UNIVERSITY OF WASHINGTON
ELECTRICAL ENGINEERING

NATION CHIAO TUNG UNIVERSITY
DEPARTMENT OF COMPUTER SCIENCE

Multiple-kernel Adaptive Segmentation and Tracking (MAST) for Robust Object Tracking

ICASSP 2016 IVMSP-L2: Video Tracking

Zheng Tang, Jenq-Neng Hwang, University of Washington, United States;
Yen-Shuo Lin, Jen-Hui Chuang, National Chiao Tung University, Taiwan

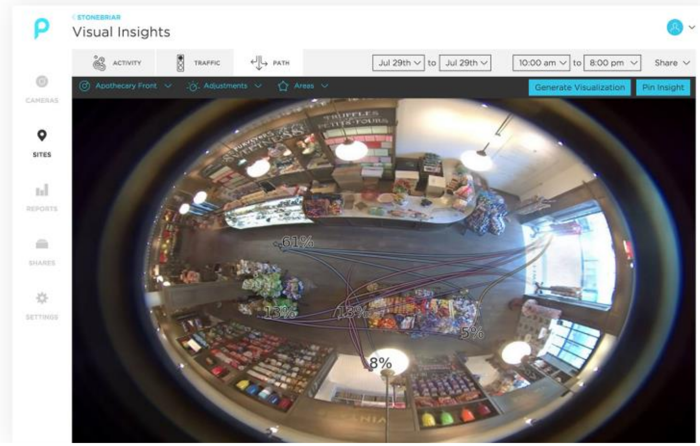
Agenda

- Introduction
- System Overview
- General Segmentation and Tracking
- Similarity Computation and Feedback Loop
- Experimental Results
- Conclusion

Agenda

- **Introduction**
- System Overview
- General Segmentation and Tracking
- Similarity Computation and Feedback Loop
- Experimental Results
- Conclusion

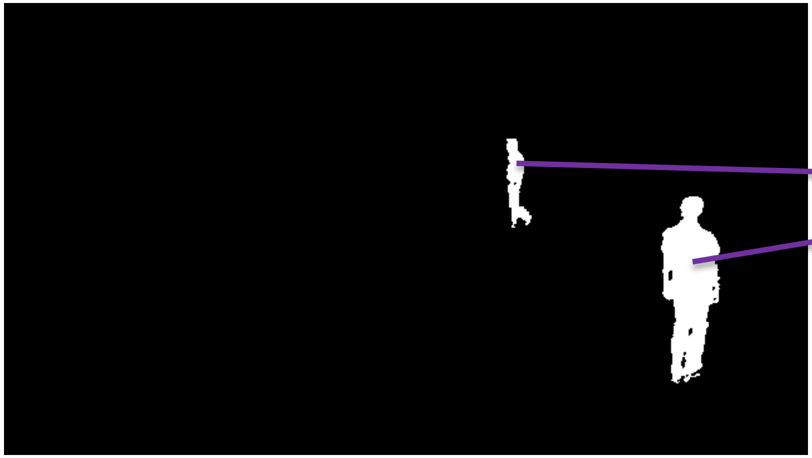
Introduction



- Intelligent video surveillance
  - Customer analysis
  - Anomaly detection
  - Suspect tracking
- Video object segmentation
  - It includes background subtraction and shadow detection.
- Video object tracking
  - It provides the information about the location of a tracked object in time.

Introduction

- Many object tracking approaches are dependent on foreground segmentation mask.
  - Example: In Kalman filter tracking, time variant matrix can consist of position, size and velocity of each foreground blob.

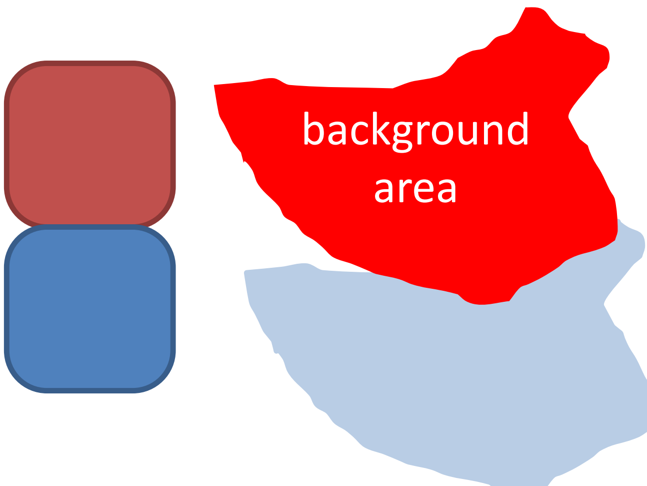


foreground blobs from segmented result

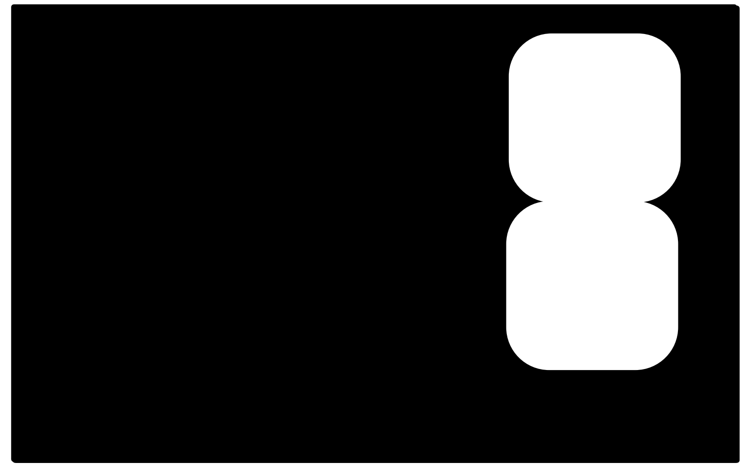
C. Chu, J. Hwang, S. Wang, and Y. Chen, "Human Tracking by Adaptive Kalman Filtering and Multiple Kernels Tracking with Projected Gradients," *Proc. ACM/IEEE Int. Conf. Distributed Smart Cameras*, 2011.

