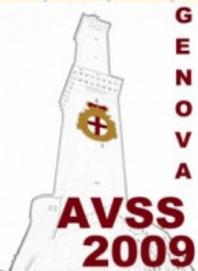


# Multimodal Information Group



## 2009 AVSS Multiple Camera Person Tracking Results Summary

<http://www.itl.nist.gov/iad/mig/tests/avss/2009/index.html>

**Jonathan Fiscus**, Travis Rose, Martial Michel,  
John Garofolo, Jerome Ajot

Sept. 2-4, 2009

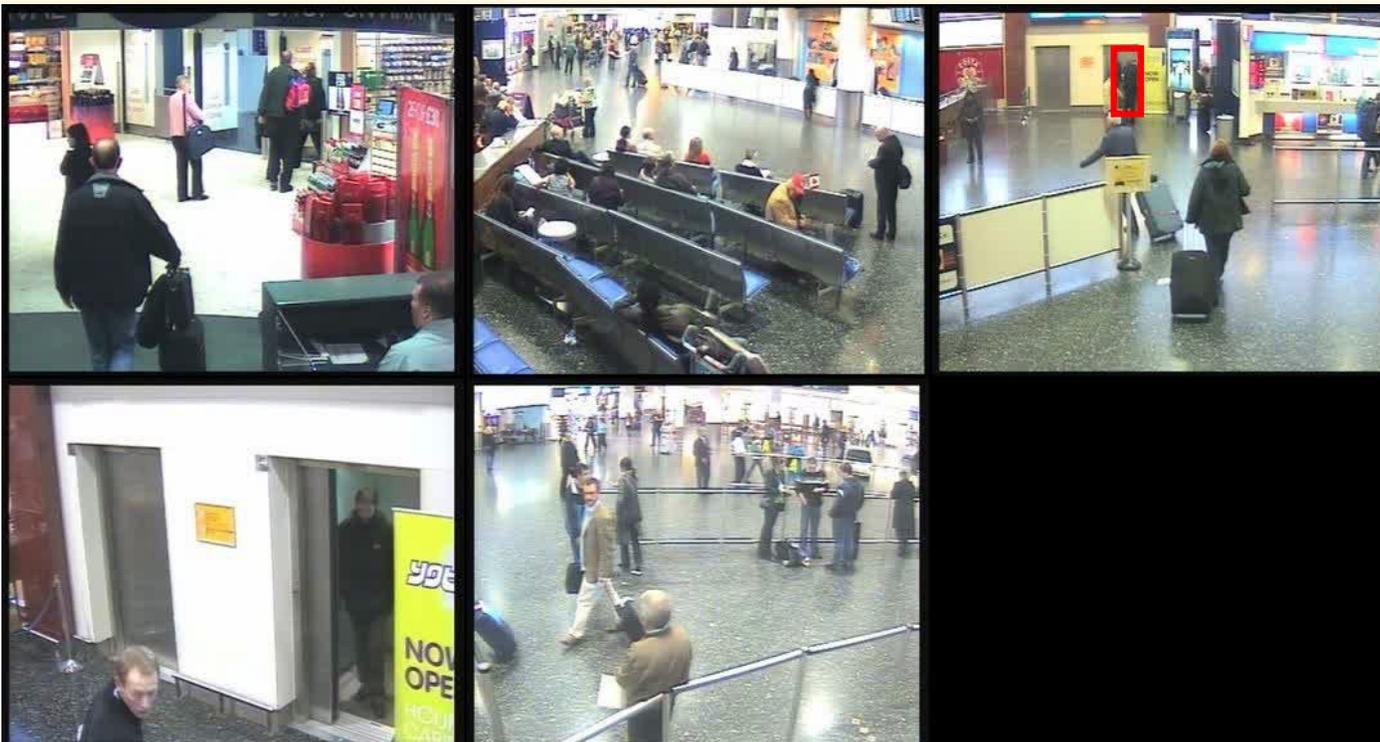
# Single Person Tracking (SPT) Task Definition

Given 5 *in situ* video frames with bounding box data specifying a person to be tracked, track the person in 5, 2 or 1 camera views by outputting bounding boxes

## Side information

Schematic of the venue with approximate camera locations

Single video frame of ROTOKIN calibration target from each camera



## Key Challenges:

- Tracking through visual occlusion
- Subject re-acquisition

# Motivation for SPT

- **Single person tracking is key enabling technology**
  - The “framework” to carry high-level analysis of behaviors and people interacting with each other and objects
  - Precursor for geospatial tracking
- **Realistic data**
  - Tracking in complex scene
  - Non-overlapping views are the norm
  - Intrinsic/Extrinsic camera calibration is impractical for large networks

