

## Observing Humans and their Activities




PRMA Seminar Mohan Trivedi, July 10, 2008

### Observing Humans and their Activities in Smart Spaces and Safer Automobiles

Mohan M. Trivedi

LISA: Laboratory for Intelligent and Safe Automobiles  
University of California at San Diego  
<http://cvrr.ucsd.edu/LISA>

July 10, 2008



## Observing Humans and Their Activities

PRMA Seminar Mohan Trivedi, July 10, 2008

### Outline of the Presentation:

**Part I:**  
Observing Humans in Smart Spaces

**Part II:**  
Human-Centered Intelligent Driver Support Systems

## Observing Humans and Their Activities

PRMA Seminar Mohan Trivedi, July 10, 2008

### Part I: Observing Humans in Smart Spaces:



- Introduction, Research Scope and Challenges
- "Smart" or Intelligent Spaces: Motivation
- Mutli-perspective Vision Systems
- Observing Human Movement
- Observing Articulated Human Body
- Observing and Tracking Body Parts
- Observing Gestures
- Understanding Human Activities and Interactions
- Human-Vehicle Interactions
- Discussions




## Observing Humans in Smart Spaces: Useful References

PRMA Seminar Mohan Trivedi, July 10, 2008

1. Systems and Architecture:
  - Mohan M. Trivedi, Kohsia S. Huang, Ivana Mikic, Dynamic Context Capture and Distributed Video Arrays for Intelligent Spaces, *IEEE Trans on Systems, Man and Cybernetics, Part A*, Jan 2009. ( )
  - Mohan M. Trivedi, Tarak L. Gandhi, Kohsia S. Huang, Distributed Interactive Video Arrays for Event Capture and Enhanced Situational Awareness, *IEEE Intelligent Systems*, Special Issue on AI in Homeland Security, Sept.-Oct. 2008. ( )
2. Body, Head and Hands:
  - Ivana Mikic, Mohan M. Trivedi, Edward Hunter, Pamela Cosman, Human Body Model Acquisition and Tracking using Voxel Data, *International Journal of Computer Vision*, pp. 199-223, Jul. 2003. ( )
  - Shinko Y. Cheng, Mohan M. Trivedi, Articulated Human Body Pose Inference from Voxel Data Using a Kinematically Constrained Gaussian Mixture Model, *CVPR Workshop on Evaluation of Articulated Human Motion and Pose Estimation*, June 2007. ( )
3. Facial Landmarks and Affects:
  - Junwen Wu, Mohan M. Trivedi, Simultaneous Eye Tracking and Blink Detection with Interactive Particle Filters, *EURASIP Journal on Advances in Signal Processing*, vol. 2008, 2008. ( )
  - Joel McCall, Mohan M. Trivedi, Facial Action Coding Using Multiple Visual Cues and a Hierarchy of Particle Filters, *IEEE Workshop on Vision for Human Computer Interaction in Conjunction with IEEE CVPR 2006*, Jun. 2006. ( )
4. Pointing and Gestures:
  - Kohsia S. Huang, Mohan M. Trivedi, 3D Shape Context Based Gesture Analysis Integrated with Tracking using Omni Video Array, *IEEE CVPR Workshop on Vision for Human-Computer Interaction (VHCI)*, Jun 2005. ( )
  - Erik Murphy-Gluchorion and Mohan M. Trivedi, Head Pose Estimation in Computer Vision: A Survey, *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 2008. ( )
5. Activities and Interactions:
  - Tarak Gandhi, Mohan M. Trivedi, Person Tracking and Reidentification: Introducing Panoramic Appearance Map (PAM) for Feature Representation, *Machine Vision and Applications: Special Issue Surveillance Systems*, August 2007. ( )
  - Sangho Park, Mohan M. Trivedi, Multi-person interaction and activity analysis: a synergistic track- and body-level analysis framework, *Machine Vision and Applications: Special Issue on Surveillance Systems*, August 2007. ( )
  - Sangho Park and Mohan M. Trivedi, Understanding Human Interactions with Track and Body Synergies (TBS) Captured from Multiple Views, *Computer Vision and Image Understanding*, Feb 2008. ( )

## Challenges in "Observing Humans"

PRMA Seminar Mohan Trivedi, July 10, 2008

### "Observing Humans":

#### Challenges

Body Covering

Body is Dynamic

Body is Deformable and Articulated

Body can carry Objects

Body is Expressive

Body is Purposeful



## Observing Humans and their Activities

PRMA Seminar Mohan Trivedi, July 10, 2008

• Watching and understanding people and their activities are essential for developing an "awareness" of a place, situation and its context

• However, this poses a tremendous challenge for computer vision systems.

• One way to make "observing humans" tractable is to consider an infrastructure, which can offer multiple perspectives and multiple modalities for observations.