Introduction

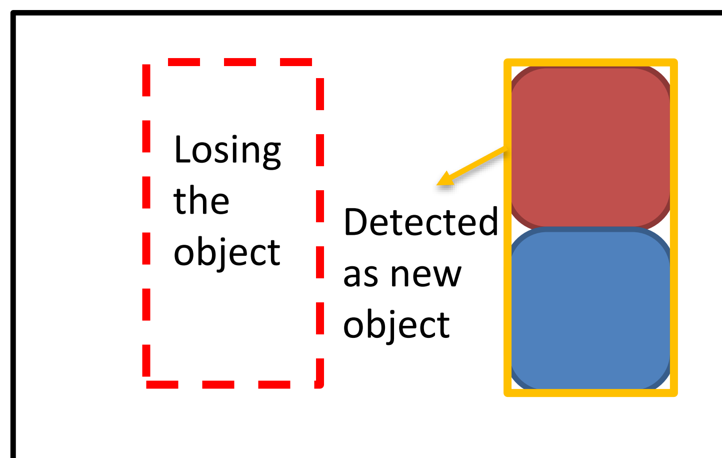
- Object merging



Segmentation result

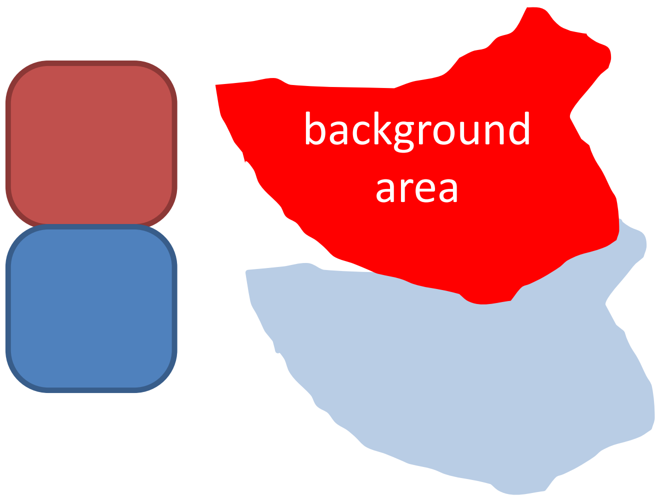


Tracking result

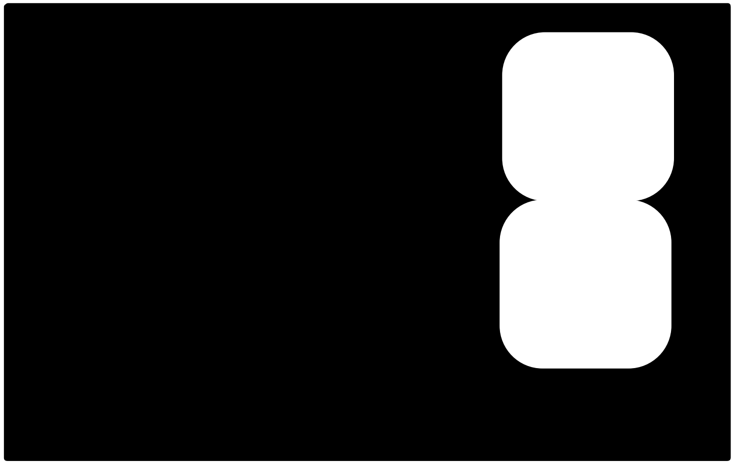


Introduction

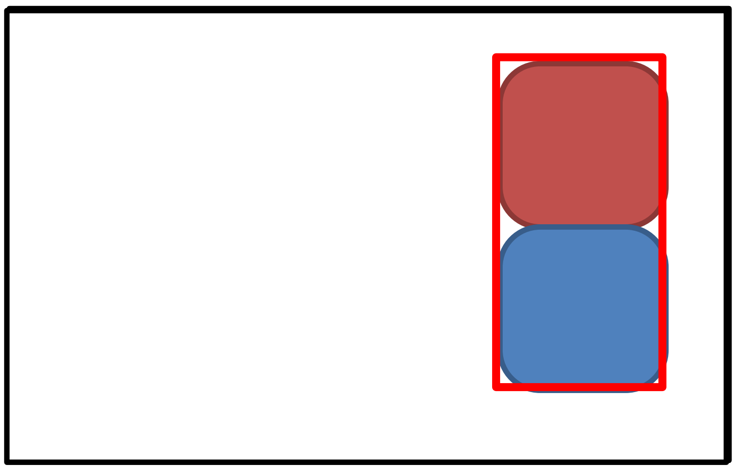
- Motivation



Segmentation result



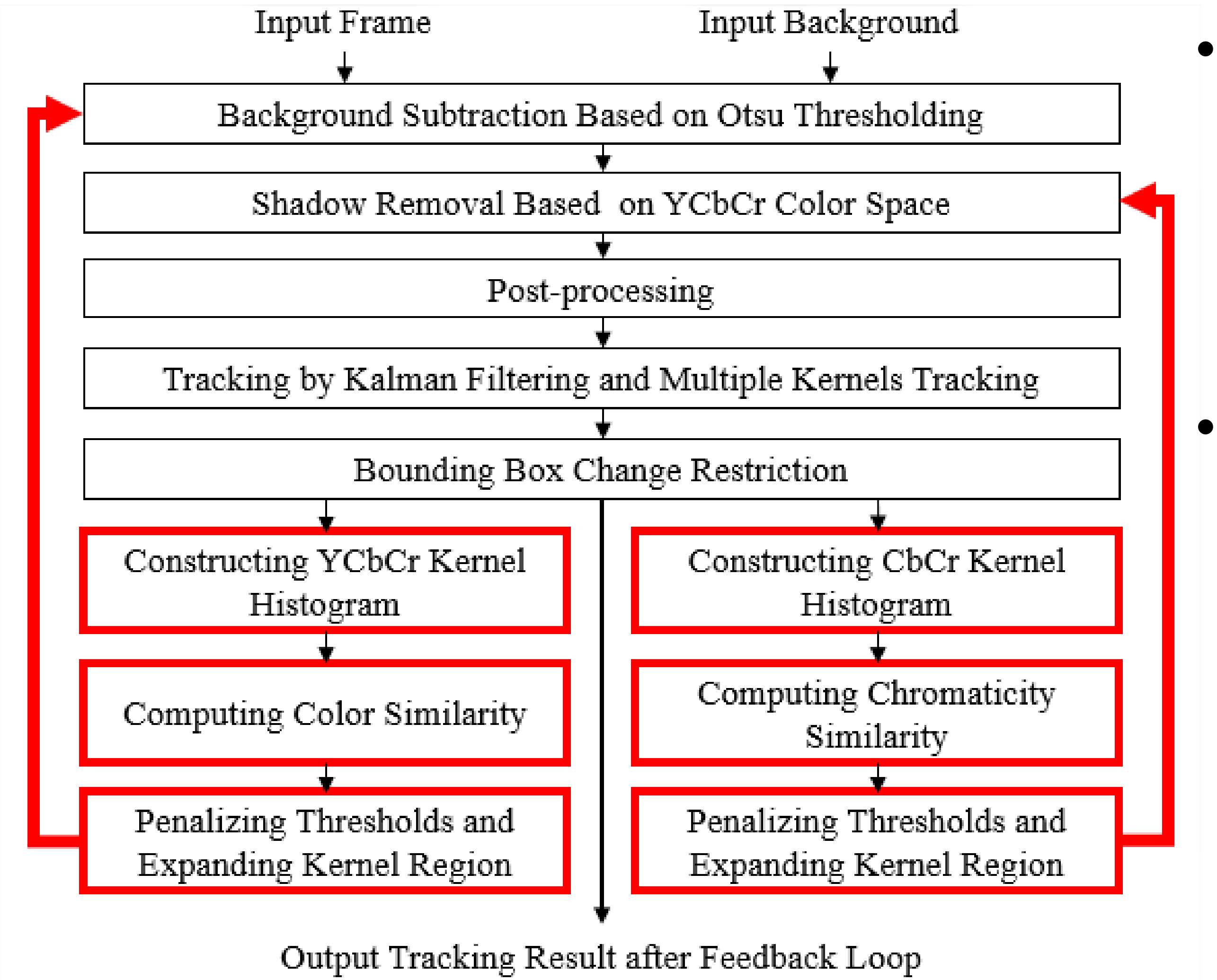
Tracking result



Agenda

- Introduction
- **System Overview**
- General Segmentation and Tracking
- Similarity Computation and Feedback Loop
