Automated Visual Surveillance: Detection of Criminal Behaviour

Jean-Christophe Nebel

Digital Imaging Research Centre

Kingston University

Context

To prevent crime, the UK government has invested in visual surveillance systems.

5 millions Closed Circuit Television (CCTV) cameras are operating in the UK

1 camera per 12 inhabitants

~20% of the world's allocation

Challenge

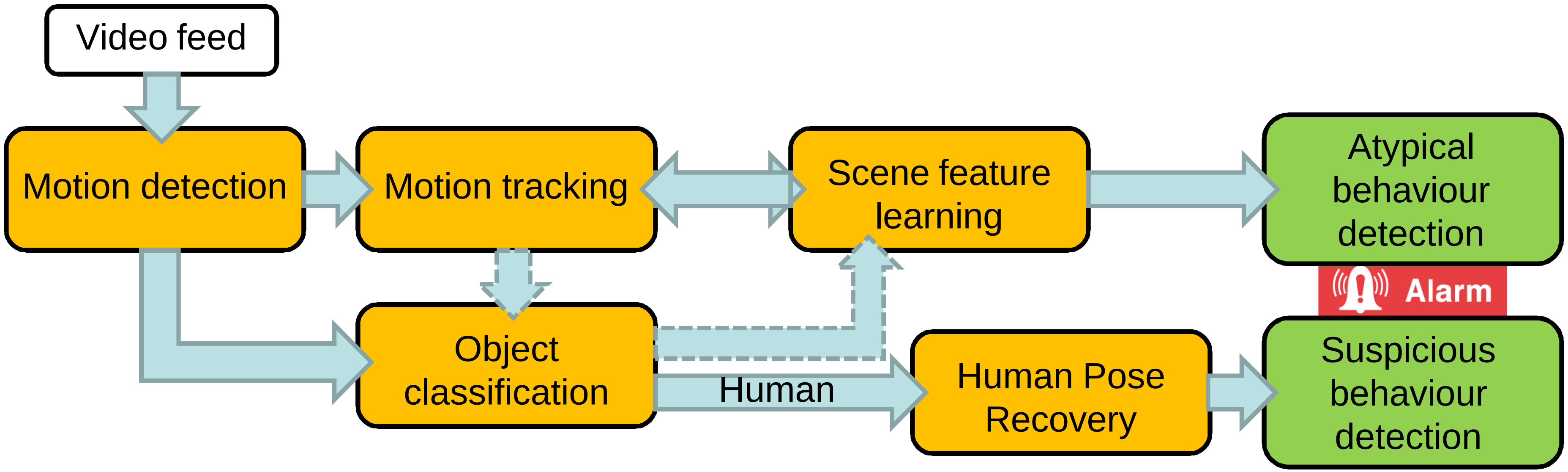
It is not possible to rely only on continuously monitoring by human personnel to detect suspicious events.

After 22 min, CCTV operators miss 95% of scene activity!

Automated Visual Surveillance

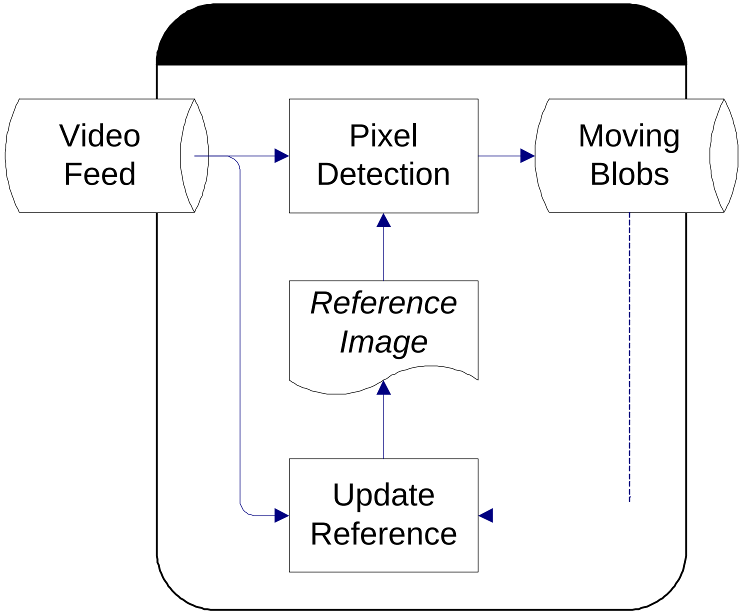
Tools for assisting operators

Intelligent systems able to detect unusual or suspicious behaviour to trigger an alarm