→

Intelligent or "Smart" Spaces are such infrastructure.

M. Trivedi, K. Huang, I. Mikic, Dynamic Context Capture and Distributed Video Arrays for Intelligent Spaces, *IEEE Trans. on Systems, Man and Cybernetics, Part A*, Jan 2005.



## Intelligent Environments

PRMA Seminar Mohan Trivedi, July 10, 2008

### Intelligent Environments can:

- Develop and maintain awareness of the space, activity, and events
- Adapt to the dynamic changes in their surroundings
- Interact in a natural, efficient and flexible manner with the users
- Support "distributed" interactivity

### • Archive and "Recall" Events

Why Intelligent Spaces??

- Basic Artificial Intelligence Problem/s
- Challenging Engineering Systems Problem/s
- Implications to Cognitive Science and Psychology
- Improved Human-"Machine" Interactions
- Enhanced Human-Human Interactions
- Life Experience Enhancer: Improved Safety

M. Trivedi, K. Huang, I. Mikic, Dynamic Context Capture and Distributed Video Arrays for Intelligent Spaces, *IEEE Trans. on Systems, Man and Cybernetics, Part A*, Jan 2005.



## Awareness

PRMA Seminar Mohan Trivedi, July 10, 2008

### Situational Awareness involves:

Where? → Localization, Tracking

When? → Timing, Synchronization

What? → "Pattern" recognition

Who? → "Human" Identification, Biometrics

Why? → "Intent" and "context" analysis

M. Trivedi, K. Huang, I. Mikic, Dynamic Context Capture and Distributed Video Arrays for Intelligent Spaces, *IEEE Trans. on Systems, Man and Cybernetics, Part A*, Jan 2005.



## Intelligent Environments and Humans

PRMA Seminar Mohan Trivedi, July 10, 2008

Awareness of Humans and their Activities Involves  
Detection of Human:

- Presence
- Location
- Movements
- Gait Patterns
- Posture
- Gesture
- Identity
- Facial Landmarks: Eyes, Lips,
- Affects
- Attention and Gaze
- Activity
- Intentions
- Interactions

M. Trivedi, K. Huang, I. Mikic, Dynamic Context Capture and Distributed Video Arrays for Intelligent Spaces, *IEEE Trans. on Systems, Man and Cybernetics, Part A*, Jan 2005.



## Audio-Video Interfaces for Intelligent Environments (DiMI Grant 1998)

PRMA Seminar Mohan Trivedi, July 10, 2008

### AVIARY Audio-Video Interactive Appliances Rooms, & sSystems

[from DiMI Grant  
Proposal, 1998]



#### Research Goals

Develop novel intelligent systems which can:

- Develop and maintain awareness of their environment
- Acquire and respond to the voice and visual inputs from the users in a robust manner
- Adapt to the dynamic changes in their surroundings
- Interact in a natural and flexible manner with the users



M. Trivedi, K. Huang, I. Mikic, Dynamic Context Capture and Distributed Video Arrays for Intelligent Spaces, *IEEE Trans. on Systems, Man and Cybernetics, Part A*, Jan 2005.



## Research Evolution

PRMA Seminar Mohan Trivedi, July 10, 2008

**AVIARY**

Audio  
Video  
Interactive  
Appliances,  
Rooms, and  
sSystems  
(1998 – 2004)

**Indoor:**  
Human  
3-DTracking,  
Event Capture,  
Virtual Views

**DIVA**

Distributed  
Interactive  
Video  
Arrays  
(1999 -- )

**Outdoor/Indoor:**  
Vehicle + Human  
Tracking,  
Shadow Analysis,  
Camera Handoff,  
Event Capture

**MICASA**

Multimodal  
Interfaces and  
Context  
Aware  
SpAces  
(2001 -- )

**Indoor:**  
Body Modeling-Tracking  
Face Detection,  
Recognition  
Voxelization  
Gesture Analysis

**LISA**

Laboratory  
Intelligent &  
Safe  
Automobiles  
(2001 -- )

**Mobile:**  
Body Parts,  
Head+Face  
Posture  
Surround,  
Affects,  
Intent



## Research Evolution

PRMA Seminar Mohan Trivedi, July 10, 2008

**MICASA**

Multimodal  
Interfaces and  
Context  
Aware  
SpAces  
(2001 -- )

**Indoor:**  
Body Modeling-Tracking  
Face Detection,  
Recognition  
Voxelization  
Gesture Analysis

**LISA**

Laboratory  
Intelligent &  
Safe  
Automobiles  
(2001 -- )

**Mobile:**  
Body Parts,  
Head+Face  
Posture  
Surround,  
Affects,  
Intent

**SHIVA**

Systems for  
Human  
Interaction  
Visualization and  
Analysis  
(2006 -- )

**Indoor:**  
Multimodal  
Multiple-Levels  
Attention and Context  
Events, Interactions  
Intentions and  
Long term observations  
Semantic summarization



## Tracking, Identification and Activity Analysis

PRMA Seminar Mohan Trivedi, July 10, 2008

### Key Features:

Apply different types of camera arrays to provide *multiple signal-level resolutions*.