- Experimental Results
- Conclusion

System Overview

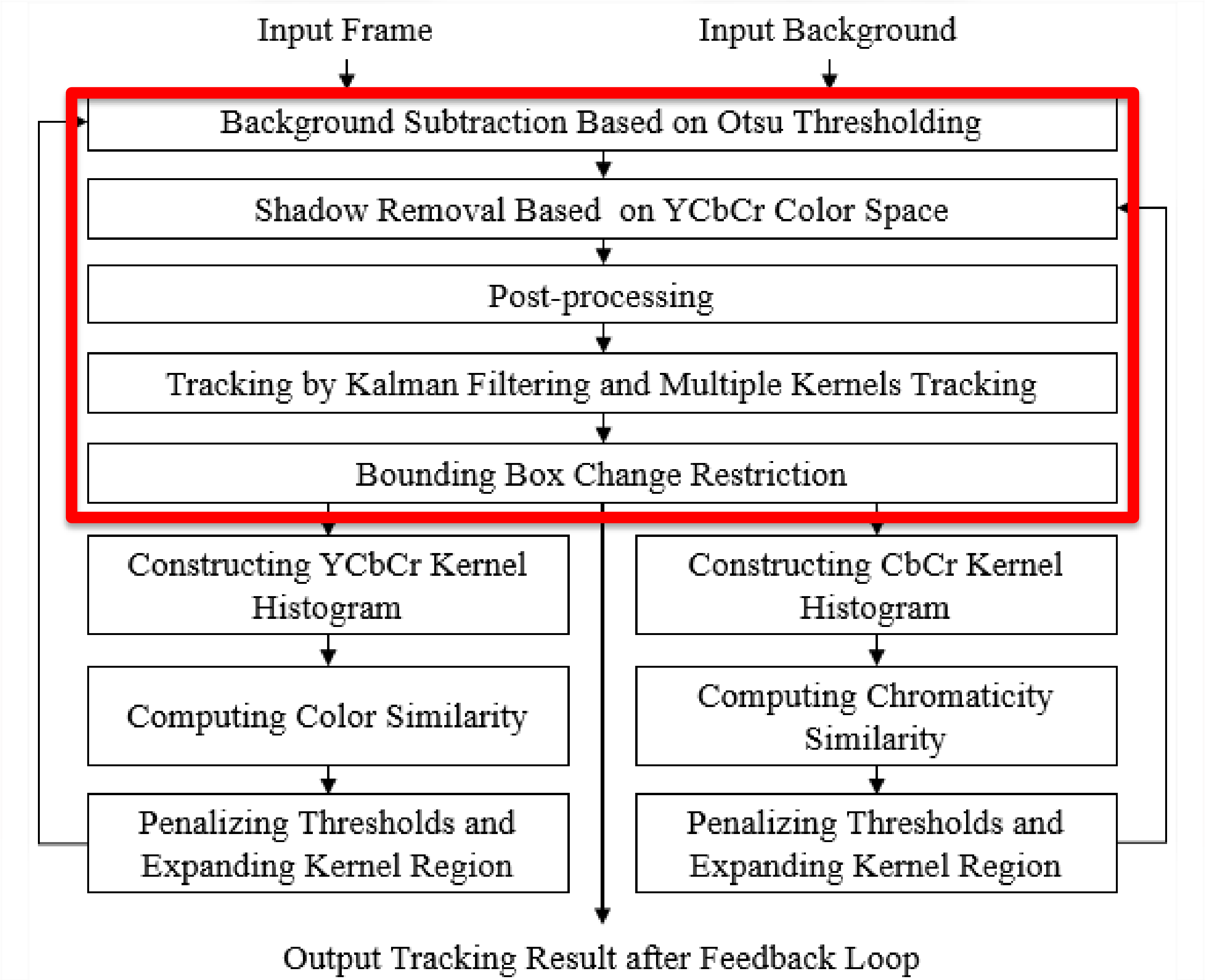


- The segmentation block can be substituted by any method based on thresholding.
- The tracking block can be substituted by any method based on segmentation results.

Agenda

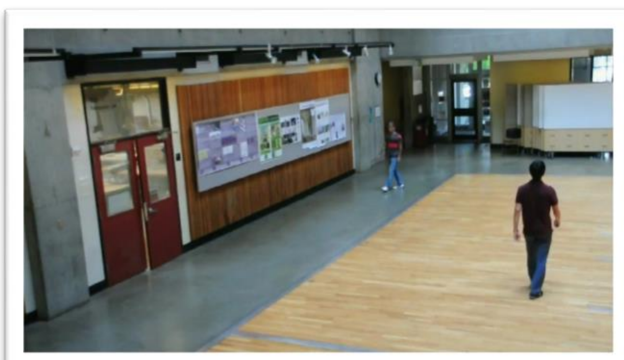
- Introduction
- System Overview
- **General Segmentation and Tracking**
- Similarity Computation and Feedback Loop
- Experimental Results
- Conclusion

