Multi-camera networks

Andrea Cavallaro

Multi-camera networks

Multi-camera networks

Multi-camera networks

Context

EU FP7 project APIDIS

Autonomous production of images via distributed and intelligent sensing

www.apidis.org

UK EPSRC project MOTINAS

Multi-modal object tracking in networks of audiovisual sensors

www.spevi.org

EU Erasmus Mundus Joint Doctorate ICE

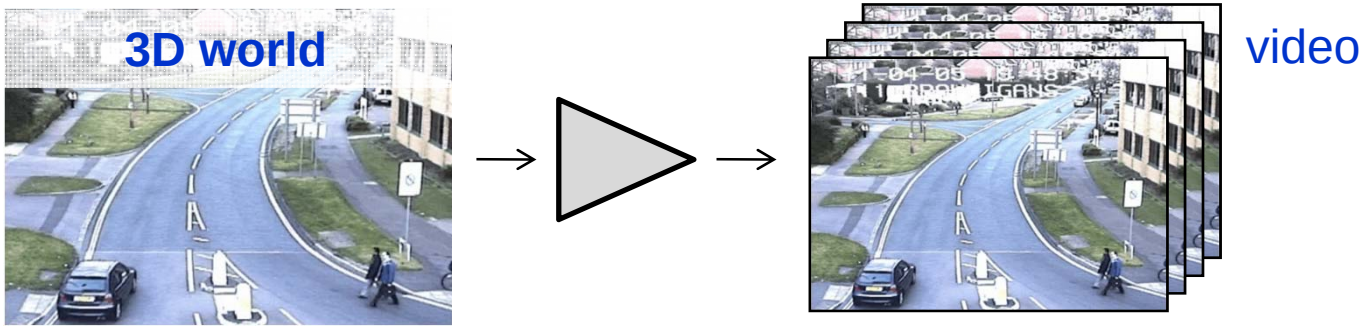
Interactive and Cognitive Environments

www.icephd.org

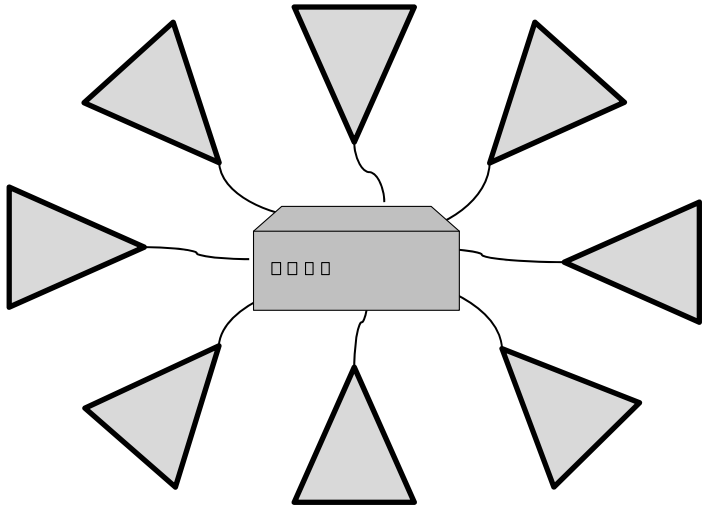
Motivation

Traditional & desired solutions
Deployment examples
Scale

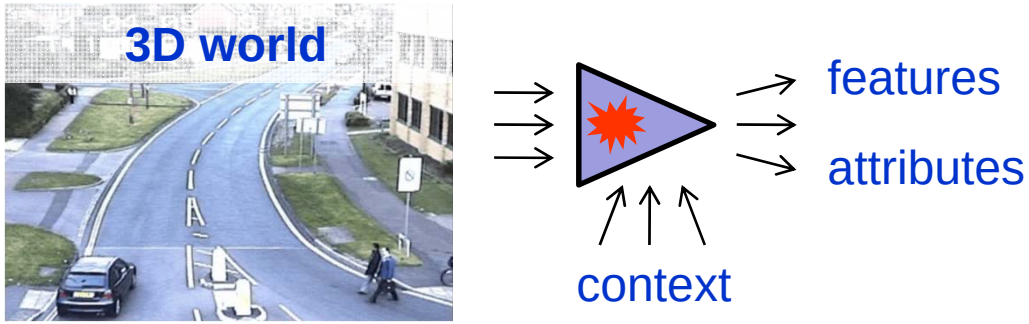
Traditional solution



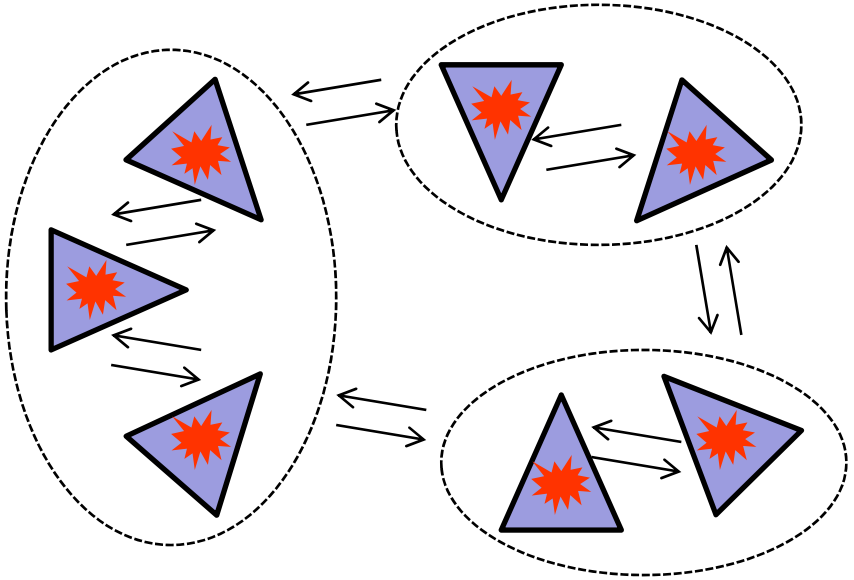
- Traditional multi-camera systems
  - stream images to a powerful centralised (processing) facility
  - need accurate calibration
  - need precise synchronisation
  - have limited scalability



Desired solution



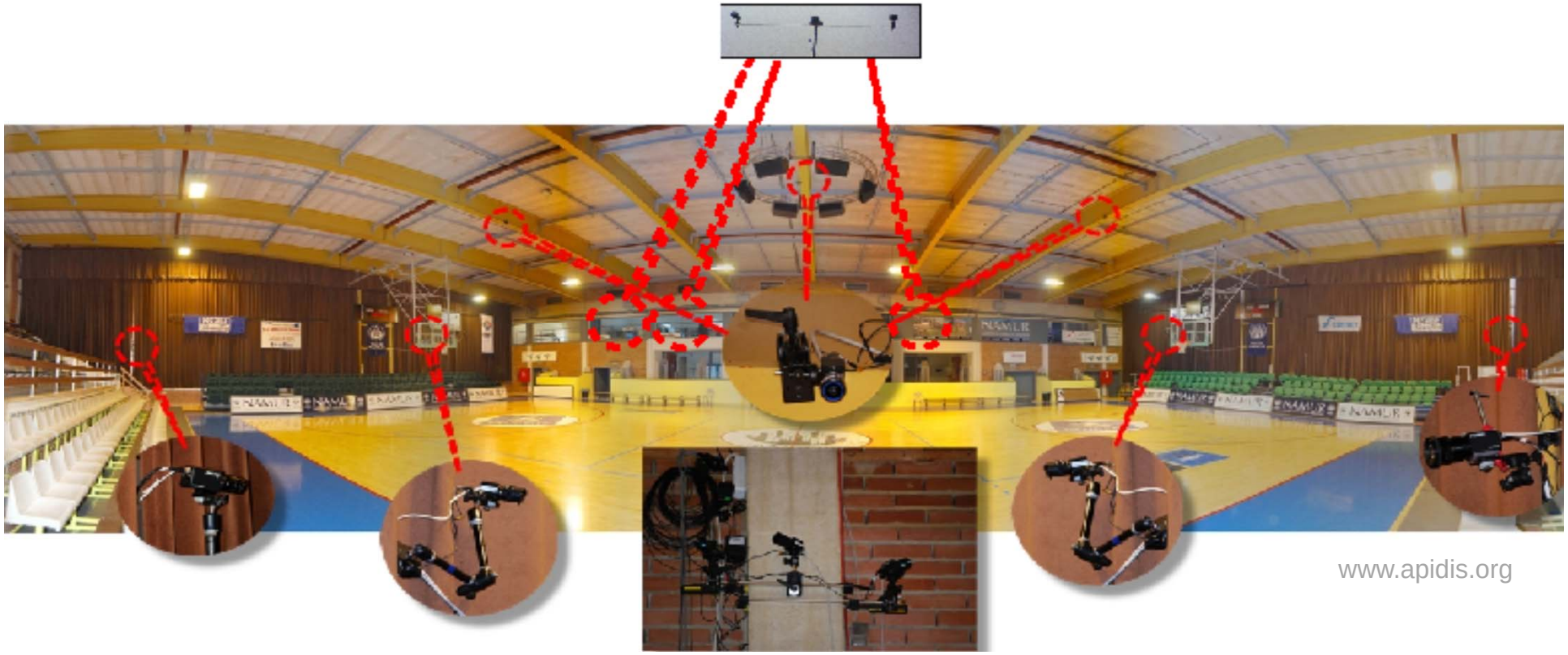
- Current trend
  - from centralised to distributed
  - processing at the nodes
  - exchange of metadata
  - scalability
  - multimodal sensors



Deployment example 1

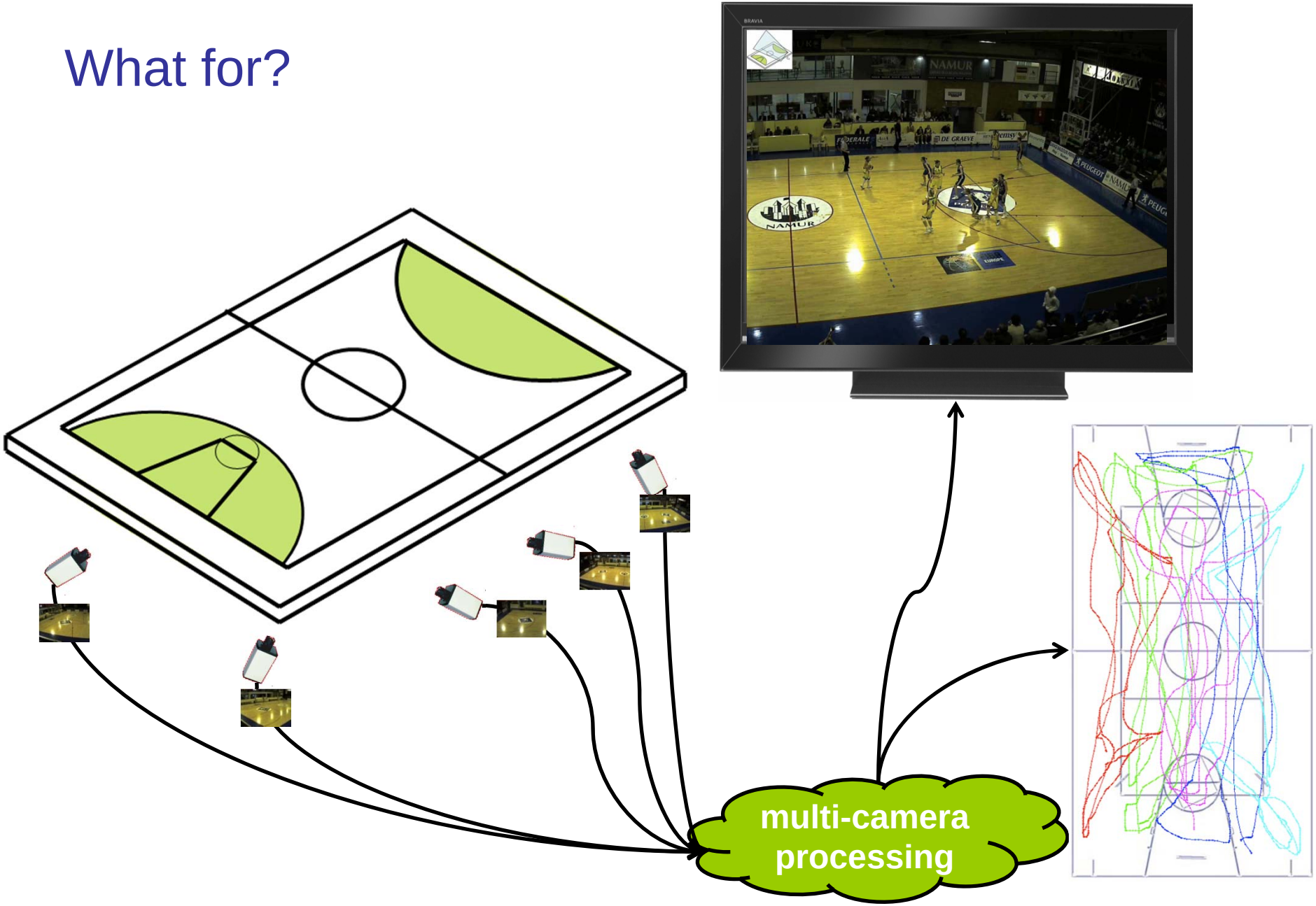
www.apidis.org

Deployment example 2



www.apidis.org

What for?