Motion Detection

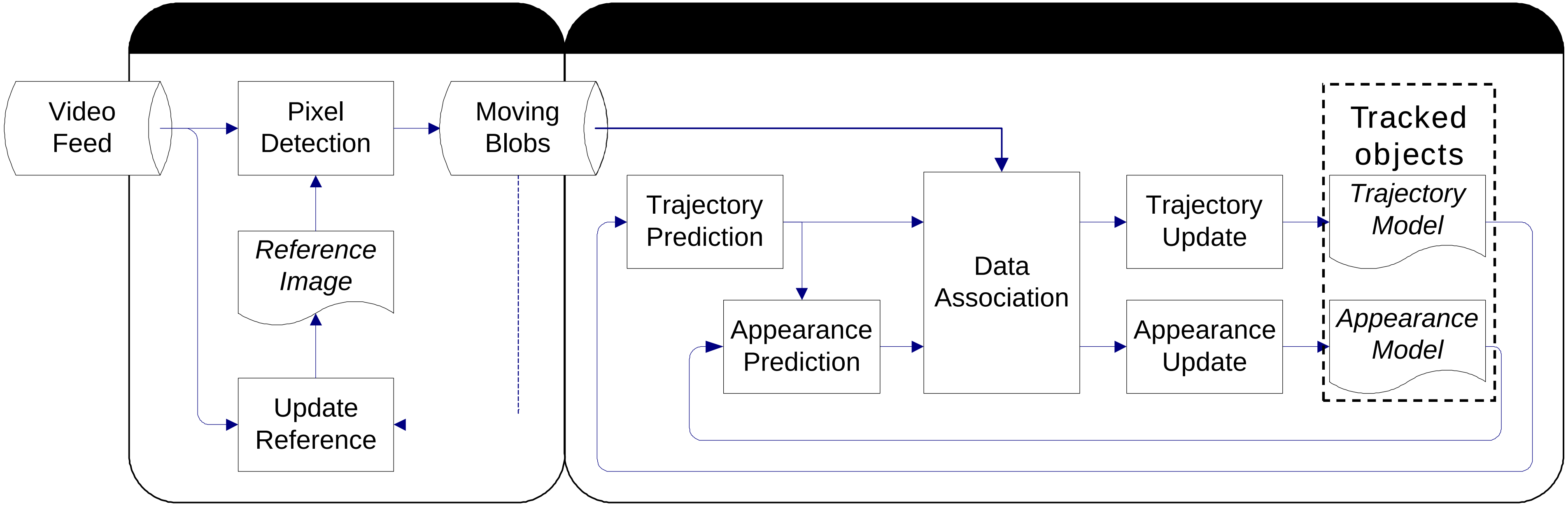
For fixed camera, it is based on subtraction between reference image and current image.



Reference Image
Incoming Image
Detected Blobs

Motion Tracking

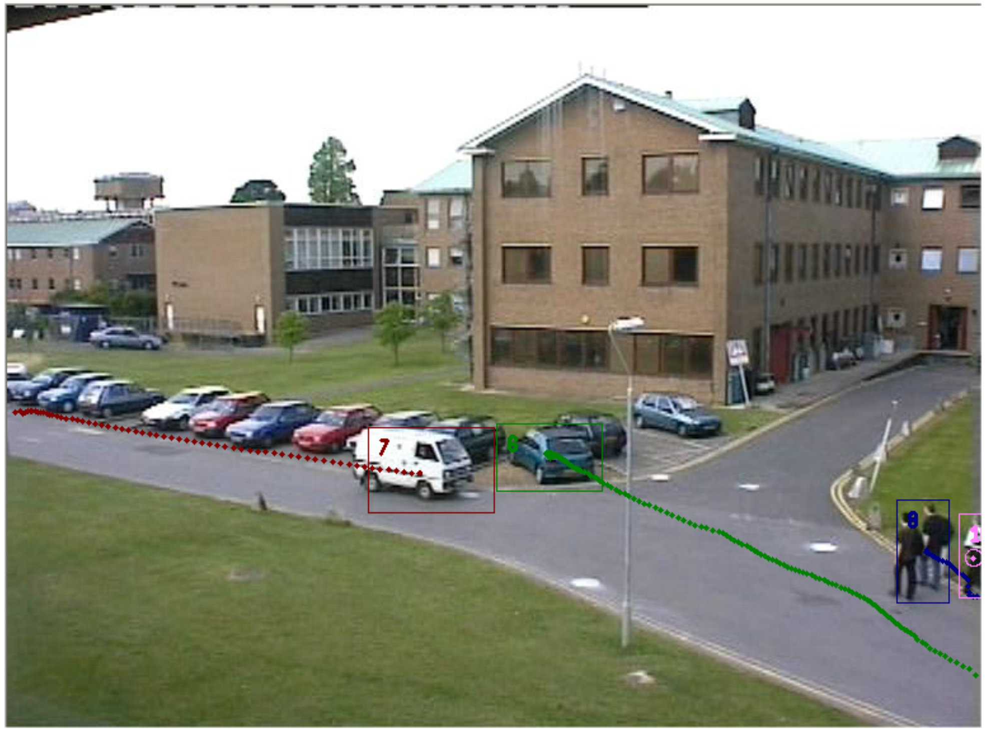
Typical object tracking methodology for fixed camera uses motion detection and blob tracking.