System Overview



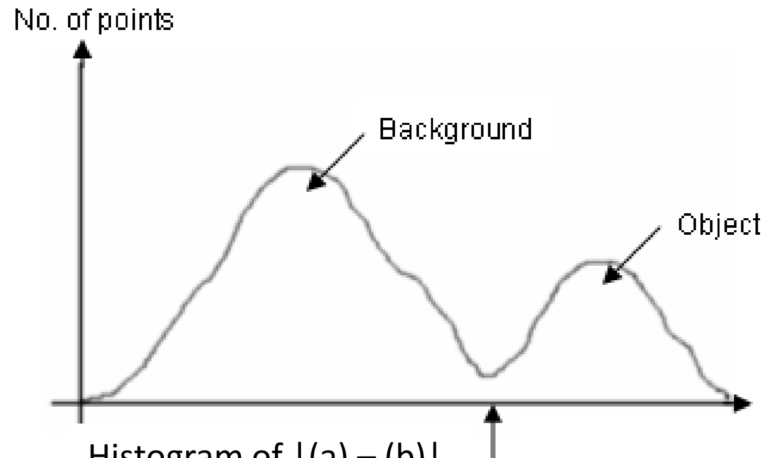
Background Subtraction (Otsu Thresholding)

(a) Background Image



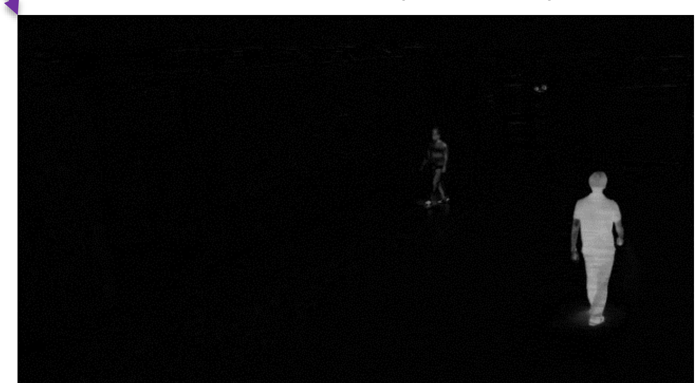
(b) Current image

|(a) – (b)|



Histogram of |(a) – (b)|

Finding the optimum global threshold that maximizes the between-class variance (measurement of separability)



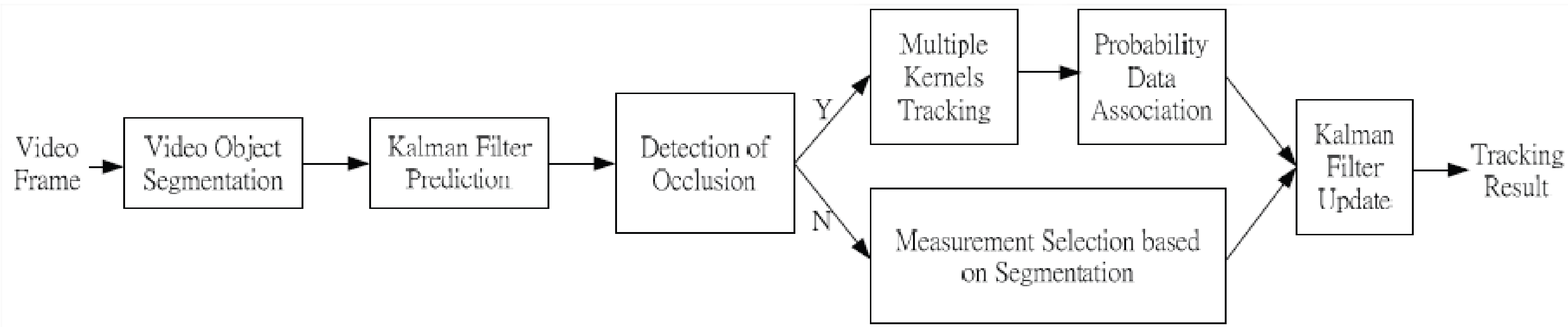
Shadow Removal (YCbCr)

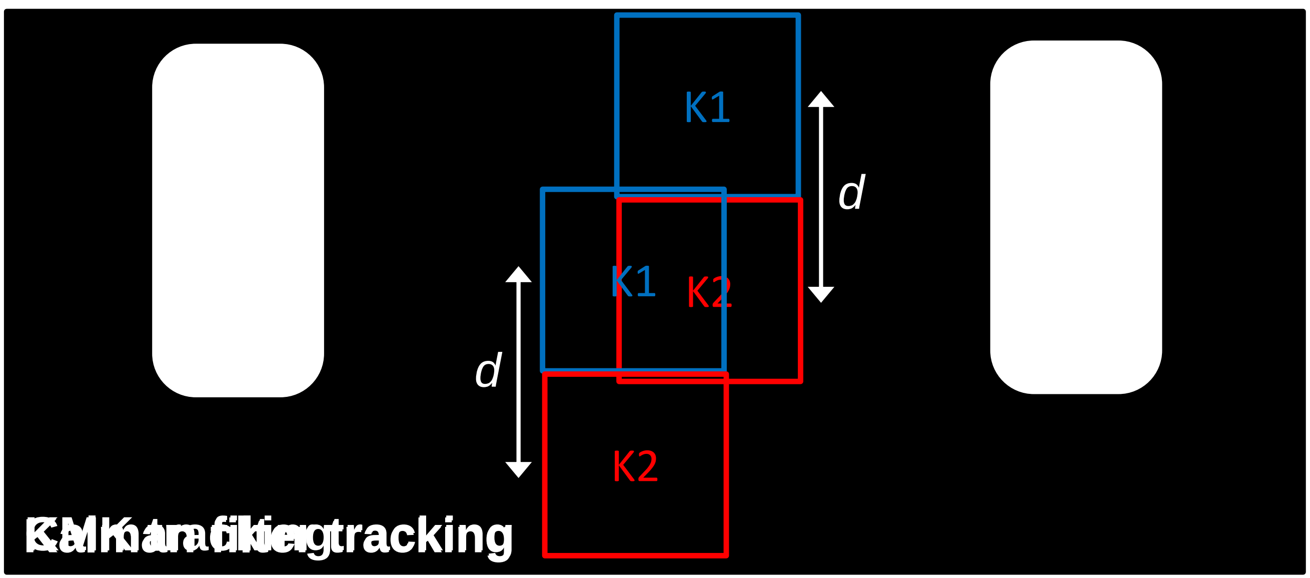
- Identifier of shadow point

$$\mathrm{SInd}(x,y)=\begin{cases}1, & \left(\alpha \le Y^{I}(x,y)/Y^{B}(x,y) \le \beta\right) \\ & \wedge\left(\left|Cb^{I}(x,y)-Cb^{B}(x,y)\right| \le \tau_{Cb}\right) \\ & \wedge\left(\left|Cr^{I}(x,y)-Cr^{B}(x,y)\right| \le \tau_{Cr}\right) \\ 0, & otherwise\end{cases}$$

  - I: Current frame
  - B: Background
  - α, β: Threshold parameters for Y channel
  - τ_{Cb}: Threshold parameter for Cb channel
  - τ_{Cr}: Threshold parameter for Cr channel

