configuration of the group: moving & deforming shape**
- **"Shape Activity": model activity performed by a group of moving & interacting "objects" by its shape dynamics**
- **"Abnormal Activity": change in learned shape dynamical model, which could be slow or sudden and whose parameters are unknown**

Landmark Shape

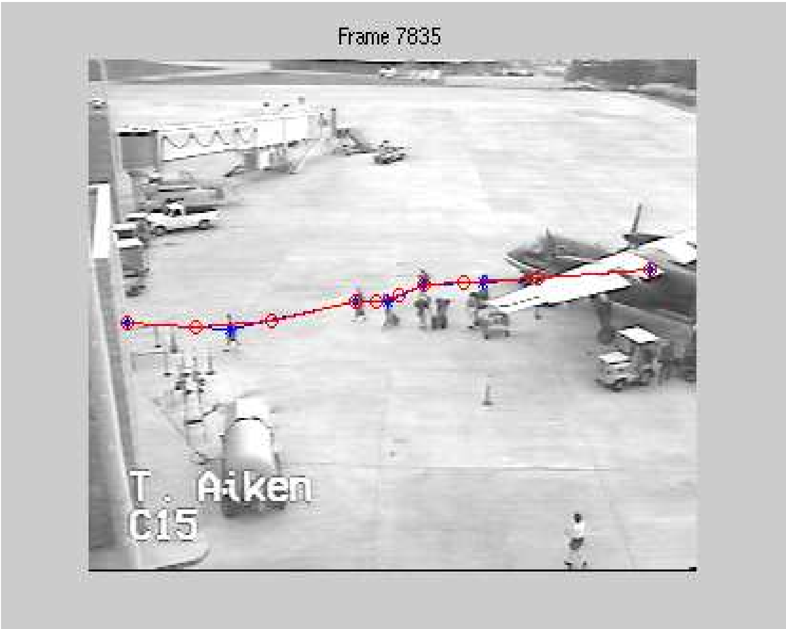
- **Shape: geometric information that remains when location, scale & rotation effects are filtered out [Kendall]**

- **Shape of k landmarks in 2D**
  - Represent the X and Y coordinates of the k points as a k-dimensional complex vector: **Configuration**
  - Translation Normalization: **Centered Configuration**
  - Scale Normalization: **Pre-shape**
  - Rotation Normalization: **Shape**
- **Landmarks in 3D: represent by a $k \times 3$ matrix**

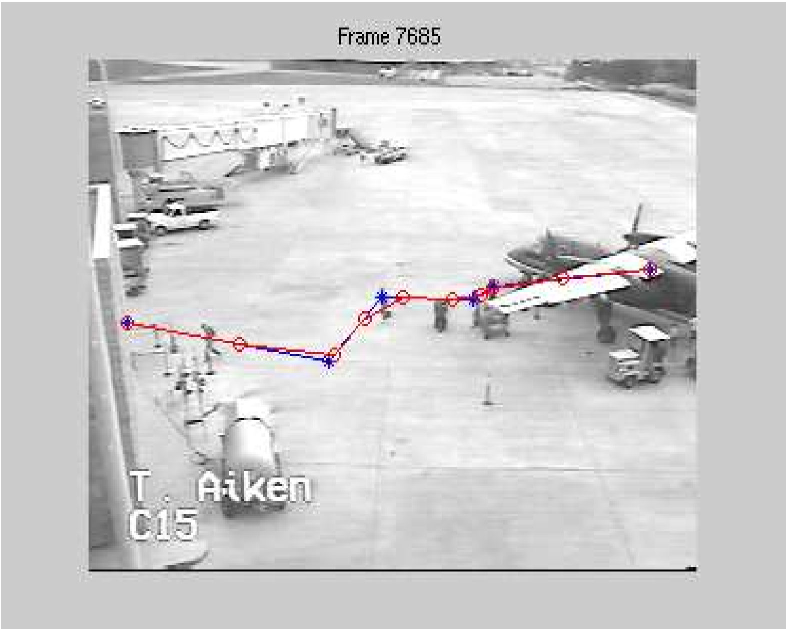
Example: Group of Robots

Example: Group of Passengers Deplaning

Each person forms a "landmark"



