*"When multimedia meets surveillance and forensics in people security"**

Prof. Rita Cucchiara
Facoltà di Ingegneria "Enzo Ferrari"

Università di Modena e Reggio Emilia (Italy)

* Slides presented at ACM Multimedia 2010:
R. Cucchiara "When Multimedia meets surveillance and forensics in people security" Keynote at MIFOR2010 "Multimedia in Forensics Security and Intelligence" Workshop in conjunction with ACM Multimedia 2010, October 29 Firenze, Italy

(When) multimedia meets surveillance and forensics in people security

People security
The degree of protection of people against damage, danger, criminal or terroristic actions

Multimedia
The use of different media contents (text audio images, video, animation interactivity…) to convey information for users

Digital forensics
A form of digital investigation that can be entered in a legal court

Forensic video analysis
The scientific examination, comparison and/or evaluation of video in legal matters <u>for investigation</u>

Surveillance
The <u>manual/automatic monitoring</u> of situation, behaviour, activity to generate alarm and record meaningful situation

Video Surveillance
The acquisition and processing of visual data about the external word
to detect target along the time and the space,
to recognize interesting and dangerous situations,
to <u>generate alarms and record data</u> on the scene

Digital Forensics

A form of digital investigation that can be entered in a legal court [Car09]

Example uses of digital investigations:

- Catching online predators or pedophiles
- Linking people to physical crimes, (missing persons, murders..)
- Detecting corporate fraud, computer intrusion, drug dealing, and terrorism cases
- Analyzing digital data as a support for investigation

[Car09]Brian D. Carrier “Digital Forensics Works” IEEE Computers April 2009

Digital Forenscis challanges

Challanges in digital forensics [Car09]

1. **Data Preservation:** “no distructive test”

2. **Evidence gathering** : define and achieve reliable tools of search with a good recall and an acceptable precision

3. **Reconstruction:** how reconstruct events and knowledge. An open research problem is to determine how to assign certainty values to digital event reconstruction.

- *Differently of physical forensic sciences, in digital forensics the procedures are not well assessed neither the results.*

What Multimedia can do?

Multimedia approaches for data management in digital forensics

- Support for **preservation**
  - Watermarking,secure coding..
- Support for tools of **evidence gathering**
  - Interactive and natural interfaces
  - Search and retrieval
- Support for **reconstruction**
  - Research in multimedia data analysis

“Multimedia forensics?”

Digital Forensic and Computer Forensics

DIGITAL FORENSICS

- "The digital revolution introduces digital trace of our activity in Real Life" [Franke09]
  - Computer activities
  - Digital Social Interactions
  - Digital Evidence of analog Processes

COMPUTER FORENSCIS

- "When computer are involved criminal activities" [Kruse01]:
  - Tools for committing crimes
  - Substrate where crimes are committed

[Francke09] Franke K., Shriari S. Computational Forensics an Overview 2009

[Kruse01] Kruse, W., Heiser, J.: Computer Forensics: Incident Response Essentials. Addison Wesley, Reading (2001)

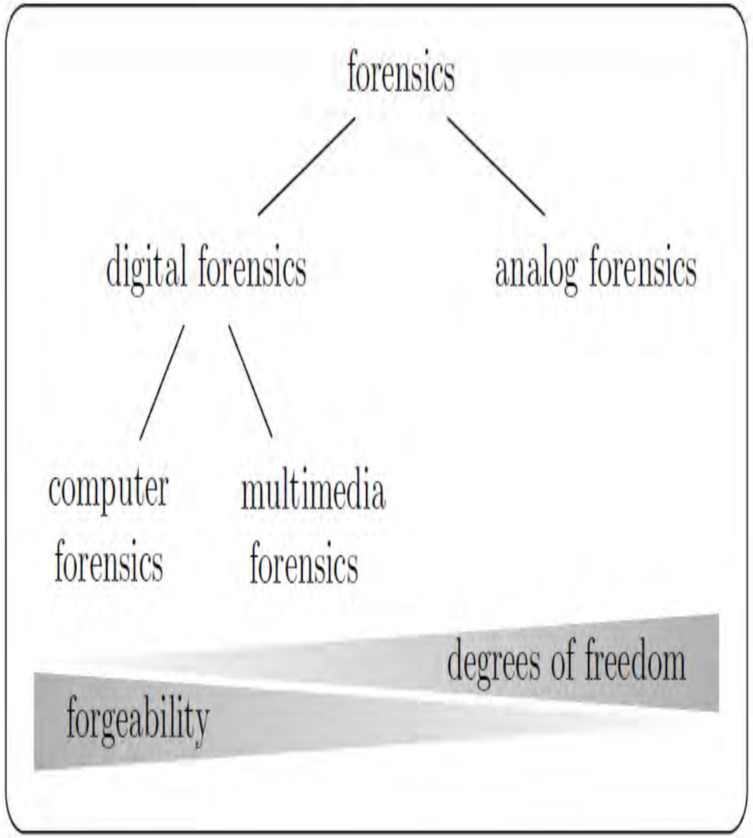
Multimedia Forensic

MULTIMEDIA FORENSICS [Bohme09]

- Many data are collected through sensors
- Create a digital counterpart of reality
- Digital data can be probative elements in many investigation (e.g. video, audio, photos..)

“Data must be authentic and reliable”

ONTOLOGY ON FORENSICS



[Boheme09] Böhme, R., Freiling, F. C., Gloe, T., and Kirchner, M. 2009. Multimedia Forensics Is Not Computer Forensics. In Proc. of the 3rd international Workshop on Computational Forensics

Computational Forensic is not Computer Forensic

Computational Forensic refers to applying computer aided techniques to digital data understanding:

- Assist in basic and applied research and data mining
- Establish or prove the scientific basis of a particular investigative procedure
- Support the forensic examiner in their daily case work.

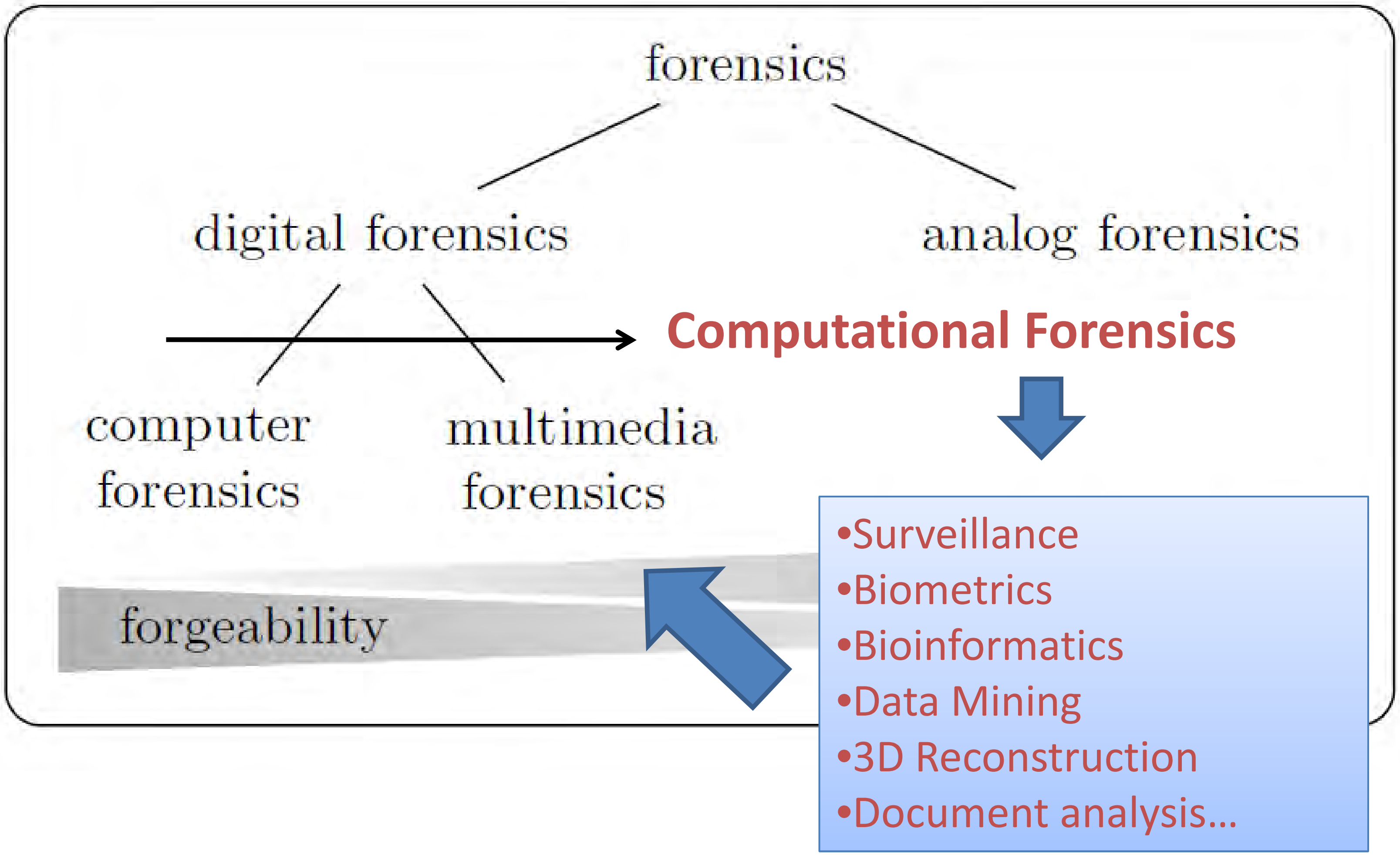
"Modern crime investigation shall profit from the hybrid-intelligence of humans and machines"

[Francke09] Franke K., Shriari S. Computational Forensics an Overview 2009

Computational Forensics Techniques

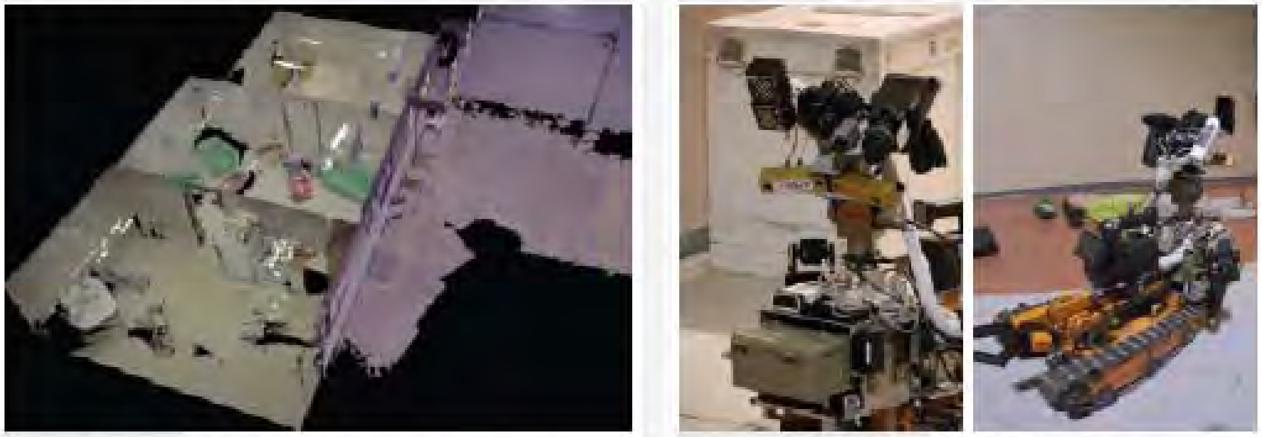
- **Signal / Image Processing :** one-dimensional signals and 2-dimensional images are transformed for the purpose of better human or machine processing,
- **Computer Vision :** images are automatically recognized to identify objects,
- **Computer Graphics / Data Visualization** : two-dimensional images or three-dimensional scenes are synthesized from multidimensional data for better human understanding,
- **Statistical Pattern Recognition** : abstract measurements are classified as belonging to one or more classes, e.g., whether a sample belongs to a known class and with what probability,
- **Data Mining** : large volumes of data are processed to discover nuggets of information, e.g., presence of associations, number of clusters, outliers in a cluster,
- **Robotics :** human movements are replicated by a machine, and
- **Machine Learning** : a mathematical model is learnt from examples.