# Experimental Design

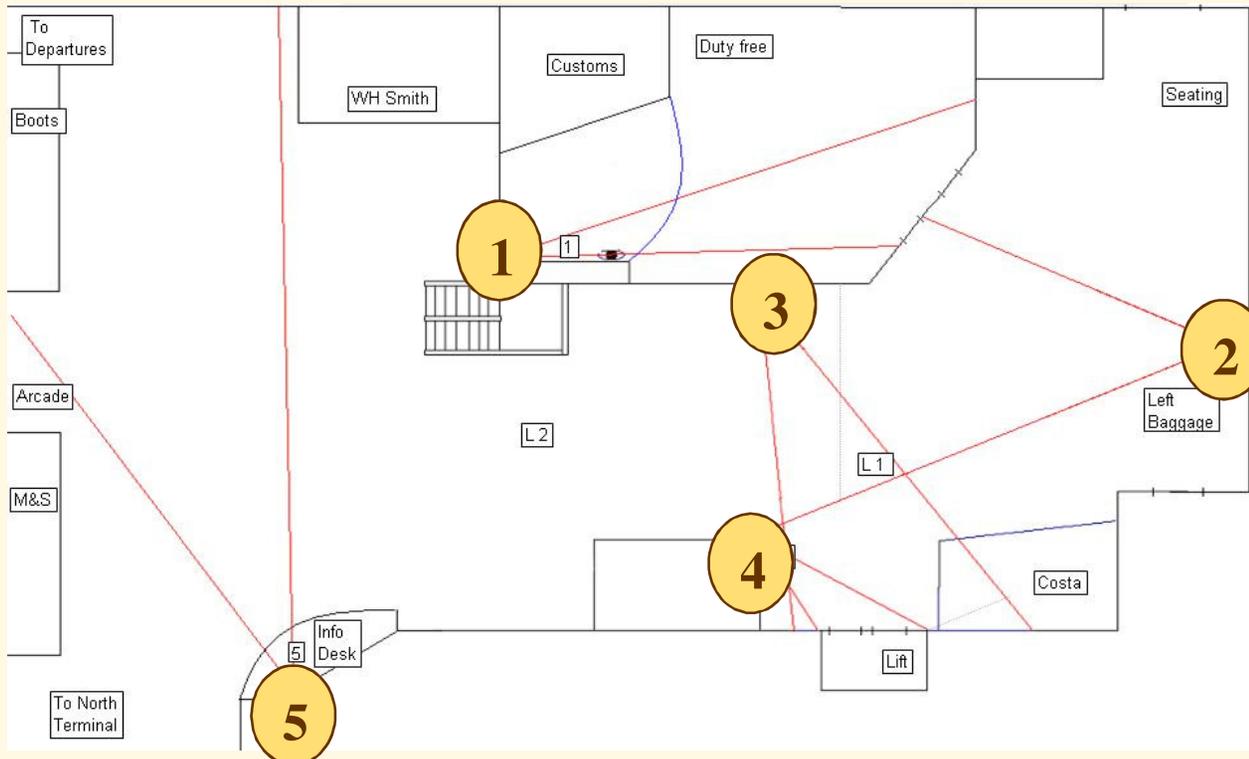
- **Tracking training frames specify person to track**
  - 5 video frames from a single (primary) camera view
  - Bounding box annotation of non-occluded image
- **Three camera set conditions:**
  - Multiple Camera SPT (MCSPT) : Track in 5 cameras
  - Camera Pair SPT (CPSPT): Track in 2 cameras
    - MCSPT has a limited set of primary/secondary camera pairs
    - Simplification to increase the number of trials
  - Single Camera SPT (SCSPT): Track in 1 camera
    - Failure to re-acquire the subject in MCSPT/CPSPT conditions will mask within-frame tracking performance.
    - Simplification to test within frame tracking in many more camera views

# Experimental Design

## Test Selection Procedure

- **Data:**
  - Home Office's i-LIDS Multiple Camera Tracking Training Corpus
    - The corpus is publically available
    - The AVSS 2009 evaluation is potentially not "blind"
  - 105, 5-camera single person tracking sets collected on twelve days
  - Subject annotation every 5 frames
- **Divided into two subsets:**
  - AVSS 2009 Development set (2/3 of data)
  - AVSS 2009 Evaluation test set (1/3 of data)
- **Division criteria:**
  - Analysis of camera involvements for each tracking trial
  - Development and evaluation sets were drawn from separate recording days