Motion Tracking

Detection of moving objects -> Blobs

Blob matching -> Trajectories



Object classification

Machine learning technique based on **supervised** learning

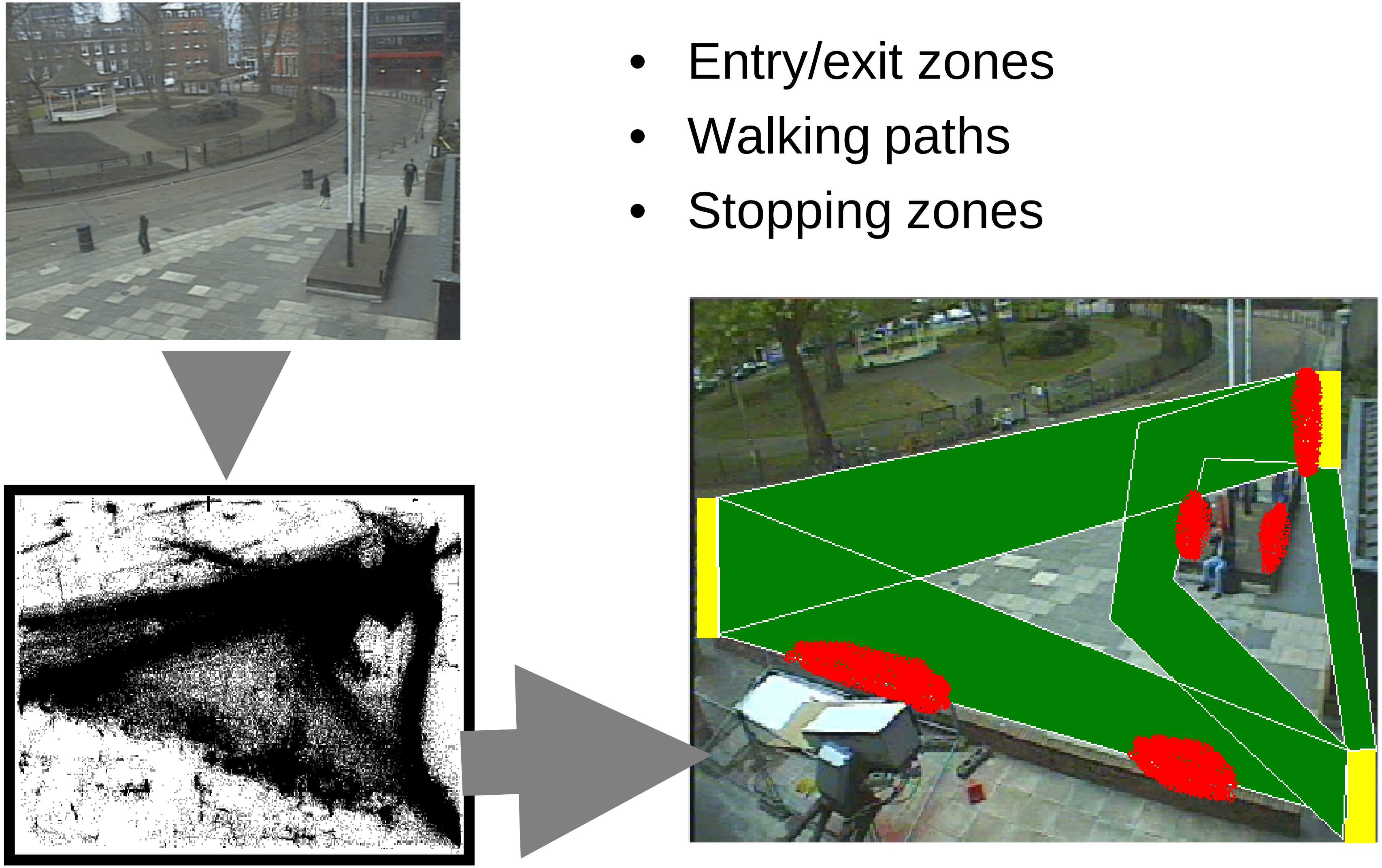
Learning scene features

Since a CCTV camera records 24/7, a huge amount of data is available.

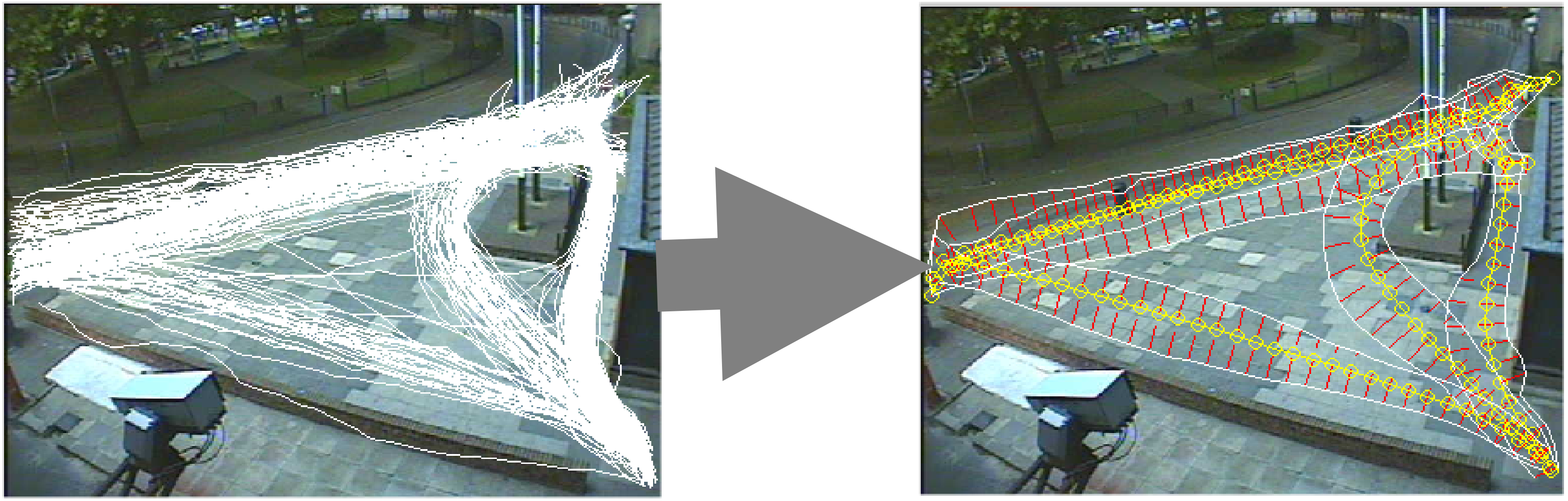
Unsupervised learning algorithms can extract scene features.