Ontology of Forensic Revisited



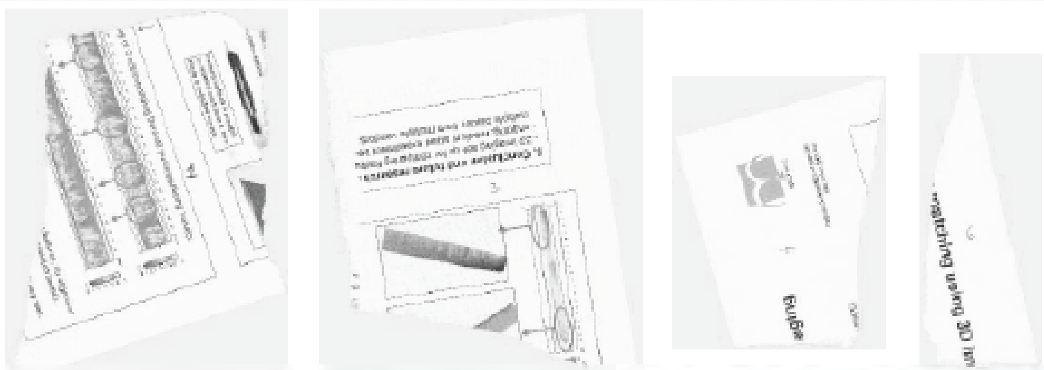
Examples

Anthropology by Balleri et al. 2007

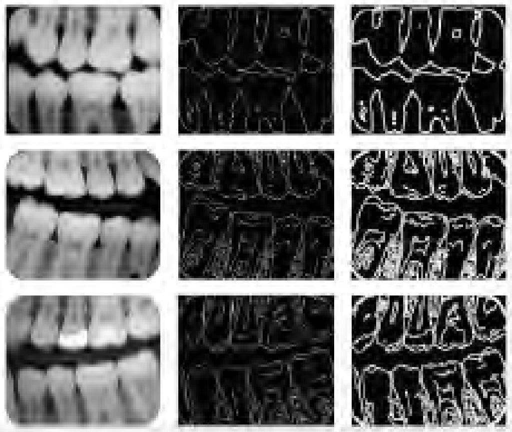


Crime Scene Reconstruction by Gryz et al. 2007

Examples



Reconstruction of Shredder and Ripped-Up Documents by Ukowich et al. 2007 and de Smet 2007.



Forensic Odontology by Nomir et al. 2007

Adversary Modeling in Multimedia Surveillance, Mohan Kankanhalli.

Graffiti-ID: Identifying Gang Graffiti Images, Anil Jain et al (Michigan State University, US);

Temporal Normalization of Videos Using Visual Speech, Usman Saeed et al (Institute Eurecom, FR);

Video Surveillance and Multimedia Forensics: an Application to Trajectory Analysis, Simone Calderara et al

Single View Geometry and Active Camera Networks Made Easy, Federico Pernici et al

Multi-target Tracking in Time-lapse Video Forensics, Paul Koppen et al (University of Amsterdam, NL);

Videntifier Forensic: A New Law Enforcement Service for Automatic Identification of Illegal Video Material, Herwig Lejsek et al (Reykjavik University, IS

Image Spam Clustering - An Unsupervised Approach, Chengcui Zhang et al (University of Alabama at Birmingham, US);

Design and Deployment of a Digital Forensics Service Platform for Internet Videos, Wen Hui et al (University of Science and Technology Beijing, CN);

Digital Forgery Estimation into DCT Domain - A Critical Analysis, Sebastiano Battiato et al (University of Catania, IT);

A New Approach for JPEG Resize and Image Splicing Detection, Qingzhong Liu (New Mexico Tech, US)

Exposing Digital Video Forgery by Ghost Shadow Artifact, Jing Zhang (Tianjin University,

Multimedia surveillance

- The integration of multimedia technologies and sensor networks constitutes the fundamental infrastructure of new generation of multimedia surveillance systems,
- where many different media streams (audio,video, text, 3D graphics, sensor data..) concur to provide an automatic analysis of the controlled environment and a support for human interpretation of the scene [Cuc05].

- Multimedia surveillance systems to
  - Enlarge the view
  - Enhance the view
  - Explore new views   for human security employers

[Cuc05]R. Cucchiara "Multimedia surveillance systems" Proc of VSSN'05 at ACM Multimedia Singapore 2005

Multimedia (Video) surveillance

On-line analysis

Real time response

Many fixed conditions and pre-defined constraints

Correlation between cameras
Consistent labeling
Recognition

Multimedia (Video) forenscis

Off-line analysis on logged data

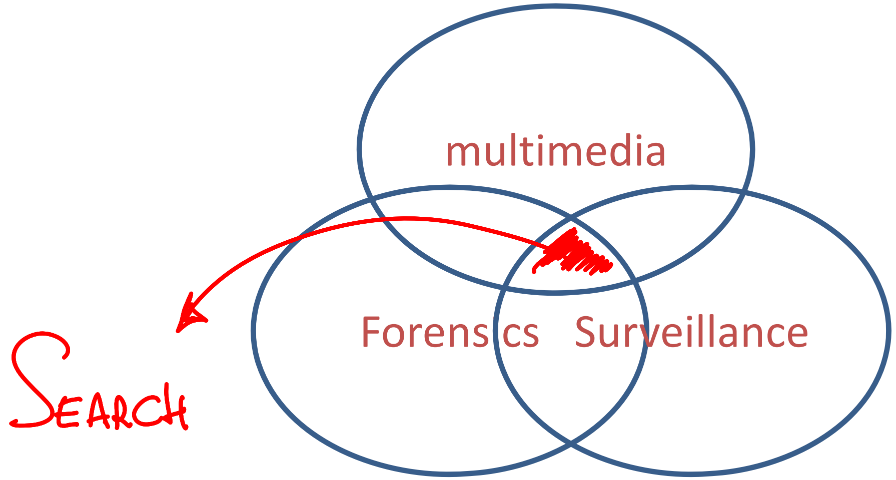
Fast processing in large sets of data

Often undefined camera settings and constraints

Correlation between objects
Re-identification
Mining....

Noise and unceraininity on (movement) data
High variability of visual data
Need of multimedia data management
and analysis tools ***for people security***

People search



People search for forensics and surveillance;
Why?

1) *search people for answering a specific query* (i look for people with sun glasses and a blue jaked with a red luggage..)

2) *search people similar to a given shape* (similarity search, CBIR)

3) *search people moving in an area or with a given behaviour* (people data and metadata annotation for search and mining)

From detecting
to reasoning on people

People detection

People segmentation

Localization

Face detection & recognition

People Identification

Identity assessment
Biometry

People tracking

People Re-identification

Search
Soft-biometry

Motion analysis

Activity analysis

Posture,
Gesture,
Gate,
Trajectories..

Action (& interaction) analysis

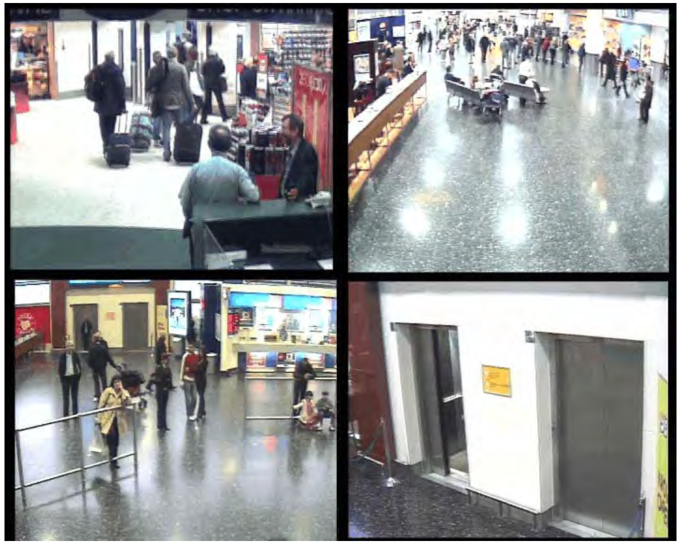
With environment
With moving objects
With people

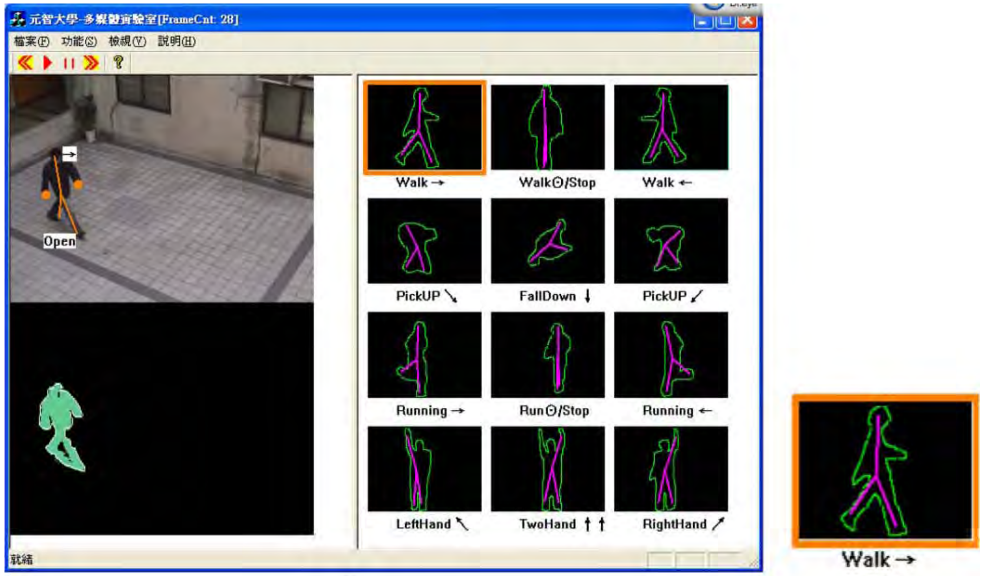
Bheavior analysis

Understanding
Modeling

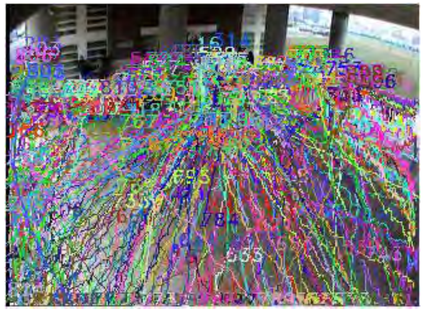
People segmentation

- Typical approach of surveillance with static cameras





(a) (b)



People segmentation

PEOPLE DETECTION AND TRACKINGI WTH STATIC CAMERAS

C. Wren, A. Azarbayejani, T. Darrell, and A.P. Pentland, “Pfinder: real-time tracking of the human body,” *IEEE Trans. PAMI*, (19) 7, 1997.