Multi-camera networks: scale

In-body
health
medicine

Personal
video conferencing
gaming

Home
energy efficiency
domotics

Building
service optimisat.
security

Facility
goods tracking
safety

City
traffic control
planning

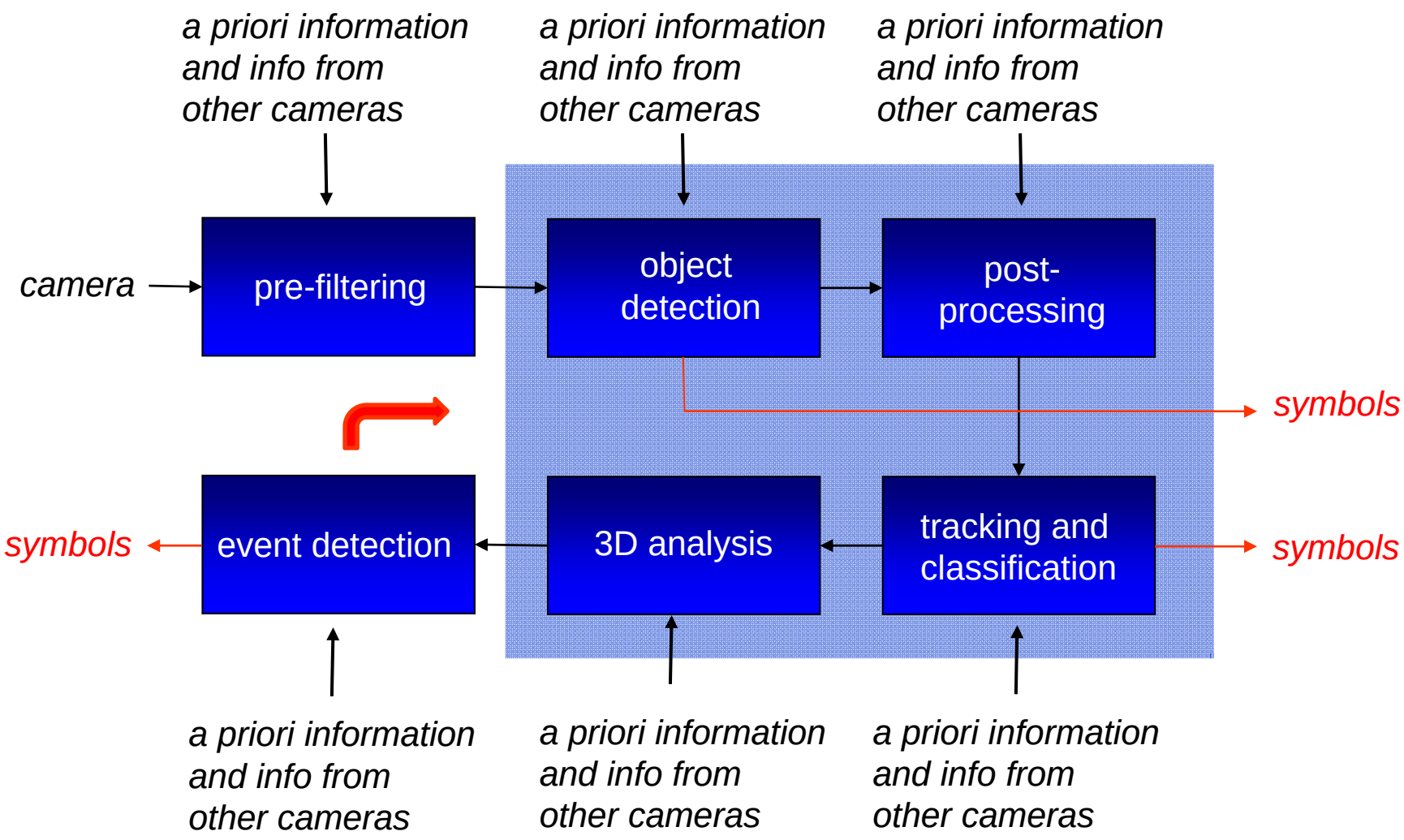
Inter-city
traffic monitoring
transit control

In the wild
fauna monitoring
fire detection

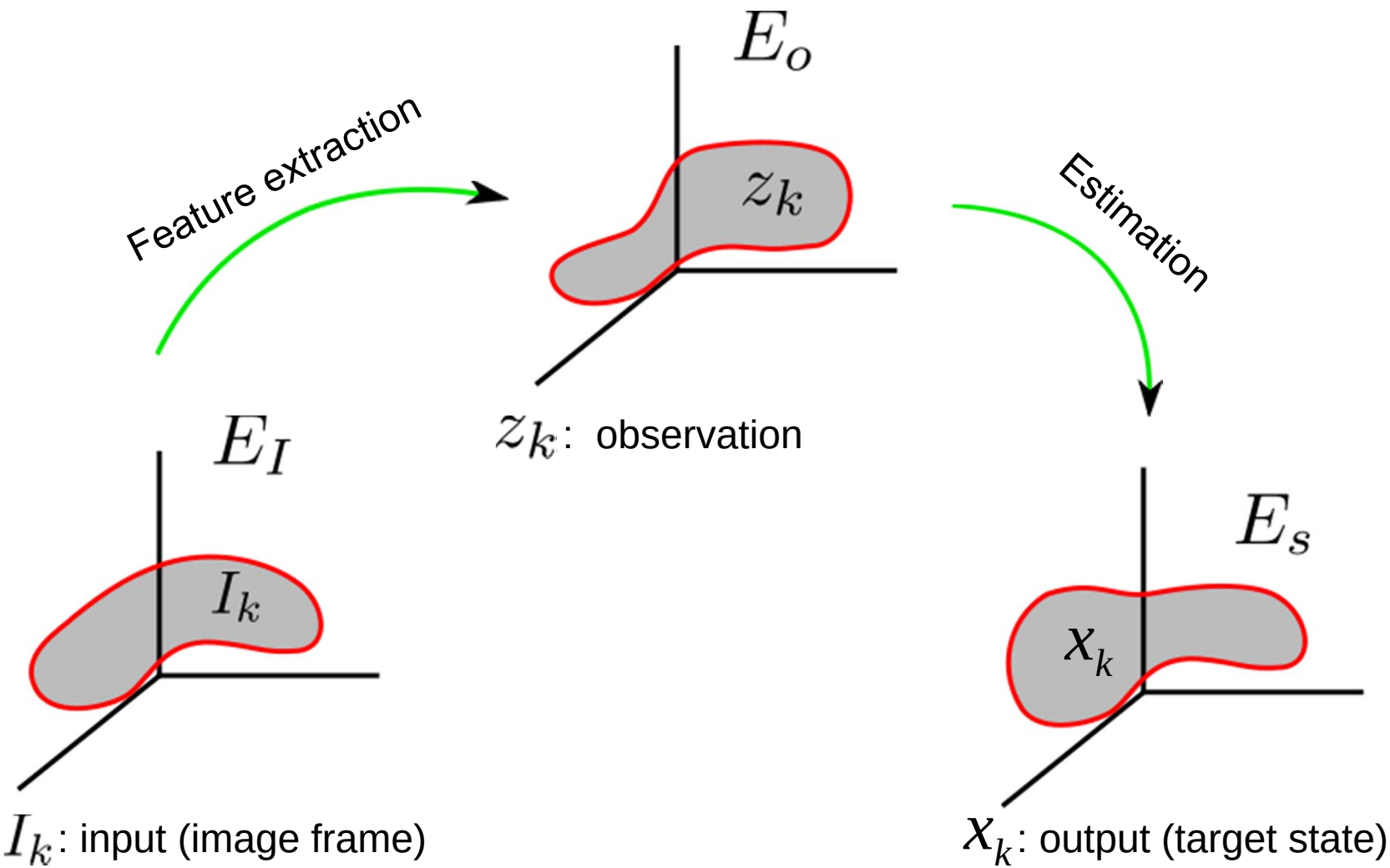
Multi-camera networks

Processing pipeline
States and observations
Bayesian approaches
Context: on-line learning
User feedback

Queen Mary
University of London

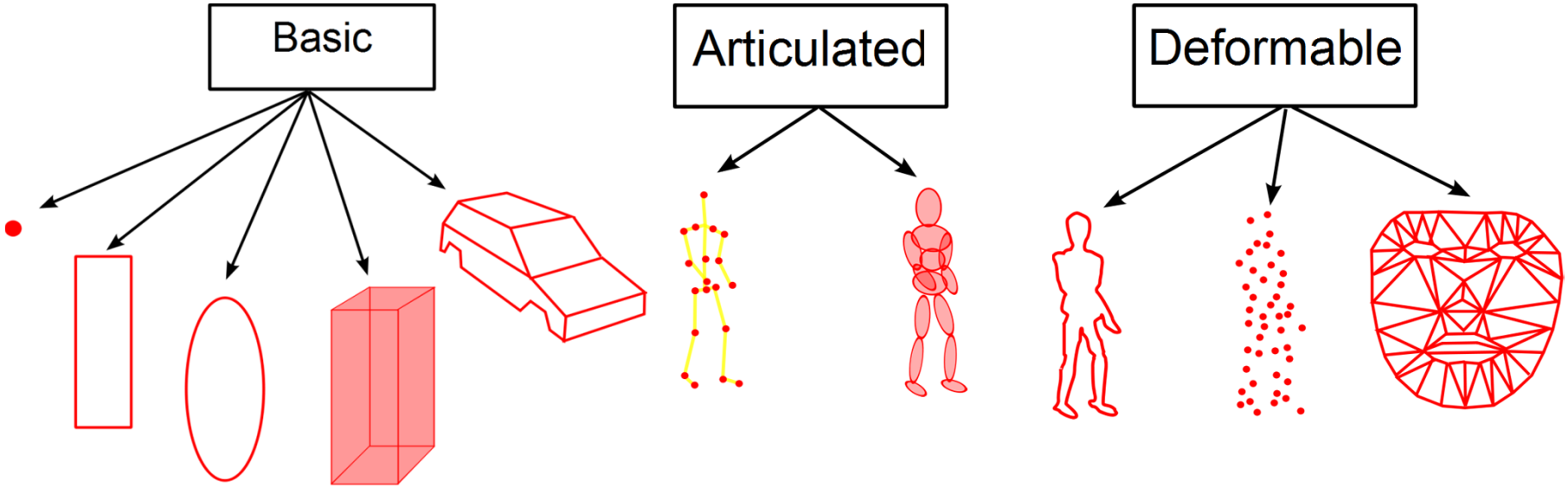

a priori information and info from other cameras
a priori information and info from other cameras
a priori information and info from other cameras
camera
pre-filtering
object detection
post-processing
symbols
symbols
event detection
3D analysis
tracking and classification
symbols
a priori information and info from other cameras
a priori information and info from other cameras
a priori information and info from other cameras

Information flow



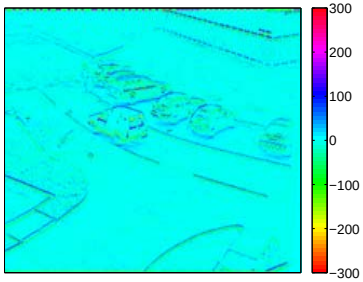
Target state

- The state x_k is a vector of parameters
  - Shape and position
  - Appearance
  - Higher order moments (e.g. speed, acceleration)

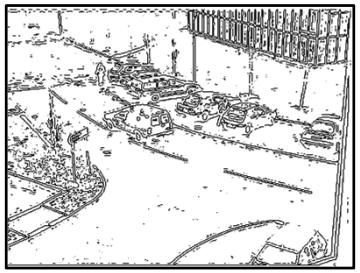


Observation

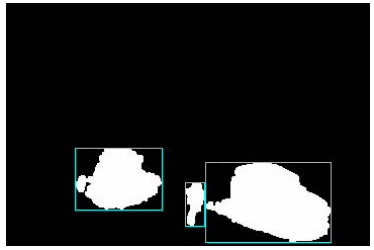
- The observation vector z_k represents object information extracted from the image
  - Low-level information
    - e.g. pixel colour, gradient



  - Mid-level features
    - e.g. location of edges

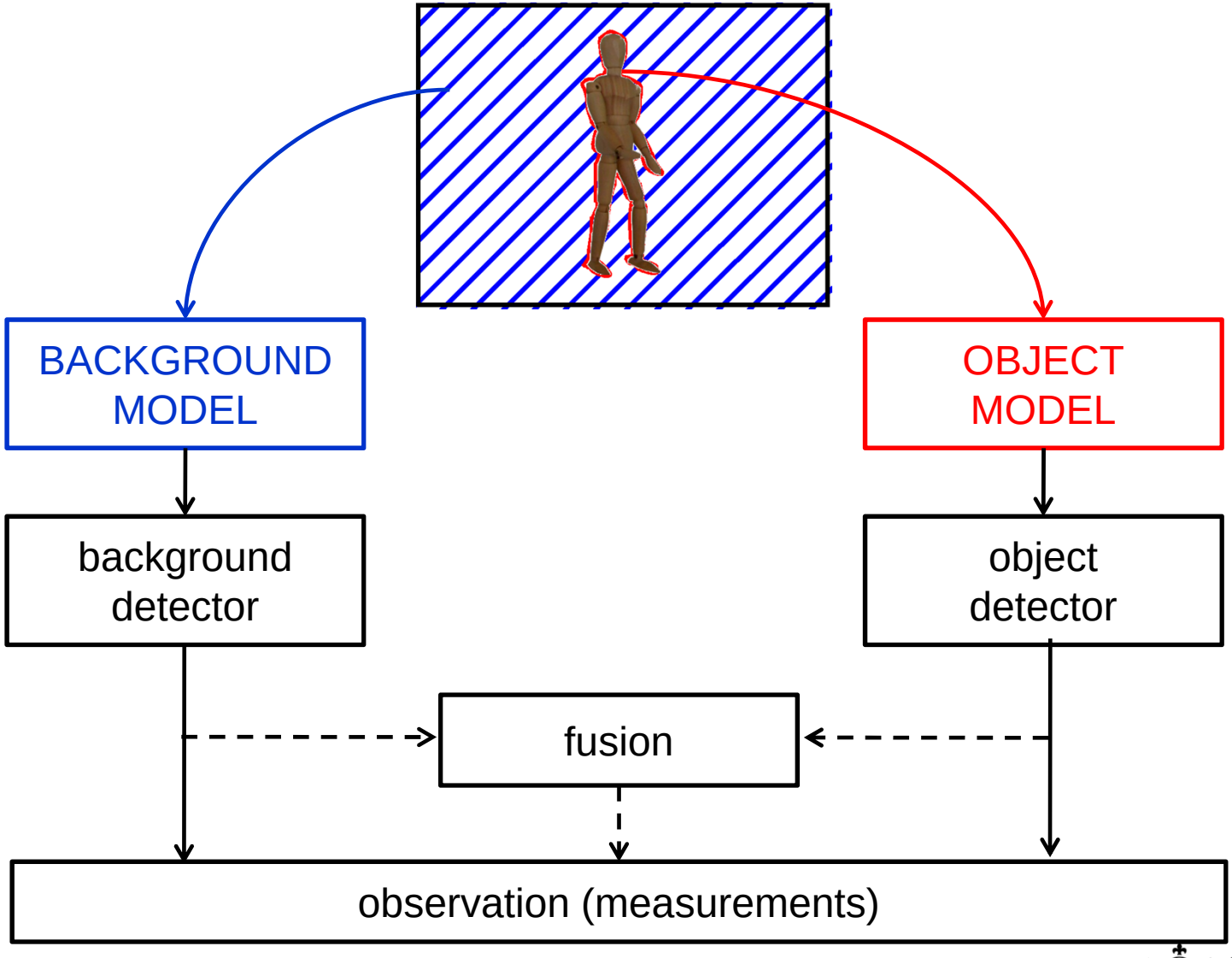


  - High-level features
    - e.g. object detection



  - Combinations of the above

Object detection strategies



Multi-target observation and state: example

- Observations

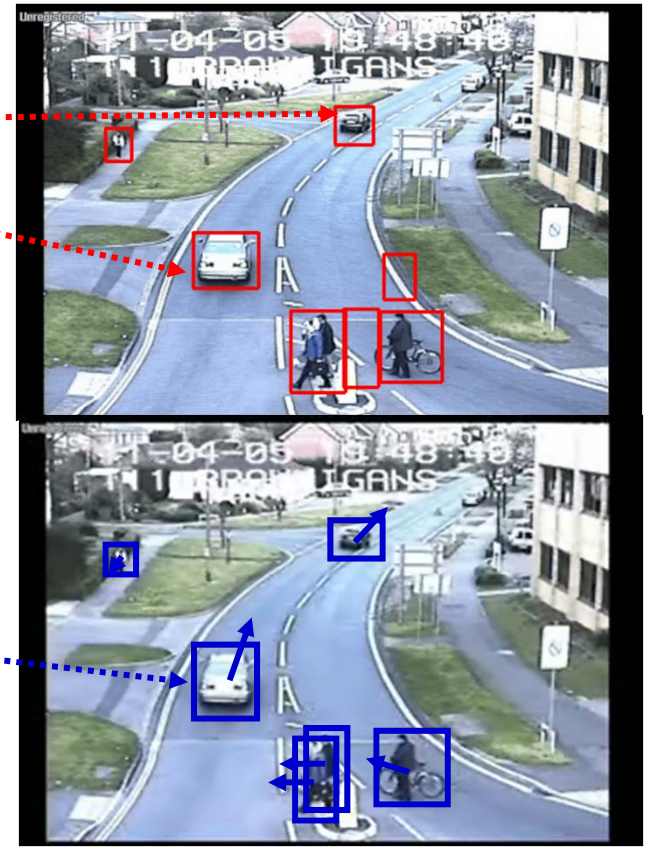
$$Z_k = \{z_{k,1}, \ldots, z_{k,N(k)}\}$$

$$z = [y_1, y_2, w, h]$$

- State

$$X_k = \{x_{k,1}, \ldots, x_{k,M(k)}\}$$

$$x = [y_1, y_2, \dot{y}_1, \dot{y}_2, w, h]$$



Some basics first…

Single-target Particle Filter

- State $\mathbf{x}_k = \mathbf{f}_k(\mathbf{x}_{k-1}, \mathbf{u}_k)$
- Observation $\mathbf{z}_k = \mathbf{h}_k(\mathbf{x}_k, \mathbf{n}_k)$