Ability to derive semantic information at *multiple levels of abstraction*.

Ability to be *"attentive"* to specific events and activities.

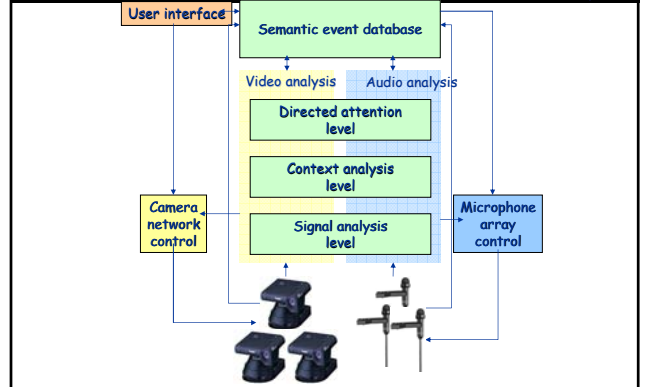
Ability to *actively shift the focus of attention* at different "semantic" resolutions.

M. Trivedi, K. Huang, I. Mikic, Dynamic Context Capture and Distributed Video Arrays for Intelligent Spaces, *IEEE Trans. on Systems, Man and Cybernetics, Part A*, Jan 2005.



## Tracking, Identification and Activity Analysis

PRMA Seminar Mohan Trivedi, July 10, 2008



M. Trivedi, K. Huang, I. Mikic, Dynamic Context Capture and Distributed Video Arrays for Intelligent Spaces, *IEEE Trans. on Systems, Man and Cybernetics, Part A*, Jan 2005.



## Multiple Abstractions

PRMA Seminar Mohan Trivedi, July 10, 2008

Simultaneous 3D tracking of multiple blobs

Face recognition

Capture of "interesting" events

Face orientation estimation

Affect Analysis

IMR Video



M. Trivedi, K. Huang, I. Mikic, Dynamic Context Capture and Distributed Video Arrays for Intelligent Spaces, *IEEE Trans. on Systems, Man and Cybernetics, Part A*, Jan 2005.



## Video Arrays for Robust Tracking and Event Capture

PRMA Seminar Mohan Trivedi, July 10, 2008

Head/Face Tracking (Walking)

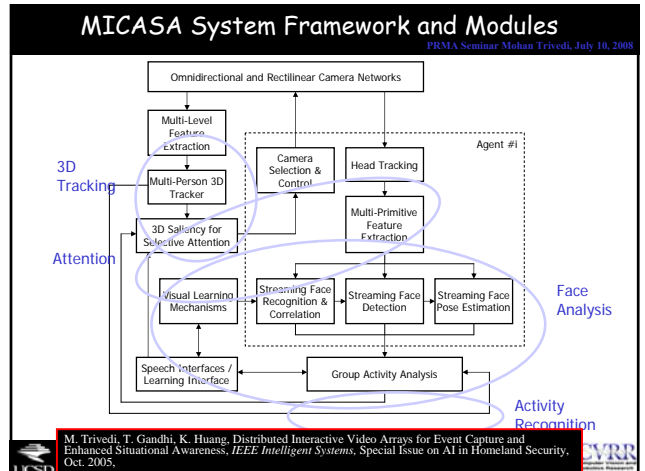
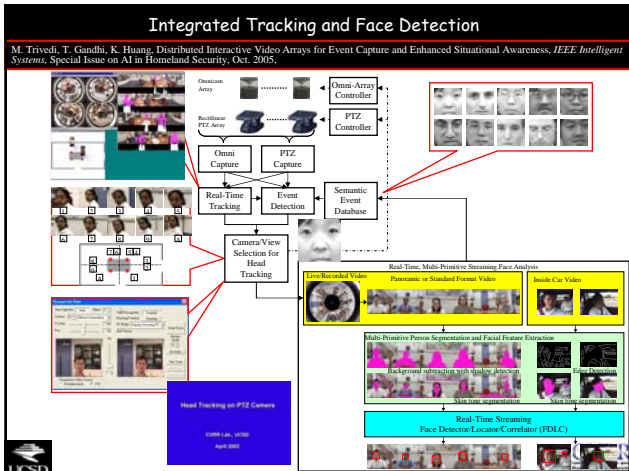
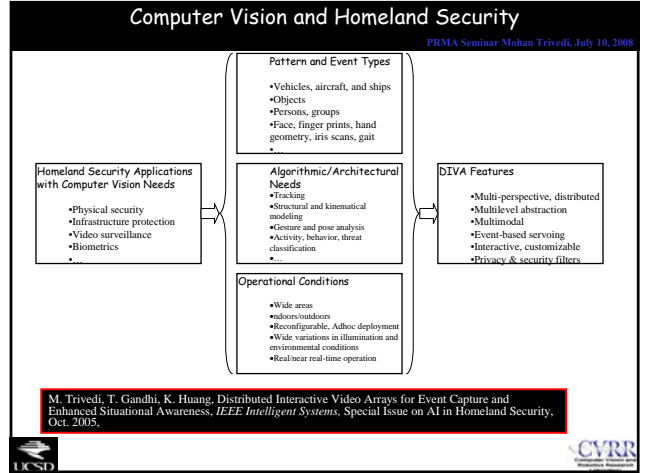
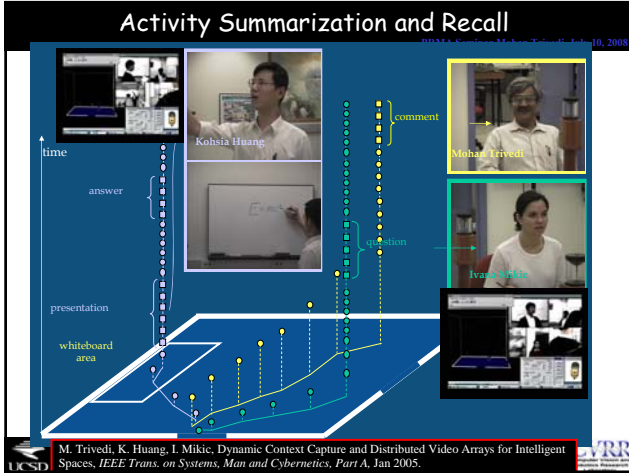
Head/Face Tracking (Sitting)

