

Security and surveillance: performance evaluation

Andrea Cavallaro

Queen Mary, University of London

andrea.cavallaro@elec.qmul.ac.uk
<http://www.elec.qmul.ac.uk/staffinfo/andrea>

SSPE 2007



"If you can not measure it,
you can not improve it."

Lord Kelvin (1883)

SSPE 2007



Ideal measure for video analytics ...

Number of false alarms / camera / day

Evaluation in video surveillance research

- Examples
 - PETS workshop series
 - UK i-LIDS: imagery Library for Intelligent Detection Systems
 - ETISEO: French Techno-Vision evaluation network
 - CAVIAR project: Context Aware Vision using Image-based Active Recognition
 - US ARDA VACE programme
 - VERAEE project: Video Event Recognition Algorithm Assessment Evaluation
 - CLEAR workshops: Classification of Events, Activities and Relationships
 - CHIL project: Computers In the Human Interaction Loop
 - AMI project: Augmented Multi-part Interaction
 - Evaluation:
 - distributed datasets
 - ground-truth
 - metrics for comparing results

Performance evaluation and validation

- Performance evaluation: still an open problem
 - evaluation metrics and datasets are not yet taken up by the research community at large
 - computer vision algorithms need be evaluated with large test corpora containing significant statistical data variability as the accuracy of their results is highly data dependent
- Lack of
 - commonly accessible datasets and
 - commonly accepted evaluation protocols

SSPE 2007



Objective of this symposium

- To provide a forum for the discussion of
 - algorithms
 - results
 - protocols
 - datasets
- To promote evaluation
- To promote data collection and (free) distribution

SSPE 2007

BMVA symposium
**Security and surveillance:
performance evaluation**
(SSPE 2007)



Surveillance Performance Evaluation Initiative

- SPEVI
 - Surveillance Performance Evaluation Initiative – www.spevi.org
 - One-stop web site collecting existing datasets
 - Pointers to other evaluation programmes / datasets
 - Server hosting received datasets and ground-truth
 - Free distribution (for research)
 - Requested citation acknowledgment
 - If you want to contribute to this collection of datasets please contact info@spevi.org

SSPE 2007



Round table discussion

SSPE 2007



Datasets / tasks

- Are we evaluating too many things simultaneously?
 - Too many variables
- Is a dataset too heterogeneous?
 - Can we generalize the results obtained so far?
- Are we happy with the current definition of events?
- Do we want to evaluate all the event types together?
 - should we focus on subsets of events and move on progressively
- Is the dataset too heterogeneous?
 - Can we generalize the results obtained so far?

Measures / ground truth

- Give meaning to measures
 - what is the “value” of these numbers?
 - e.g., compare with a naïve result
 - what is the “value” of a difference of (e.g.) 0.1?
- How many measures do we need?
 - remove redundant measures
- Is the ground truth really “truth”?
 - statistical analysis / more annotators / confidence level
- Are we happy with the current demarcation of regions?

- **Person and vehicle detection and tracking: experiments on the CLEAR dataset**
M. Taj, E. Maggio, A. Cavallaro
CLEAR 2007, Baltimore, Maryland (USA), May 8-9 2007
- **Evaluation of multi-sensor surveillance event detectors**
F. Ziliani and A. Cavallaro
ICDP 2006, London, UK, 13-14 June, 2006
- **Event detection in underground stations using multiple heterogeneous surveillance cameras** [invited paper]
A. Cavallaro
ISVS 2005, Lake Tahoe, Nevada, 5-7 December 2005.
- **Performance evaluation of event detection solutions: the CREDS experience**
F. Ziliani, S. Velastin, F. Porikli, L. Marcenaro, T. Kelliher, A. Cavallaro,
P. Bruneaut (AVSS 2005), Como (I), 15-16 September 2005