- Objective
  - to estimate unknown state $\mathbf{x}_k$ based on a sequence of observations $\mathbf{z}_k, k = 0,1,\ldots$
  - to approx. the posterior distribution $p(\mathbf{x}_k \mid \mathbf{z}_{1:k}) \approx \sum_{i=1}^{L} w_k^i \, \delta(\mathbf{x}_k - \mathbf{x}_k^i)$
- Solution (Bayesian)
  - Prediction step
    - Based on state equation
  - Update step
    - Based on likelihood function

Sequential Monte Carlo (Particle Filter)

- Monte Carlo approximation of the Bayes recursion
  - Approximate the posterior *pdf*

$$p(x_k \mid z_{1:k}) \approx \sum_{i=1}^{L} \omega_k^{(i)} \, \delta(x_k - x_k^{(i)})$$

  - Samples propagated using state and observation equations

$$\omega_k^{(i)} \propto \omega_{k-1}^{(i)} \frac{g(z_k \mid x_k^{(i)}) \, \cancel{f(x_k^{(i)} \mid x_{k-1}^{(i)})}}{\cancel{q(x_k^{(i)} \mid x_{k-1}^{(i)}, z_k)}}$$

- Sequential Importance Sampling (SIS)
  - Sample according to dynamics

$$q(x_k \mid x_{k-1}, z_k) = f(x_k \mid x_{k-1})$$

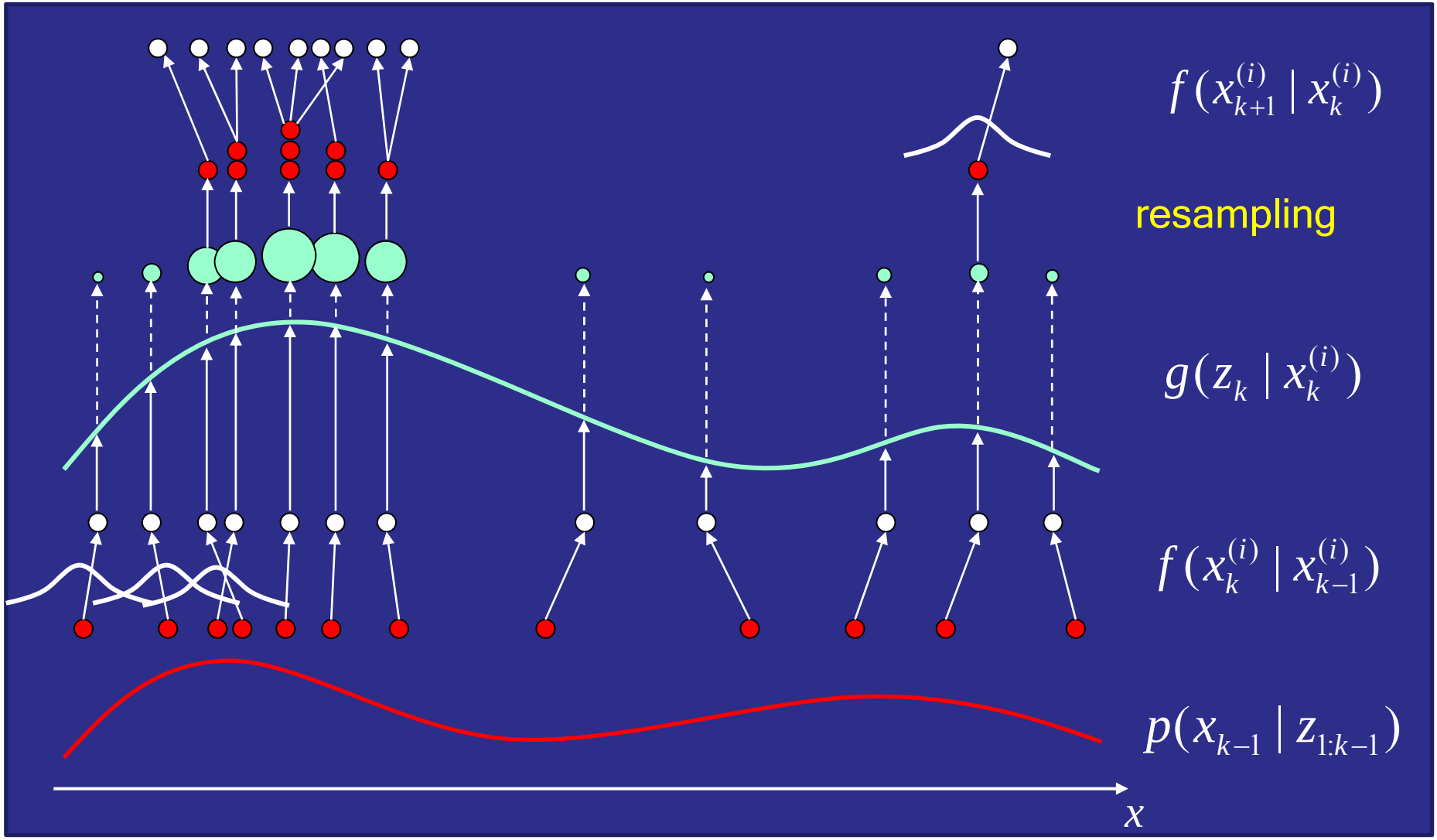
Resampling

- SIS problem
  - The sampling *does not depend on the measurement*
  - After some steps particles have *small weights*

- Solution
  - To eliminate particles with **small weight**
  - To propagate multiple times particles with **large weight**

- Sequential Importance Resampling (SIR)
  - Sample proportionally to $\omega_{k-1}^{(i)}$
  - New weight $\omega_k^{(i)} \propto g(z_k \mid x_k^{(i)})$ likelihood

CONDENSATION



Back to the multi-object case…

Multi-target observation and state

- Observations

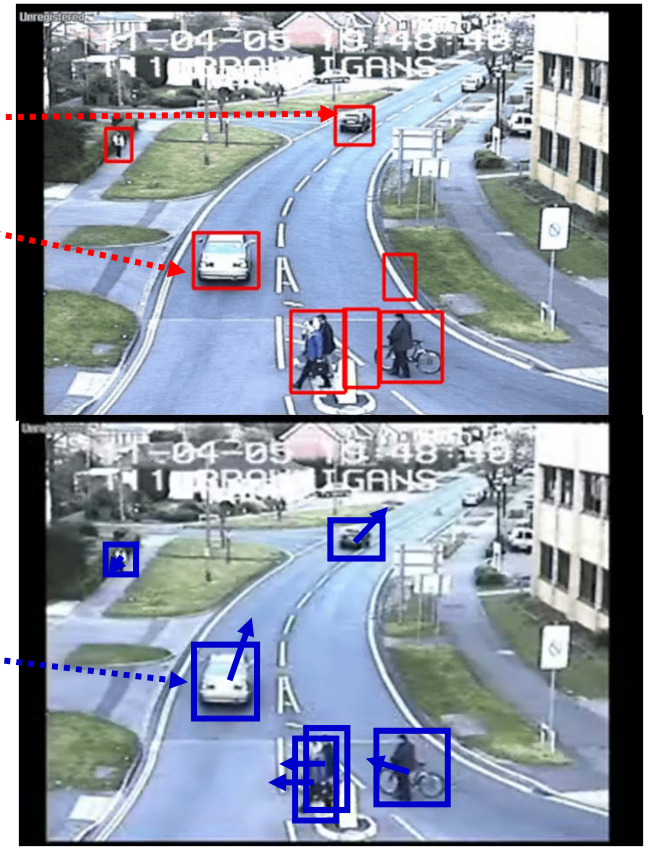
$$Z_k = \{z_{k,1}, \ldots, z_{k,N(k)}\}$$

$$z = [y_1, y_2, w, h]$$

- State

$$X_k = \{x_{k,1}, \ldots, x_{k,M(k)}\}$$

$$x = [y_1, y_2, \dot{y}_1, \dot{y}_2, w, h]$$



Example

Data association filters

- Objective
  - to associate detections

- Approaches
  - Nearest Neighbour (NN) filter
    - one hypothesis per trajectory
  - Joint Probabilistic Data Association Filter (JPDAF)
    - evaluate all associations between two frames
    - handles clutter
  - Multiple Hypotheses Tracker (MHT)
    - evaluates hypotheses over multiple frames
  - *gating* to reduce computational cost
  - *extending the Bayes recursion* to multiple targets (using Random Finite Sets)

Filtering in the multi-target space

Multi-target Bayes filter

Single target

Random vectors

$$\mathbf{x}_k = \mathbf{f}\left(x_{k-1}, \mathbf{u}_k\right)$$
$$\mathbf{z}_k = \mathbf{g}\left(x_k, \mathbf{n}_k\right)$$

Multiple targets

Random Finite Sets (RFS)

$$\Xi_k = S_k\left(X_{k-1}\right) \cup \Gamma_k$$
$$\Omega_k = \Theta_k\left(X_k\right) \cup \mathrm{K}_k$$

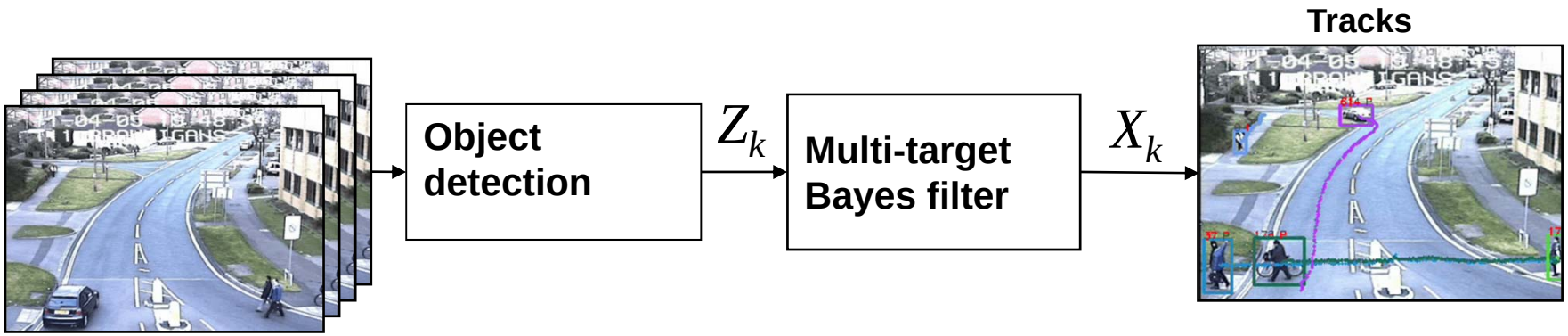
- Recursively propagates the multi-target state
  - Prediction

$$p(X_k \mid Z_{1:k-1}) = \int f(X_k \mid X_{k-1}) p(X_{k-1} \mid Z_{1:k-1}) \mu_s\left(dX_{k-1}\right)$$

  - Update

$$p(X_k \mid Z_{1:k}) = \frac{g(Z_k \mid X_k) p(X \mid Z_{1:k-1})}{\int p(Z_k \mid X_k) p(X_k \mid Z_{1:k-1}) \mu_s\left(dX_k\right)}$$

Multi-target Bayes filter



- Model target *birth*, *death*, *clutter* and *missing detections*

- Monte Carlo approximation
  - Dimensionality of the state grows with the number of targets
  - Number of particles is exponential with the number of targets
  - Not feasible in real-world applications! ☹

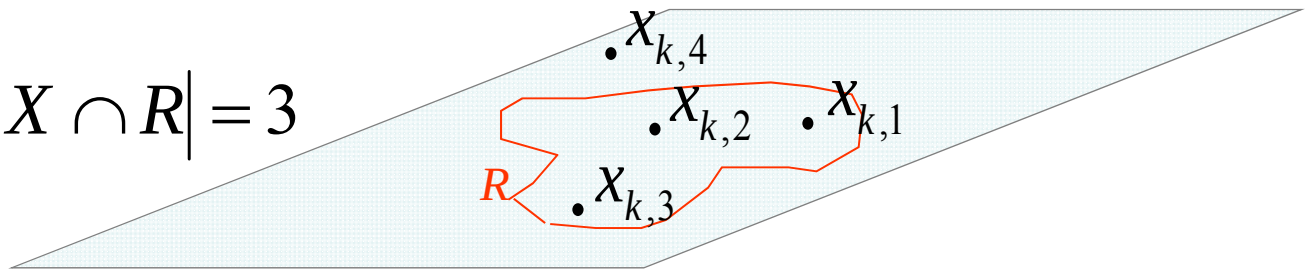
- Solution
  - to propagate the first order moment of the RFS only [Mahler]

Probability Hypothesis Density (PHD)

- The first order moment of a RFS → PHD
  - The integral on any R gives the expected number of targets in R

$$\int_R D_\Xi(x)dx = E\left[\left|\Xi \cap R\right|\right]$$

$$\left|X \cap R\right| = 3$$



$$\begin{array}{lccccccc}
\textbf{\textit{Multiple-target Bayes filter}} & \longrightarrow & p(X_{k-1}|Z_{1:k-1}) & \longrightarrow & p(X_k|Z_{1:k-1}) & \xrightarrow[\uparrow Z_k]{} & p(X_k|Z_{1:k}) & \longrightarrow \cdots \\
 & & \downarrow & & \downarrow & & \downarrow & \\
\textbf{\textit{PHD Filter}} & \longrightarrow & D_{k-1|k-1} & \longrightarrow & D_{k|k-1} & \xrightarrow{} & D_{k|k} & \longrightarrow \cdots
\end{array}$$