Learning scene features: zones

- Entry/exit zones
- Walking paths
- Stopping zones



Learning scene features: routes

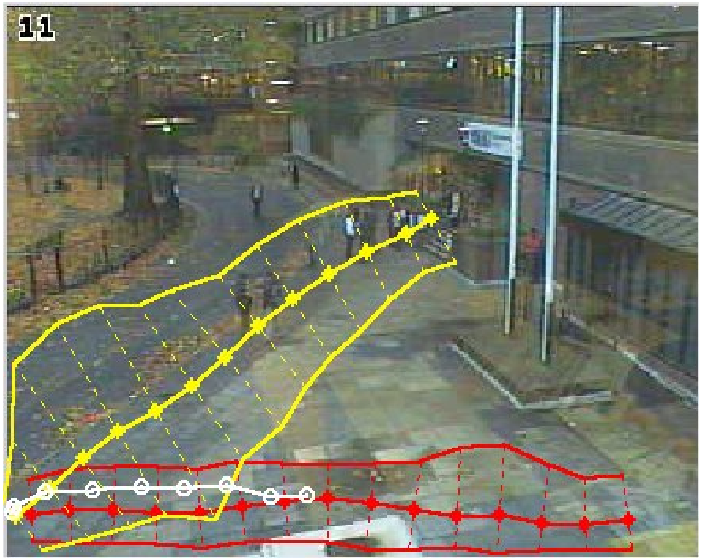


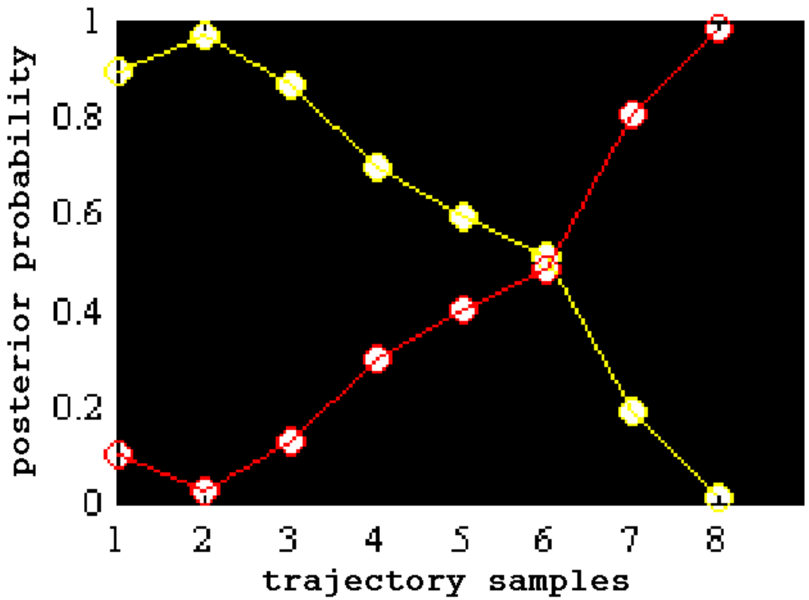
After manual labelling of scene features, activities can be automatically annotated.

Pedestrian #58
enters the scene from Goswell rd, at 12:51:04
goes through route Goswell rd-Main Entrance
exits the scene to the Main Entrance, at 12:51:18

Long term trajectory prediction

- Trajectory modelling using HMM
- Evaluation of observed position:
  - White: Trajectory
  - Yellow, Red: candidate paths