Kalman Filter Tracking with Constrained Multiple-Kernel (CMK) Tracking

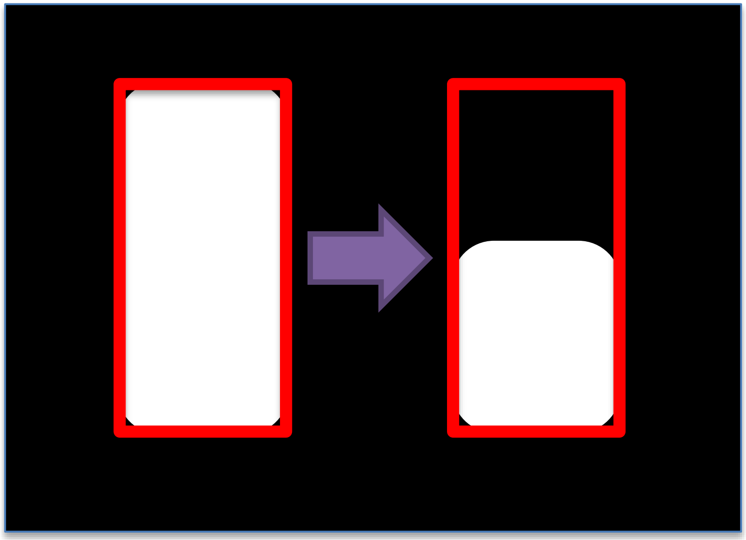




C. Chu, J. Hwang, S. Wang, and Y. Chen, "Human Tracking by Adaptive Kalman Filtering and Multiple Kernels Tracking with Projected Gradients," *Proc. ACM/IEEE Int. Conf. Distributed Smart Cameras*, 2011.

Bounding Box Change Restriction

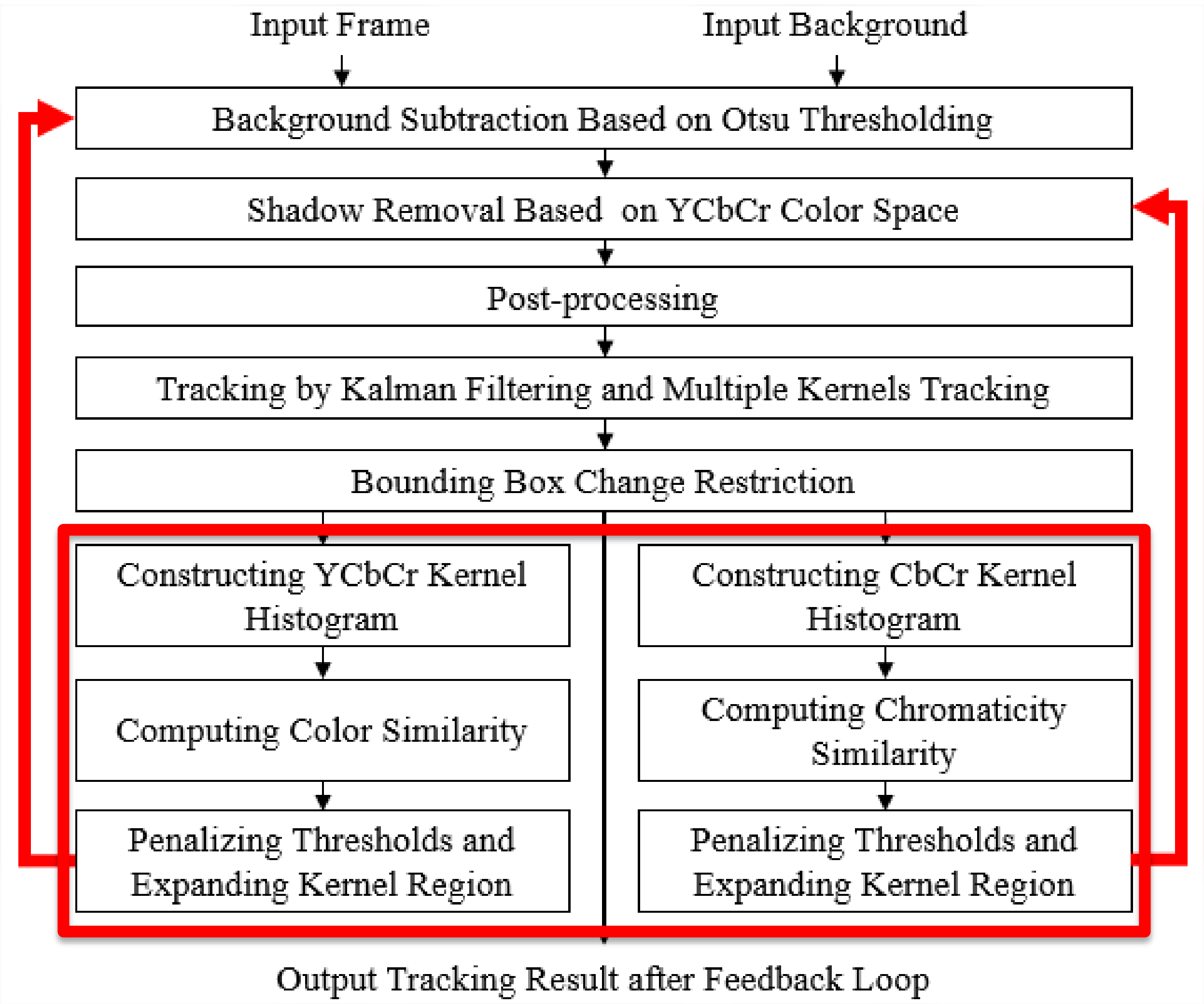
- Some constraints are imposed to prevent sudden change of the bounding boxes caused by noise or segmentation failure.
- CMK tracking is applied when the segmented foreground blob is not reliable.
- The constraints include:
  - limited size-change ratio
  - limited width-change ratio
  - limited height-change ratio