Recursive PHD filter

- Prediction

$$D_{k|k-1}(x_k) = \gamma(x) + \int \phi(x_k, x_{k-1}) D_{k-1|k-1}(x_{k-1}) dx_{k-1}$$

  – PHD state transition

Birth intensity

$$\phi(x_k, x_{k-1}) = e(x_k) f(x_k \mid x_{k-1})$$

- Update

Survival probability

Single target transition

$$D_{k|k}(x_k) = \left[p_M(x_k) + \sum_{z \in Z_k} \frac{\psi_z(x_k)}{\kappa(z) + \langle \psi_z, D_{k|k-1}(x_k) \rangle} \right] D_{k|k-1}(x_k)$$

Missing detection

Clutter intensity

Normalization

  – PHD likelihood

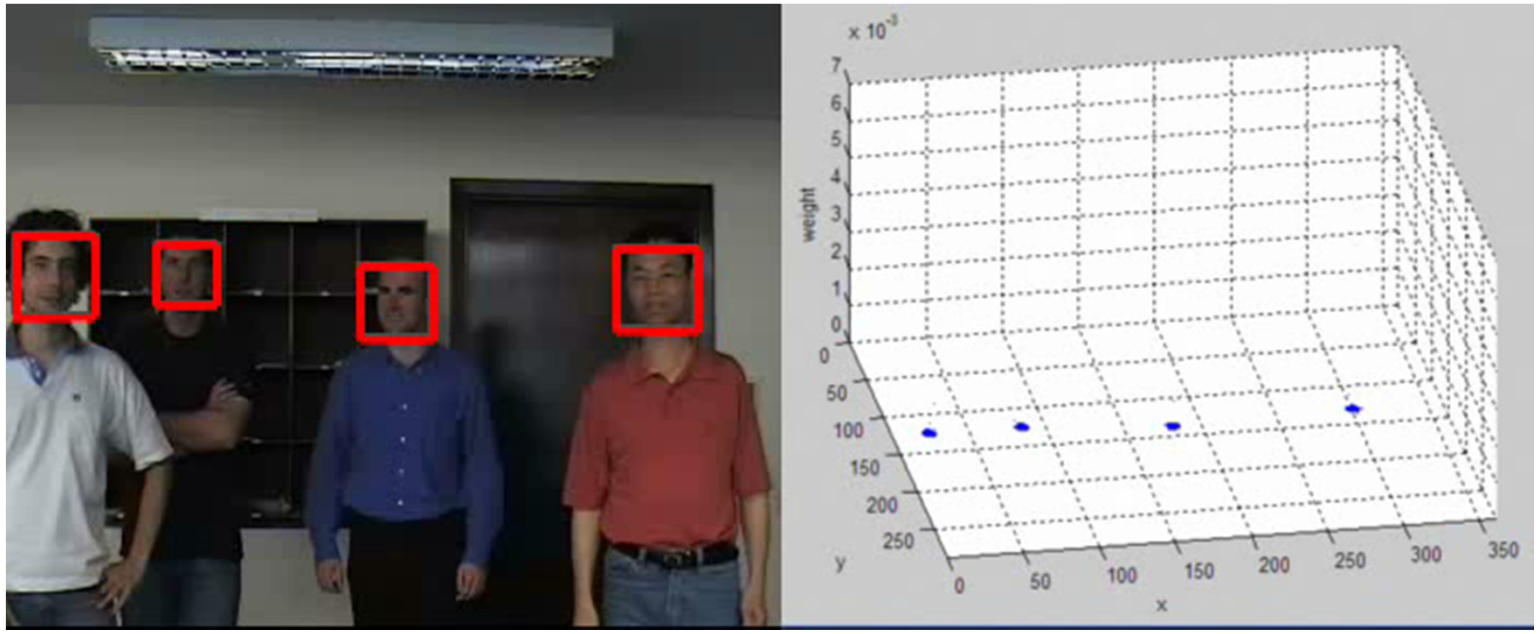
$$\psi_z(x) = (1 - p_M(x)) g(z \mid x)$$

Single target likelihood

Example: PHD propagation

- **<u>Input</u>**: Face detector

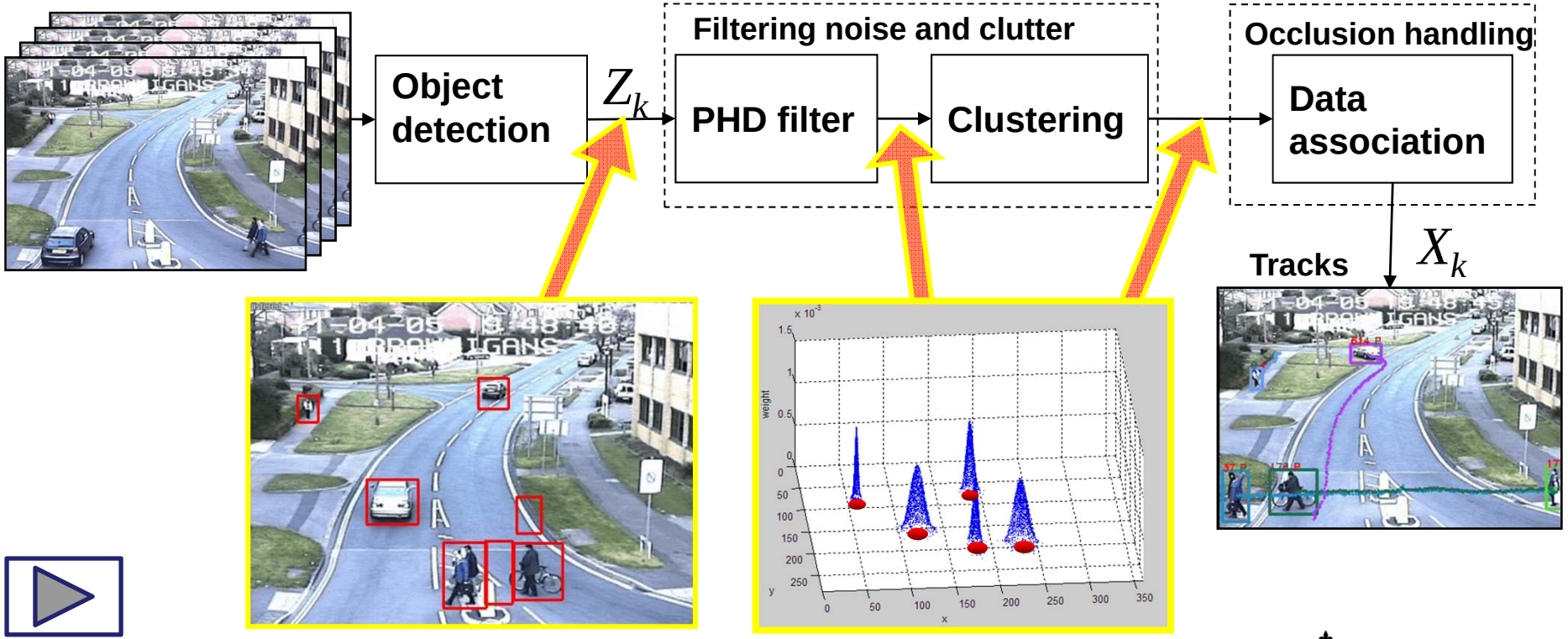
- **<u>Output</u>**: Monte Carlo approximation of the PHD



converges to 1

Tracking with the PHD filter

- **Advantage**: Linear complexity with the target number
- **Drawback**: Data association is not integrated in the Bayesian recursion



Results – CLEAR/i-LIDS dataset

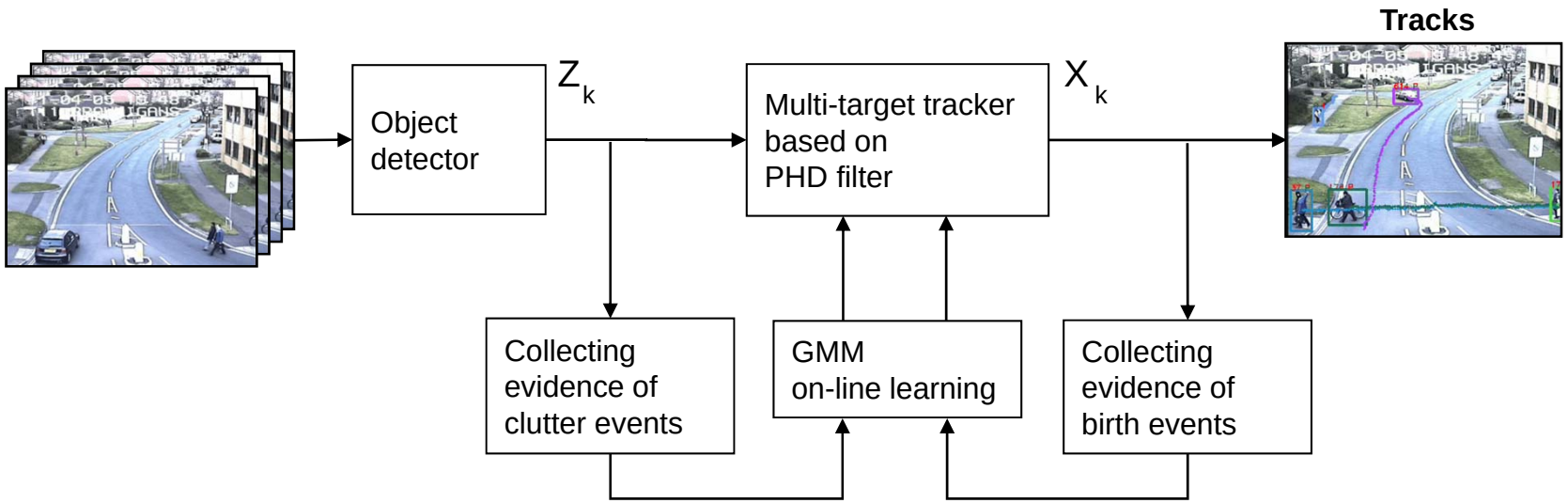
Detections PHD output

Efficient multi-target visual tracking using Random Finite Sets
E. Maggio, M. Taj, A. Cavallaro
IEEE Trans. on Circuits and Systems for Video Technology, August 2008, pp.1016-1027

Queen Mary
University of London

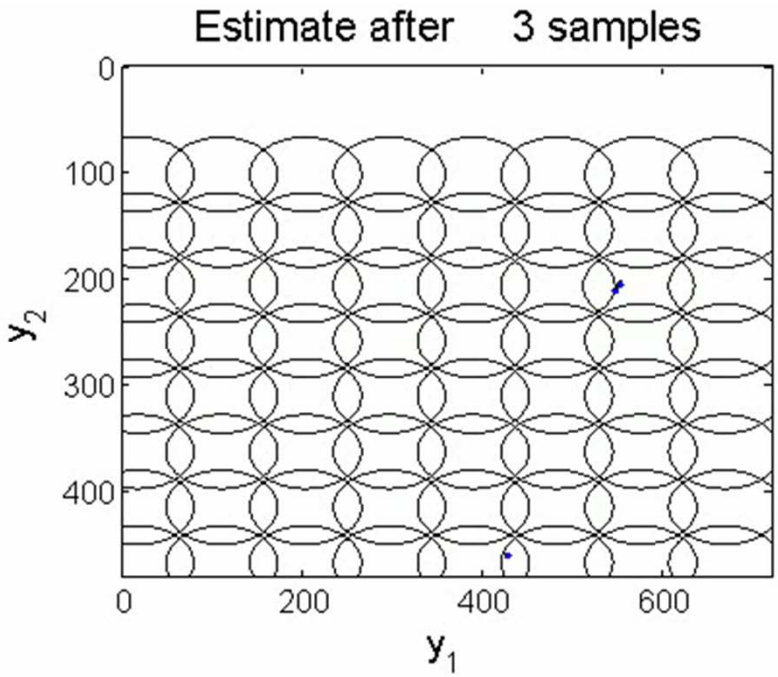
Tracking with context modelling

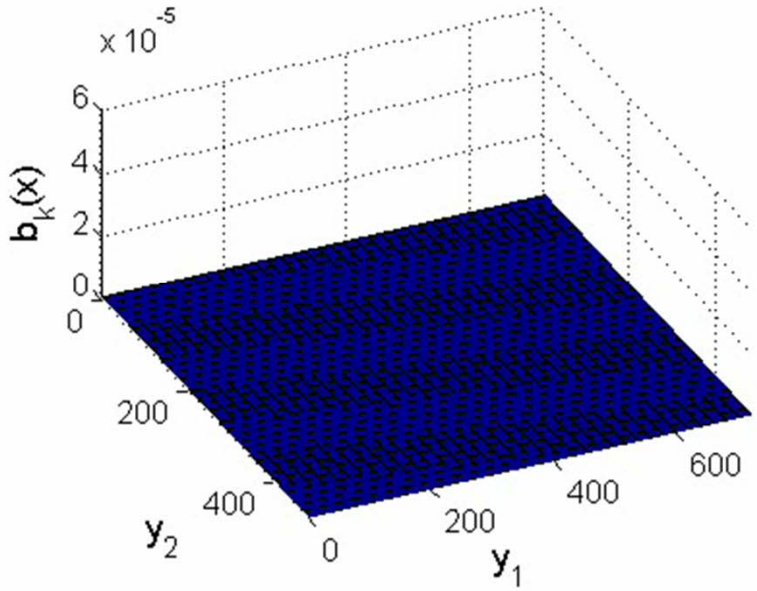
Using scene context in the PHD recursion



- Goal: context adaptive filtering
  - Birth intensity -- **to model** entry areas
  - Clutter intensity -- **to model** detection errors
- Data collection
  - Birth events from the output of the tracker
  - Clutter events via user feedback
- Density estimation