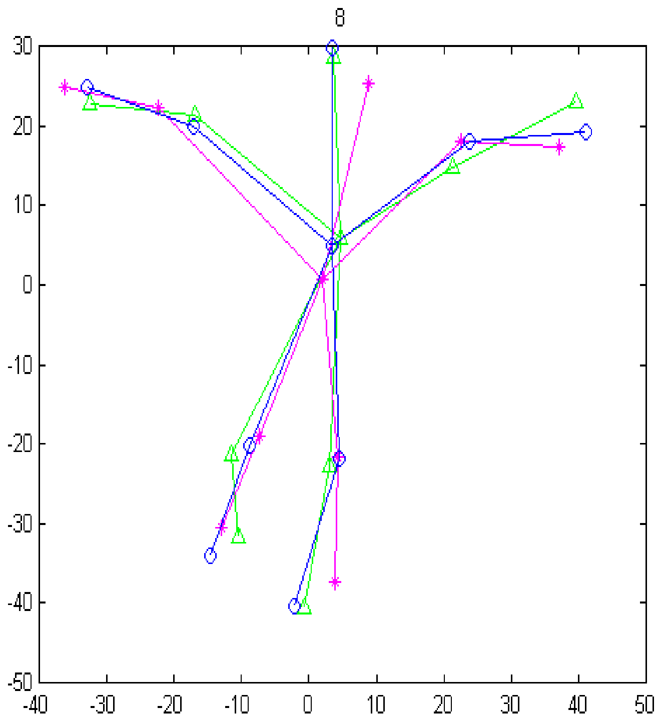
A 'normal activity' frame



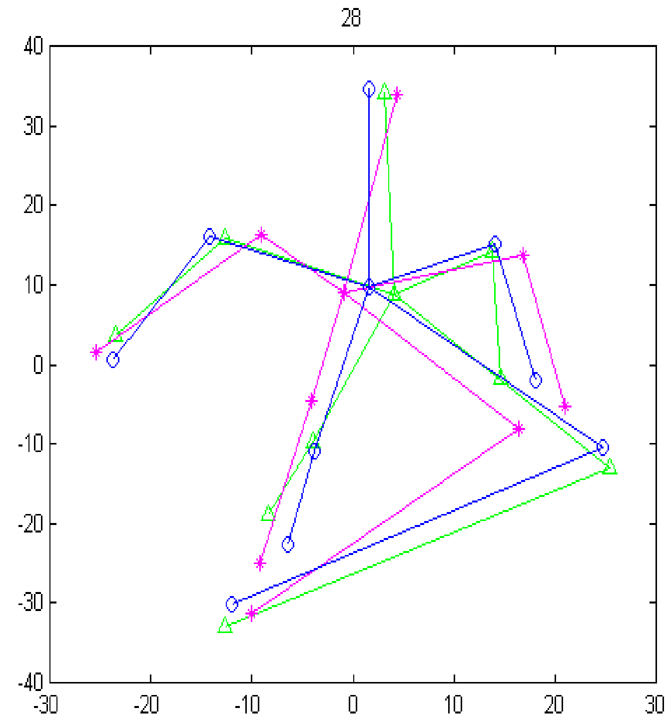
Abnormality

Human Actions

Each body part forms a landmark



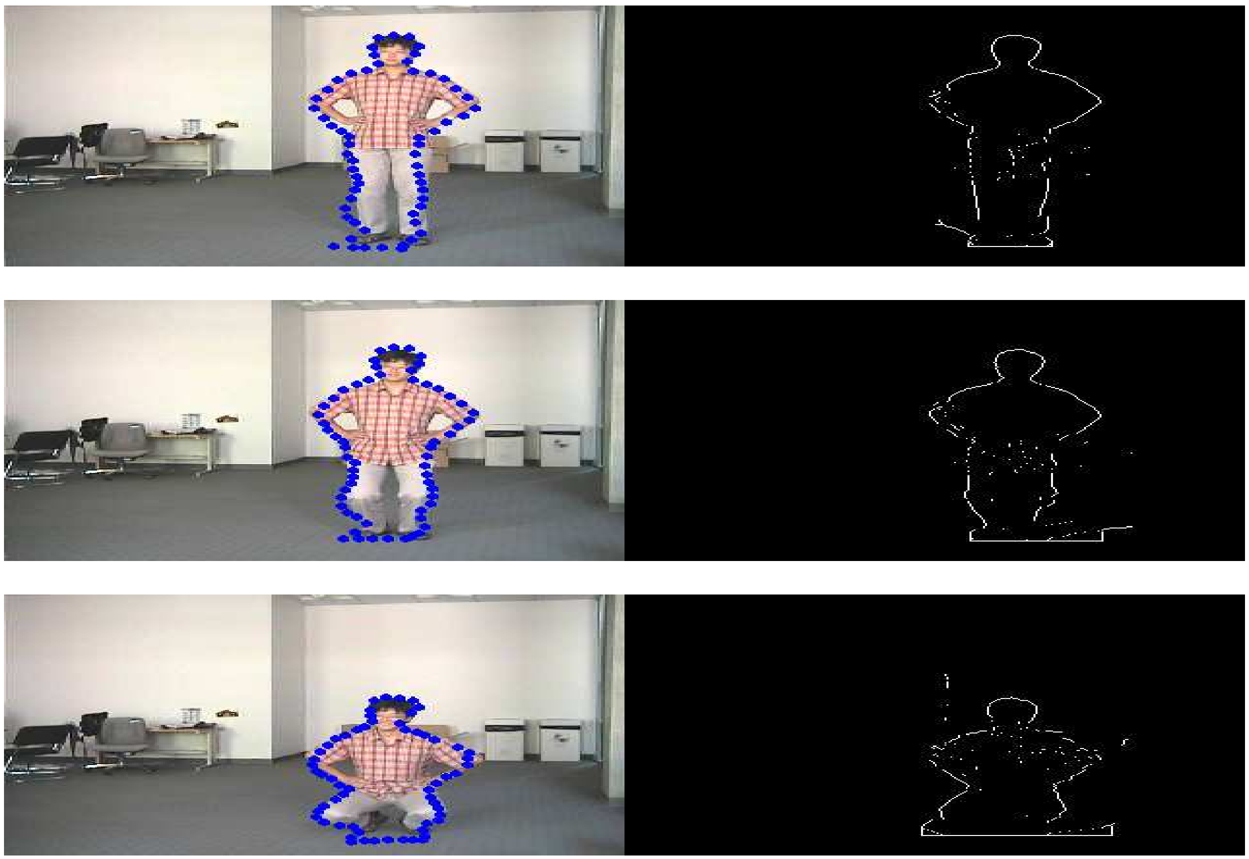
Normal action



Abnormality

Human Actions (ongoing work)

Uniformly sampled points on outer contour form landmarks



Motivation

Track global motion (scale/rotation/translation in 2D) & shape. Use shape dynamics to represent/recognize activity

- **Modeling & recognition invariant to global motion, e.g.:**
  - **Global scale change of activity**, e.g. person taller/shorter
  - **Scaled orthographic camera motion. Valid model for:**
    * Distant PTZ camera rotated to align with line of sight
    * Random jitter of UAV looking down at activity
    * Activity close to any camera's principal axis, little out-of-plane rotation
- **Use estimated global motion to control a PTZ camera or a UAV to "follow" a "moving" activity**