C. Stauffer and W.E.L. Grimson, “Learning patterns of activity using real-time tracking,” *IEEE Trans. PAMI*, (22) 8, 2000.

I. Haritaoglu, D. Harwood, and L.S. Davis, “W4: real-time surveillance of people and their activities,” *IEEE Trans. PAMI*, (22) 8, 2000.

Tao Sawneir, Kumar. Object tracking with bayesian estimation of dynamic layer representation IEEE Trans on PAMI 24,1 2002

R. Cucchiara, C. Grana, M. Piccardi, A. Prati “Detecting Moving Objects, Ghosts and Shadows in Video Streams“, IEEE *Trans on PAMI*, 2003

PEOPLE TRACKING WITH MULTIPLE STATIC CAMERAS

A. C. Sankaranarayanan, A.Veeraraghavan, and R.Chellappa, Object Detection, Tracking and Recognition for Multiple Smart Camera Proceedings of the IEEE | Vol. 96, No. 10, October 2008

S. Calderara, A. Prati, R. Cucchiara, “Bayesian-competitive Consistent Labeling for People Surveillance“ on IEEE Trans on PAMI, feb. 2008

Saad M. Khan and Mubarak Shah; Tracking Multiple Occluding People by Localizing on Multiple Scene Planes; IEEE TRANS. ON PAMI, VOL. 31, NO. 3, MARCH 2009

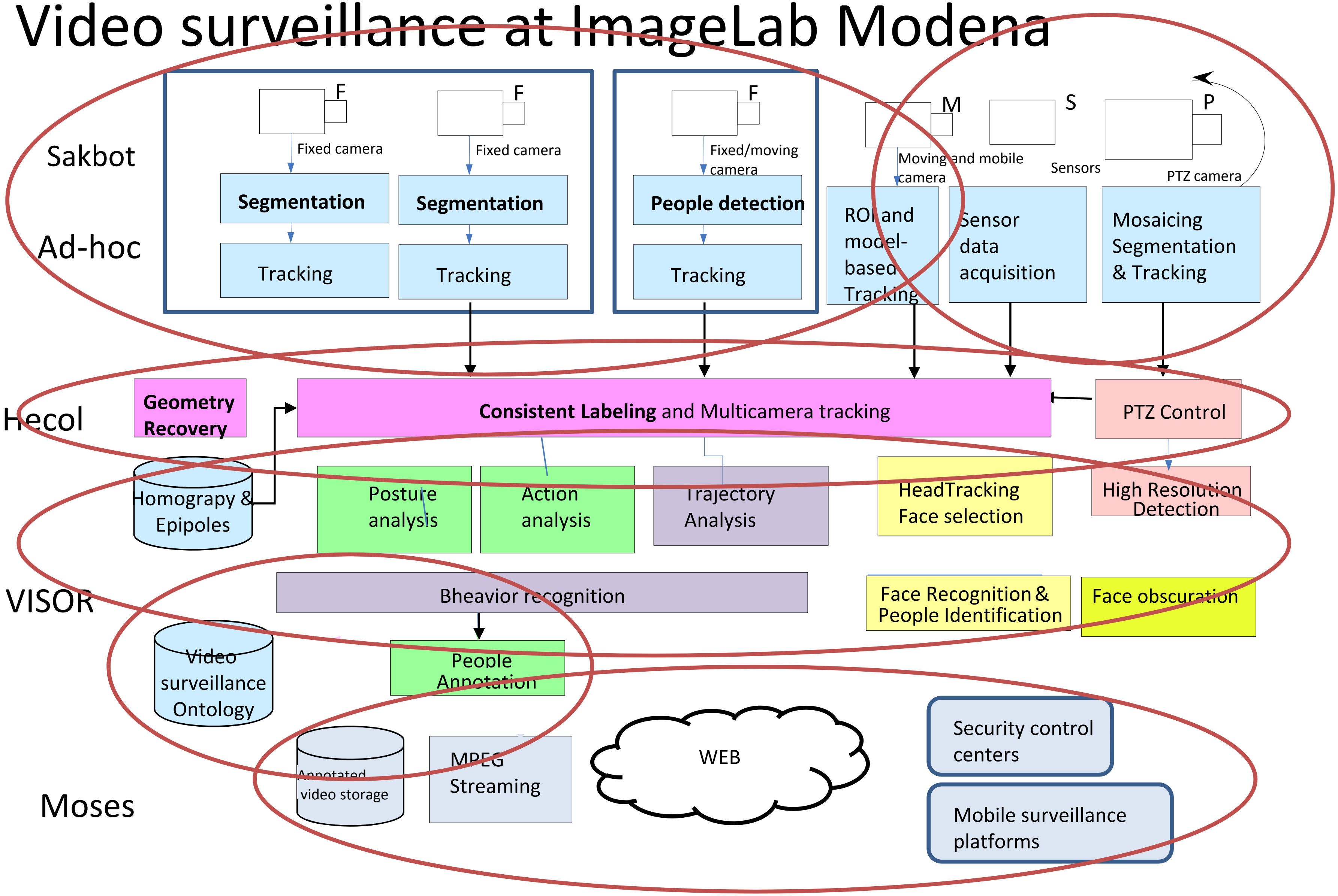
OCCLUSION DETECTION

Nguyen, H.T. Smeulders, A. Fast occluded object tracking by a robust appearance filter IEEE Trans on PAMI, 2004

Tao Zhao Nevatia, R. Tracking multiple humans in complex situations IEEE Trans. PAMI 2004

With forest of cameras
With moving and mobile cameras
..
In crowd…
(Shah’s Talk ACM MM2010)