Video

Reference:

K. Huang, and M. Trivedi, "Video Arrays for Real-Time Tracking of Persons, Head and Face in an Intelligent Room," *Machine Vision and Applications, Omni-directional Vision, Special Issue*, 2003.





### DIVA Face Capture

PRMA Seminar Mohan Trivedi, July 10, 2008

15:12:31  
Dug

15:13:36  
Pov

15:14:07  
Mohan

15:15:01  
Tasha

15:16:00  
Mohan

15:16:02  
Chin

15:16:57  
Mohan

15:14:02  
Kobin +

15:18:13  
Kobin

15:14:02  
Chin

15:19:10  
Shahid

15:45:14  
Chin

15:20:50  
Jeff

15:46:55  
Chin

15:21:12  
Steve

15:54:22  
Kobin +

15:26:57  
Dug

15:59:45  
Ofer

16:01:44  
Kobin

M. Trivedi, T. Gandhi, K. Huang. Distributed Interactive Video Arrays for Event Capture and Enhanced Situational Awareness. *IEEE Intelligent Systems*, Special Issue on AI in Homeland Security, Oct. 2005.

### Driver Head-Pose and View Estimation with a single Omni-video Stream

PRMA Seminar Mohan Trivedi, July 10, 2008

Source omnidirectional video → Unwarped panoramic video

Unwarped panoramic video → Unwarped perspective video on driver seat

Unwarped perspective video on driver seat → Driver's head detection/tracking

Driver's head detection/tracking → Face/gaze direction estimation

Face/gaze direction estimation → Driver's view: 30° right

**Challenges:**

- Drastic illumination changes, both on brightness and color.
- High frame rate (30fps) to capture detailed dynamics.

K. Huang, M. Trivedi, T. Gandhi, "Driver's View and Vehicle Surround Estimation using Omnidirectional Video Stream," *Proc. IEEE Intelligent Vehicles Symposium*, June 2003.

### Human Body Modeling and Tracking: Research Overview

PRMA Seminar Mohan Trivedi, July 10, 2008

```

    graph TD
      A[Motion tracking] --> B[Handling large frame to frame displacements]
      B --> C[Hybrid voxel labeling procedure (distance & shape)]
      D[Body model acquisition and initialization] --> E[Automated acquisition and initialization]
      E --> F[Template-based initialization; Refinement using Bayesian nets]
      G[Human body modeling] --> H[Modeling framework]
      H --> I[Twist-based articulated model]
      J[System framework and architecture] --> K[Choice of input data (voxels vs. pixels)]
      K --> L[Voxels]
  
```

I. Mikic, M. Trivedi, E. Hunter, P. Cosman. Human Body Model Acquisition and Tracking using Voxel Data. *International Journal of Computer Vision*, Jul. 2003.