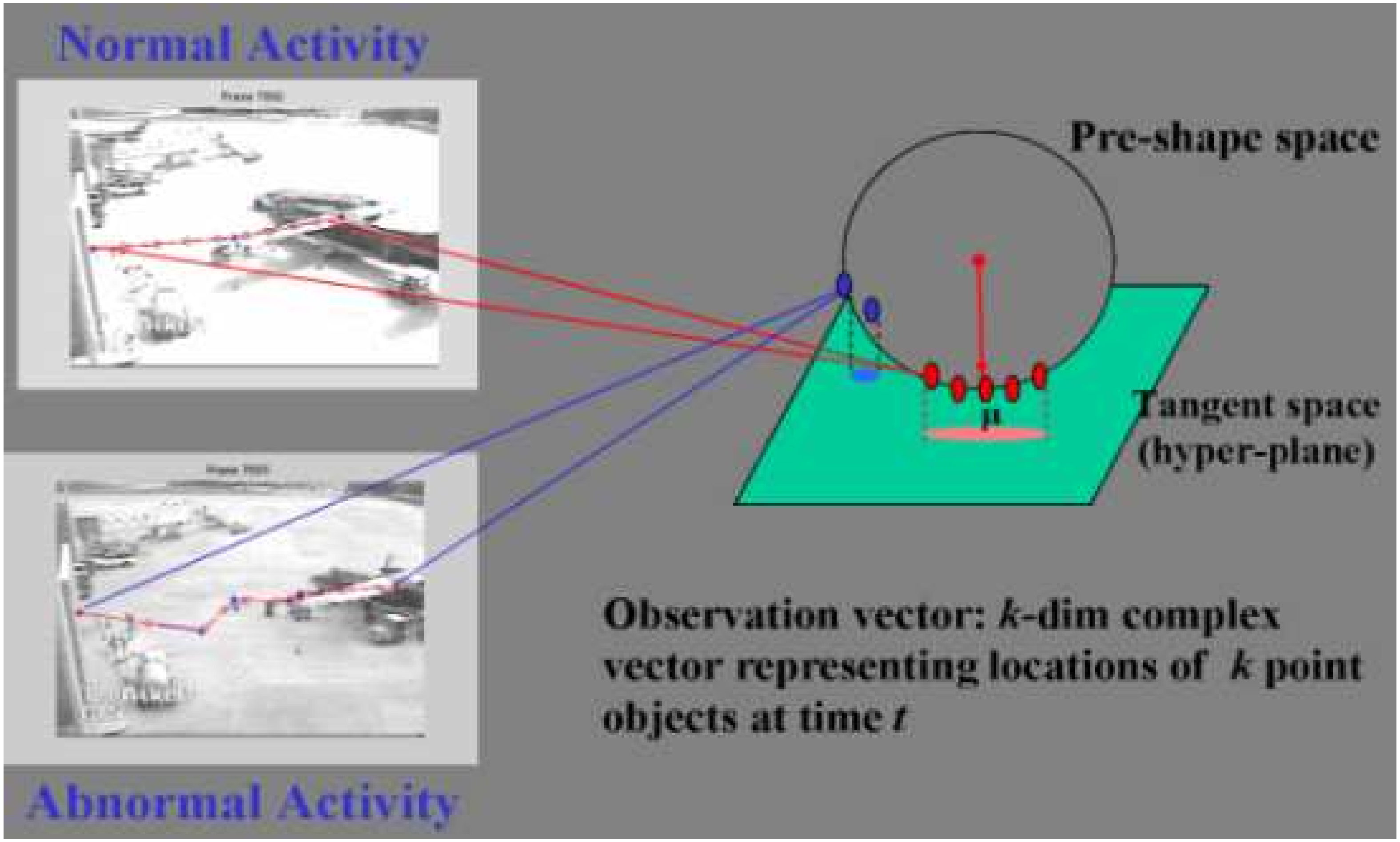
A Common Framework for...

- **Tracking Groups of Moving/Interacting "Objects"**
- **Abnormal Activity Detection & Tracking**
  - Suspicious behavior detection (people groups), Lane change detection (vehicle groups), Abnormal human action detection
- **Sequence Identification & Tracking**
  - Sequence of human actions, track & summarize video
- **Activity Segmentation & Tracking**
  - Video coding + summarization, Unsupervised learning
- **Sensor independent approach**
  - Observations may be obtained using any sensor, e.g. audio, infra-red, radar, fuse different sensors

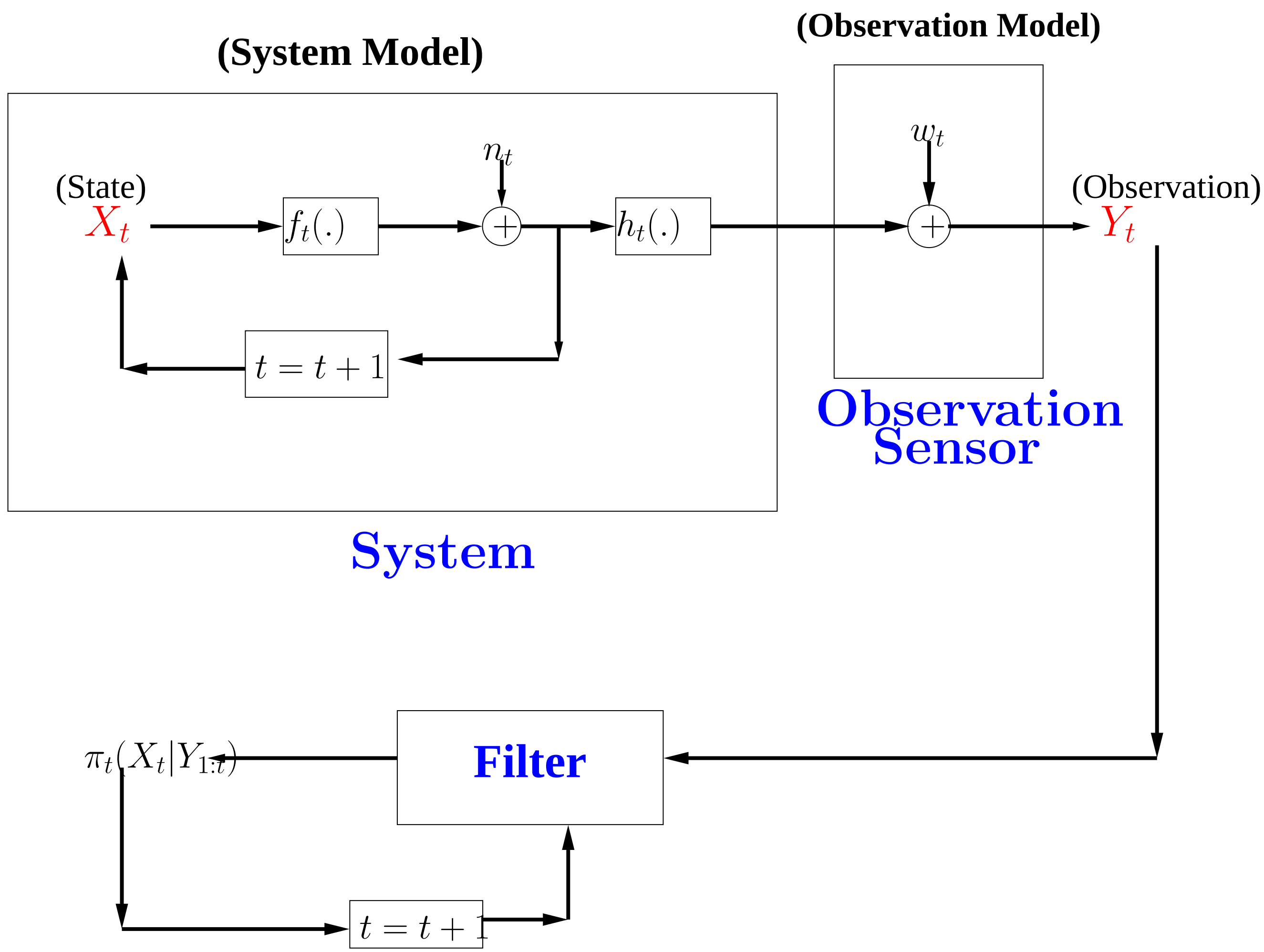
Existing Work

- **Joint tracking and event recognition**
  - DBN (or FS-HMM) tracked using a Rao-Blackwellized PF, e.g. Condensation for gesture tracking/recognition, figure tracking/recognition, traffic monitoring
  - Assume p.w. constant mode, sample from prior on mode, compute posterior, e.g. [Zhou et al]
- **Tracking groups of moving/interacting objects**, e.g. data association (JPDAF), Condensation, robot formation control.
- **Activity/Action Recognition**, e.g. space-time shapes, shape based factorization, view invariant approaches, multiple levels of zoom, DBN, co-occurrence statistics: tracks obtained by other means.