# LGW Schematic



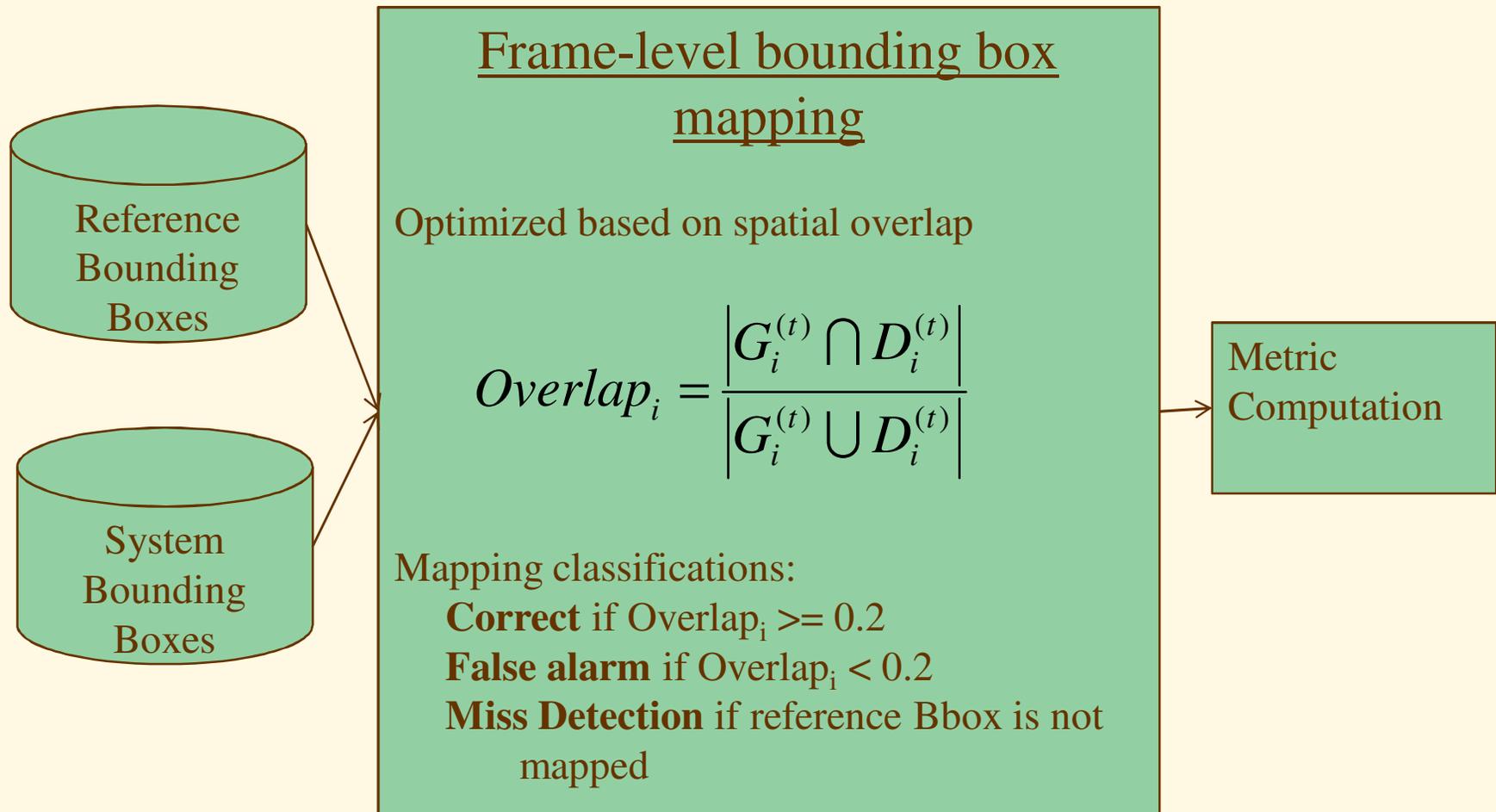
# AVSS 2009 MCT Data Set

	Development MCSPT	Test Set Characteristics		
		MCSPT	CPSPT	SCSPT
Number of Tracking Trials	70	35	26	53 (Derived from CPSPT)
Total Camera hrs. of video	~28.1	~15.2	~4.3	~4.4
Camera hrs. of annotated video	~12.5	~6.9	~1.4	~0.8
Number of annotated frames	43,426	23,705	11,253	13,157
Percent of annotated frames where the subject was not occluded	58%	53%	48%	52%

# 2009 Evaluation Participants

Site	Evaluation Task		
	Multiple Camera	Camera Pair	Single Camera
Athens Inst. Of Tech. (AIT)			X
Brno Univ. (Brno)	X	X	
GE Global Research Center (GEGRC)			X
Kingston Univ. (KuDir)	X		

# Evaluation Procedure: CLEAR MOTA and MOTP Metrics



# Tracking Metrics

## Multiple Object Tracking Accuracy (MOTA)

$$1 - \frac{\sum_{t=1}^{N_{frames}} (c_m (miss^{(t)}) + c_f (falseAlm^{(t)}))}{\sum_{t=1}^{N_{frames}} N_G^{(t)}}$$

Perfect: MOTA = 1.00

Range: [1:-∞)

{negative MOTA possible due to false alarms }

## Multiple Object Tracking Precision (MOTP)

$$\frac{\sum_{i=1}^{N_{correct}} \sum_{t=1}^{N_{frames}^i} \left[ \frac{|G_i^{(t)} \cap D_i^{(t)}|}{|G_i^{(t)} \cup D_i^{(t)}|} \right]}{\sum_{t=1}^{N_{frames}} N_{correct}^{(t)}}$$

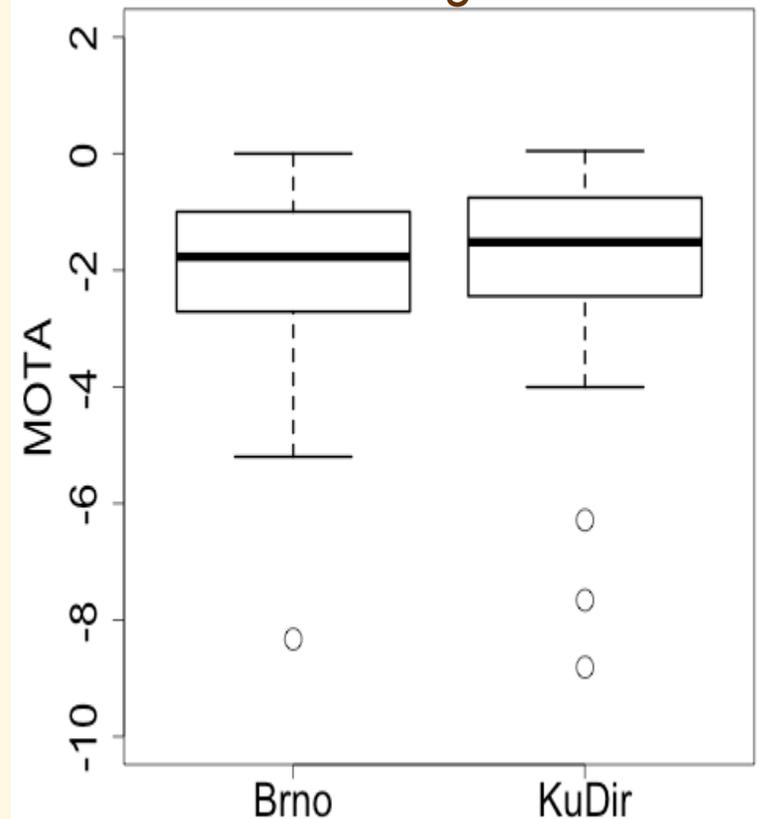
Perfect: MOTP = 1.00

Range: [1:0]

# MCSPT Results

			Brno	KuDir
MOTA	Test Set Average		-1.183	-1.40
	Track averaged	Mean	-2.052	-2.072
		Median	-1.770	-1.517
MOTP (Primary camera)	Test set average		NA	NA
	Track averaged	Mean	0.22	0.24
		Median	0.24	0.17