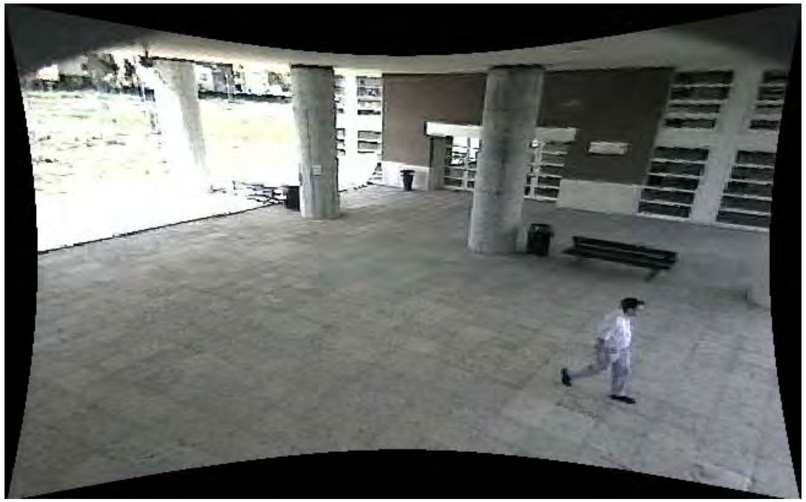
Video surveillance at ImageLab Modena

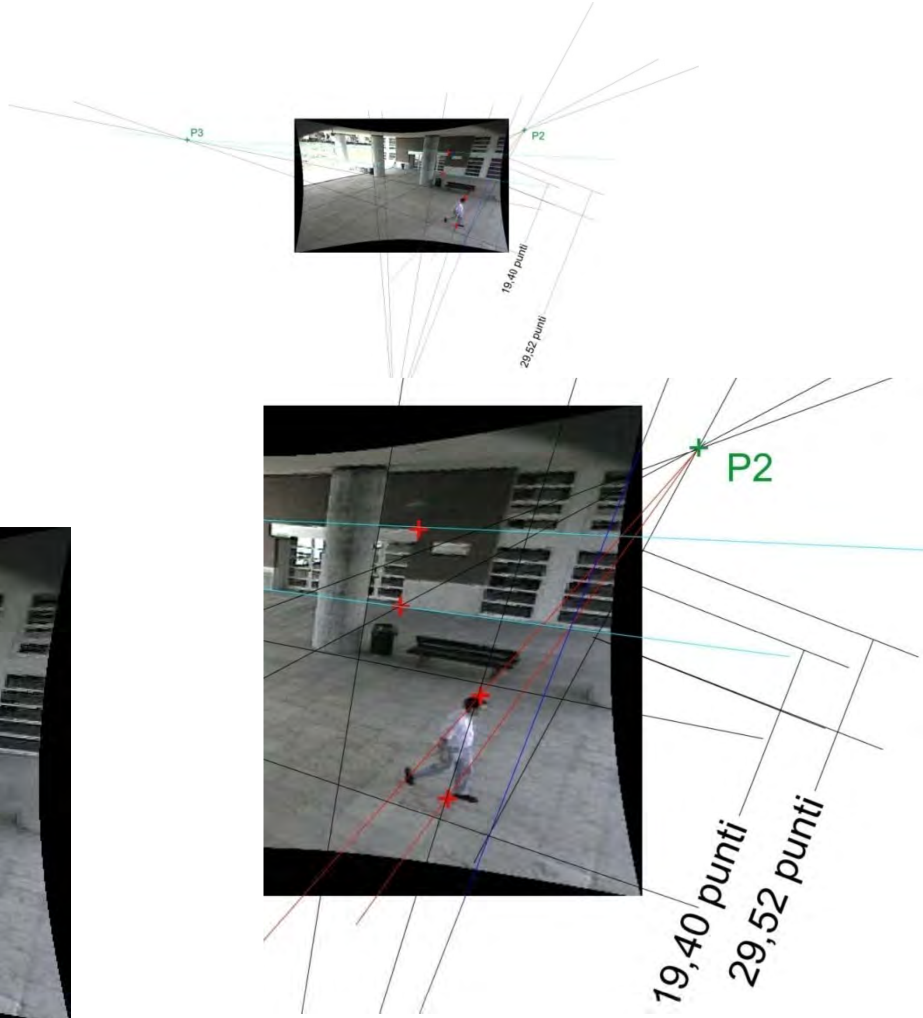


Experiments in Surveillance

Esperiments in forensics in Modena…

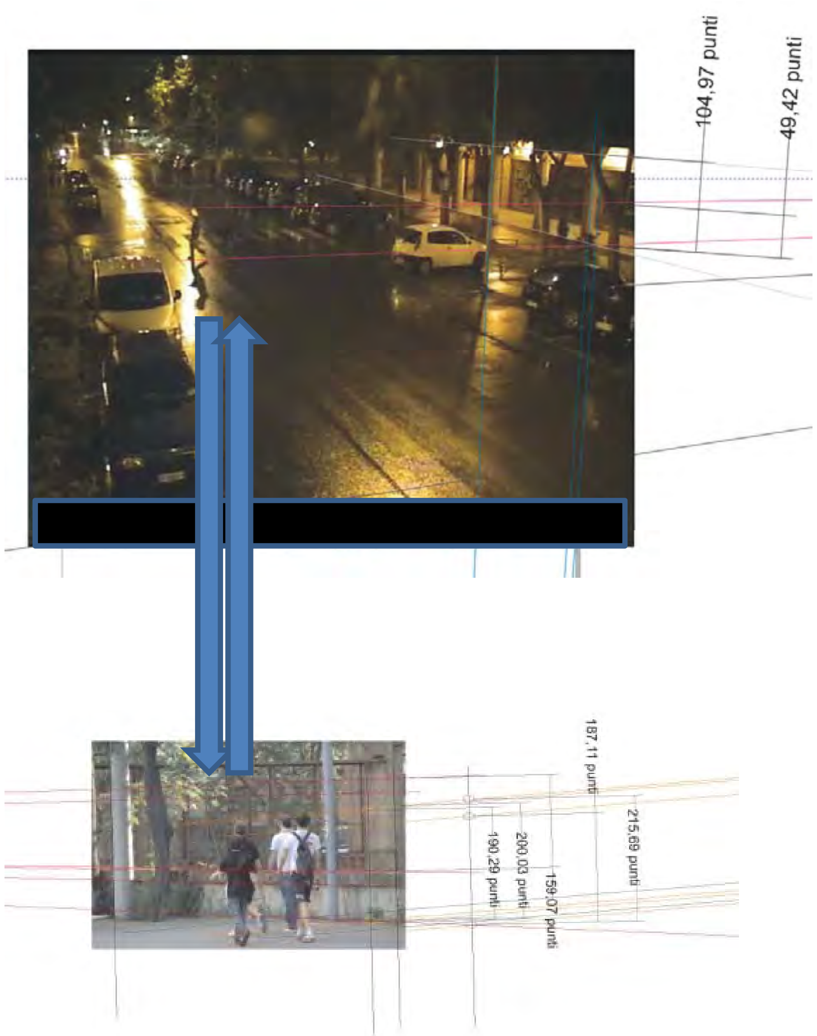
A tool for soft-biometry measure

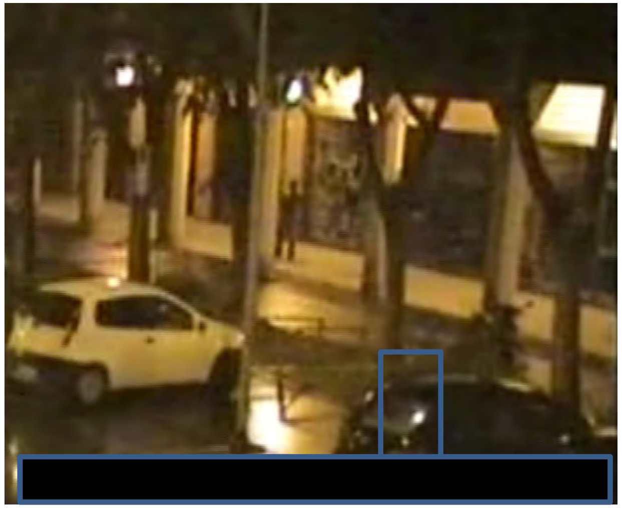




Heigh cm: 177,43

Example





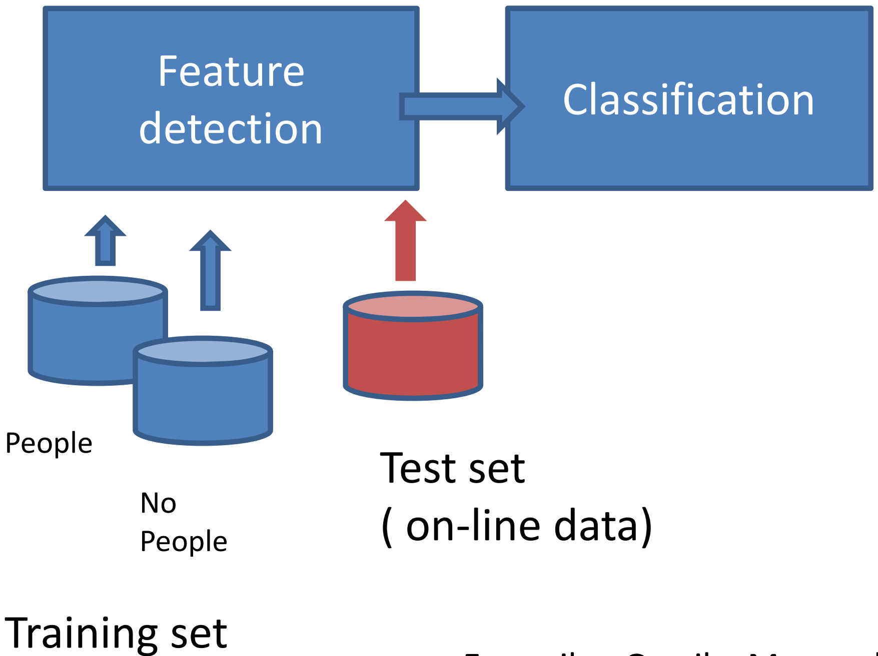
Data correlation
for manual identification

Support of investigation

Pedestrian detection

- When camera is moving
- When the environment is too complex
- When the background is not available…

→People (pedestrian) detection with machine learning approach in still images (and video)



- **Many features**
- **Many classifiers:**
  - **SVMS**
  - **Boosting classifiers with Sliding windows search**

Enzweiler, Gavrila, Monocular Pedestrian Detection: Survey and Experiments Trans on PAMI 2009

Cascade classifiers Biblio……

Ref.	Object	feature	weak cl.	strong cl.	# layers	training set	avg. # windows
Lienhart *et al.* (2003)	Face	enh. HLF	stumps or CART	AdaBoost/ Gentle Ad.	10-20	MIT+CMU[1] (PT=5000, NT=3000)	≃426K* (ps=1,ss=1.1) ≃243K* (ps=1,ss=1.2)
Viola & Jones (2004)	Face	HLF	single perceptron	AdaBoost	32	MIT+CMU[1] (PT=4916, NT=10000)	≃145K* (ps=1.5,ss=1.25)
Lehmann *et al.* (2004)	Face and other	HLF	semi-naive Bayes cl.	AdaBoost	32	MIT+CMU[1],UIUC[2],... (PT=UNK, NT=UNK)	UNK
Froba-Ernst (2004)	Face	mLBP	single perceptron	AdaBoost	4	MIT+CMU[1],BioID[3] (PT=6000, NT=2000)	≃197K* (ps=1,ss=1.25)
Wu *et al.* (2005)	Face	HLF	single perc.+FDA	AdaBoost/ AsymmB.	21-22	MIT+CMU[1] (PT=5000, NT=5000)	≃197K* (ps=1,ss=1.25)
Zhu *et al.* (2006)	Pedes.	variable size HOG	linear SVM	AdaBoost	30	INRIA[5] (PT=2418)	12800 (ps=8*,ss=1.25*)
Zhang *et al.* (2007)	Pedes. and other	HOG	linear SVM	n/a	4 resol. 4* scales	INRIA[5],VOC 2006[6] (PT=1208, NT=563)	UNK
Verschae *et al.* (2008)	Face and eyes	HLF and and mLBP	stumps	AdaBoost	10	MIT-CMU[1], others, own (PT=5000, NT=3600, PV=flip PT, NV=1500)	≃230K* (ps=6-3-1,ss=1.2)
Pais. *et al.* (2008)	Pedes.	LRF,HOG, COV.MAT.	single perceptron	AdaBoost/ SVM	20-29	INRIA[5], [16] (PT=2419+4800, NT=UNK+5000)	17280 (ps=4,ss=1.25)
Tuzel *et al.* (2008)	Pedes.	Cov.Mat.	Linear logistic regressor	LogitBoost	30	INRIA[5], [16] (PT=2419+4800, NT=10000+5000)	≃28K* (ps=6,ss=1.2)
Zhang *et al.* (2009)	Face	SRF	single perceptron	AdaBoost	17	MIT+CMU[1],others (PT=4858, PV=511)	≃550K* (ps=1,ss=1.25)
Hota *et al.* (2010)	Car	HLF+HOG	single perceptron	AdaBoost/ SVM	UNK	MIT CBCL[7] (PT=516, NT=500)	UNK