Defining Dynamics in Shape Tangent Space



State Space Model, Tracking



- **Observation (Y_t):**
  Observed object locations after centering
- **State, X_t:**
  [Shape z_t, "Velocity Coefficients" c_t, Log Scale s_t, Rotation θ_t]

- **Observation Model:**
  Observation = h_t(Shape,Scale,Rotation) + observation noise

$$h_t(X_t) = z_t e^{s_t + j\theta_t}$$

  – Can use edge image as observation as in Condensation - incorporates clutter probability (ongoing work)
- **System model, f_t: Dynamics of shape, scale, rotation**
  – **Gauss-Markov model on shape "velocity coefficients"**

$$c_t = A_c c_{t-1} + n_{c,t}$$

– **Parallel transport c_t to tangent space at z_{t-1}, $T_{z_{t-1}}$**

$$\text{“velocity”}: \quad v = U(z_{t-1})c_t, \quad U = \text{basis for } T_{z_{t-1}}$$

– **Move on current tangent plane by “velocity” and project back to shape space: shape at next time**

$$z_t = \sqrt{1 - v^* v}\; z_{t-1} + v$$

– **Random walk model for log scale, rotation, translation (if needed)**

- **Goal: Track observed landmark locations to estimate posterior shape & shape velocity distribution, $\pi_t(x_t|Y_{1:t})$**
  – **Use a particle filter: computationally efficient & provably stable (for large N) solution for nonlinear, multimodal, large dim state tracking.**

Particle Filter (PF) [Gordon et al'93]: Basic Idea

- **Sequential Monte Carlo method, approx. true filter as number of Monte Carlo samples ("particles"),** $N \to \infty$

- Given π^N_{t-1}, **perform importance sampling & weighting,** followed by **resampling** to approx. the Bayes' recursion to get π^N_t

$$\pi^N_{t-1} \xrightarrow[\text{Imp. Samp.}]{\tilde{x}^i_t \sim q_t} \pi^N_{t|t-1} \xrightarrow[\text{Weight}]{Y_t \,\downarrow\; w^i_t \propto g_t(Y_t|\tilde{x}^i_t)} \bar{\pi}_t \xrightarrow[\text{Resample}]{} \pi^N_t$$

- Using $q_t(x_t|x^{(i)}_{t-1})$ as importance sampling density at t

Abnormal Activity Detection: Change Detection

- **“Normal Activity”: Modeled as a landmark shape dynamical model**
  - **Partially observed system(observations are noisy & nonlinear functions of state), satisfying HMM property**
- **“Abnormality”: Change w.r.t. learned shape dynamics**
  - **Parameters of changed system unknown**
  - **Change can be slow or sudden**
- **Detect changes in shape using the PF estimate of posterior of shape and/or shape velocity**

Slow v/s Sudden Change

- **Slow change: small change magnitude per unit time, "tracked" by the tracker.**
  - Error b/w estimate of posterior using the tracker with unchanged system model and the true posterior is small.

- **Sudden change: mostly "filtered out" ("loses track")**
  - Duration much smaller than "response time" of filter.

- **Quantify "rate of change", r, w.r.t. a filter: For an additive change with magnitude b per unit time, $r^2 = b^T \Sigma_{sys}^{-1} b$.**

Existing Work

Abnormal activity detection provides the problem definition: **Given the observations $Y_1, Y_2, ...Y_t$, detect, as quickly as possible, if a change occurred in the dynamics of the state X_t**

- **Change parameters unknown**
  - **Cannot use CUSUM** (or its modifications [Azimi et al]).
  - **Generalized CUSUM** intractable [Andrieu et al'04].
  - **Residue statistics [Basseville] for fault detection**, e.g.
    - Tracking Error (TE) [Bar-Shalom]
    - negative log of Observation Likelihood (OL)
    - Score function [Basseville]
- **Slow or sudden change**
  - TE, OL, Score function **miss slow changes**

Notation

- **Prior state distribution:**
  Given no observations, $X_t \sim p_t(.)$

- **Superscripts: ^{0} (unchanged system), ^{c} (changed system)**
  e.g. $X_t^0 \sim p_t^0(.)$, $\quad X_t^c \sim p_t^c(.)$

- **Prediction distribution:**
  For $\tau < t$, $\quad X_t | Y_{1:\tau} \sim \pi_{t|\tau}(.)$

- **Posterior (or Filtering distribution):**
  For $\tau = t$, $\quad X_t | Y_{1:t} \sim \pi_{t|t}(.) \stackrel{\triangle}{=} \pi_t$

Slow change detection, Unknown parameters

- **Fully observed state (no observation noise & h_t^{-1} exists):**
  - **negative Log Likelihood of state of unchanged system**

$$-\log p_t^0(X_t) = -\log p_t^0(h_t^{-1}(Y_t))$$

    * Most commonly used when have a set of i.i.d. observations, compute average LL, e.g. [Kulhavi,CDC'00]

- **Partially observed state (significant observation noise):**
  - Why not use Min. Mean Square Error estimate of this ?
  - **Our detection statistic is exactly this MMSE estimate [Vaswani,ACC'04]:**

$$\mathbf{ELL(Y_{1:t})} \triangleq \mathbb{E}[-\log \mathbf{p_t^0(X)|Y_{1:t}}]$$

Computing ELL

- **Linear and Gaussian system model**:

$$X_0 \sim \mathcal{N}(x; 0, \sigma_0^2), \quad X_t = AX_{t-1} + n_t, \quad n_t \sim \mathcal{N}(0, \sigma_n^2)$$