Distribution of MOTAs by Tracking Trials



# MCSPT: Primary/Secondary Camera Subject Re-acquisition

Heuristically define a “Successful” subject re-acquisition occurred when Prob. of Detection is  $> 0.2$  (anywhere) in secondary camera

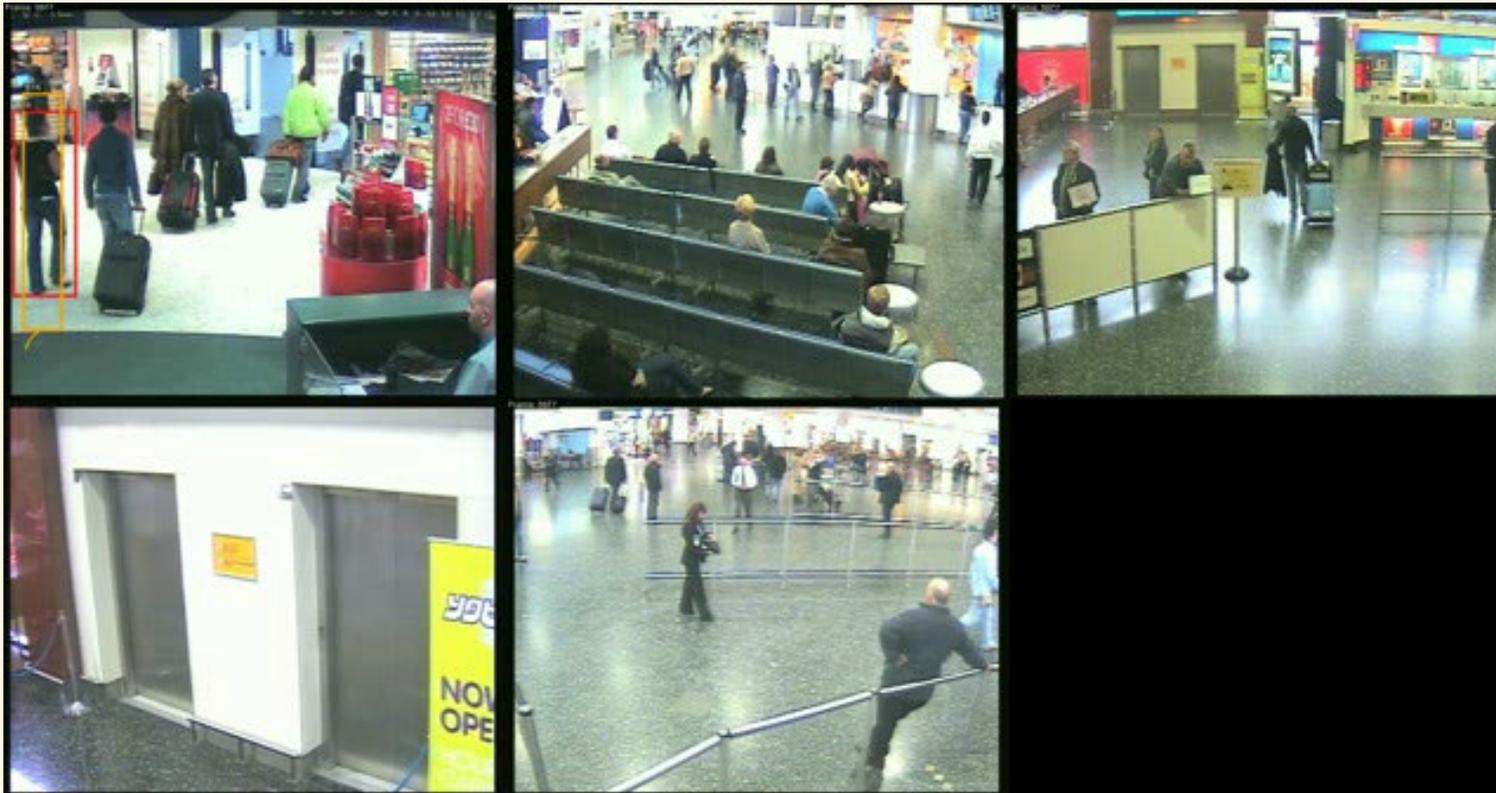
		Ground truth Re-acquisitions				
		Secondary Camera				
		1	2	3	4	5
Primary Camera	1		9			
	2			8		
	3				7	
	4			1		
	5			9		

Counts of Successful Re-acquisitions at $P_{det} \geq 0.2$										
Brno						KuDir				
Secondary Camera						Secondary Camera				
1	2	3	4	5		1	2	3	4	5
	1						0			
		2						0		
			0						0	
		0						0		
		0						0		

# MCSPT Subject Re-Acquisition Example (Brno)

Trial MOTA: -2.24

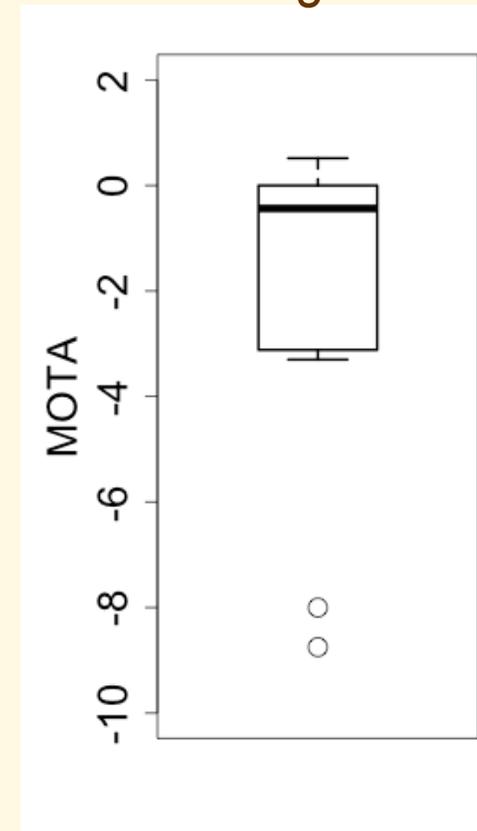
MOTP: 0.30



Non-Occ. Ref  
Occ. Ref  
Correct Det.  
Missed Detect  
False Alarm

# CPSPT Results

Distribution of MOTAs by Tracking Trials



			Brno
MOTA	Test Set Average		-1.08
	Track averaged	Mean	-2.16
		Median	-1.38
MOTP (Primary Cameras)	Test set average		NA
	Track averaged	Mean	0.21
		Median	0.23