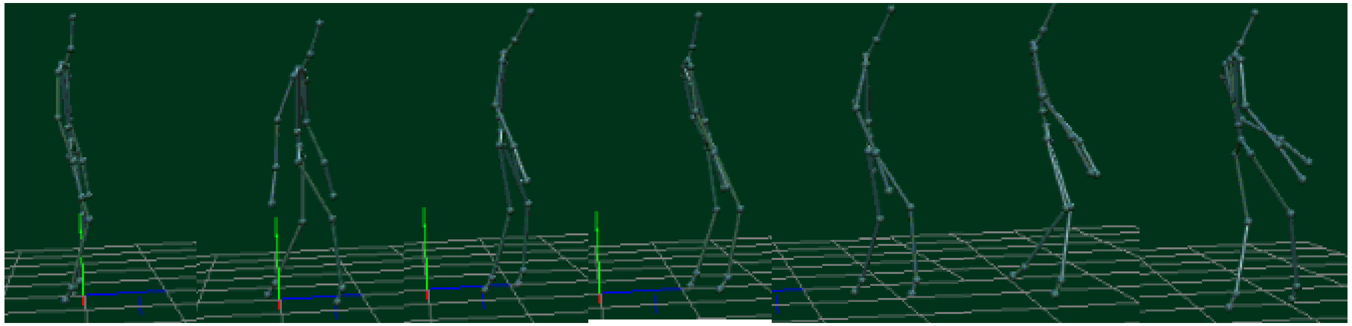
Atypical behaviour detection

HMMs can reveal atypical activities*.

**Trajectory characteristics are time-dependent*

Criminal behaviour detection*

Classification of motions of humans seen as articulated bodies



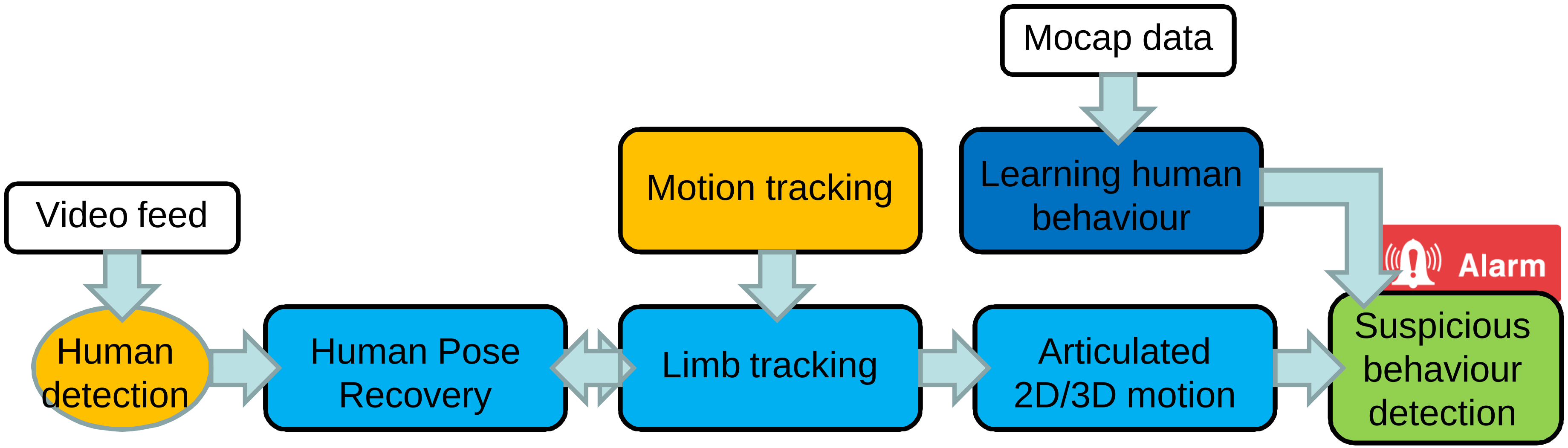
Drunken walk

*Work in progress

Criminal behaviour detection pipeline

Recovery of human posture

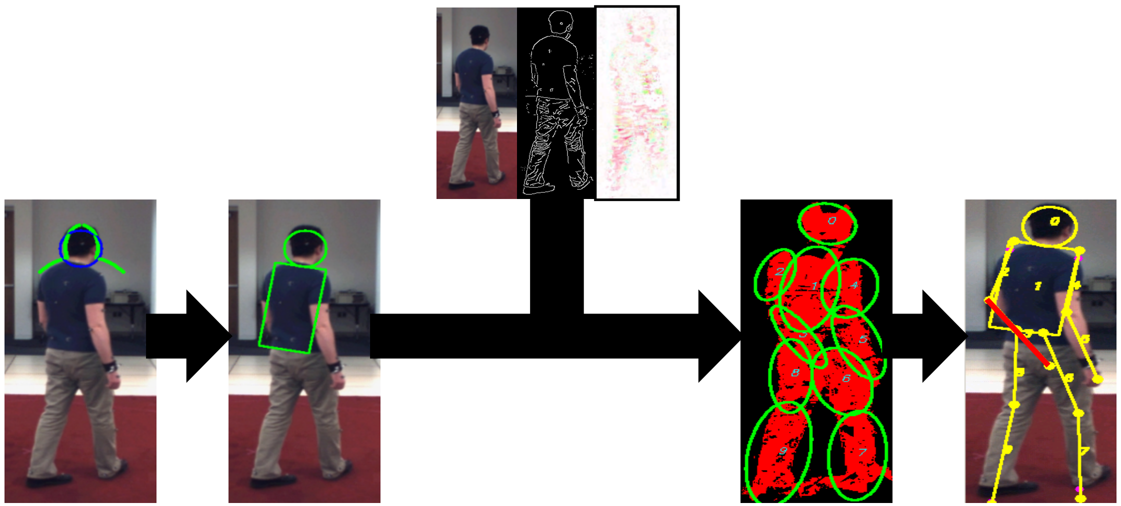
Comparison of posture sequence with known activities



Human pose recovery

Detection of body parts inside detected blob.