Example: learning birth intensity





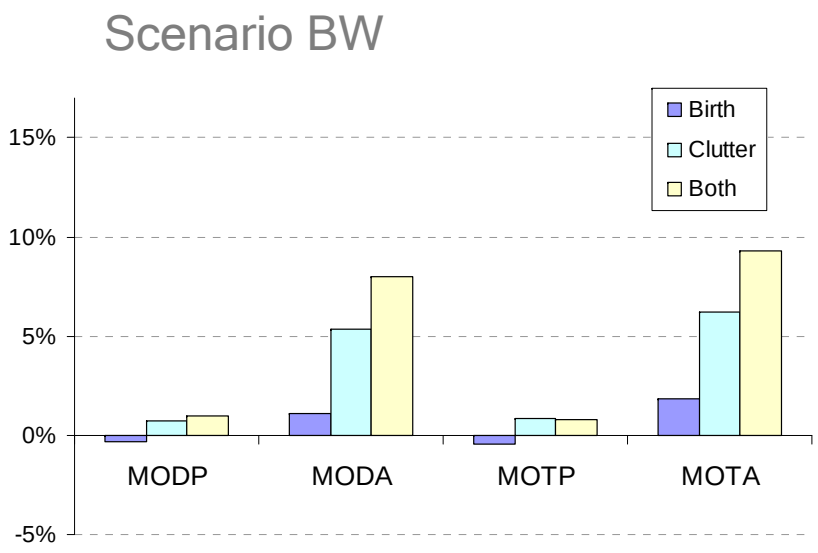
Birth and clutter intensities

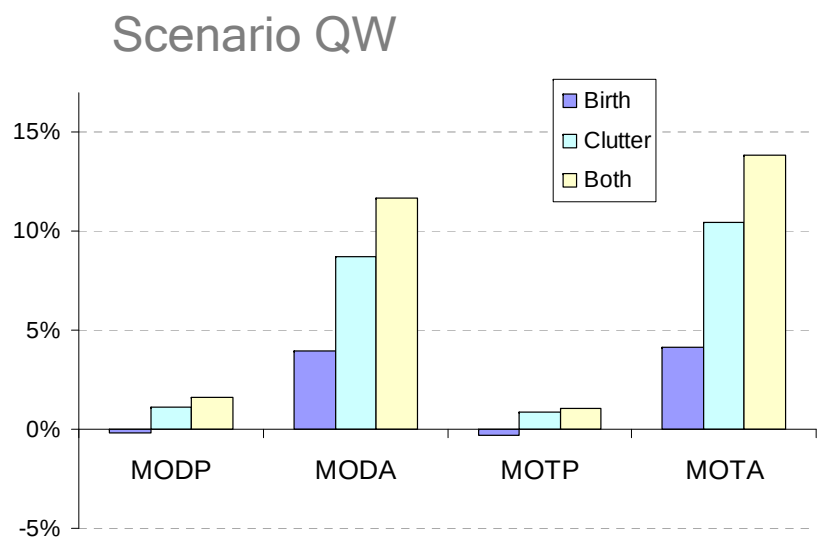
$\gamma(x)$ $\longrightarrow$

$\kappa(z)$ $\longrightarrow$

Comparison: w. and w.o. context learning

- Performance improvement on CLEAR dataset
- Reference: PHD tracker with uniform birth and clutter intensities