### 3D Voxel Reconstruction

PRMA Seminar Mohan Trivedi, July 10, 2008

I. Mikic, M. Trivedi, E. Hunter, P. Cosman. Human Body Model Acquisition and Tracking using Voxel Data. *International Journal of Computer Vision*, Jul. 2003.

### Twist-based Body Model and the Kalman Filter Tracker

L. Mikic, M. Trivedi, E. Hunter, P. Cosman, Human Body Model Acquisition and Tracking using Voxel Data, International Journal of Computer Vision, Jul. 2003.

Kalman filter state:  
23 parameters  
23 measurement points:

Kalman filter equations:  

$$x_{k+1} = Fx_k + u_k$$

$$z_k = H(x_k) + w_k$$

$$z_k^i(\theta_p, R_p, t_0) = R_0(R_1(R_2(\dots(R_m z'(0) + t_m) + \dots) + t_2) + t_1) + t_0$$

wrt world coordinate system      wrt torso coordinate system

$R_0, t_0$  ↔ Orientation and position of the torso       $\theta_p = (\theta_{i1}, \theta_{i2}, \dots, \theta_{in})$

$T_i = e^{\xi_i \theta_i} = \begin{bmatrix} R_i & t_i \\ 0 & 1 \end{bmatrix}$  → exponential map associated with the rotation  $i$

$$z_i^{LFoot}(0) = \begin{bmatrix} 0 & d_H \lambda_{torso}^{(2)} & -(\lambda_{torso}^{(2)} + 2\lambda_{LThigh}^{(2)} + 2\lambda_{LCalf}^{(2)}) \end{bmatrix}$$

UCSD      CVRR

### Body Modeling, Movement, Posture and Gait

L. Mikic, M. Trivedi, E. Hunter, P. Cosman, Human Body Model Acquisition and Tracking using Voxel Data, International Journal of Computer Vision, Jul. 2003.

UCSD      CVRR

### Body and Movements

PRMA Seminar Mohan Trivedi, July 10, 2008

L. Mikic, M. Trivedi, E. Hunter, P. Cosman, Human Body Model Acquisition and Tracking using Voxel Data, International Journal of Computer Vision, Jul. 2003.

UCSD      CVRR

### Pose Estimation from Voxel Images

PRMA Seminar Mohan Trivedi, July 10, 2008

Articulated Body Model

- Assumption: Voxel image of subject is tri-variate Gaussian mixture distributed
- Each Gaussian represents body segment
- Arrangement (Pose) of segments described by mean and covariance matrix parameters

$$\log p(x|\theta) = \prod_{l=0}^L \alpha_l^{y_l} p_l(x|\theta)^{y_l}$$

Estimation Procedure (EM Algorithm)  
 - E-step: Used to estimate mixing probabilities  
 - M-step: Unconstrained Likelihood maximization  
 - Project EM Iterates onto feasible space with constraint equations:  $C(\theta)=0$   
 - Repeat until convergence

Hand Voxel Reconstruction → E-Step: Voxel association mixing probabilities ( $\alpha_k$ )

a priori known Segment dimensions, Joint locations → M-Step: Unconstrained likelihood maximization ( $\mu_k, R_{kx}$ )

Projection Step: ( $\mu_k, R_{kx}$ ) → Hand Pose Parameters

S. Y. Cheng, M. Trivedi, Articulated Human Body Pose Inference from Voxel Data Using a Kinetically Constrained Gaussian Mixture Model, CVPR Workshop on Evaluation of Articulated Human Motion and Pose Estimation, June 2007, [Best Paper Award]

UCSD      CVRR

### Kinematic Constraints by DOF

PRMA Seminar Mohan Trivedi, July 10, 2008

- Spherical Joint Constraints (3 DOF)
 
$$\mu_i - R_{oi}a_{ij} = \mu_j - R_{oj}a_{ji}$$
- Two-rotation DOF joint (2 DOF)
 
$$0 = \langle R_{oi} \frac{a_{ij}}{\|a_{ij}\|}, R_{oj} \frac{a_{ji}}{\|a_{ji}\|} \rangle$$
- Revolute Joint Constraints (1 DOF)
 
$$R_{oi}\hat{q}_{ij} = R_{oj}\hat{q}_{ji}$$

Form equation and use Newton-Raphson procedure to solve for pose parameters  $\mathbf{c}(\theta) = \mathbf{0}$

*Points joined by dotted lines are forced together by satisfying these equations.*

S. Y. Cheng, M. Trivedi, Articulated Human Body Pose Inference from Voxel Data Using a Kinetically Constrained Gaussian Mixture Model, CVPR Workshop on Evaluation of Articulated Human Motion and Pose Estimation, June 2007. [Best Paper Award]