  - $A < 1$ & $\sigma_0^2 = \frac{\sigma_n^2}{1-A^2}$ **(stationary):** $p_t^0(x) = \mathcal{N}(x; 0, \sigma_0^2)$

$$-\log p_t^0(X) = \frac{X^2}{2\sigma_0^2} + const$$

$$ELL(Y_{1:t}) = \frac{1}{N}\sum_{i=1}^{N} \frac{(x_t^i)^2}{2\sigma_0^2}, \quad x_t^i \sim \pi_{t|t}(x)$$

  - $A = 1$ **(nonstationary)**: $ELL(Y_{1:t}) = \frac{1}{N}\sum_{i=1}^{N} \frac{(x_t^i)^2}{2(\sigma_0^2 + t\sigma_n^2)}$
    * **Problem: variance of p_t^0 increases with t: longer to detect a change for large t**

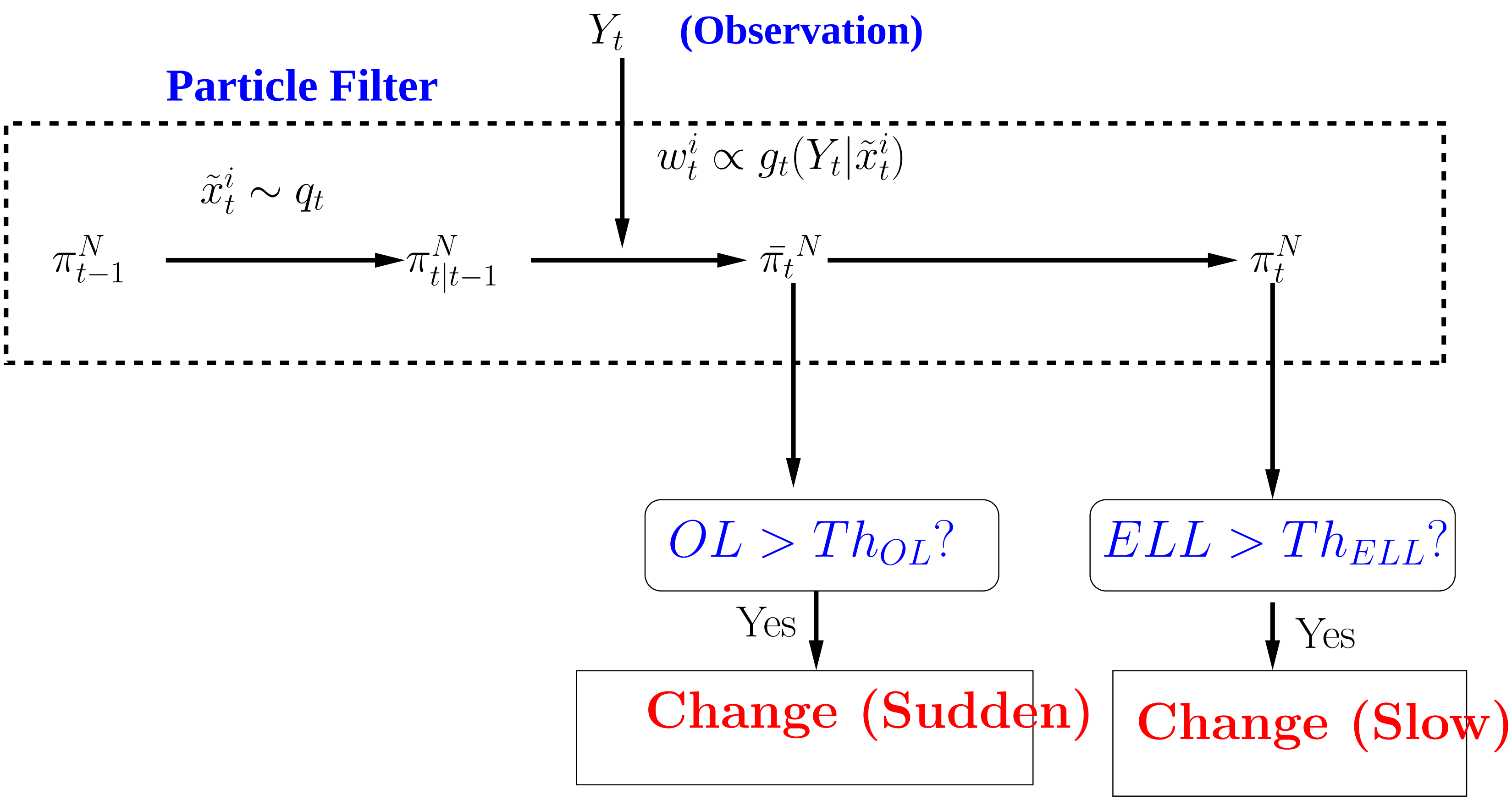
- **Nonlinear, Gaussian system:** linearize system model equation at each t, to get a Gaussian approx. to p_t^0
- **Training sequence available:** learn a p.w. constant $p_t^0(x)$

- **Handle increasing variance: Replace p_t^0 by Δ-step ahead prediction, $\pi^0_{t|t-\Delta}$**
  - **Variance bounded, Use to detect multiple changes**
  - Approx $\pi^0_{t|t-\Delta}$ as:
    * Approx. PF estimate of $\pi^0_{t-\Delta|t-\Delta}$ by a Gaussian mixture
    * Apply linearized system model Δ times to approx $\pi^0_{t|t-\Delta}$

- **Detection Threshold:** $\mathbf{Th_{ELL}} = \mathbb{E}_{\mathbf{Y^0_{1:t}}}[\mathbf{ELL^0}] + \mathbf{k}\sqrt{\mathbf{Var(ELL^0)}}$

$$\mathbb{E}_{\mathbf{Y^0_{1:t}}}[\mathbf{ELL^0}] = \mathbf{h(p_t^0)} = \text{differential entropy of } X_t^0$$

Change Detection Algorithm [ACC'04]