# CPSPT: Primary/Secondary Camera Subject Re-acquisition

Heuristically define a “Successful” subject re-acquisition occurred when Prob. of Detection is  $> 0.2$  (anywhere) in secondary camera

		Ground truth Re-acquisitions				
		Secondary Camera				
		1	2	3	4	5
Primary Camera	1		3			
	2			3	3	2
	3		3		3	3
	4		3			
	5			3		

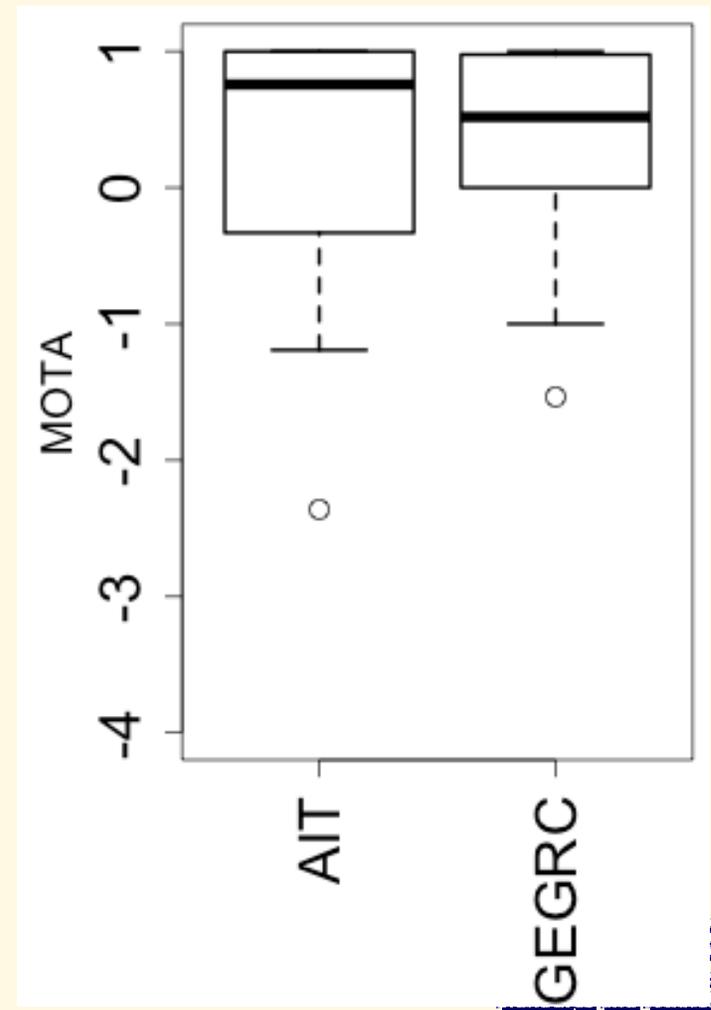
Counts of Successful  
Re-acquisitions  
Brno

		Secondary Camera				
		1	2	3	4	5
Primary Camera	1		0			
	2			1	0	0
	3		0		0	0
	4		0			
	5			0		