### Hand Pose Estimation Using Kinematic Gaussian Mix. Model

CVPR 2008

S. Y. Cheng, M. Trivedi, Articulated Human Body Pose Inference from Voxel Data Using a Kinetically Constrained Gaussian Mixture Model, CVPR Workshop on Evaluation of Articulated Human Motion and Pose Estimation, June 2007. [Best Paper Award]

### Human Body is Multiscale = Need Multi-resolution Models

PRMA Seminar Mohan Trivedi, July 10, 2008

S. Y. Cheng, M. Trivedi, Articulated Human Body Pose Inference from Voxel Data Using a Kinetically Constrained Gaussian Mixture Model, CVPR Workshop on Evaluation of Articulated Human Motion and Pose Estimation, June 2007. [Best Paper Award]

### Human Body and Hand Modeling and Tracking

CVPR 2008

S. Y. Cheng, M. Trivedi, Articulated Human Body Pose Inference from Voxel Data Using a Kinetically Constrained Gaussian Mixture Model, CVPR Workshop on Evaluation of Articulated Human Motion and Pose Estimation, June 2007. [Best Paper Award]

### NOVA: Human Body Voxelization

PRMA Seminar Mohan Trivedi, July 10, 2008

- Voxel-panorama mapping on omni cameras

$$\theta = \arctan 2(y - Y, x - X)$$

$$x_p = (\theta / 2\pi) \cdot W$$

$$y_p = \frac{z - Z}{\sqrt{(x - X)^2 + (y - Y)^2}} + y_H$$

- Body voxelization on omni video array

K. Huang, M. Trivedi, 3D Shape Context Based Gesture Analysis Integrated with Tracking using Omni Video Array, CVPR Workshop on Vision for Human-Computer Interaction, Jun2005.

### Spatial-Temporal Body Gesture Recognition

PRMA Seminar Mohan Trivedi, July 10, 2008

- Spatial-temporal context dynamics modeling: VQ+DHMM

DHMM:  $(\pi_{s_{t-1}}, A_{s_{t-1}}, B_{s_{t-1}})$

For: kick, stretch, walk, bow, pick, point, and wave-hand

For: stand-up and sit-down

TTL: State Transition Length

- Temporal likelihood accumulation

$$L = \pi_i a_{ij} b_{jk} a_{kl} \dots a_{mn} b_{nk}, \quad 1 \leq i, j, k, m, n \leq N$$

K. Huang, M. Trivedi, 3D Shape Context Based Gesture Analysis Integrated with Tracking using Omni Video Array, CVPR Workshop on Vision for Human-Computer Interaction, Jun2005.

PRMA Seminar Mohan Trivedi, July 10, 2008

K. Huang, M. Trivedi, 3D Shape Context Based Gesture Analysis Integrated with Tracking using Omni Video Array, CVPR Workshop on Vision for Human-Computer Interaction, Jun2005.

### Moods, Affects, and Emotions

PRMA Seminar Mohan Trivedi, July 10, 2008

**Initialization**

**Feature Tracking**

**Feature Selection**

Bayesian Estimation and Affect Classification

K. Huang, M. Trivedi, 3D Shape Context Based Gesture Analysis Integrated with Tracking using Omni Video Array, CVPR Workshop on Vision for Human-Computer Interaction, Jun2005.

## Research Objectives

PRMA Seminar Mohan Trivedi, July 10, 2008

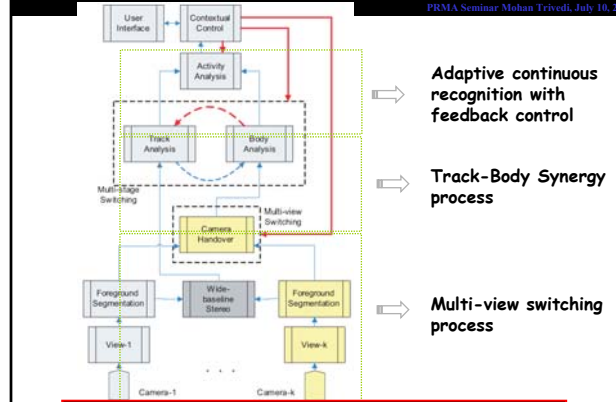
- A new framework for human interactions and activity analysis
  - a two-stage multi-view approach.
- A distributed vision system that synergistically integrates track- and body- level processing
  - for understanding human activity.
- Efficient handling of occlusion problem
- Data-driven bottom-up process + context-driven top-down process
  - for continuous understanding of human activity

S. Park and M. Trivedi, Understanding Human Interactions with Track and Body Synergies (TBS) Captured from Multiple Views. *Computer Vision and Image Understanding*, Feb 2008.