ELL v/s OL (or TE): Slow & Sudden Changes

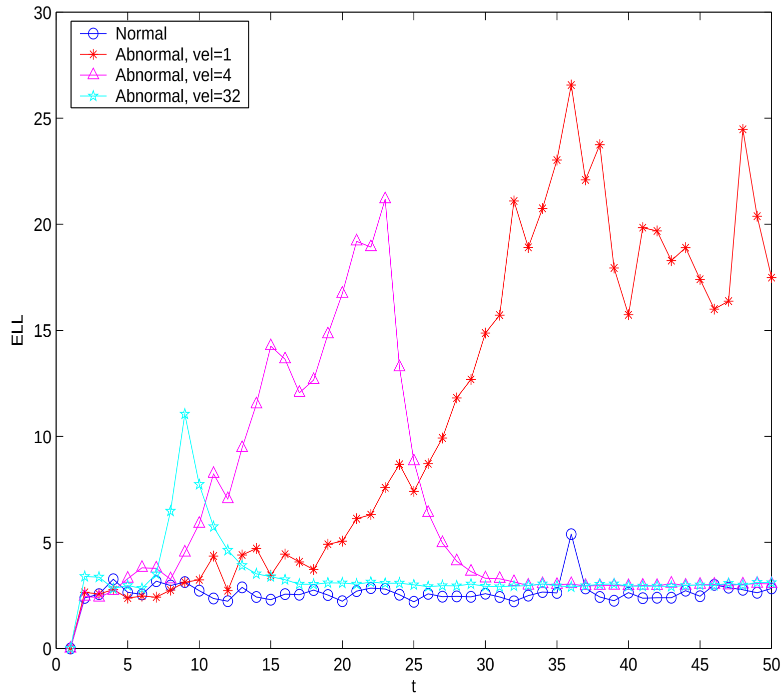
- **OL (or TE) rely on loss of track to detect a change**
- **ELL detects based on "tracked part of the change"**
- **ELL detects change before loss of track: very useful**
- **Slow Change:**
  - PF: stable under mild assumptions, tracks slow change well
  - **Loss of track small: OL, TE fail or take longer**
  - Estimated posterior close to true posterior of changed system
  - **ELL detects as soon as change becomes "detectable"**
- **Sudden Change: PF loses track**
  - **OL (or TE) detect immediately, ELL fails/takes longer**

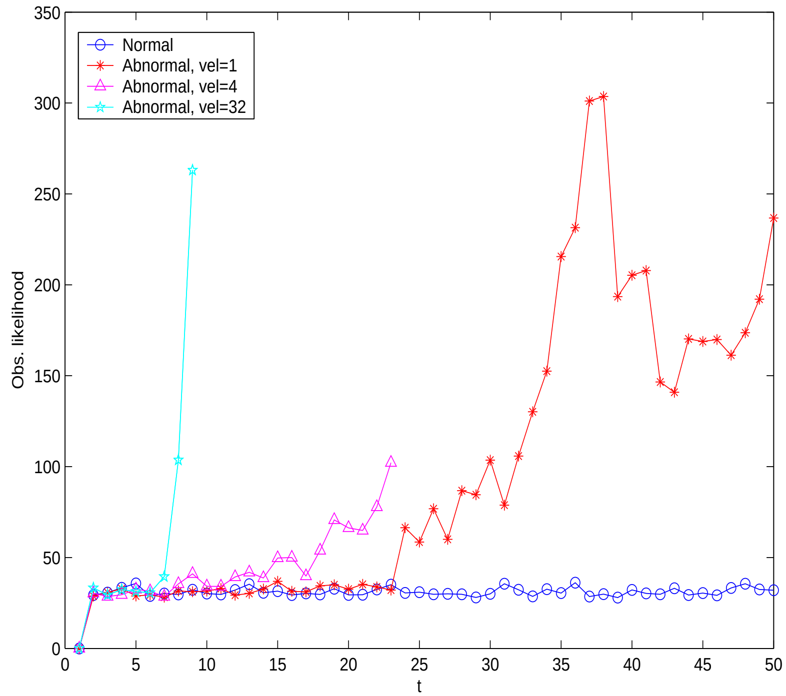
NonStationary Shape Activity Models

- **Full NonStationary SA Model (discussed earlier):**
  - **Track & detect abnormality, Sequence id, Activity segmentation**
  - Markov model on shape velocity: “moves” current shape
- **Simpler Special cases:**
  - **Strict Sense Stationary SA(SSA) & Constant Mean SSA (CMSA):** Abnormality Detection [Trans. IP, Oct’05]
  - **Piecewise CMSA: Activity Sequence Id**
    * Slow mean shape change: approx. as piecewise constant
    * Sequence of CMSAs with nonstationary transition period

Group of People: Abnormality Detection Using SSA

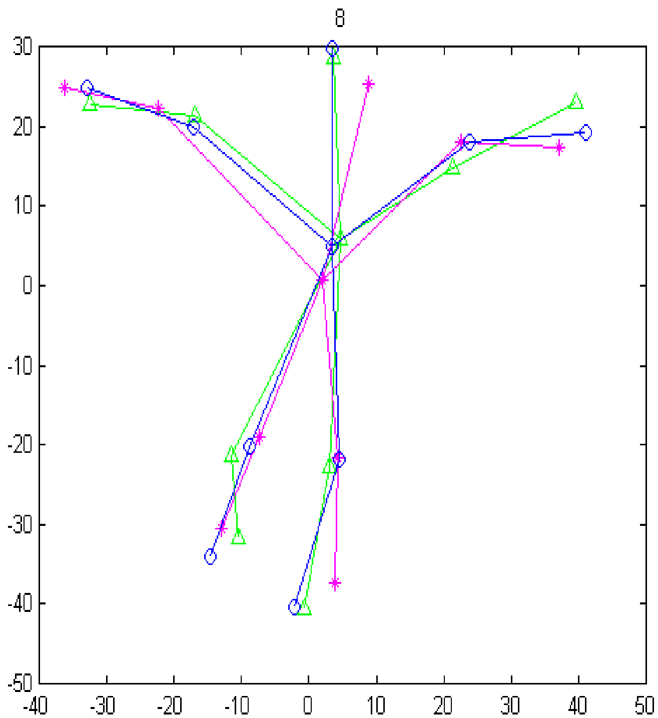
- Abnormality (one person walking away) begins at $t = 5$.
- **ELL detects abnormality faster than OL**