# SCSPT Results

Distribution of MOTAs by Tracking Trials

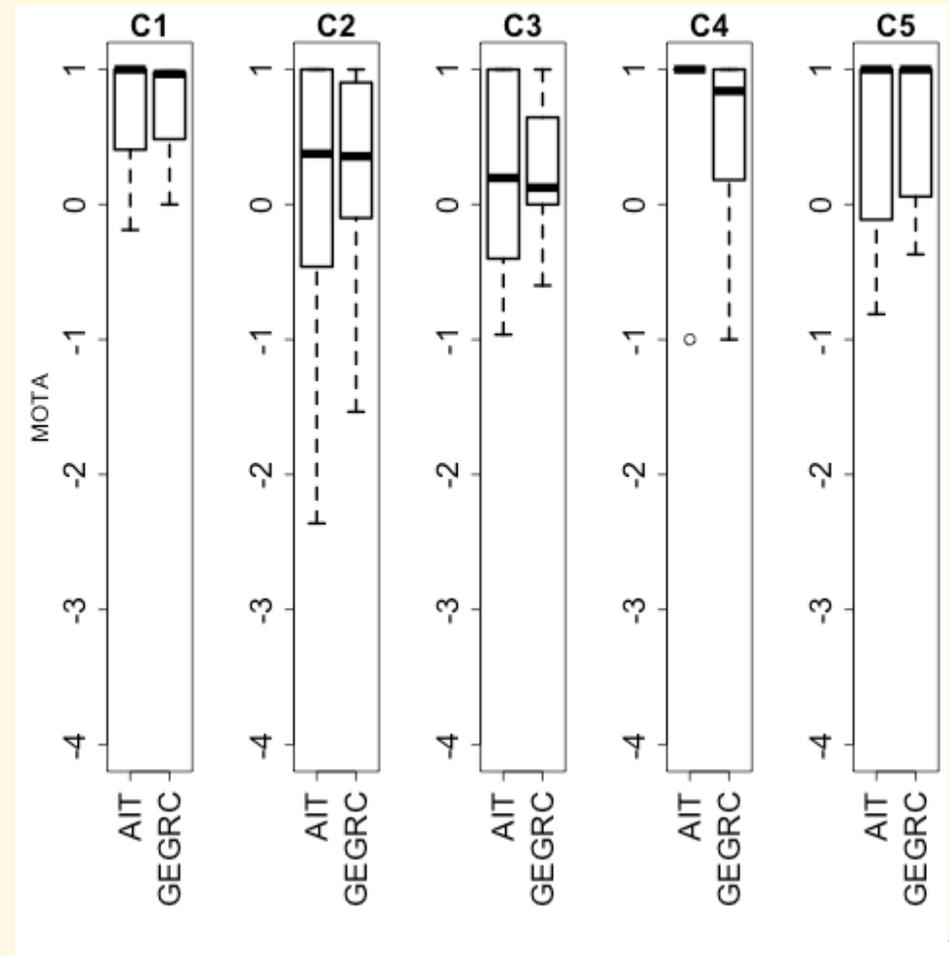
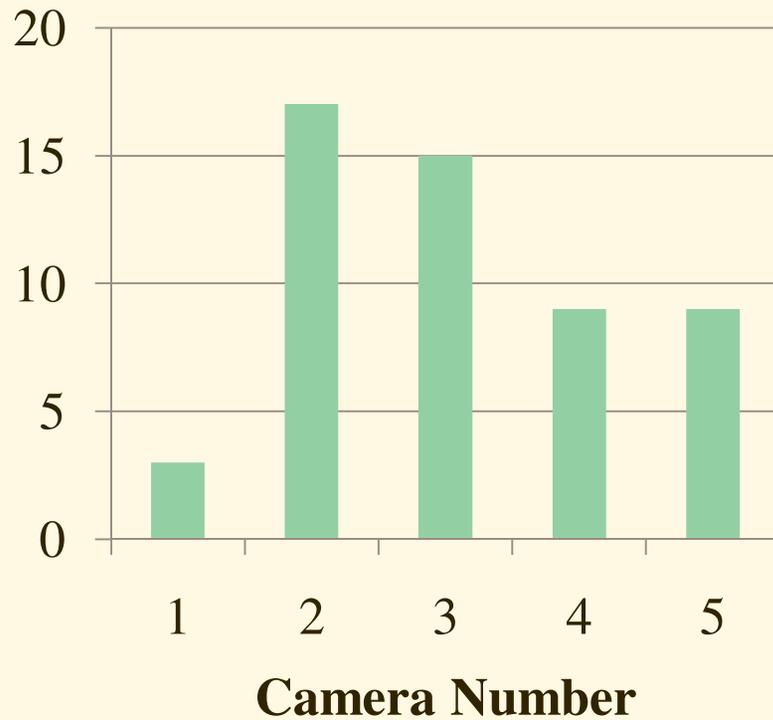
			AIT	GEGRC
MOTA	Test Set Average		-0.196	0.24
	Track averaged	Mean	0.30	0.38
		Median	0.75	0.51
MOTP	Test set average		NA	NA
	Track averaged	Mean	0.39	0.34
		Median	0.41	0.40



# SCSPT Results by Camera

Distribution of Tracking Trial MOTAs as a function of Camera

## Number of Tracking Trials per Camera

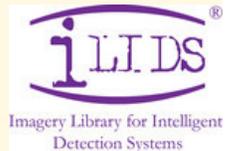
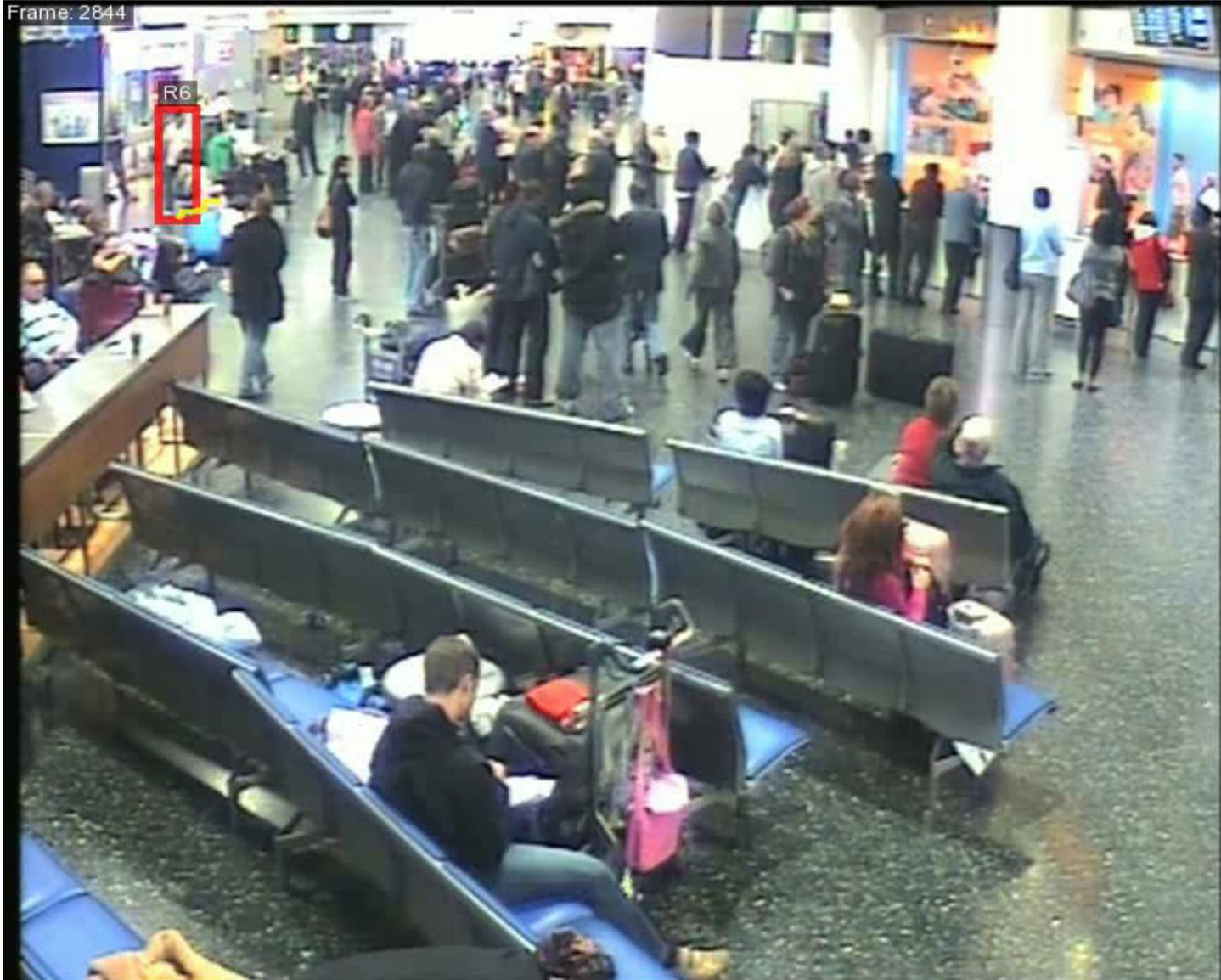