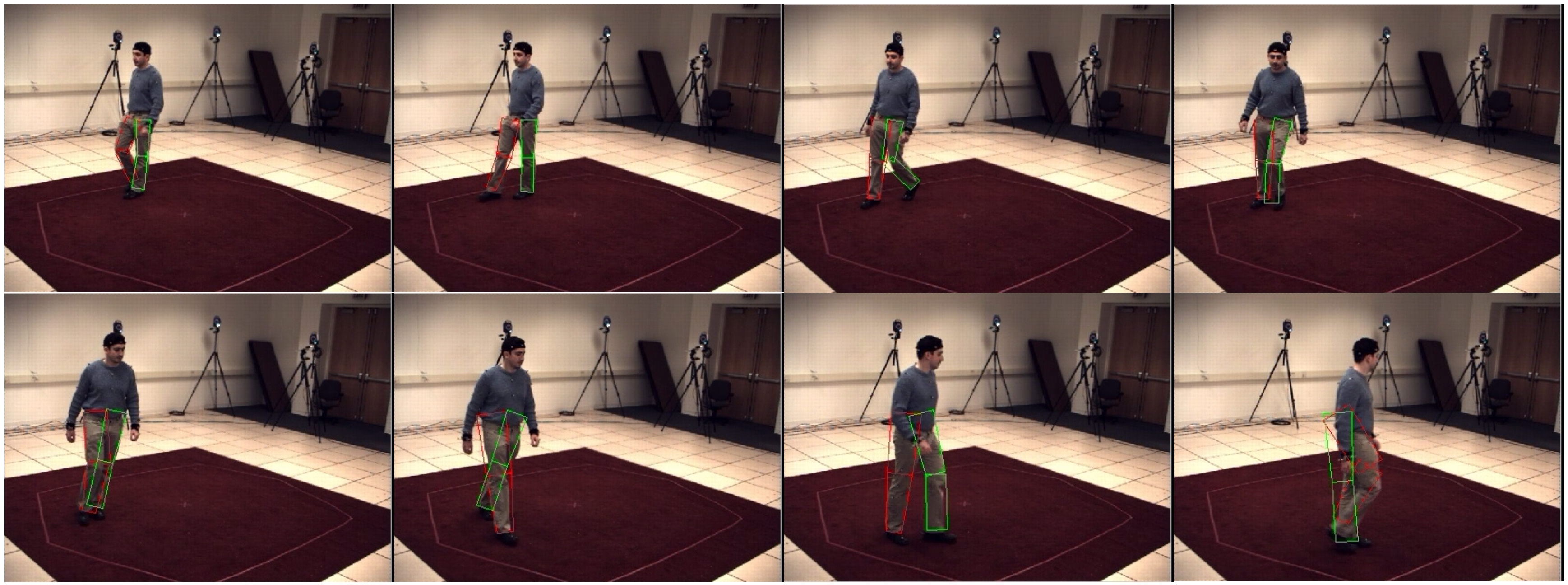
- Head detection using omega model
- Torso fitting using head and colour
- Clustering cues from colour, edges and optical flow
- Skeleton production and evaluation



Limb tracking

Tracking body parts cannot be achieved using a linear motion model (e.g. Kalman filter).

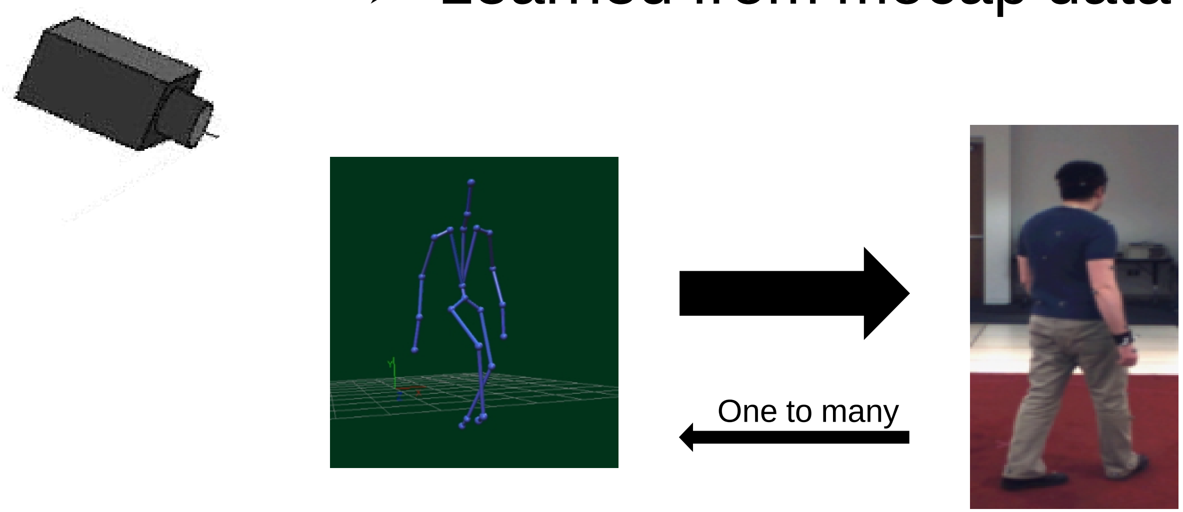
Particle filter tracking based on a 2D articulated model using colour and edges as cues.