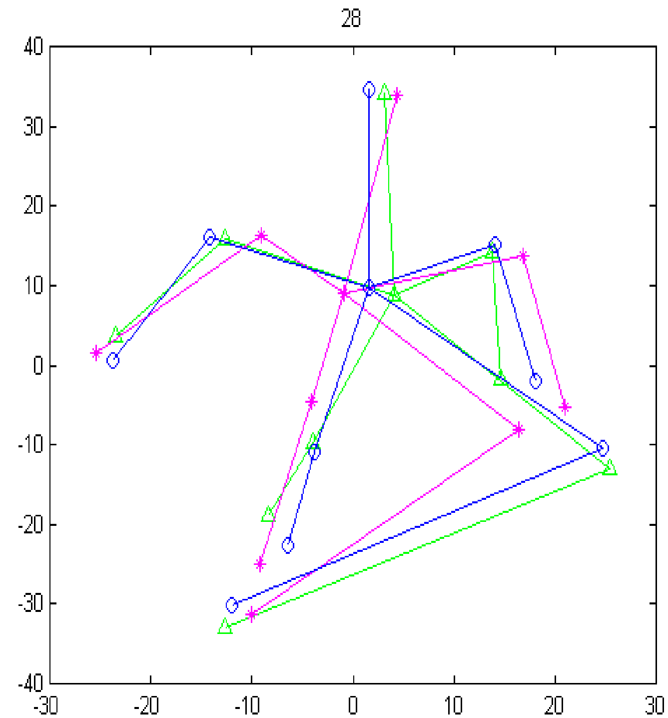


ELL **OL**

Human Actions: Tracking Using NSSA



Normal action

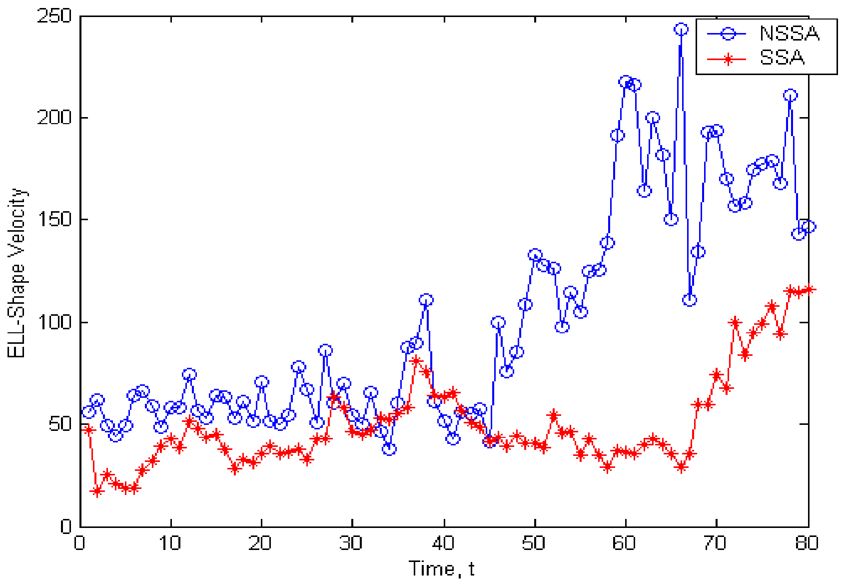


Abnormality Tracked

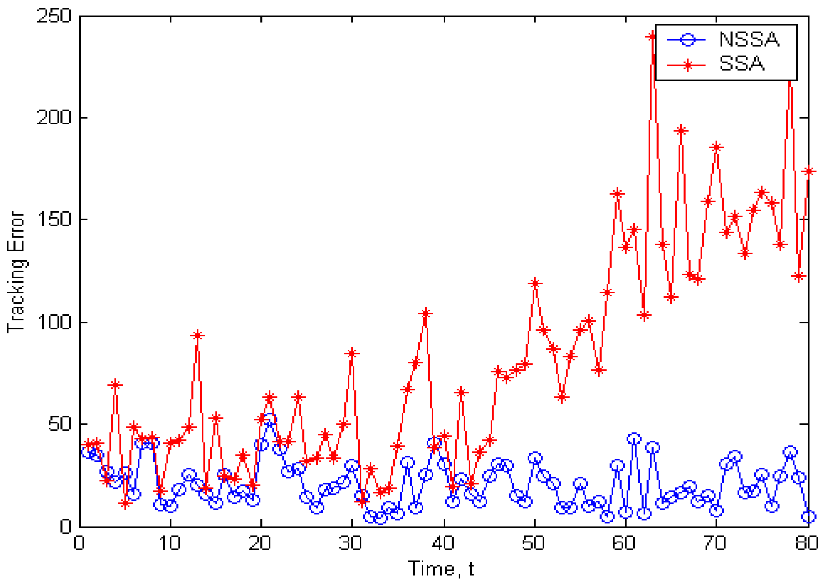
Green: Observation, Blue: Ground Truth, Magenta: Tracked

Human Actions: Abnormality Detection Using NSSA,SSA

- Abnormality begins at $t = 20$
- SSA cannot track, only detects using TE
- NSSA detects using ELL & also does not lose track