## System Architecture for Track-Body Synergy

PRMA Seminar Mohan Trivedi, July 10, 2008



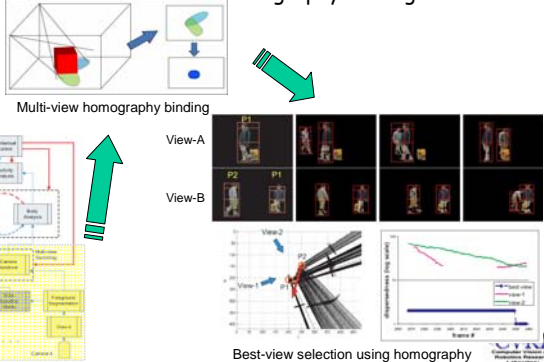
S. Park, M. Trivedi, Multi-person interaction and activity analysis: a synergistic track- and body-level analysis framework. *Machine Vision and Applications: Special Issue Surveillance Systems*, August 2007.



## Multi-view Multi-level Adaptive Switching

S. Park and M. Trivedi, Understanding Human Interactions with Track and Body Synergies (TBS) Captured from Multiple Views. *Computer Vision and Image Understanding*, Feb 2008.

- Multi-view switching using homography binding



Best-view selection using homography



## Experimental Analysis

PRMA Seminar Mohan Trivedi, July 10, 2008

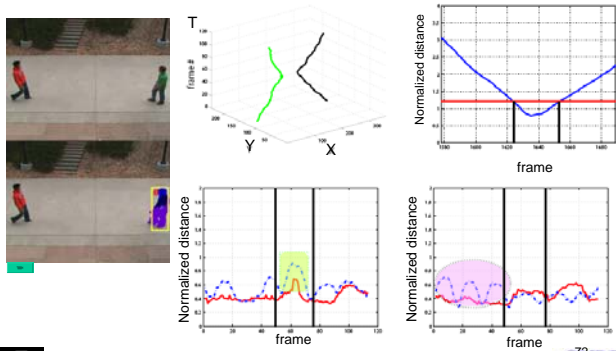
- Single-view tracking of people interactions
  - Outdoor surveillance
- Multi-view recognition of single-person activity
  - Indoor detailed view
  - Walk left, walk right, drop down, pick up
- Multi-view two-person interaction + object
  - Indirect handover of object
  - Direct handover of object

S. Park and M. Trivedi, Understanding Human Interactions with Track and Body Synergies (TBS) Captured from Multiple Views. *Computer Vision and Image Understanding*, Feb 2008.



## Interaction Analysis: *kicking*

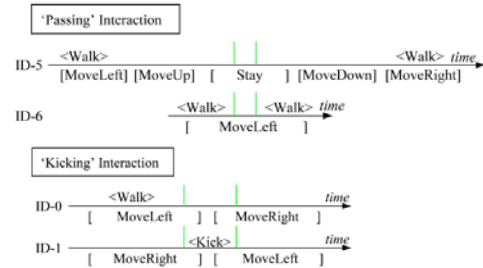
PRMA Seminar Mohan Trivedi, July 10, 2008



S. Park, M. Trivedi, Multi-person interaction and activity analysis: a synergistic track- and body-level analysis  
*International Journal of Intelligent and Robotic Systems, August 2007*

## Semantic Representation of two person interactions

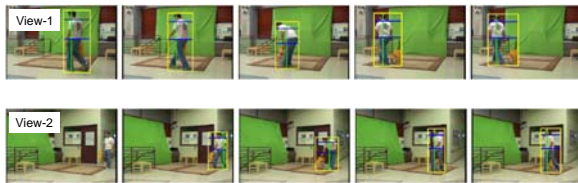
PRMA Seminar Mohan Trivedi, July 10, 2008



S. Park, M. Trivedi, Multi-person interaction and activity analysis: a synergistic track- and body-level analysis  
*framework, Machine Vision and Applications: Special Issue Surveillance Systems, August 2007.*

## Multi-view single-person activity

PRMA Seminar Mohan Trivedi, July 10, 2008



Confusion Matrix for View-1

	Walk-Left	Drop-Object	Pick-Object	Walk-Right
Walk-Left	83.3	0.0	0.0	16.7
Drop-Object	0.0	100.0	0.0	0.0
Pick-Object	0.0	0.0	100.0	0.0
Walk-Right	0.0	0.0	0.0	100.0

Confusion Matrix for View-2

	Walk-Left	Drop-Object	Pick-Object	Walk-Right
Walk-Left	80.0	0.0	20.0	0.0
Drop-Object	0.0	100.0	0.0	0.0
Pick-Object	0.0	0.0	25.0	75.0
Walk-Right	14.3	14.3	0.0	71.4

S. Park and M. Trivedi, Understanding Human Interactions with Track and Body Synergies (TBS) Captured from Multiple Views. *Computer Vision and Image Understanding, Feb 2008.*

## Hierarchical Composition of Activity

PRMA Seminar Mohan Trivedi, July 10, 2008

A hierarchy of activity vocabulary associated with the current experimental design.