# Longest Correct Track

GEGRC-SCSPT

MOTA=0.90

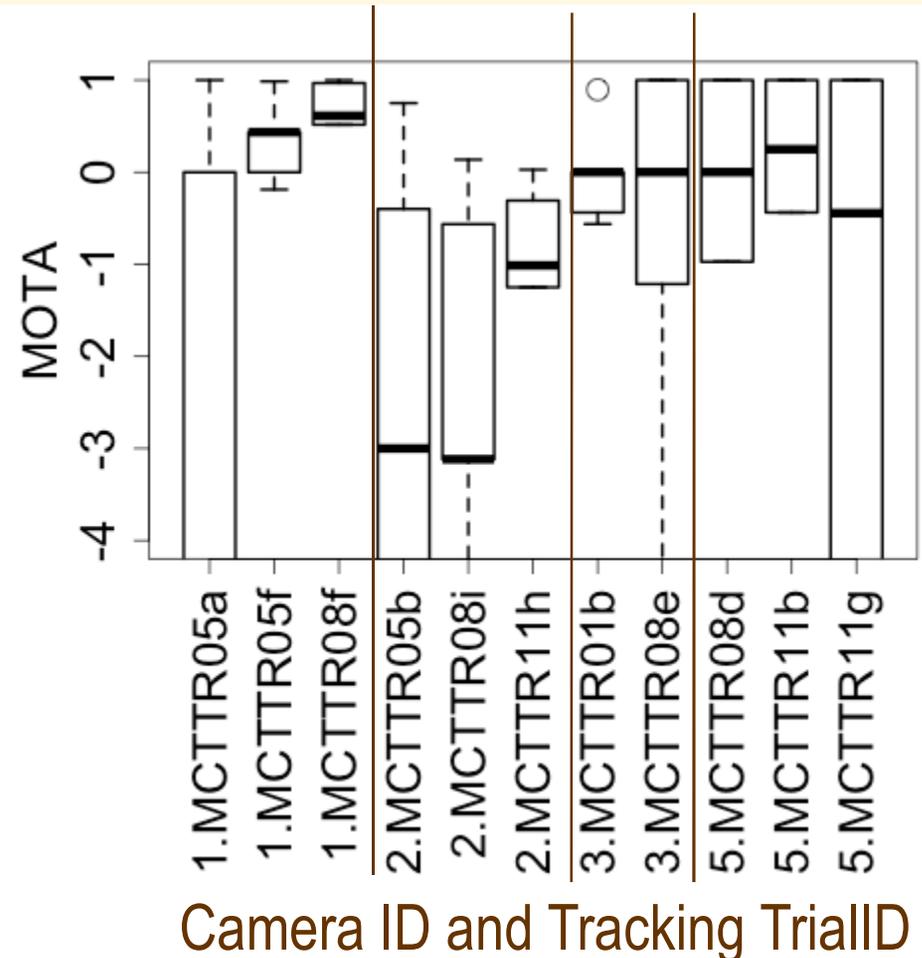
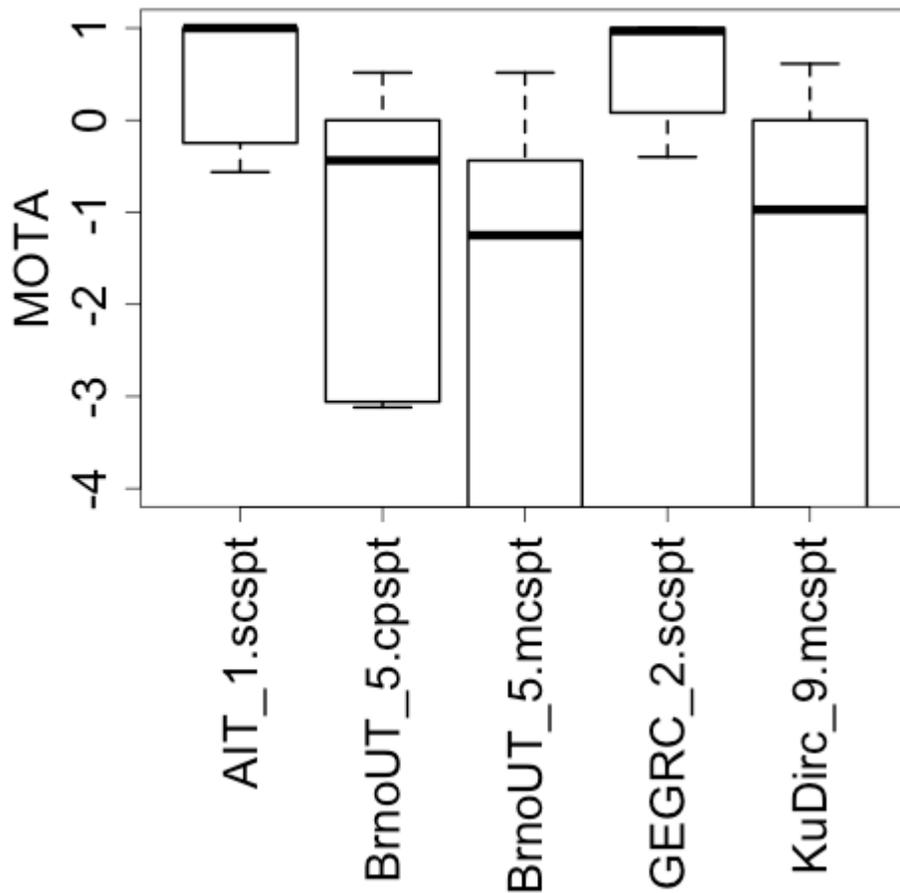
413 Ground Truth Frames