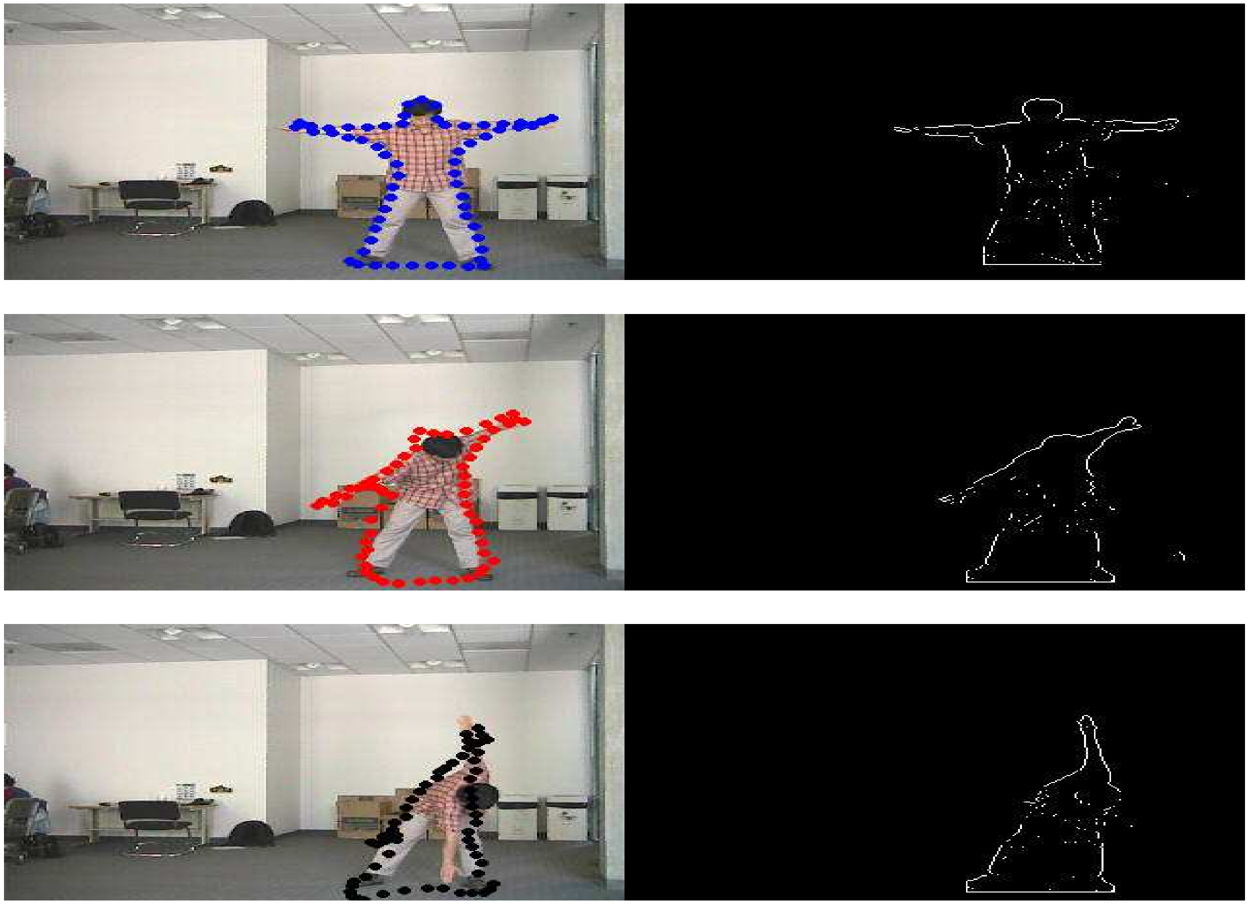
ELL



Tracking Error (TE)

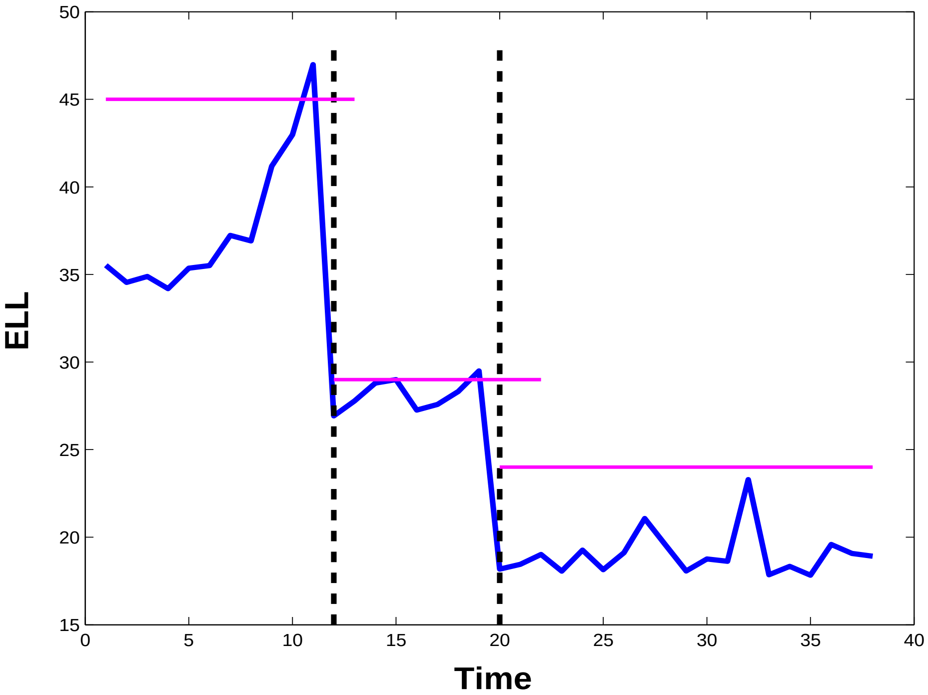
Human Actions: Tracking a Sequence Using PCMSA

- Ongoing collaboration with Song & RoyChowdhury at UCR

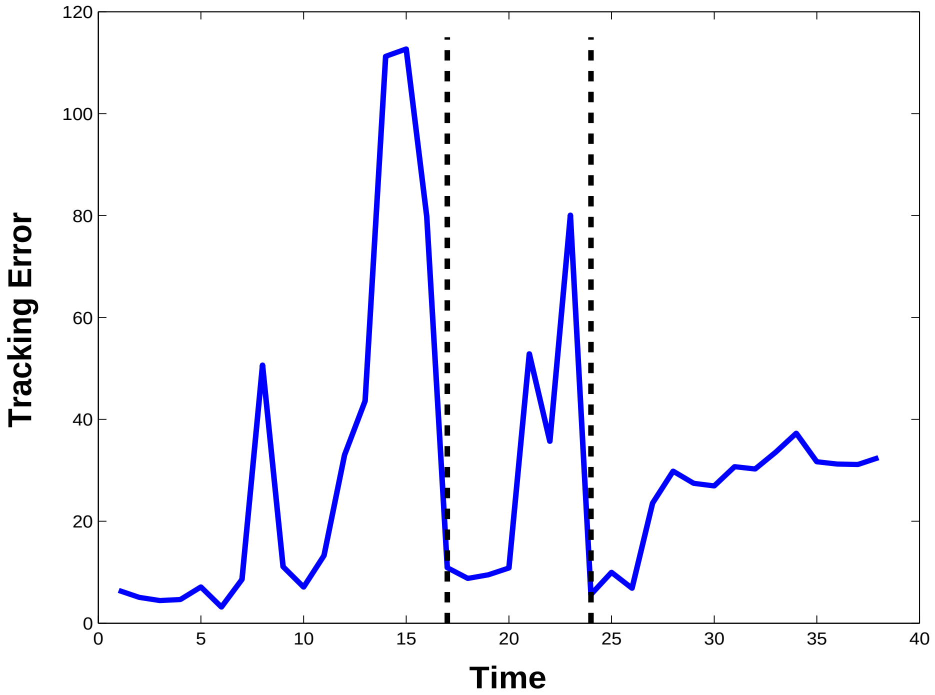


Human Actions: Sequence Identification Using PCMSA

- Detect change in activity model, recognize new CMSA, begin tracking with it. **ELL detects faster than TE**



Sequence Id using ELL



Sequence Id using TE

Summary

- **Proposed a common framework for:**
  - **Tracking Groups of Moving/Interacting "Objects"**
    - "Objects": Human body parts or people or vehicles or robots
  - **Abnormal Activity Detection & Tracking**
  - **Sequence Identification & Tracking**
  - **Activity Segmentation & Tracking**

- **Ongoing/Future work:**
  - Activity segmentation using NSSA
  - PTZ camera control to "follow" activity
  - 3D landmark shape dynamics, 2D affine shape dynamics