Articulated 3D motion

2D-to-3D transform is necessary to infer the 3D motion.
However, it is an ill-posed problem
This can only be solved with:

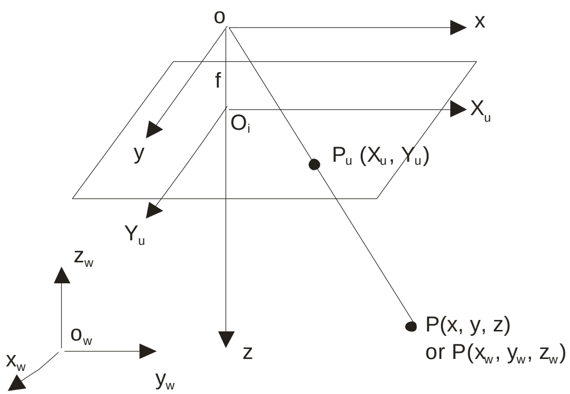
- Camera calibration parameters
- Constraints to choose among possible solutions
  - ➢ Learned from mocap data

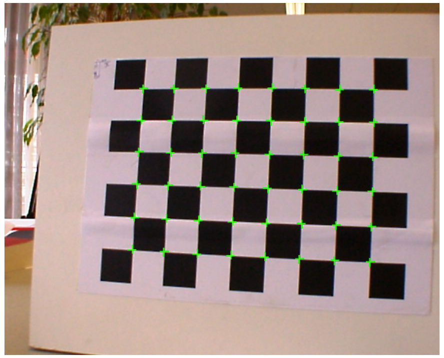


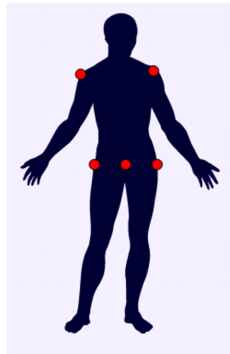
Camera auto-calibration using human body features

Pin-hole camera model proposed by Tsai [1].

Calibration can be achieved with at least 5 coplanar points







Human body can be used as a calibration target:

5 points are coplanar during mid-stance position

[1] R.Y. Tsai, An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision, Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Miami Beach, FL, pp. 364-374, 1986

Conclusions

- Automated Visual Surveillance based on Computer Vision is the future in Surveillance Technology
- Low level modules can assist operators
- High-level modules can provide automatic interpretation
- Recovery of articulated 3D motion offers a lot of potential for criminal behaviour detection

- …but it is not a solved problem yet!

DIRC Visual Surveillance Team

Academics

Dr Sergio Velastin, **Centre Director**
Dr James Orwell, **Group Leader**
Prof Tim Ellis
Prof Graeme A. Jones
Dr Dimitrios Makris
Dr Jean-Christophe Nebel

Researchers

James Annesley
Justin Cobb
Dr Robert Hung
Dr Paul Kuo
Dr Valerie Leung
Dr Christopher Madden
Dr Najla Megherbi
Dr Hossein Ragheb
Simone Boragno
Dr Maria Valera Espina
Dr James Black

Research Students

Norbert Buch
Alberto Colombo
Amin Dadgar
Neda Lazarevic-McManus
Jesus Martinez-del-Rincon
Michal Lewandowski
Kamran Virk
Fei Yin
Bei Bei Zhan

Relevant Publications by the Team (1/2)