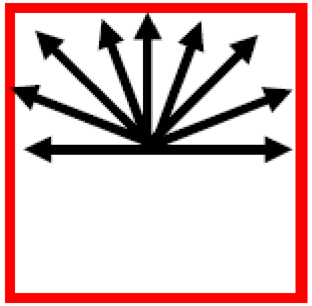
Histograms of Oriented Gradients

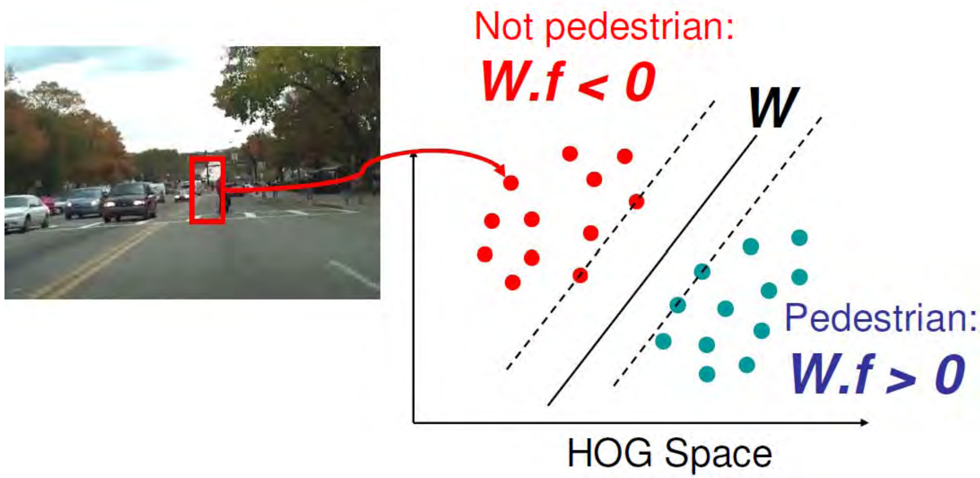
Scan Input Image

Resize window
to 64x128

Divide into overlapping
16x16 blocks

Compute histogram of
gradient over 9 directions

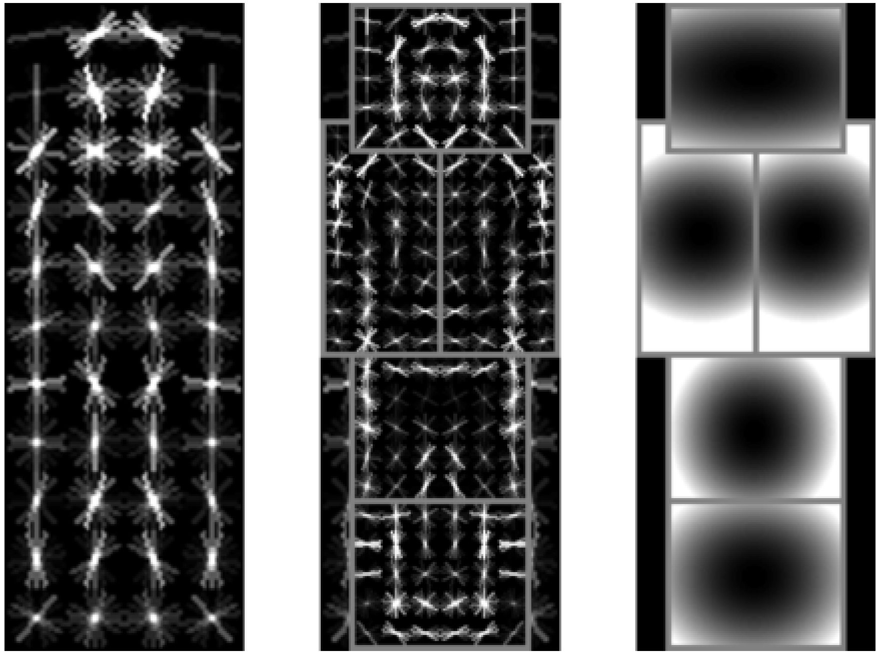




Slides: courtesy of
www.andrew.cmu.edu

Hog + Deformable Part Model

*P. Felzenszwalb, D. McAllester, D. Ramanan “**Object Detection with Discriminatively Trained part based Models**” ,TPAMI-2009*



(a) (b) (c)

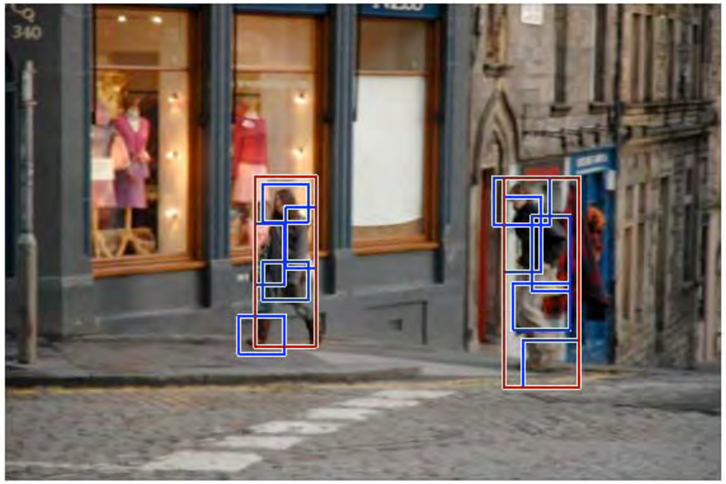
This work exploits the same HOG feature of Dalal et al.

The model of the target object is made of:

a) a coarse root filter (it corresponds to Dalal model of pedestrian)

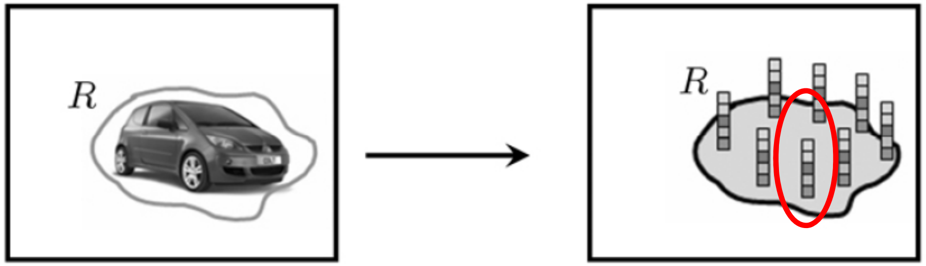
b) several higher resolution part filters

c) a spatial model for the location of each part relative to the root



Covariance and LogitBoost Classifier

*O. Tuzel, F. Porikli, and P. Meer, “**Pedestrian detection via classification on riemannian manifolds**,” IEEE Trans. on PAMI, Oct. 08*



F is a set of pixel-wise features.

For generic region matching:

$$F(x,y) = \begin{bmatrix} x & y & R(x,y) & G(x,y) & B(x,y) & |I_x(x,y)| & |I_y(x,y)| & |I_{xx}(x,y)| & |I_{yy}(x,y)| \end{bmatrix}^T$$

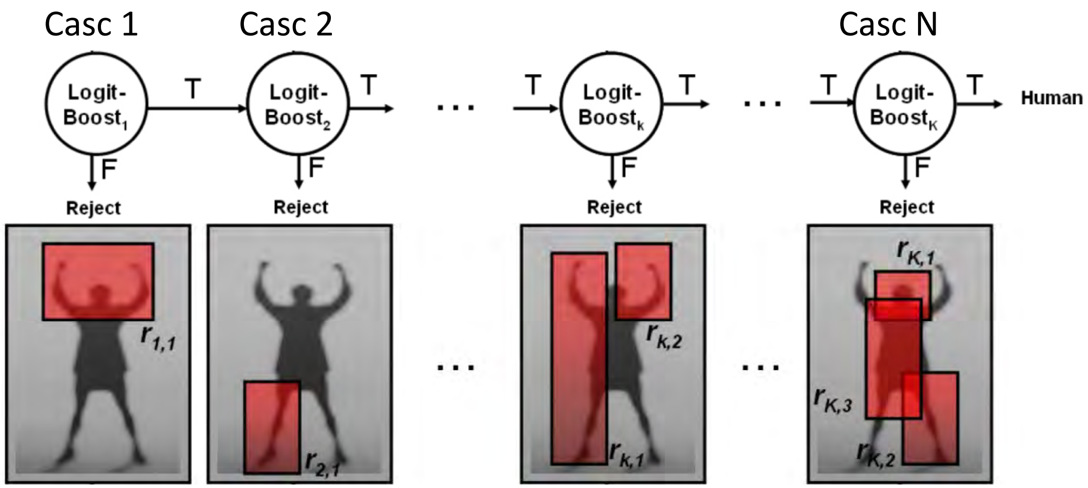
For texture classification:

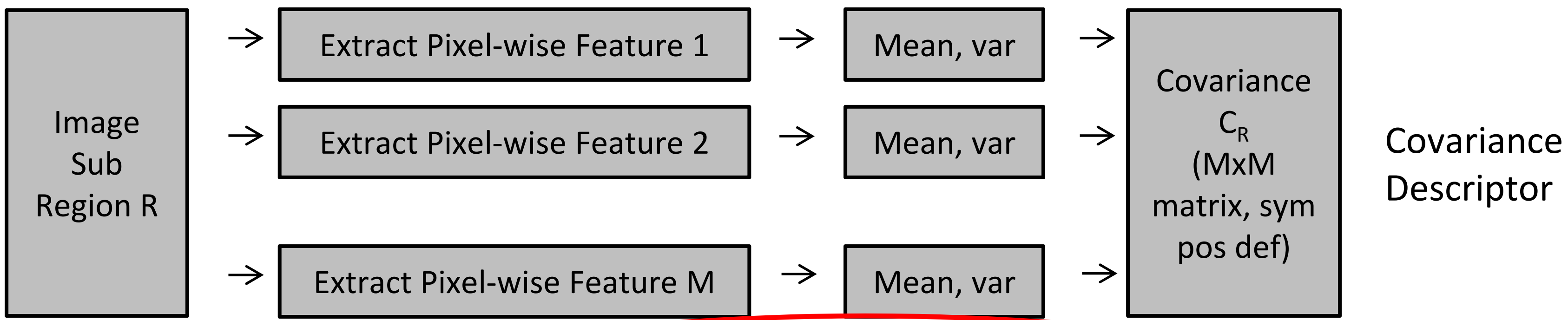
$$F(x,y) = \begin{bmatrix} I(x,y) & |I_x(x,y)| & |I_y(x,y)| & |I_{xx}(x,y)| & |I_{yy}(x,y)| \end{bmatrix}^T$$

For pedestrian detection:

$$F(x,y) = \begin{bmatrix} x & y & |I_x| & |I_y| & \sqrt{I_x^2 + I_y^2} & |I_{xx}| & |I_{yy}| & \arctan \frac{|I_x|}{|I_y|} \end{bmatrix}^T$$

Covariance and LogitBoost Classifier

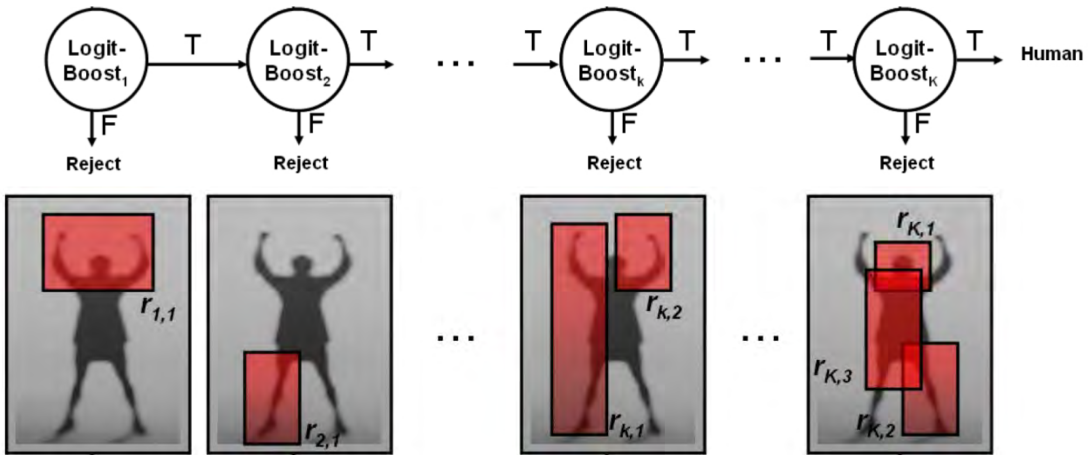




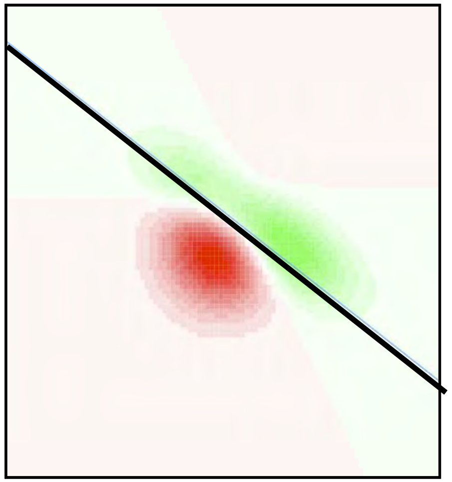
$$\left[x, y, |I_x|, |I_y|, \sqrt{I_x^2 + I_y^2}, |I_{xx}|, |I_{yy}|, \arctan\frac{|I_y|}{|I_x|} \right]^T$$