Learning scene context for multiple object tracking

E. Maggio and A. Cavallaro

IEEE Trans. on Image Processing, August 2009, pp. 1873-1884

Queen Mary
University of London

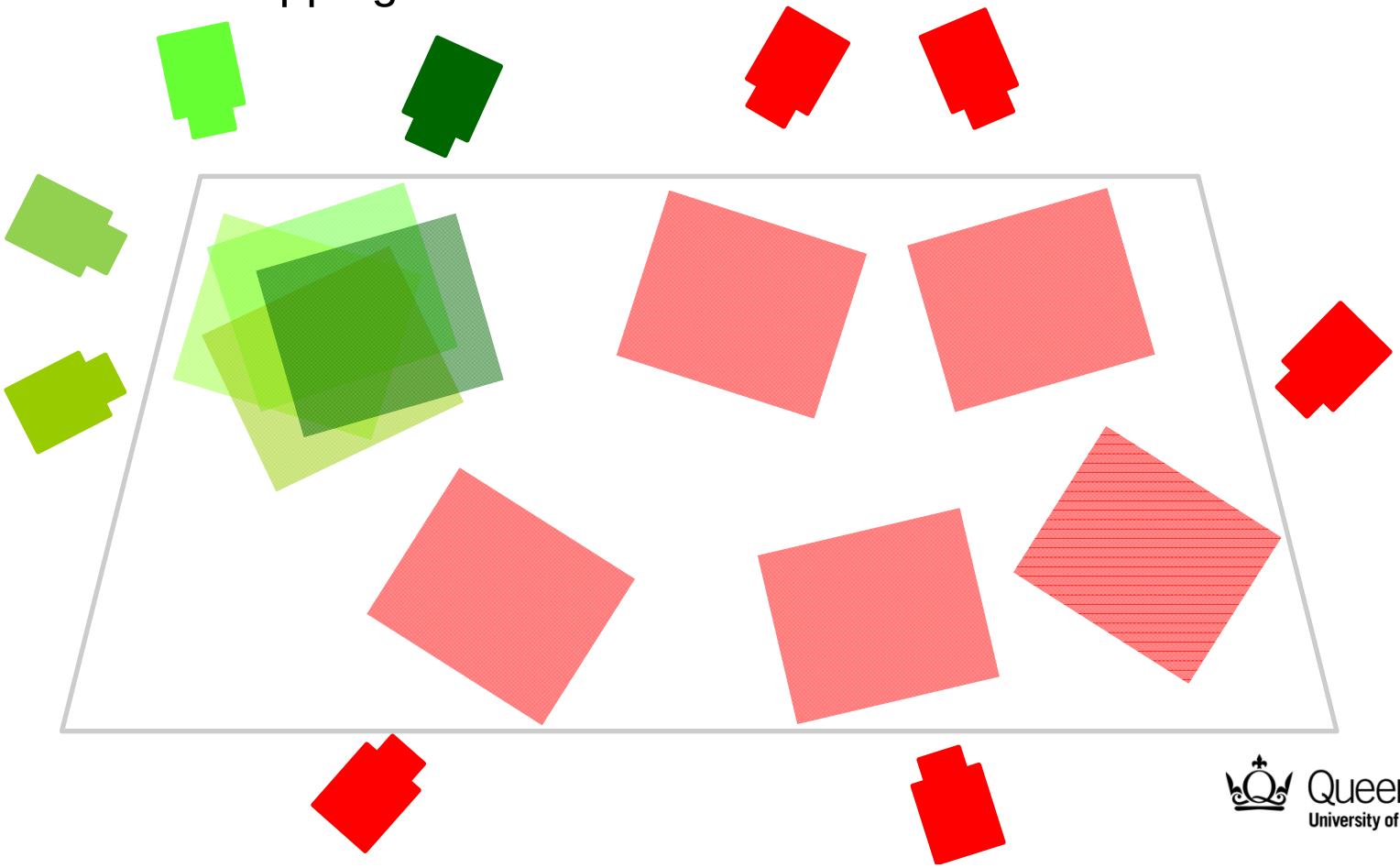
Multi-camera networks

Layout
Approaches

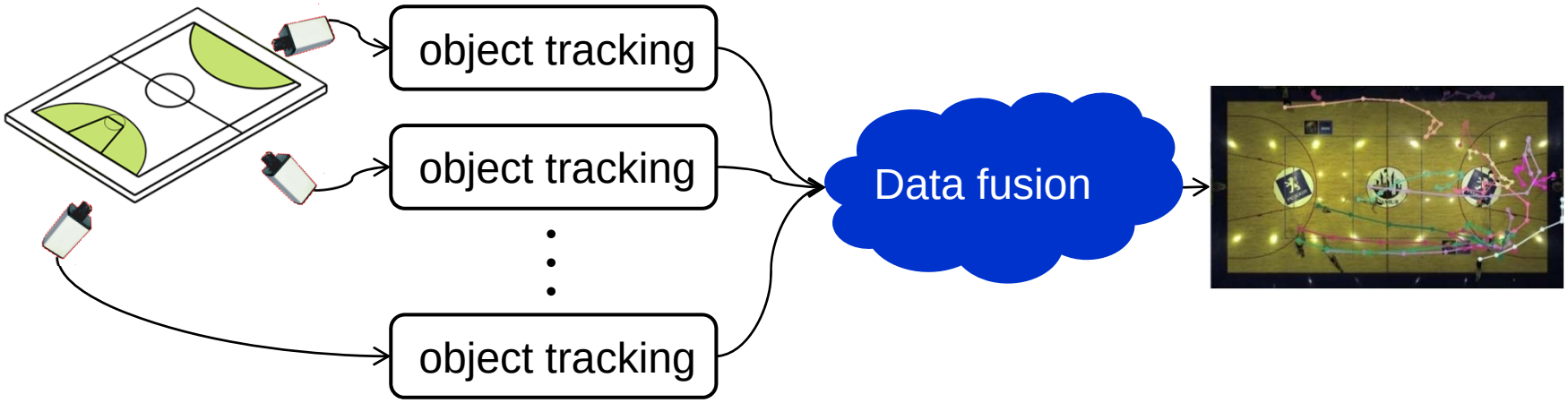
Queen Mary
University of London

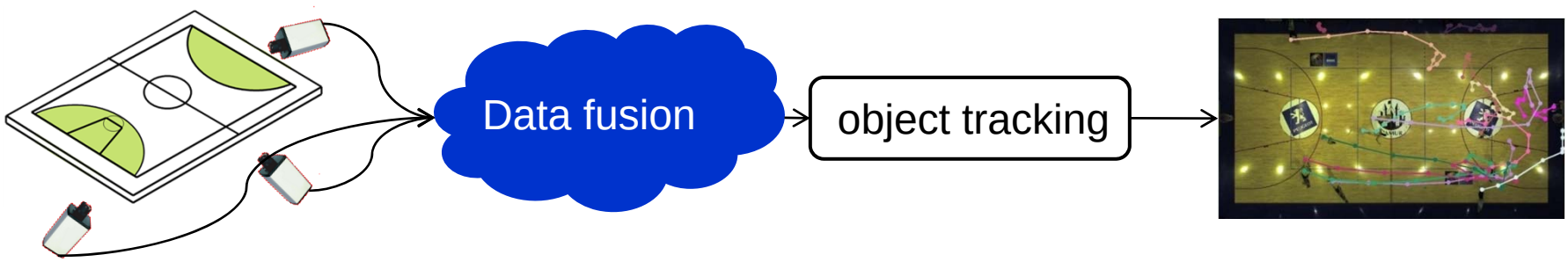
Multiple cameras

- Layout
  - Partially overlapping
  - Non-overlapping



Approaches

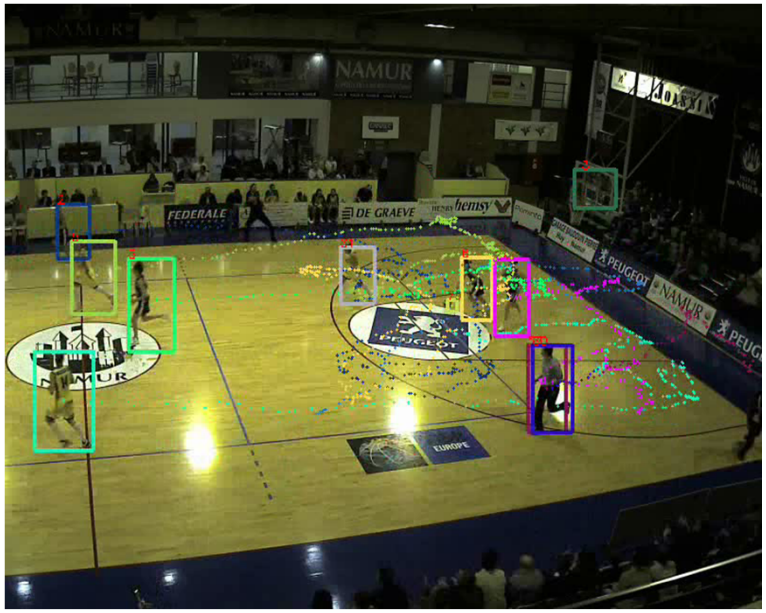

object tracking
object tracking
object tracking
Data fusion


Data fusion
object tracking

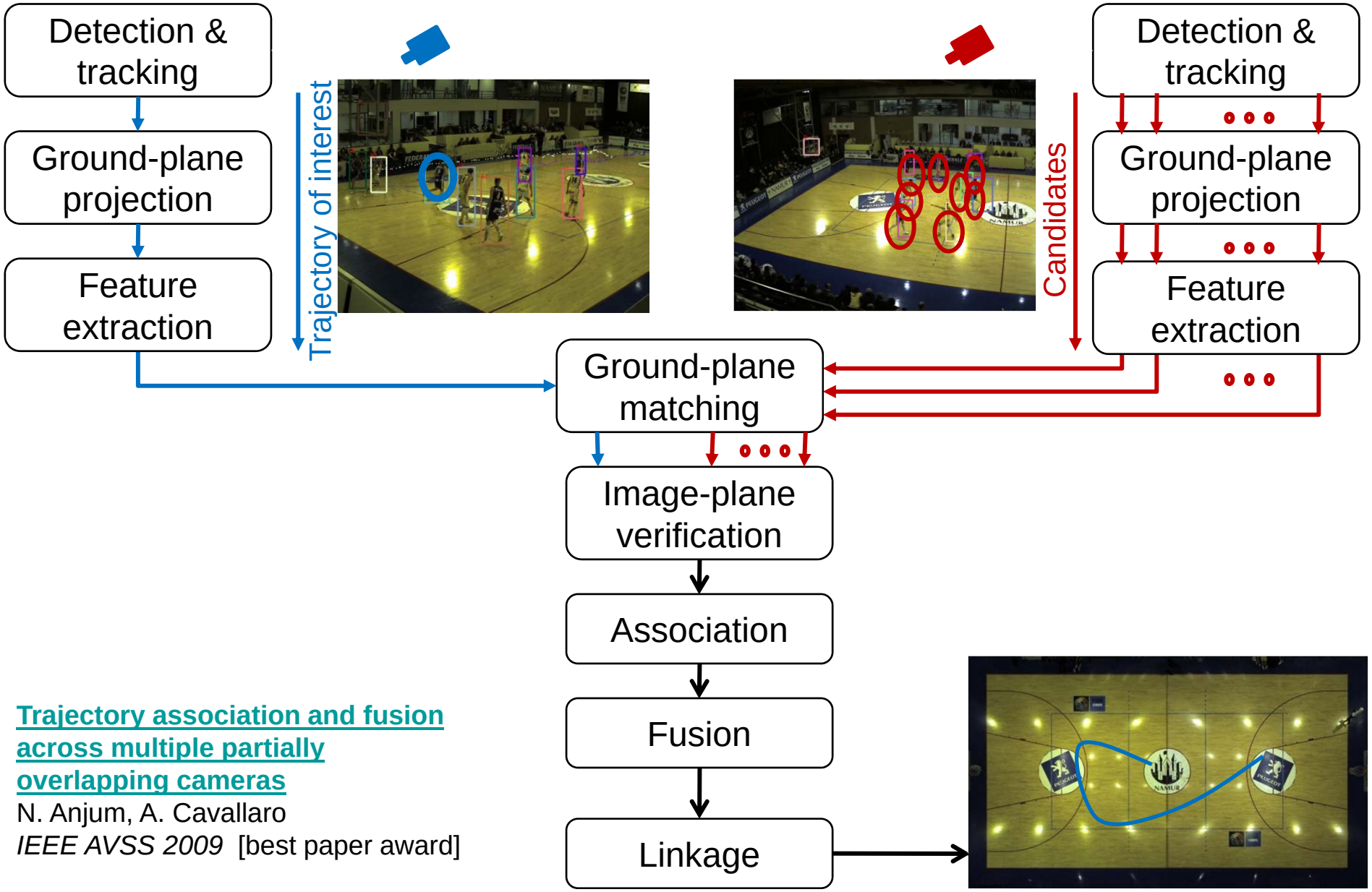
Approach 1: track and then fuse

Approach 2: fuse and then track

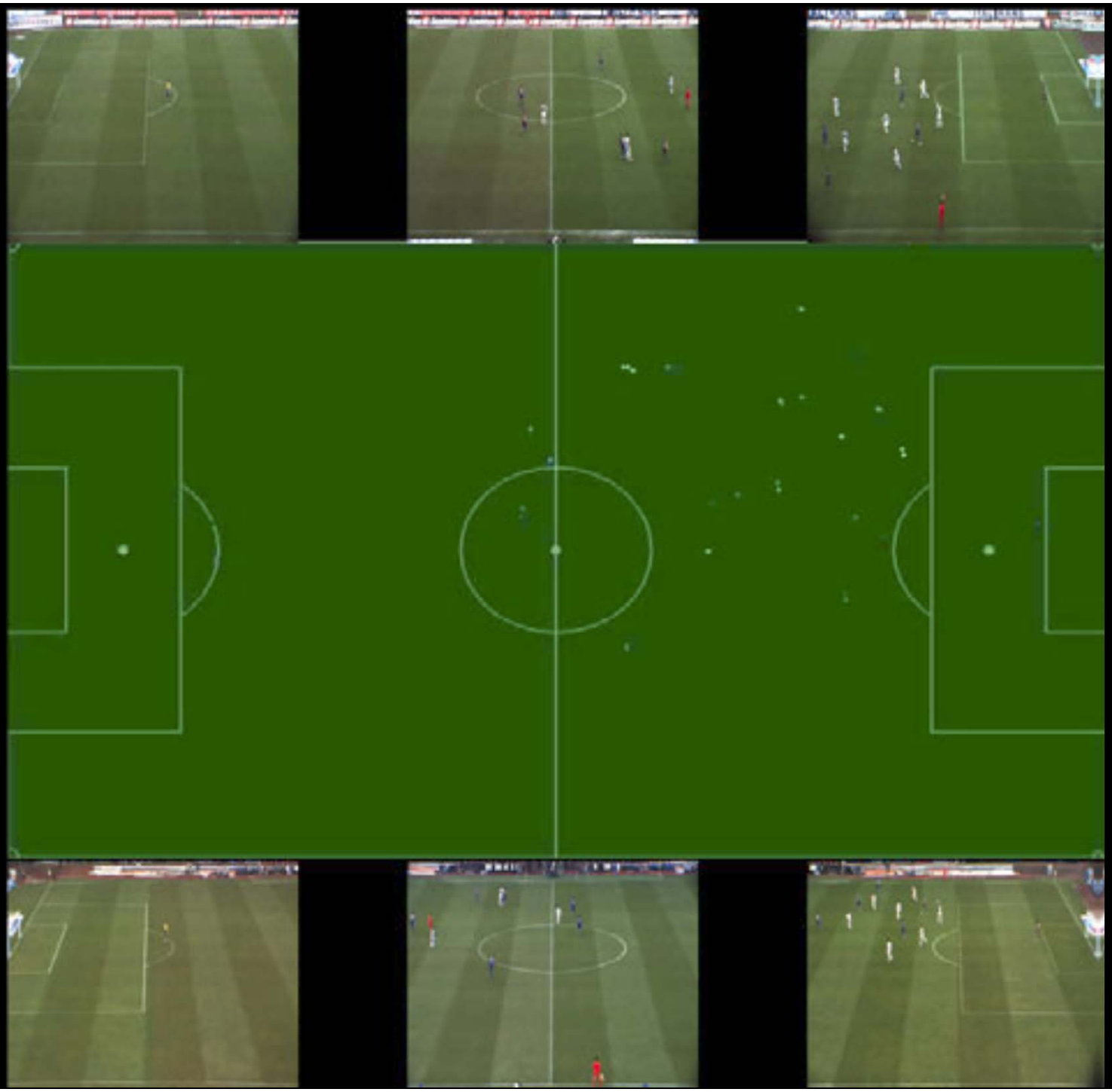
A simple example: 2 cameras

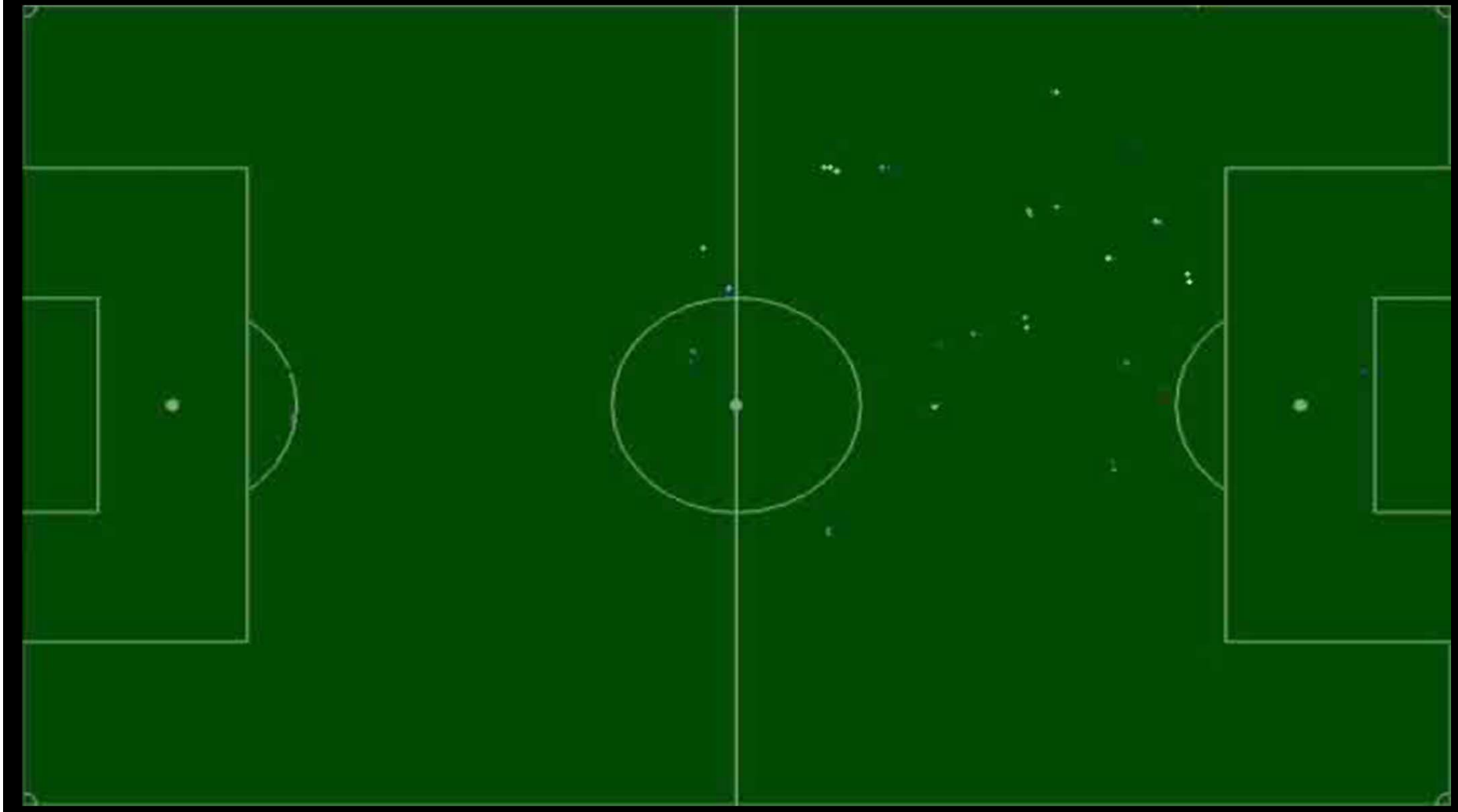


Global trajectory reconstruction



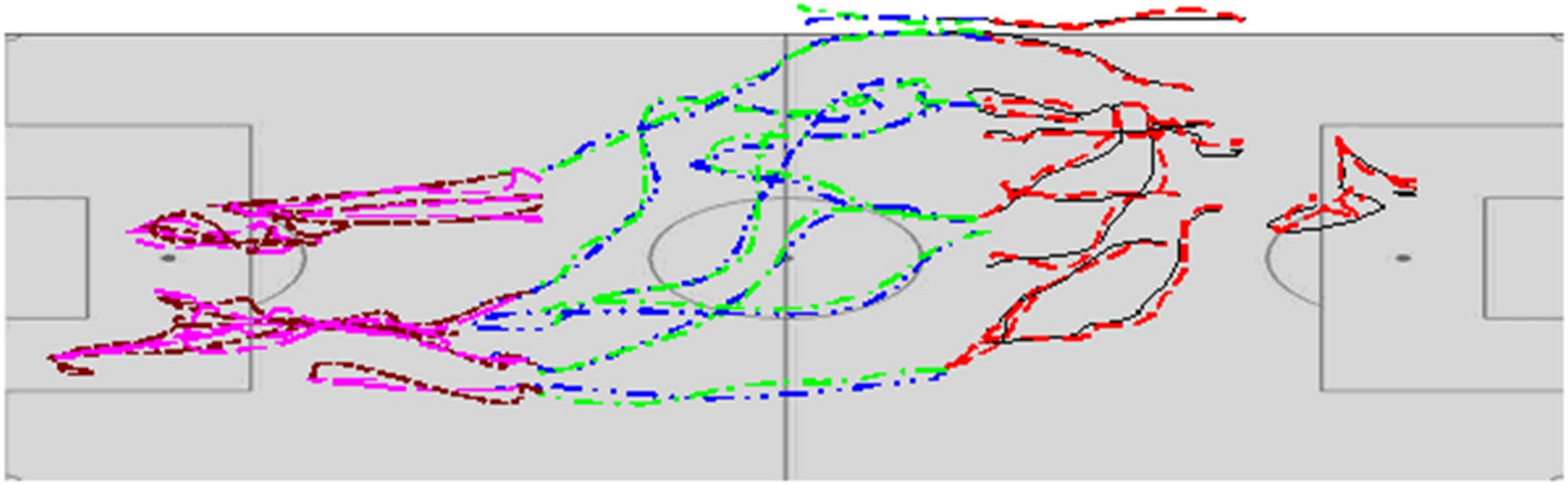
Trajectory association and fusion across multiple partially overlapping cameras
N. Anjum, A. Cavallaro
IEEE AVSS 2009 [best paper award]