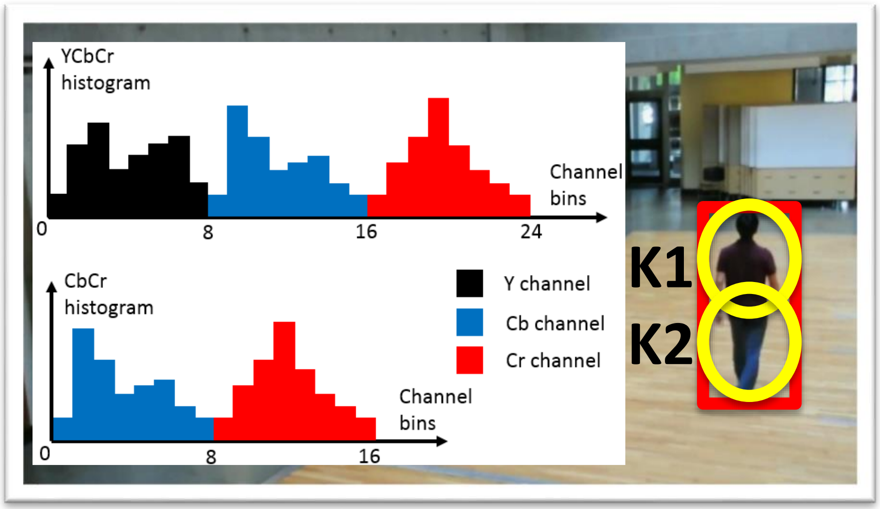
Agenda

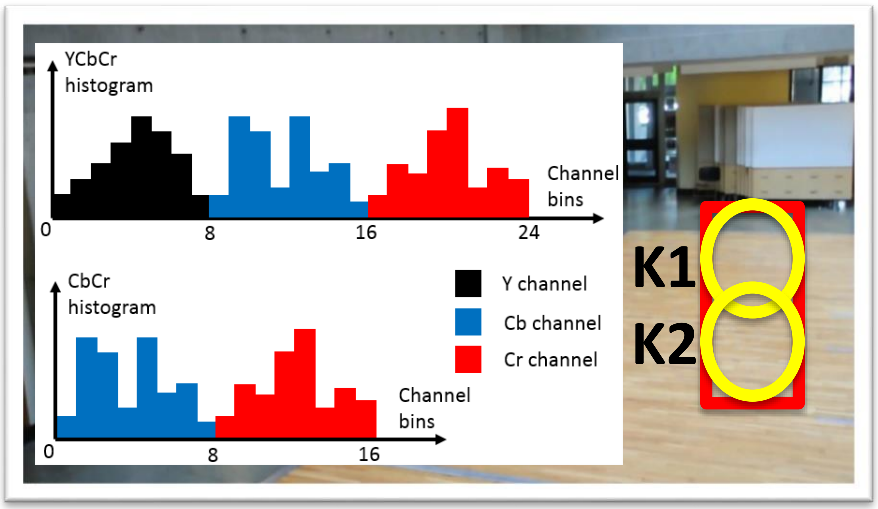
- Introduction
- System Overview
- General Segmentation and Tracking
- **Similarity Computation and Feedback Loop**
- Experimental Results
- Conclusion

System Overview



Constructing Kernel Histograms





- Two histograms are built for each kernel
  - YCbCr histogram for computing penalty for background subtraction
  - CbCr histogram for computing penalty for shadow removal
- Gaussian kernel function weight: $w = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\frac{(x-x_c)}{2\sigma_x^2}} e^{-\frac{(y-y_c)}{2\sigma_y^2}}$

Computing Color Similarity (Bhattacharyya Similarity)

- The color similarity and chromaticity similarity are computed by the reciprocals of Bhattacharyya distances between the corresponding kernel histograms:

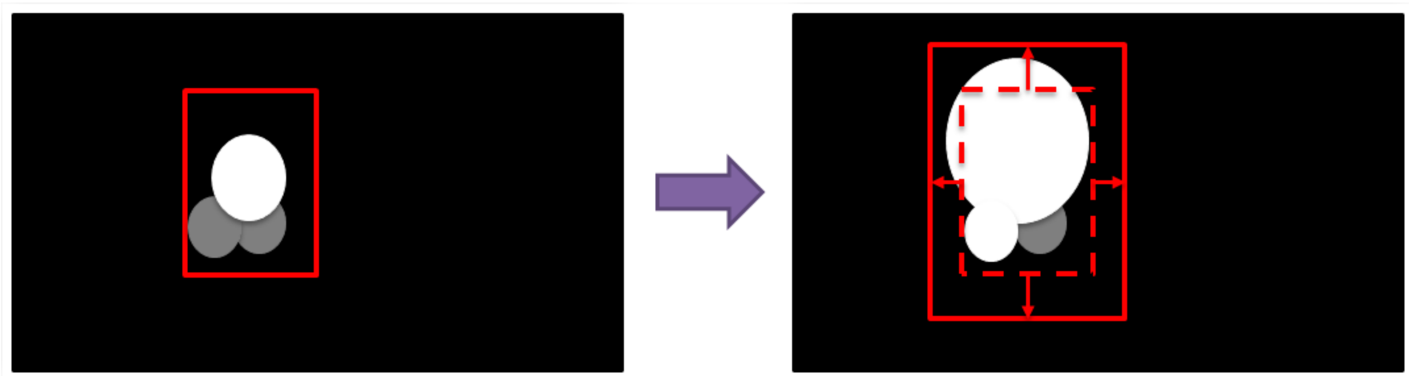
$$colorSimi = \frac{1}{\sum \sqrt{hist^{I}_{YCbCr}(x, y) \bullet hist^{B}_{YCbCr}(x, y)}}$$

$$chromSimi = \frac{1}{\sum \sqrt{hist^{I}_{CbCr}(x, y) \bullet hist^{B}_{CbCr}(x, y)}}$$

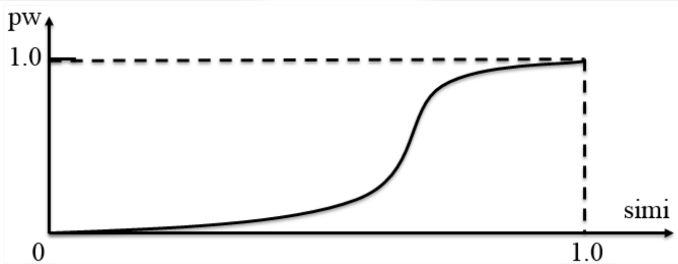
  - I: Current frame
  - B: Background

Penalizing Thresholds and Expanding Kernel Region

- Otsu's threshold in background subtraction or the chromaticity thresholds τ_{Cb} and τ_{Cr} in shadow removal will be penalized by multiplying (1 - *pw*).
- The kernel region to be re-segmented is expanded by a factor of *(1 + pw/2)*.