LogitBoost Classifier on Riemannian Manifolds

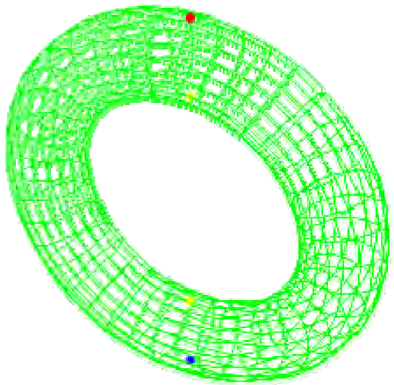
Casc 1 Casc 2 Casc N



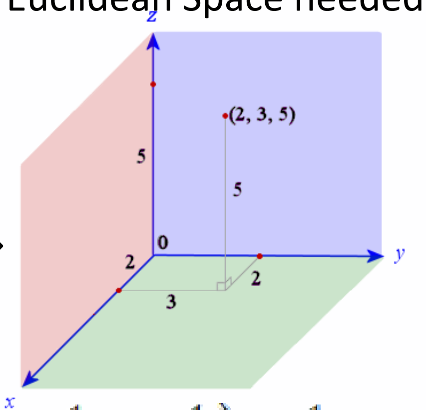
Linear Logistic Regressor



on Riemannian Manifolds



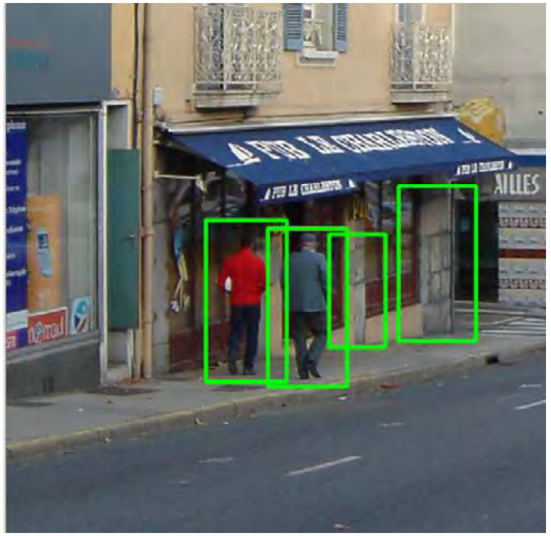
Euclidean Space needed



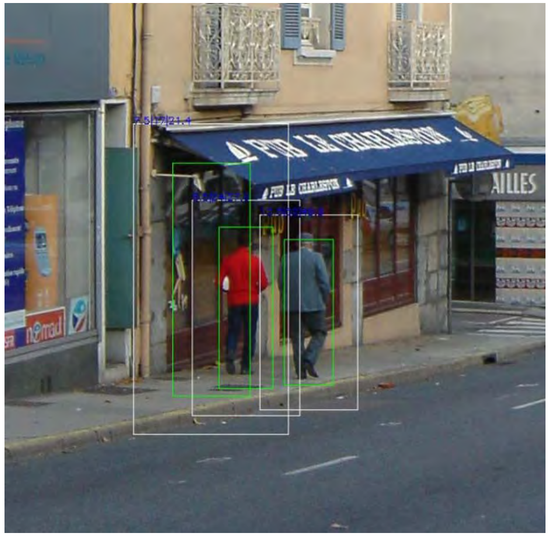
$$\log_{\mu}(Y) = \mu^{\frac{1}{2}} \log\left(\mu^{-\frac{1}{2}} Y \mu^{\frac{1}{2}}\right) \mu^{\frac{1}{2}}$$

Examples

Dalal



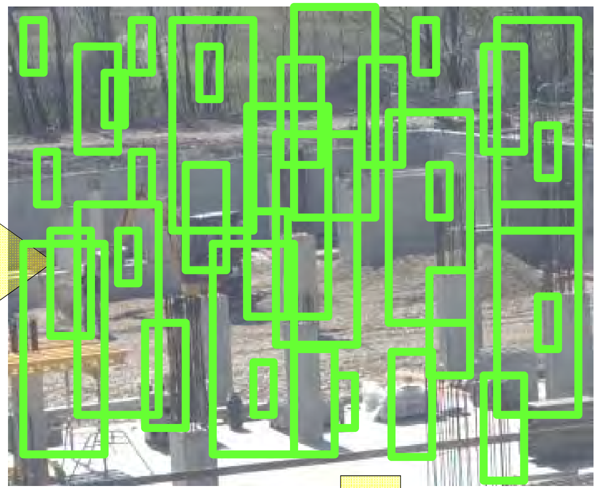
Tuzel



Felzenszwalb

Ped. Detection with Sliding Window

On each frame



Apply pedestrian classifier on each window

Exhaustive sliding window approach

Ped. Detection with Sliding Window

- Two problems

1. Accuracy
   - Many false positives
   - Localization errors

2. Computation time
   - Proportional to sliding windows size and overlap
   - Higher in Riemannian manifold

Two approaches

Exploiting other cues

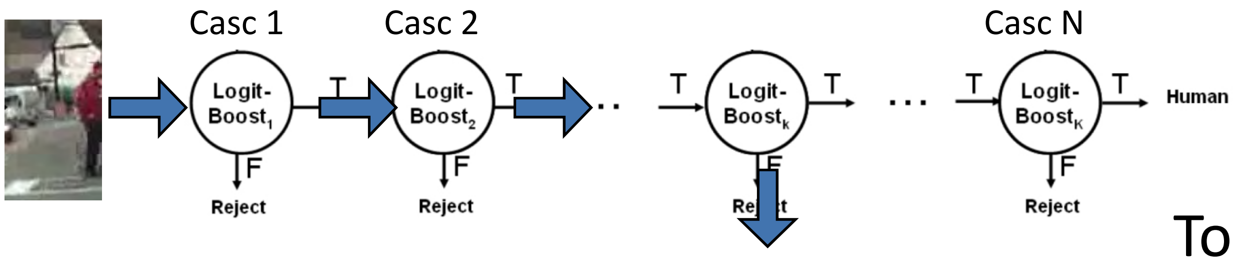
1) Learning context
2) Using relevance feedback

Exploiting statistics

1) Learning distribution
2) Multi-stage search

1) Using context and relevance feedback

LogitBoost Classifier on Riemannian Manifolds



detection on negatives (precision) increases
detection on positives (recall) decreases

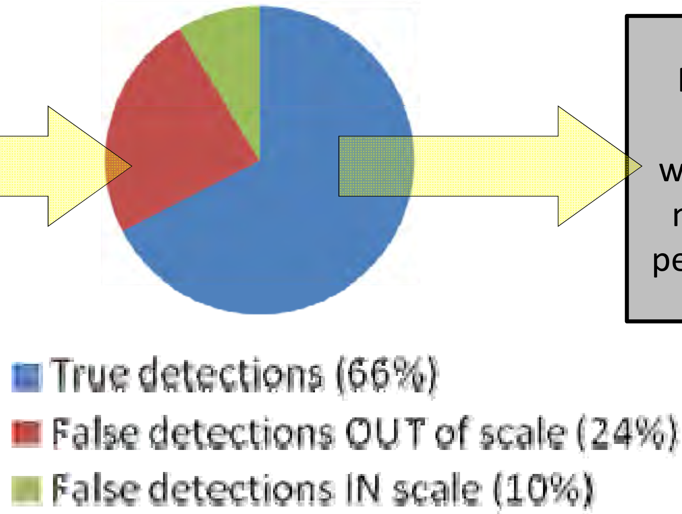
To increase precision without affecting recall..

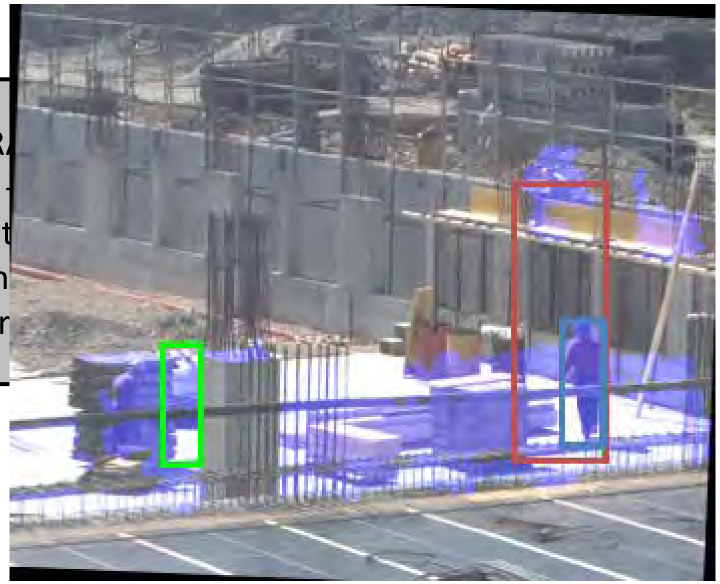
Estimate the Size of Pedestrian within the Frame

G.Gualdi, A. Prati, R. Cucchiara, "Covariance Descriptors on Moving Regions for Human Detection in Very Complex Outdoor Scenes" in *ACM/IEEE ICDSC* 2009

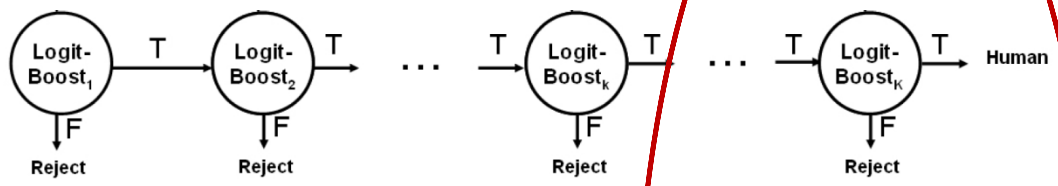
Exploit **Pedestrian detector** to infer pedestrian size in the frame