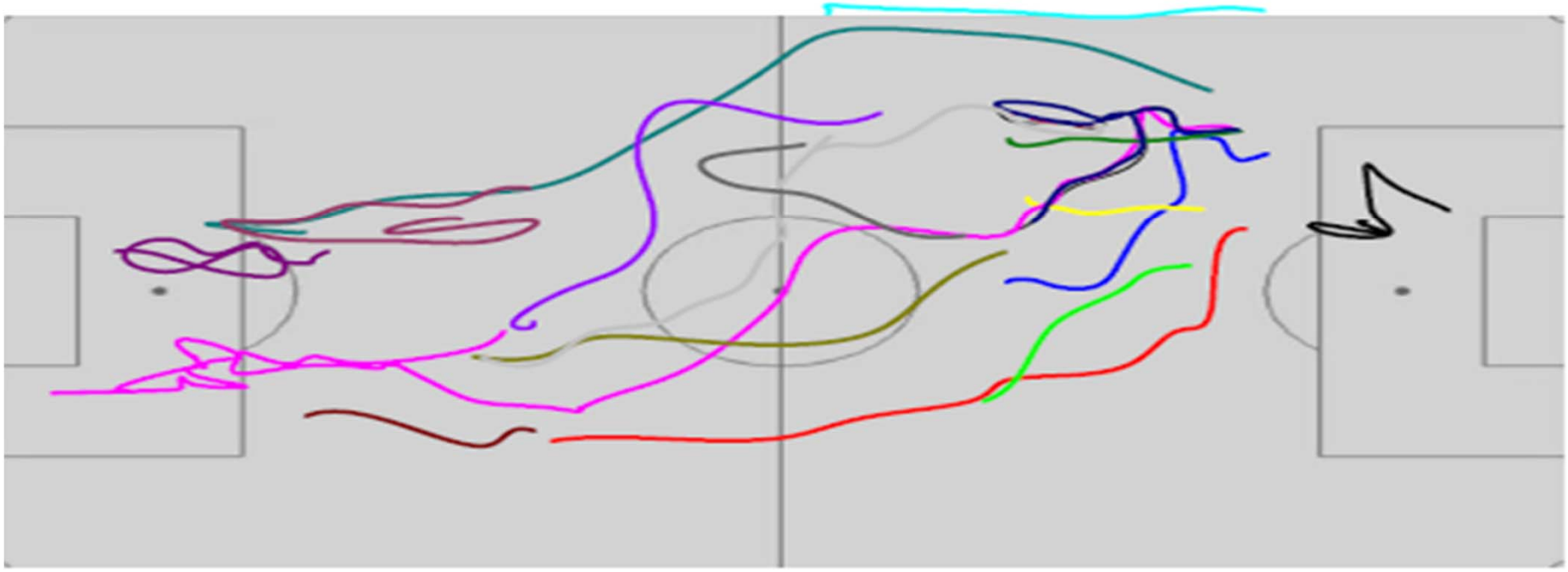




A. Cavallaro

Global trajectory reconstruction

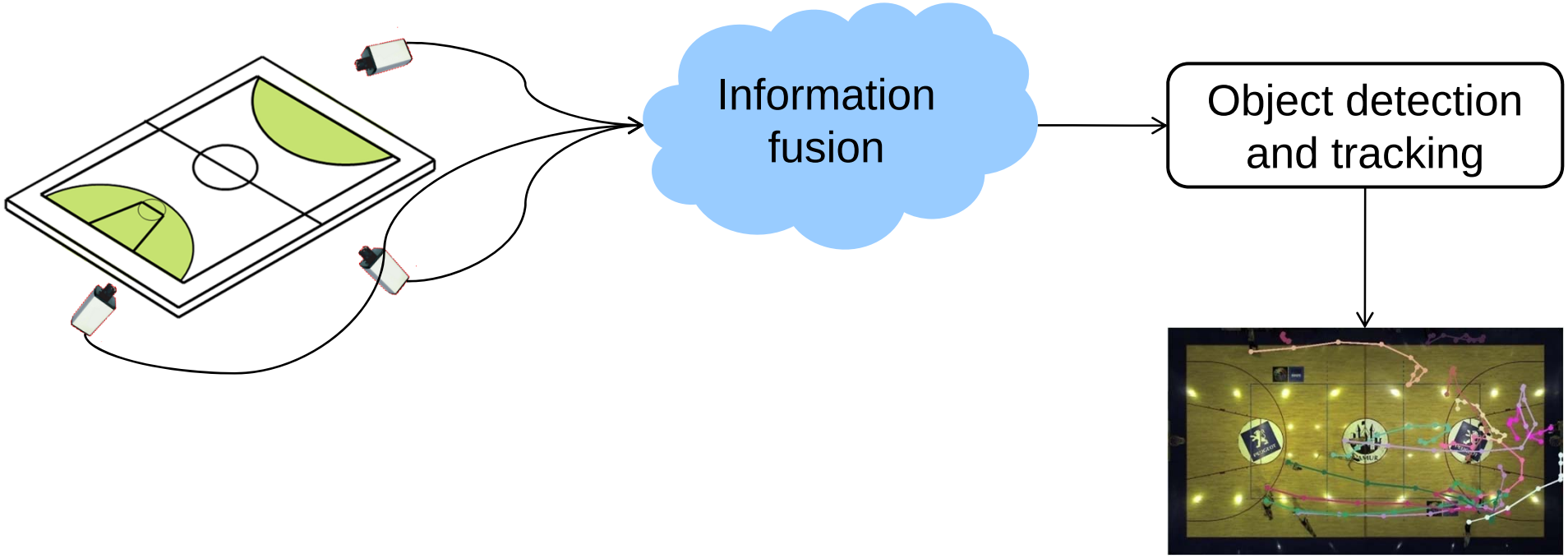




Approach 1: track and then fuse

Approach 2: fuse and then track

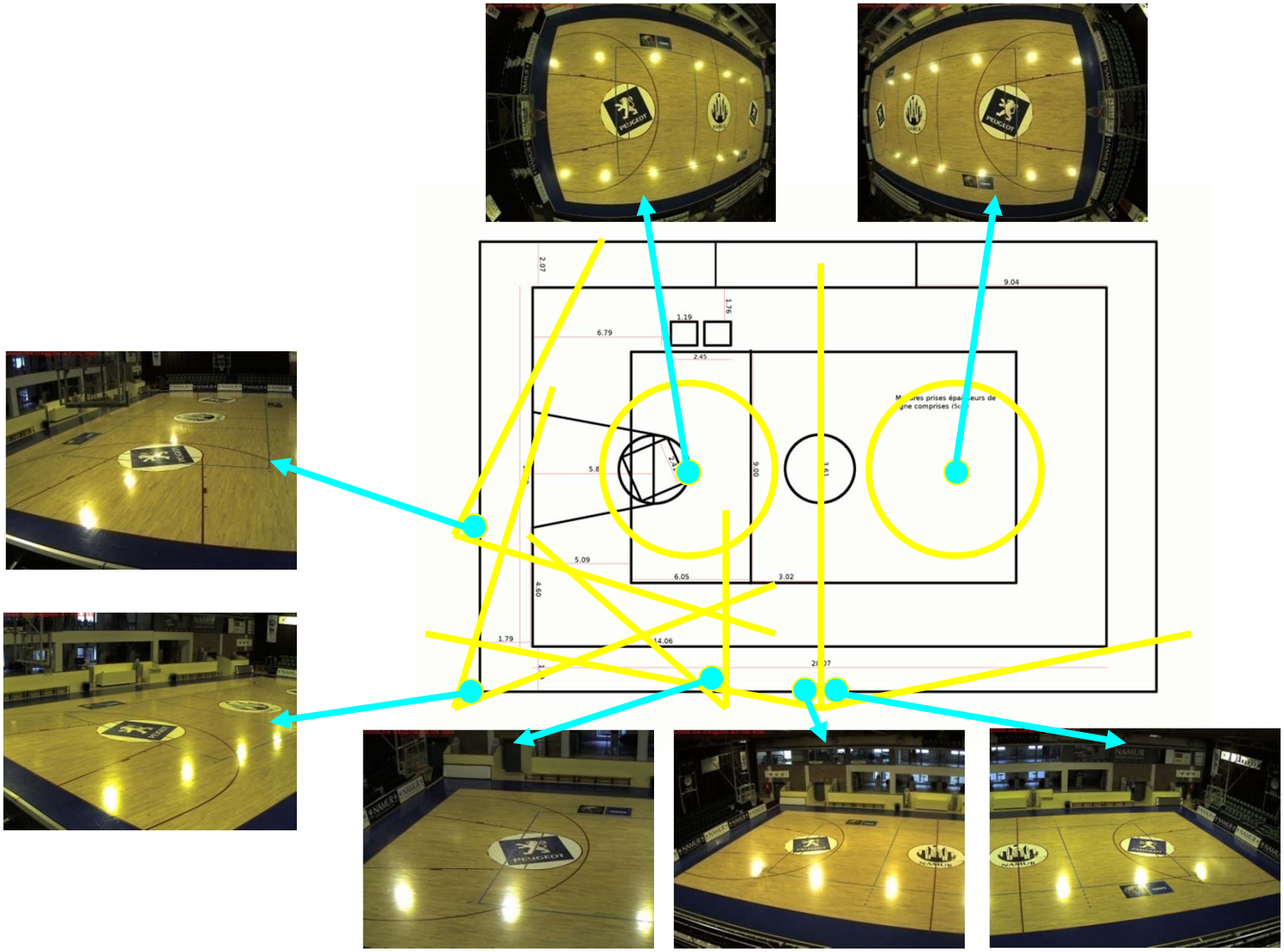
Track before detect



Multi-camera track-before-detect
M. Taj, A. Cavallaro
Proc. of ACM / IEEE ICDSC 2009

Queen Mary
University of London

Multi-camera tracking

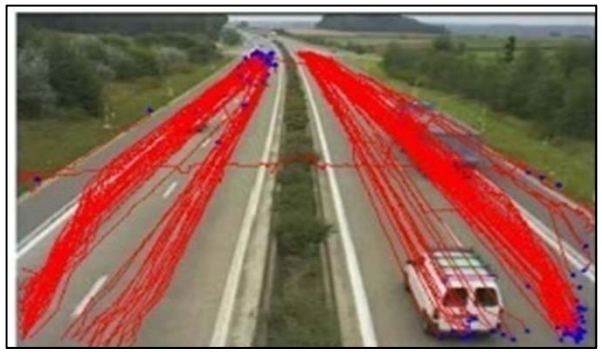


Applications

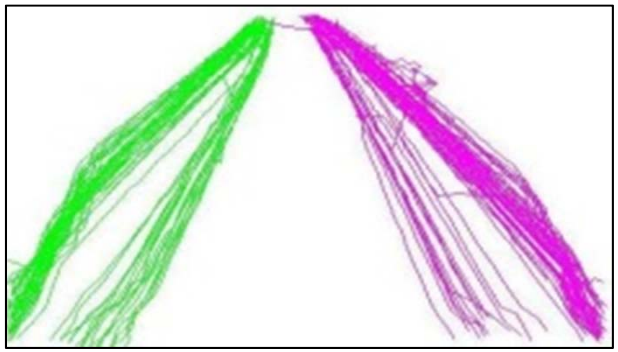
Behaviour analysis
Camera scheduling

Behaviour analysis

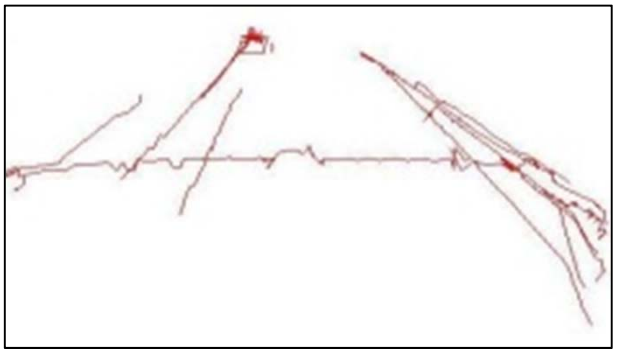
Trajectories



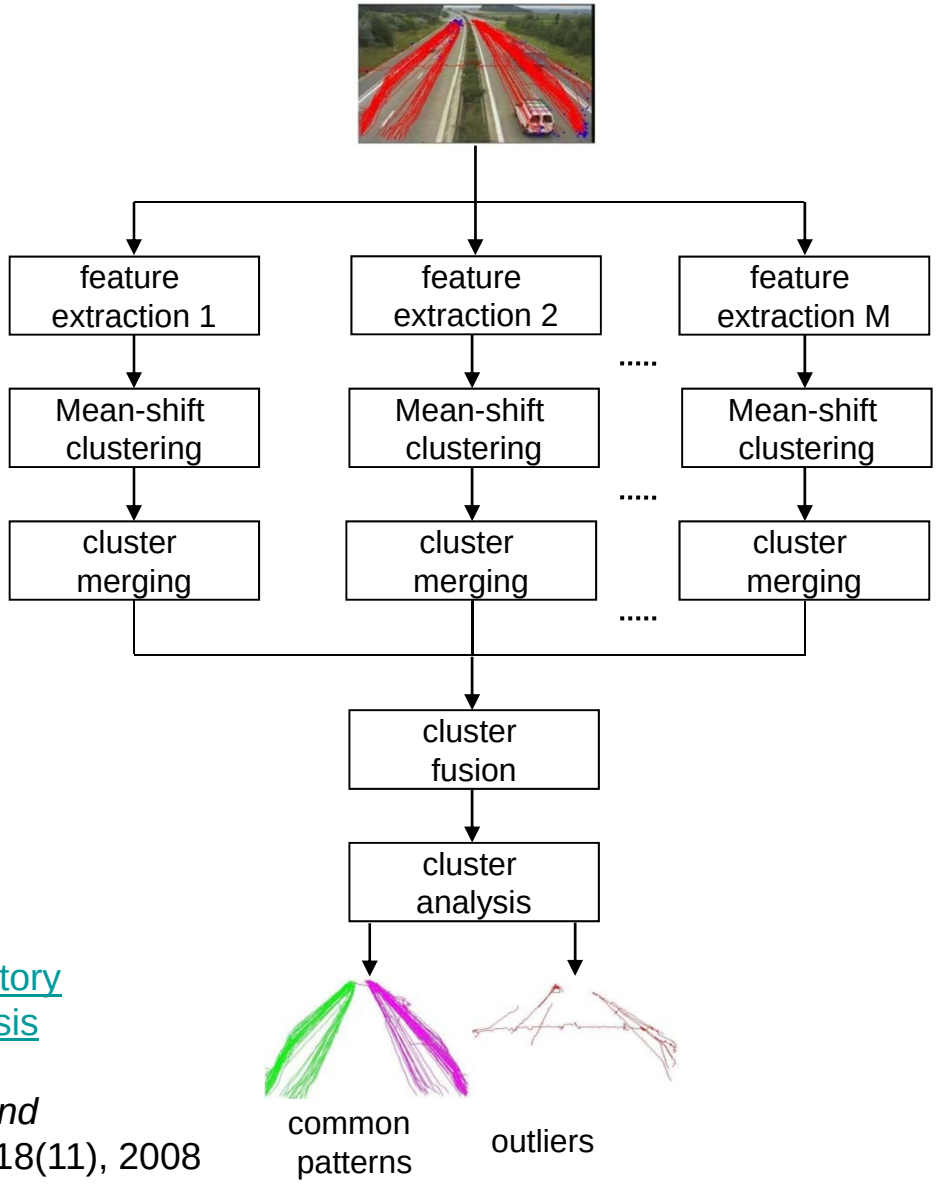
Common patterns



Outliers



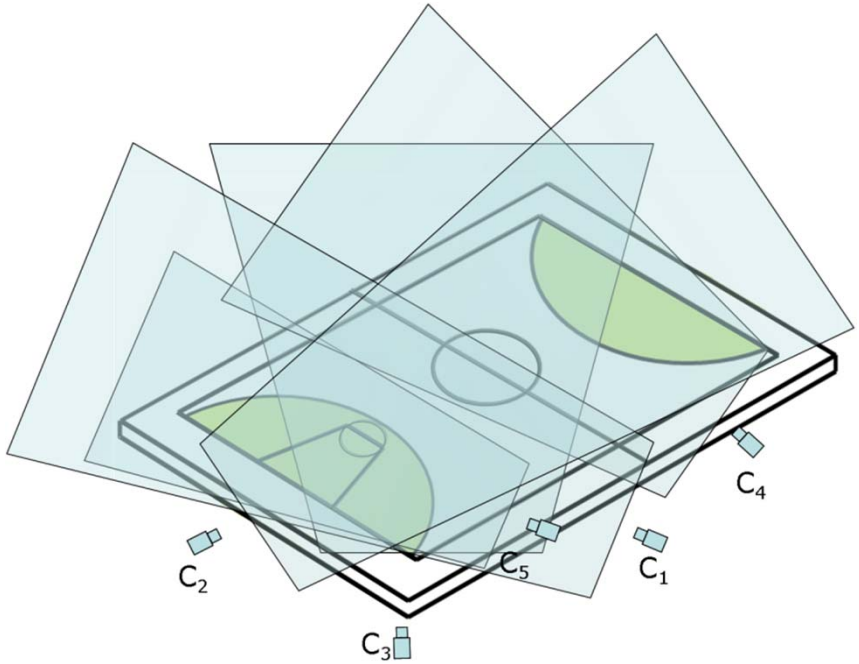
Trajectory clustering



Multi-feature object trajectory clustering for video analysis
N. Anjum, A. Cavallaro
IEEE Trans. on Circuits and Systems for Video Tech. 18(11), 2008

Content production (camera scheduling)

- Example
- Example with audio



Multi-camera scheduling for video production
F. Daniyal, A. Cavallaro
Proc. of Conference on Visual Media Production, London, 16-17 November 2011

Content and task-based view selection from multiple video streams
F. Daniyal, M. Taj, A. Cavallaro
Multimedia Tools And Applications, Vol. 46, Numbers 2-3, January 2010, pp. 235-258

Queen Mary
University of London

Multi-camera networks

Distributed calibration
Topology
Distributed tracking

Queen Mary
University of London

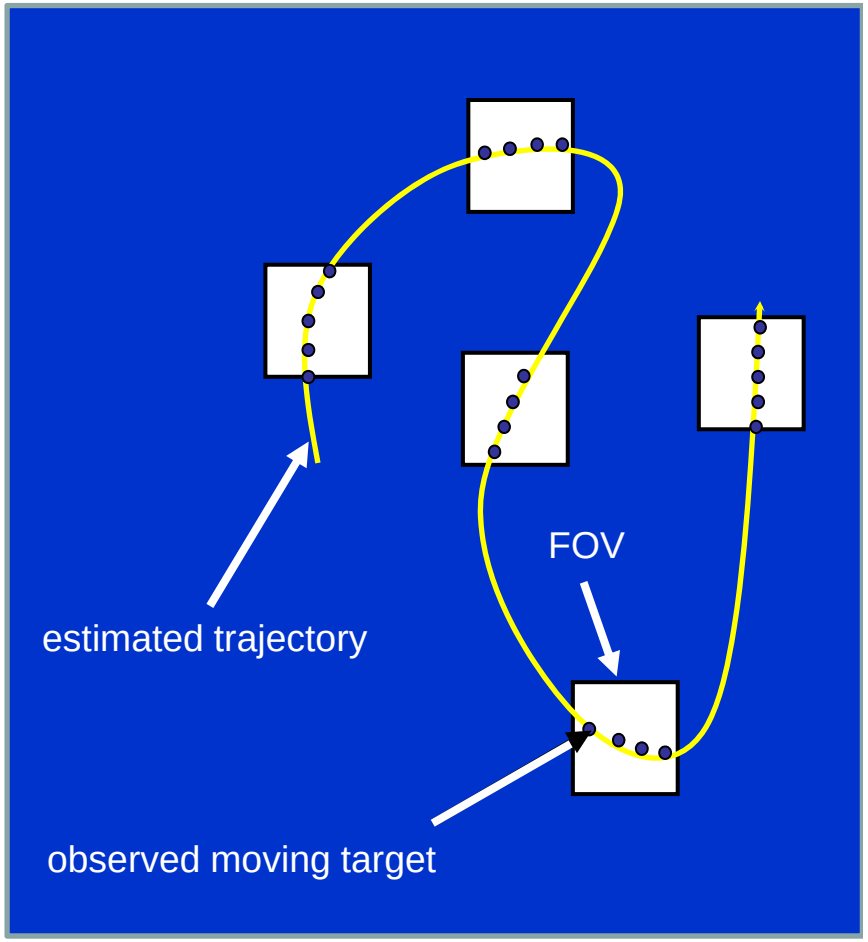
Calibration methods

- Small scale (overlapping fields of view)
  - multi-view geometry
  - epipolar geometry
  - projective transformations
  - feature detection and matching