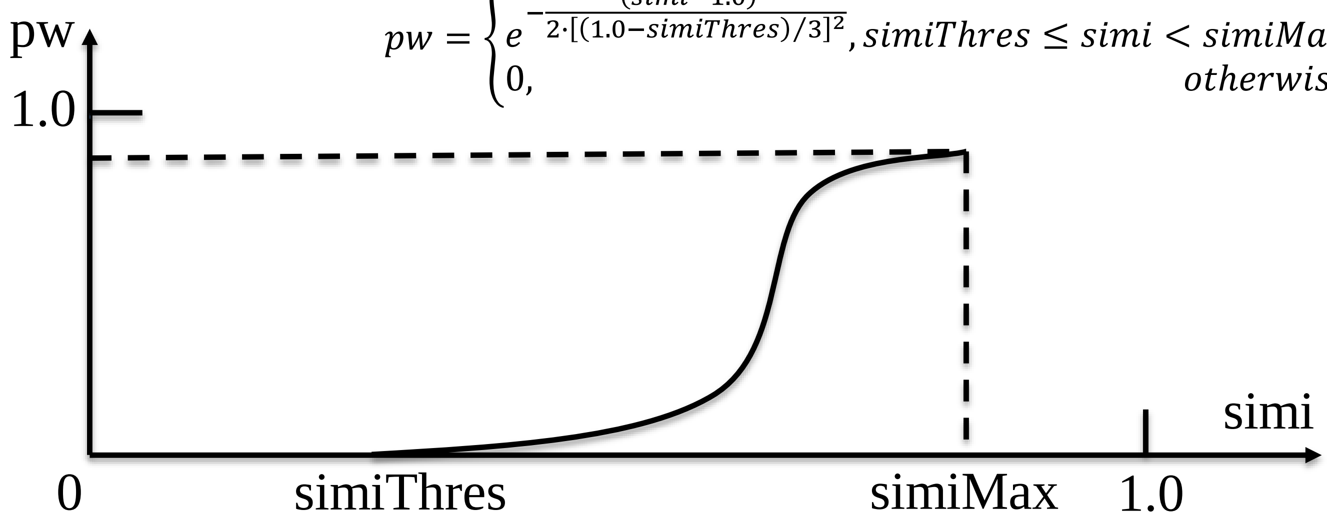
pw: penalty weight function
simi: color/chromaticity similarity



Penalizing Thresholds and Expanding Kernel Region

- The penalty weight is computed using a fuzzy Gaussian function:

$$pw = \begin{cases} e^{-\frac{(simi-1.0)^2}{2\cdot[(1.0-simiThres)/3]^2}}, & simiThres \le simi < simiMax \\ 0, & otherwise \end{cases}$$



  - *simi*: color similarity or chromaticity similarity
  - *simiThres*: threshold for the corresponding similarity
  - *simiMax*: upper limit for the similarity

Agenda

- Introduction
- System Overview
- General Segmentation and Tracking
- Similarity Computation and Feedback Loop
- **Experimental Results**
- Conclusion

Experimental Results of Segmentation

Quantitative comparison of MAST system to several state-of-the-art methods on seven measures in the *shadow* scenario of CVPR 2014 Change Detection challenge

	Recall	Spec	FPR	FNR	PWC	Prec	F
SuBSENSE [7]	0.9419	0.9920	0.0080	0.0581	1.0120	0.8646	0.8986
IUTIS-3 [12]	0.9478	0.9914	0.0086	0.0522	1.0410	0.8585	0.8984
GMM [13]	0.7960	0.9871	0.0129	0.2040	2.1951	0.7156	0.7370
CP3 [14]	0.7840	0.9832	0.0168	0.2160	2.5175	0.6539	0.7037
MAST	0.8679	0.9864	0.0136	0.1321	1.8906	0.7249	0.7884

TP: True Positive
FP: False Positive
FN: False Negative
TN: True Negative
Recall: TP / (TP + FN)
Spec (Specificity): TN / (TN + FP)

FPR (False Positive Rate): FP / (FP + TN)
FNR (False Negative Rate): FN / (TP + FN)
PWC (Percentage of Wrong Classifications):
100 * (FN + FP) / (TP + FN + FP + TN)
Precision: TP / (TP + FP)
F (F-Measure): (2 * Precision * Recall) / (Precision + Recall)

N. Goyette, P. M. Jodoin, F. Porikli, J. Konrad, and P. Ishwar, "Changedetection.net: A New Change Detection Benchmark Dataset," *IEEE Conf. Computer Vision and Pattern Recognition Workshops*, pp. 1-8, June 2012.

Experimental Results of Tracking

Comparison of average errors of tracking in terms of pixels on two video sequences, *TwoEnterShop2cor* and *ThreePastShop2cor*, in CAVIAR Dataset

Average error	MAST	Chu's method	SuBSENSE + Kalman filter tracking
TwoEnterShop2cor	10.06	26.72	17.27
ThreePastShop2cor	9.75	10.74	14.07
Overall	10.18	18.73	15.67

Average error: the distance of the centers of mass between experimental result and the ground truth in pixel

CAVIAR: Context Aware Vision using Image-based Active Recognition, EC founded CAVIAR project/IST 2001 37540, http://homepages.inf.ed.ac.uk/rbf/CAVIAR/.