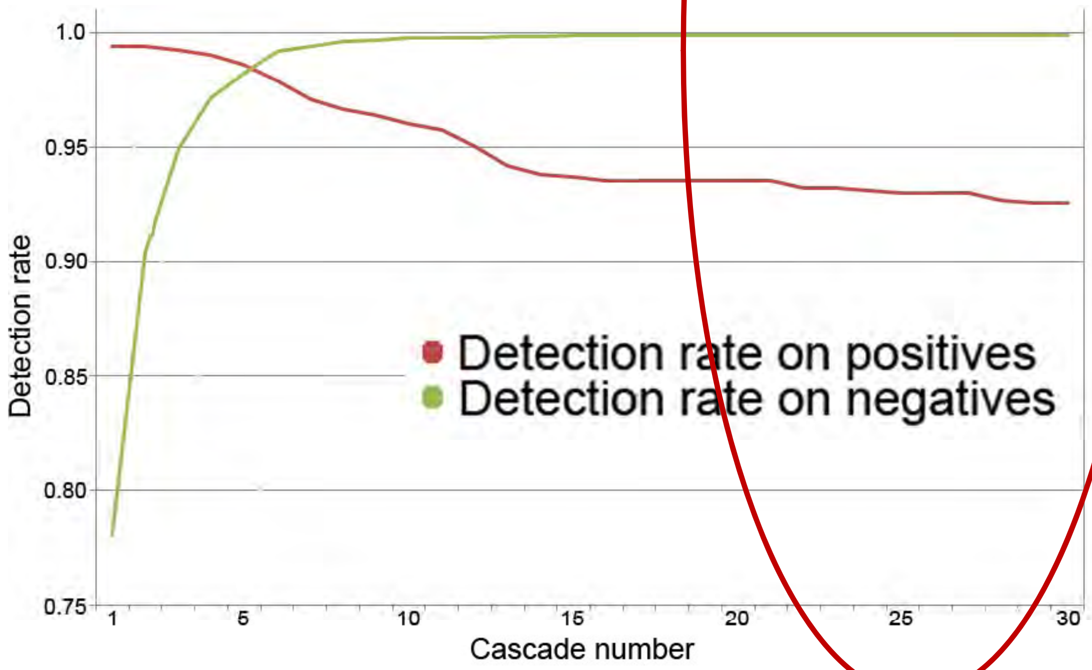
Sliding Window Pedestrian Detector

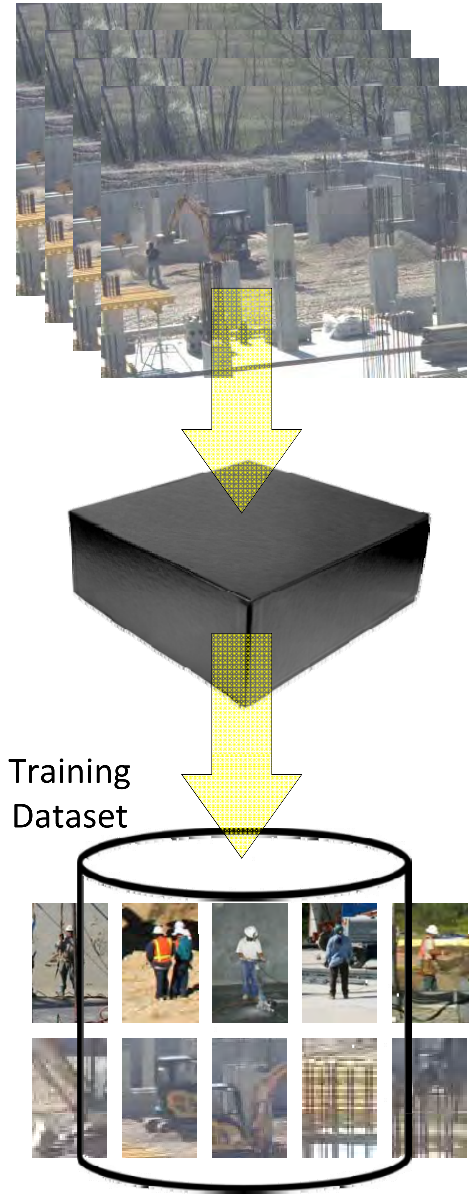




Relevance Feedback (1/2)







Relevance Feedback (2/2)

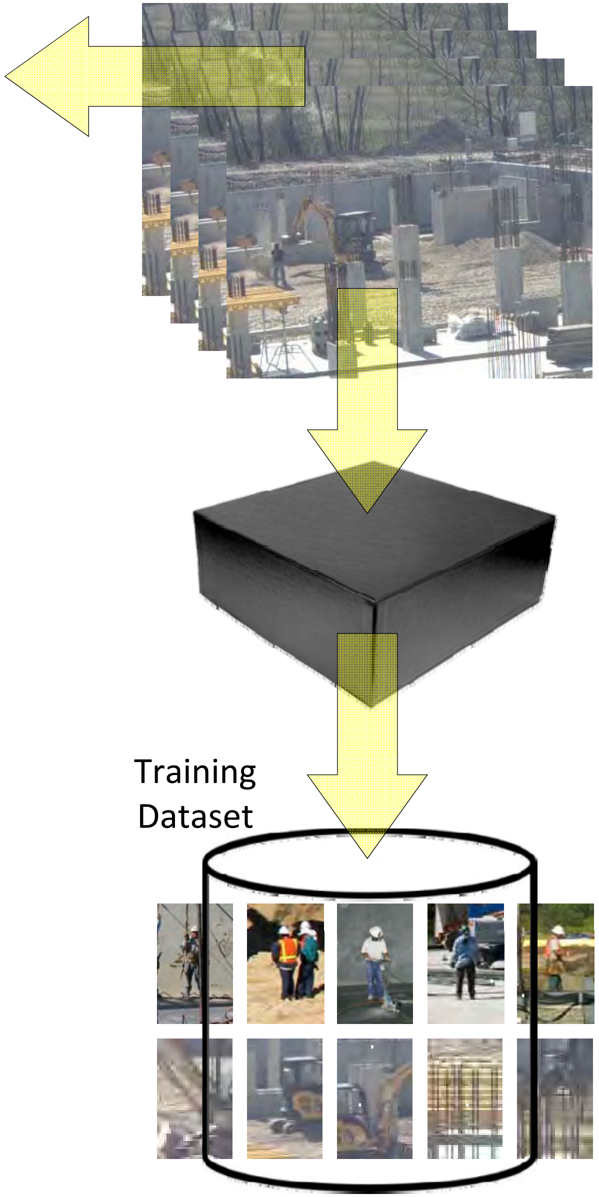
RELEVANCE
FEEDBACK

(1)
IMPLICIT
(it is automatic)

Estimate background images
=> negative training set

(2)
EXPLICIT
(it needs user assessment)