ACTION	Unit action	Single-person action	Two-person interaction
	Enter	Enter	Enter
	Leave	Leave	Leave
	Walk left	Walk left	Walk left
	Walk right	Walk right	Walk right
	Stretch	Stretch	Stretch
	Kick	Kick	Kick
	Bend down	Bend down	Bend down
	Bend up	Bend up	Bend up
		Drop object	Drop object
		Pick object	Pick object
<b>INTERACTION</b>			<b>Object exchange</b>
			Person 1 → object → person 2
			Person 1 ← object ← person 2
			Person 1 drops object and person 2 picks the object.

### Two-person: Direct handover of objects

PRMA Seminar Mohan Trivedi, July 10, 2008

View1  
View2  
Composite  
Tracker1  
Tracker2  
Selection  
Activity

P1 has P1 has P1 has P2 has P2 has P2 has P2 has  
object-1 object-1 object-1 object-2 object-1 object-1 object-1

S. Park and M. Trivedi, Understanding Human Interactions with Track and Body Synergies (TBS) Captured from Multiple Views. *Computer Vision and Image Understanding*, Feb 2008.

### "DIVA"--Event Driven Servoing

PRMA Seminar Mohan Trivedi, July 10, 2008

EVENT	ACTION
stopped car in a given area (like emergency lane)	Zoom in (Secondary camera 1)
stopped car in a given area (like emergency lane)	Primary
incident detected	Secondary
incident detected (injuries)	
flat tire car	
incident detected	
....	

Flexible Architecture. User can create new event-action tuples

UCSD

### DIVA for Tracking, Identification and Activity Analysis

PRMA Seminar Mohan Trivedi, July 10, 2008

Outdoor

M. Trivedi, T. Gandhi, K. Huang, Distributed Interactive Video Arrays for Event Capture and Enhanced Situational Awareness. *IEEE Intelligent Systems*, Special Issue on AI in Homeland Security, Oct. 2005.

UCSD

### Mobile Sentry

PRMA Seminar Mohan Trivedi, July 10, 2008

RR

UCSD

## Automatic Event Detection



- Human and vehicle on road
- Vehicle stopped on road
- Speeding human








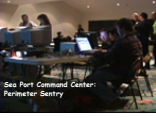




M. Trivedi, T. Gandhi, K. Huang, Distributed Interactive Video Arrays for Event Capture and Enhanced Situational Awareness, *IEEE Intelligent Systems*, Special Issue on AI in Homeland Security, Oct. 2005.






## DIVA at Super Bowl 2003

PRMA Seminar Mohan Trivedi, July 10, 2008

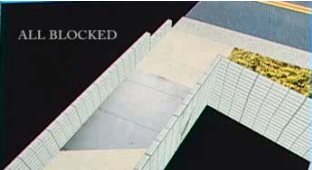
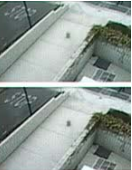
M. Trivedi, T. Gandhi, K. Huang, Distributed Interactive Video Arrays for Event Capture and Enhanced Situational Awareness, *IEEE Intelligent Systems*, Special Issue on AI in Homeland Security, Oct. 2005.

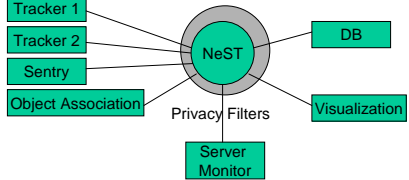











## "Blind" DIVA (Surveillance w/ Privacy)

PRMA Seminar Mohan Trivedi, July 10, 2008



M. Trivedi, T. Gandhi, K. Huang, Distributed Interactive Video Arrays for Event Capture and Enhanced Situational Awareness, *IEEE Intelligent Systems*, Special Issue on AI in Homeland Security, Oct. 2005.

## Observing Humans: Concluding Remarks

PRMA Seminar Mohan Trivedi, July 10, 2008

*Where we may be going?*

- *Keep climbing up the Semantic Ladder: Multiple Level and Scales*
- *Multi modal and Multi Sensory Analysis and Integration Frameworks*
- *Experimental and Systems Oriented Science: Need Metrics and Evaluations*

Thanks!

For more information and all publications:  
[cvrr.ucsd.edu/LISA](http://cvrr.ucsd.edu/LISA)  
and  
[cvrr.ucsd.edu](http://cvrr.ucsd.edu)