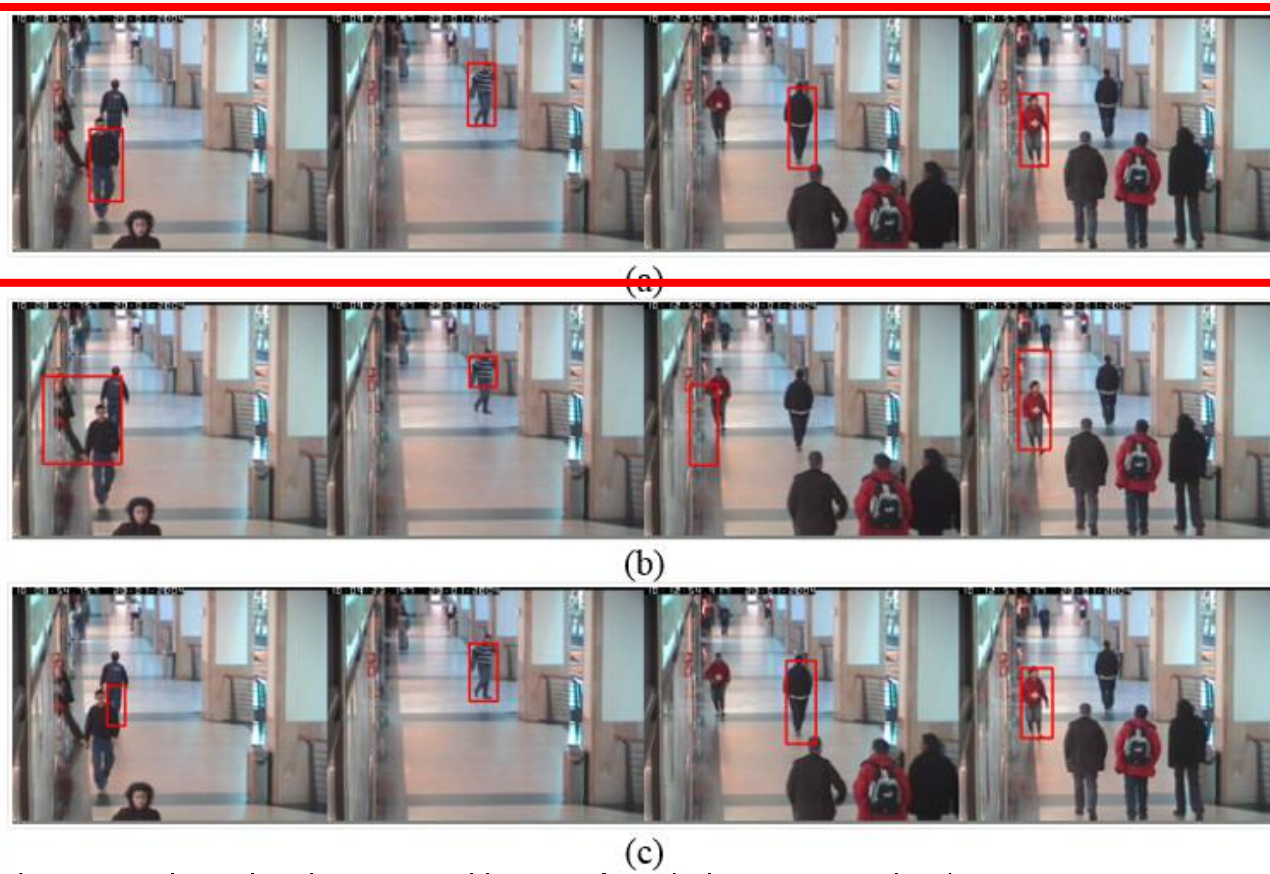
Experimental Results of Tracking

Representative frames of tracking results on two video sequences in CAVIAR Dataset

MAST

Chu's method

SuBSENSE + Kalman filter tracking



CAVIAR: Context Aware Vision using Image-based Active Recognition, EC founded CAVIAR project/IST 2001 37540, http://homepages.inf.ed.ac.uk/rbf/CAVIAR/.

Experimental Results of Tracking

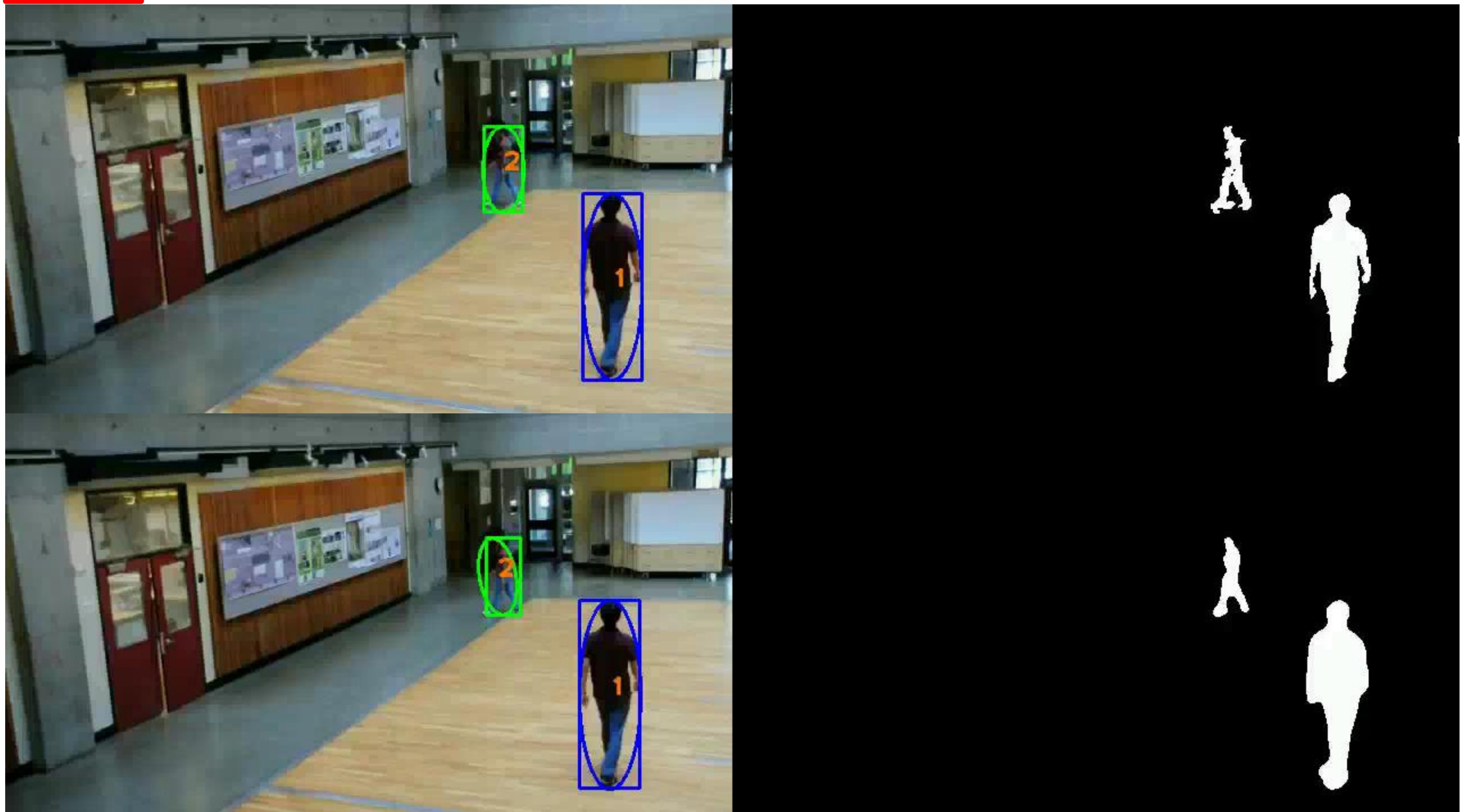
Comparison of average errors of tracking in terms of pixels on our two video sequences (video #1 and video #2)

Average error	MAST	Chu's method	SuBSENSE + Kalman filter tracking
video #1	17.88	18.26	18.90
video #2	10.30	16.10	15.95
Overall	14.09	17.18	17.43

Average error: the distance of the centers of mass between experimental result and the ground truth in pixel

Demo of Video #1

MAST Current frame | Foreground



SuBSENSE + Kalman filter tracking

Agenda

- Introduction
- System Overview
- General Segmentation and Tracking
- Similarity Computation and Feedback Loop
- Experimental Results
- **Conclusion**

Conclusion

- We proposed an adaptive segmentation and tracking system based on multiple kernels.
- The purpose is to robustly track objects when they have similar color or chromaticity with the background area.
- It is shown that MAST system can improve the performance of tracking while keeping favorable segmentation results especially when dealing with object merging problem.
- The complete demo videos can be viewed on: http://allison.ee.washington.edu/thomas/mast/

Thank you!

Q&A