1. D. Greenhill, J.R. Renno, J. Orwell, G.A. Jones, "Occlusion Analysis: Learning and Utilising Depth Maps in Object Tracking" in 'Image and Vision Computing', 26(3) Elsevier Publishing, pp. 430-441, (2008)
2. B. Zhan, N.D. Monekosso, P. Remagnino, S.A. Velastin, L Xu, "Crowd Analysis: a Survey" in 'Machine Vision and Applications', (2008)
3. N. Lazarevic-McManus, J.R. Renno, D. Makris, G.A. Jones, "An Object-based Comparative Methodology for Motion Detection based on the F-Measure" in 'Computer Vision and Image Understanding', (2008)
4. H.M. Dee, S.A. Velastin, "How close are we to solving the problem of automated visual surveillance? A review of real-world surveillance, scientific progress and evaluative mechanisms" in 'Machine Vision and Applications', (2007)
5. P. Remagnino, S.A. Velastin, G. Foresti, M Trivedi, "Novel Concepts and challenges for the next generation of video surveillance" in 'Machine Vision and Applications', 18(3-4) pp. 135-137. (2007)
6. Colombo, V Leung, J. Orwell, S.A. Velastin, "Markov Models of Periodically Varying Backgrounds for Change Detection", Visual Information Engineering 2007, IET, July, London, (2007)
7. B Vrusias, D. Makris, J.R. Renno, N. Newbold, K. Ahmad, G.A. Jones, "A Framework for Ontology Enriched Semantic Annotation of CCTV Video", Int.Workshop on Image Analysis for Multimedia Interactive Services, Santorini, Greece, (2007)
8. J.R. Renno, D. Makris, G.A. Jones, "Object Classification In Visual Surveillance Using Adaboost", Seventh International Workshop on Visual Surveillance, June 22nd, Minneapolis, (2007)
9. S Boragno, BA Boghossian, J. Black, D. Makris, S.A. Velastin, "A DSP-based system for detecting vehicles stopping in prohibited area", Advanced Video and Signal Based Surveillance, September, London, UK, (2007)
10. P. Kuo, J.-C. Nebel, D. Makris, "Camera Auto-Calibration from Articulated Motion", Advanced Video and Signal Based Surveillance , London UK, (2007)
11. Fei Yin, D. Makris, S.A. Velastin, "Performance Evaluation of Object Tracking Algorithms", 10th IEEE International Workshop on Performance Evaluation of Tracking and Surveillance (PETS2007), October, Rio de Janeiro, Brazil, (2007)
12. S.A. Velastin, BA Boghossian, MA Vicencio-Silva, "A motion-based image processing system for detecting potentially dangerous situations in underground railway stations" in 'Transportation Research Part C: Emerging Technologies', 14(2) Elsevier, April, pp. 96-113. (2006)

Relevant Publications by the Team (2/2)

13. M. Xu, T. Ellis, "Augmented tracking with incomplete observation and probabilistic reasoning" in 'Image and Vision Computing', 24(11) Elsevier, pp. 1202-1217. (2006)
14. L.M. Fuentes, S.A. Velastin, "People tracking in surveillance applications" in 'Image and Vision Computing', 24(11) November, pp. 1165-1171. (2006)
15. G. Foresti, L. Snidaro, P. Remagnino, T. Ellis, "Advanced Image and Video Processing in Active Video-Based Surveillance Systems" in 'IEEE Signal Processing Magazine', IEEE, (2005)
16. L.M. Fuentes, S.A. Velastin, "Tracking-based event detection for CCTV systems" in 'Pattern Analysis and Applications', 7(4) Springer, (2005)
17. D. Makris, T. Ellis, "Learning Semantic Scene Models from Observing Activity in Visual Surveillance" in 'IEEE Transactions on Systems Man and Cybernetics - Part B', 35(3) June, pp. 397-408.,(2005)
18. J. Black, T. Ellis, "Multi camera image tracking" in 'Image and Vision Computing', Elsevier, (2005)
19. D. Greenhill, J.R. Renno, J. Orwell, G.A. Jones, "Learning the Semantic Landscape: Embedding scene knowledge in object tracking" in 'Real Time Imaging', 11, pp. 186-203. (2005)
20. J.-C. Nebel, W.P. Cockshott, V. Yarmolenko, E. Borland, D. MacVicar, "Pre-commercial 3D Digital TV studio" in 'IEE Proceedings - Vision, Image and Signal Processing', 152(6) December, pp. 665-667. (2005)
21. J. Black, D. Makris, T. Ellis, "Hierarchical Database for a Multi-Camera Surveillance System" in 'Pattern Analysis and Applications', 7(4) Springer, December, pp. 430-446. ISBN/ISSN 1433-7541 (2004)
22. J.R. Renno, P. Remagnino, G.A. Jones, "Learning Surveillance Tracking Models for the Self-Calibrated Ground Plane" in 'Acta Automatica Sinica', Special Issue on Visual Surveillance of Dynamic Sc 29(3) pp. 381-392. (2003)
23. E. Corvee, S.A. Velastin, G.A. Jones, "Occlusion Tolerent Tracking using Hybrid Prediction Schemes" in 'Acta Automatica Sinica', Special Issue on Visual Surveillance of Dynamic Sc 23(3) pp. 356-369. (2003)
24. W.P. Cockshott, S. Hoff, J.-C. Nebel, "An Experimental 3D Digital TV studio" in 'IEE Proceedings - Vision, Image and Signal Processing', 150(1) IEE-INST ELEC ENG, HERTFORD, February, pp. 28-33, (2003)
25. D. Makris, T. Ellis, "Path Detection in Video Surveillance" in 'Image and Vision Computing', 20(12) Elsevier, October, pp. 895-903, (2002)

Abstract

Automated Visual Surveillance:
Detection of Criminal Behaviour

By Jean-Christophe Nebel

In order to prevent crime, the UK government has significantly increased its spending on visual surveillance systems over the last decade. With more than 5 millions Closed Circuit Television (CCTV) cameras installed in the country, it is not possible anymore to rely on continuously monitoring by human personnel to detect suspicious events. Computer vision and machine learning techniques are already used to facilitate the task of CCTV operators by highlighting areas of interest. However, fully automated visual surveillance systems require a higher level of complexity to detect abnormal and criminal behaviour. Towards this end, the Digital Imaging Research Centre at Kingston University, UK, has been developing new techniques based on learning semantic scene models and recovering human body postures. The presentation will focus on the latest results achieved by this research group.