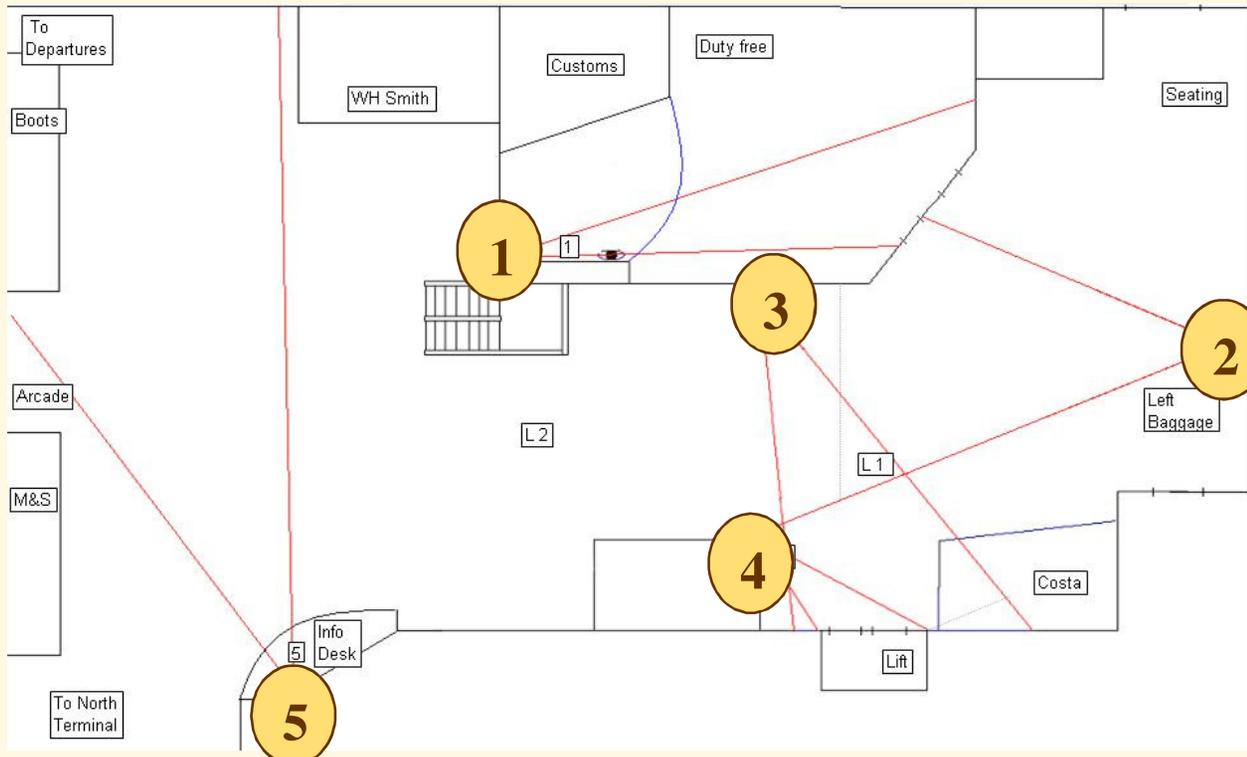
Non-Occ. Ref  
Occ. Ref  
Correct Det.  
Missed Detect  
False Alarm

# Single Camera Trial Subset for all Three Test Conditions

- 11 primary camera tracking trials were in all three test conditions



# LGW Schematic



# Conclusions

- **The task is very challenging for the brave**
  - [...] We few, we happy few, we band of brothers, For he today who sheds his blood with me Shall be my brother [...] – Henry V
  - Both occlusion and subject re-acquisition remains a challenge
- **Additional analysis**
  - Tracking as a function of person size
  - Maintaining tracks through occlusions
- **Retest evaluation conducted through the end of November**
  - Additional submissions added to the formal report
  - See me or HOSDB for data access
- **Additional AVSS MCPT evaluations**
  - AVSS '10 challenge evaluation
  - We would like to run a re-test Interim evaluations

*Thank you for your attention*

**Special thanks to:**

AVSS Organization Committee

CPNI

HOSDB – iLIDS Data Team

Participants