- Large scale (partially/non overlapping fields of view)
  - observations of an object
  - simultaneous calibration and tracking
  - simultaneous calibration and synchronization

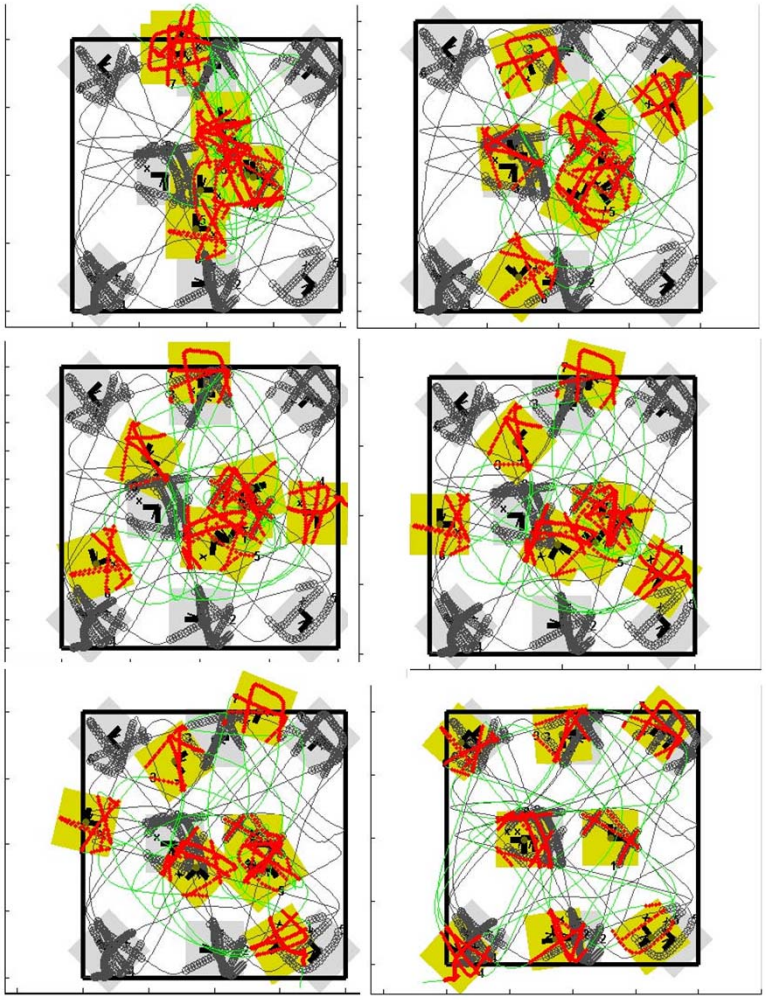
Non-overlapping cameras

- How to localize randomly placed cameras?
  - non-overlapping field of view
  - top-down view
    - optical axis is perpendicular to the ground plane, or
    - removed perspective deformation



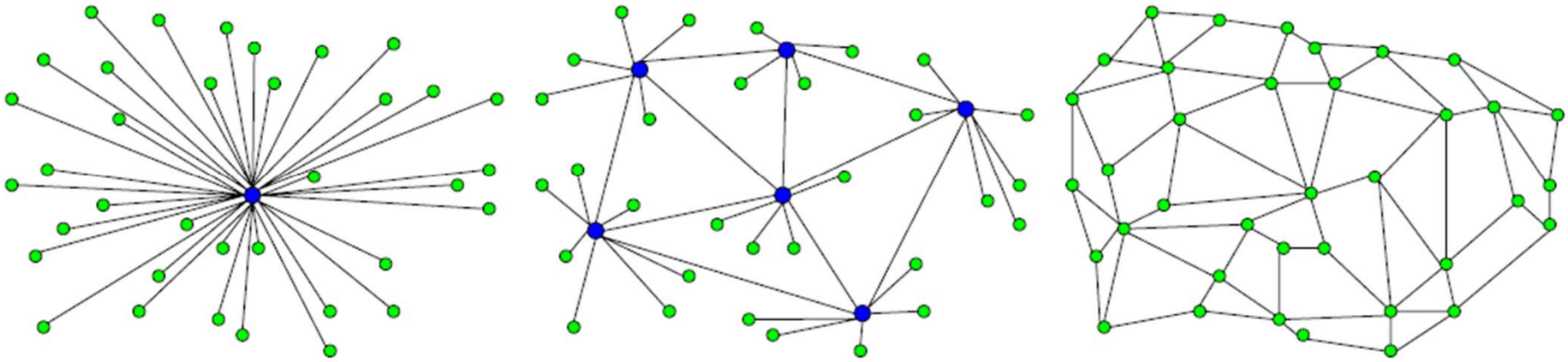
- Approach
  - Learn regression parameters from available trajectories
  - Extrapolate trajectory data to unobserved regions
  - Kalman filtering on available and estimated trajectory

Example: estimation result



Automated localization of a camera network
N. Anjum, A. Cavallaro
IEEE Intelligent Systems, to appear 2011

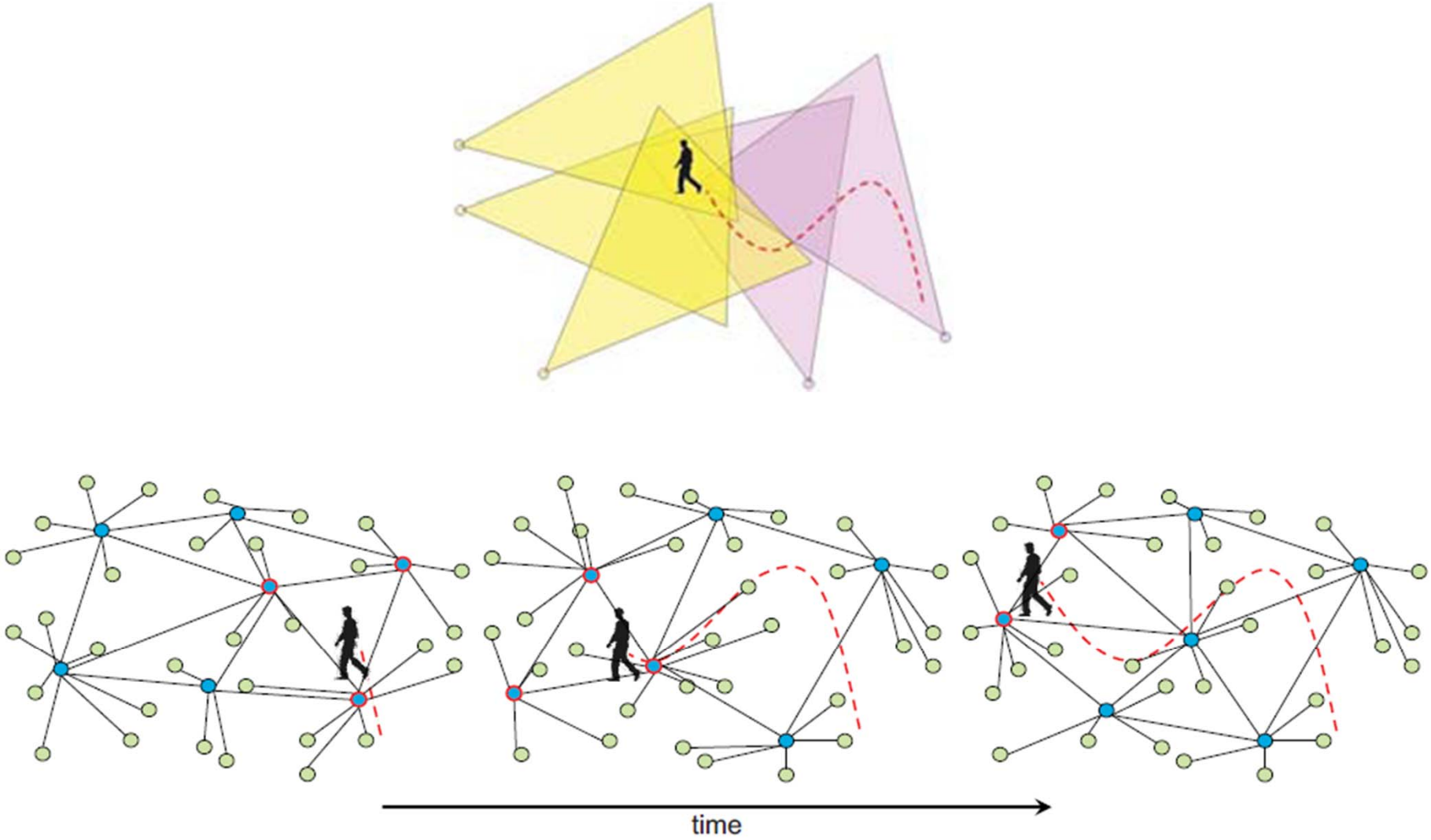
Topology



Centralized Decentralized Distributed

Distributed and decentralized multi-camera tracking
M. Taj, A. Cavallaro
IEEE Signal Processing Magazine, Vol. 28, Issue 3, May 2011

Queen Mary
University of London


time

Distributed Particle Filters (DPFs)

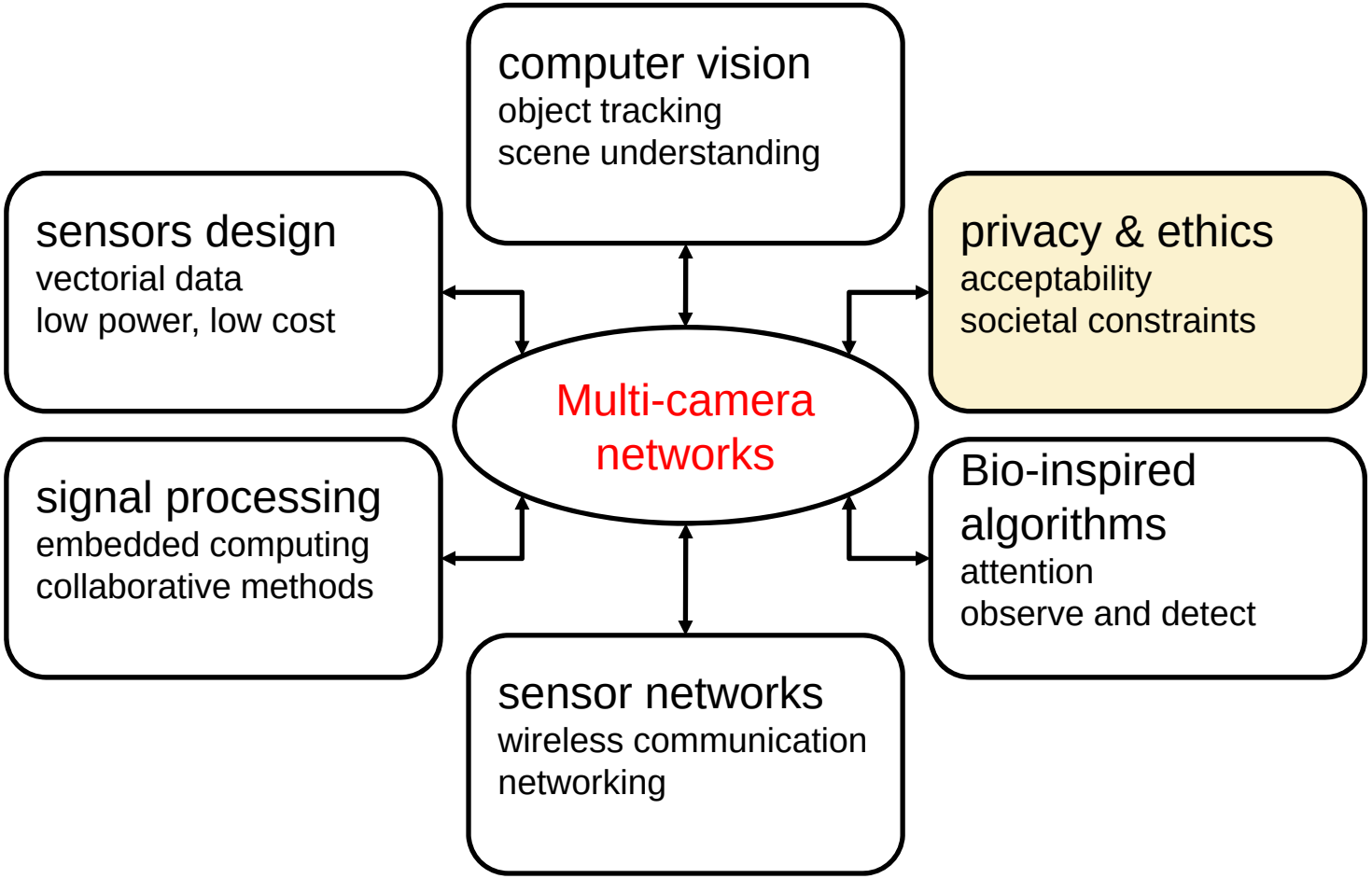
- Approach 1
  - each node executes a local Particle Filter (PF)
  - agreement on particle set reached by nodes at each time step
  - expensive!
- Approach 2
  - common posterior distribution
  - expensive!
- Solution: independence from the number of particles
  - approximating the posterior with a Gaussian Mixture Model

WiSE-MNet: an experimental environment for Wireless Multimedia Sensor Networks
C. Nastasi, A. Cavallaro
Proc. of Sensor Signal Processing for Defence, London, 28-29 September 2011

Simulator: www.eecs.qmul.ac.uk/~andrea/wise-mnet.html

Queen Mary
University of London

Convergence of multiple disciplines



To conclude …

Exciting open challenges:

- Signal processing
  - multi-layered/hybrid data exchange among cameras
  - collaborative deduction about events
  - exploitation of spatio-temporal redundancies
- Sensor networks / embedded computing
  - large amounts of data produced by imagers (bandwidth, latency)
  - network scalability issues
  - joint estimation and decision: tradeoffs
- Computer vision design
  - multi-view methods with partial processing of video locally
  - system-level constraints → influence algorithm design

Acknowledgments

Emilio Maggio

Murtaza Taj

Fahad Daniyal

Nadeem Anjum

Christian Nastasi

References and contact information

Andrea Cavallaro

Queen Mary University of London

andrea.cavallaro@eecs.qmul.ac.uk

youtube.com/smartcameras