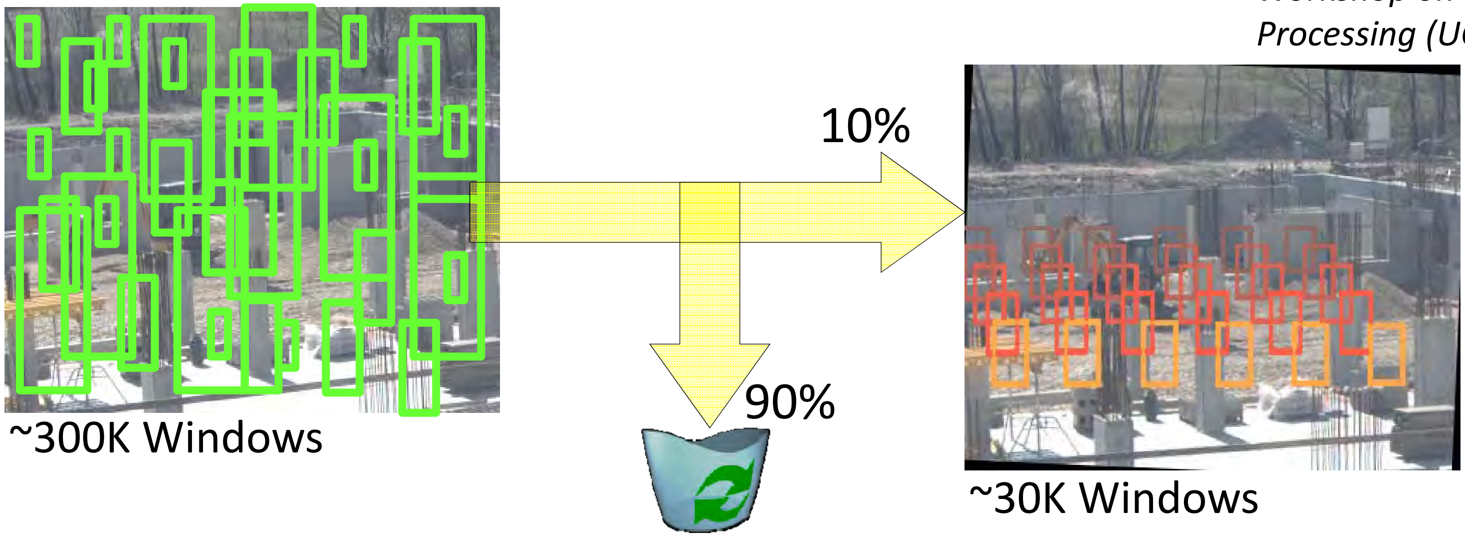
Response of the Ped Classifier

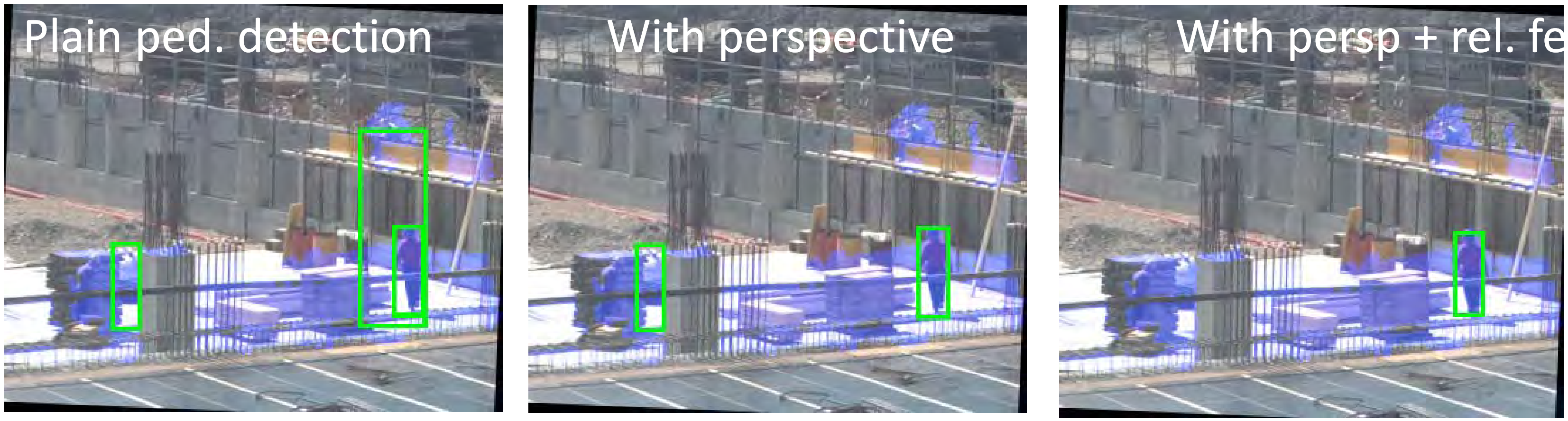


Experimental Results

Perspective Estimation

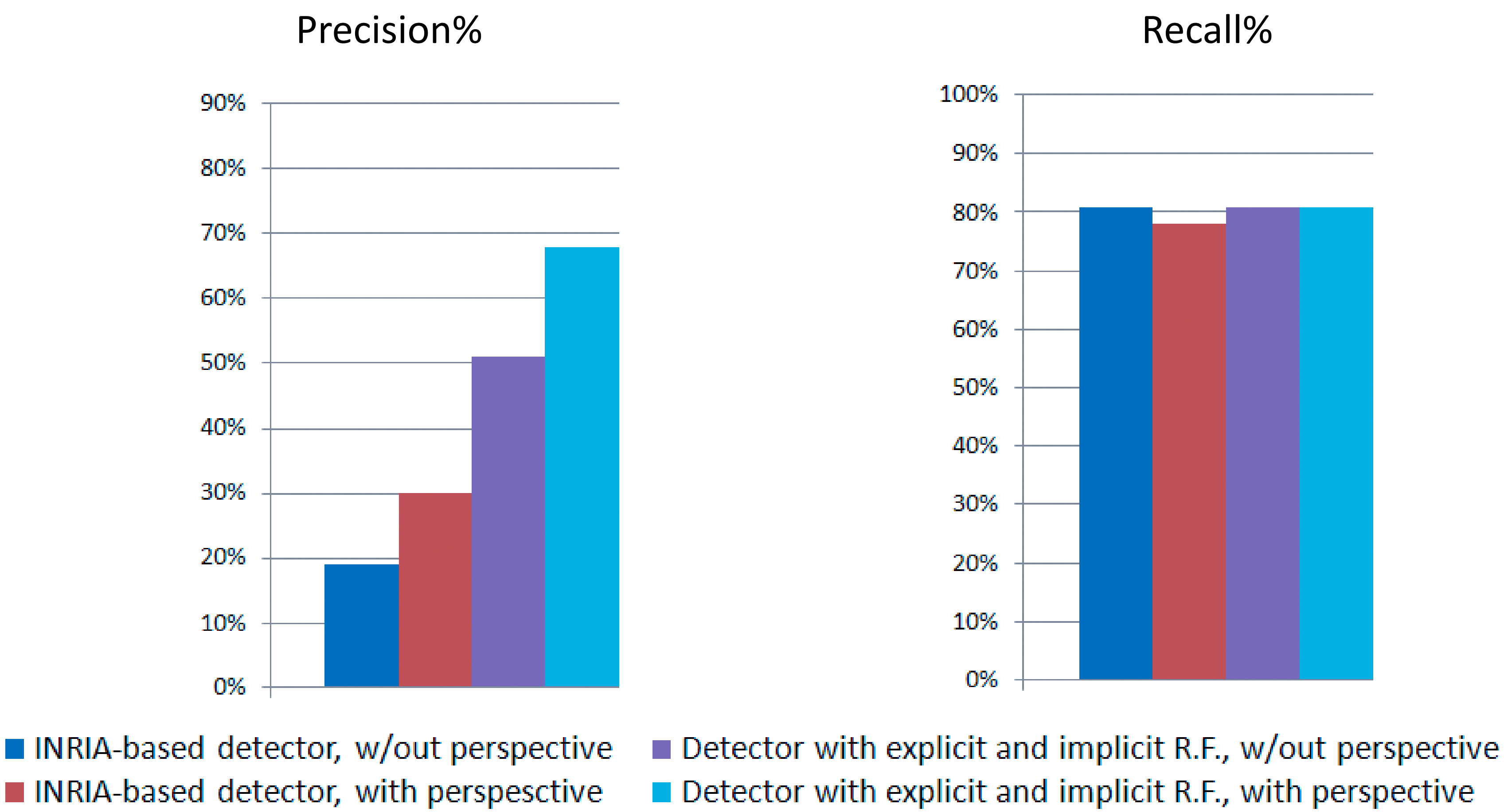
G. Gualdi, A. Prati, R. Cucchiara,"Perspective and Appearance Context for People Surveillance in Open Areas" in *Proceedings of the 2nd International Workshop on Use of Context in Video Processing (UCVP 2010)*, at CVPR 2010





Experimental Results

Perspective Estimation and Relevance Feedback



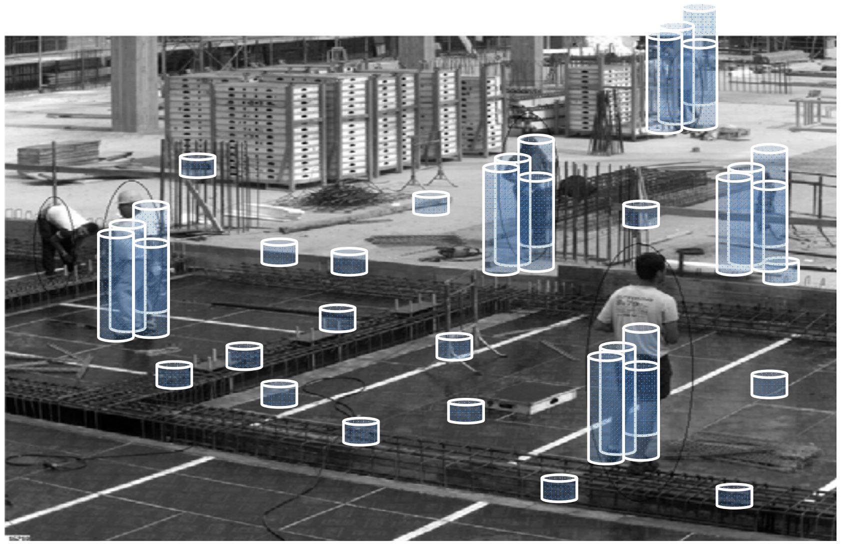
2) Use statistics: multi-stage particle windows

A probabilistic bayesian paradigm for object detection:

Detection is achieved with multi-stage search

Estimate obj. detection as a pdf

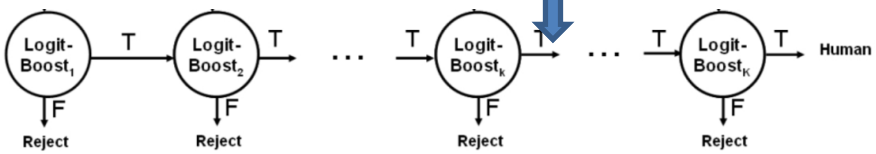
Use particle windows instead of sliding windows



G. Gualdi, A. Prati, R. Cucchiara,"Multi-stage Sampling with Boosting Cascades for Pedestrian Detection in Images and Videos" *ECCV 2010*

2) Use statistics: multi-stage particle windows

$$class\,(w, \mathbf{C}) = \begin{cases} \text{object} & \text{if } class\,(w, C_i) = \text{object}, \forall i = 1, \ldots, L \\ \text{non-object} & \text{if } \exists j \leq L | class\,(w, C_j) = \text{non-object} \end{cases}$$



Set $q_0\,(\boldsymbol{X}) = U\,(\boldsymbol{X})$

for i=1..m **do**

Draw N_i samples from $q_{i-1}\,(\boldsymbol{X})$

Assign a Gaussian kernel to each sample

Compute the measurement on each sample $s_i^{(j)}$

Obtain the measurement density function at step i:

$$p_i\,(\boldsymbol{Z}|\boldsymbol{X}) = \sum \pi_i^{(j)} \cdot \mathcal{N}\left(s_i^{(j)}, \Sigma_i^{(j)}\right)$$

Compute the new proposal distribution:

$$q_i\,(\boldsymbol{X}) = (1 - \alpha_i)\,q_{i-1}\,(\boldsymbol{X}) + \alpha_i \frac{p_i(\boldsymbol{Z}|\boldsymbol{X})}{\int p_i(\boldsymbol{Z}|\boldsymbol{X})d\boldsymbol{X}}$$

end for

The measure of each sample is a function of the **rejection level**: ***detection response***

$$R\,(w) = \frac{j_w}{L}$$

Region of support

- Often there is a basin of attraction in position and size

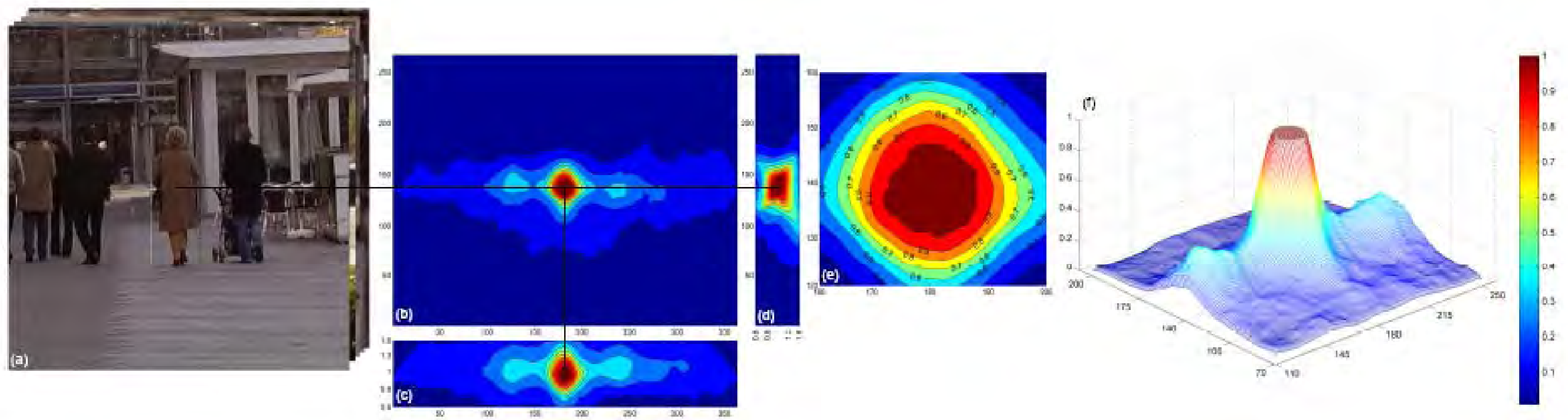


Fig. 1. Region of support and basin of attraction for the cascade of LogitBoost classifiers [21] trained on INRIA pedestrian dataset. (a) a sample taken from a set of 100 patches: there is a centered pedestrian of 50x150 pixels and an average of 2.5 other pedestrians at other positions and scales; (b-e) response of the classifier averaged on the 100 patches: (b) fixed scale w_s (equal to 50x150), sliding w_x, w_y; (c) fixed w_x (equal to x of patch center), sliding w_s and w_y; (d) fixed w_y (equal to y of patch center), sliding w_x and w_s; the central region of (b) is enlarged in (e) and plotted in 3D in (f).

Face detection

- Also with viola and jones face detector

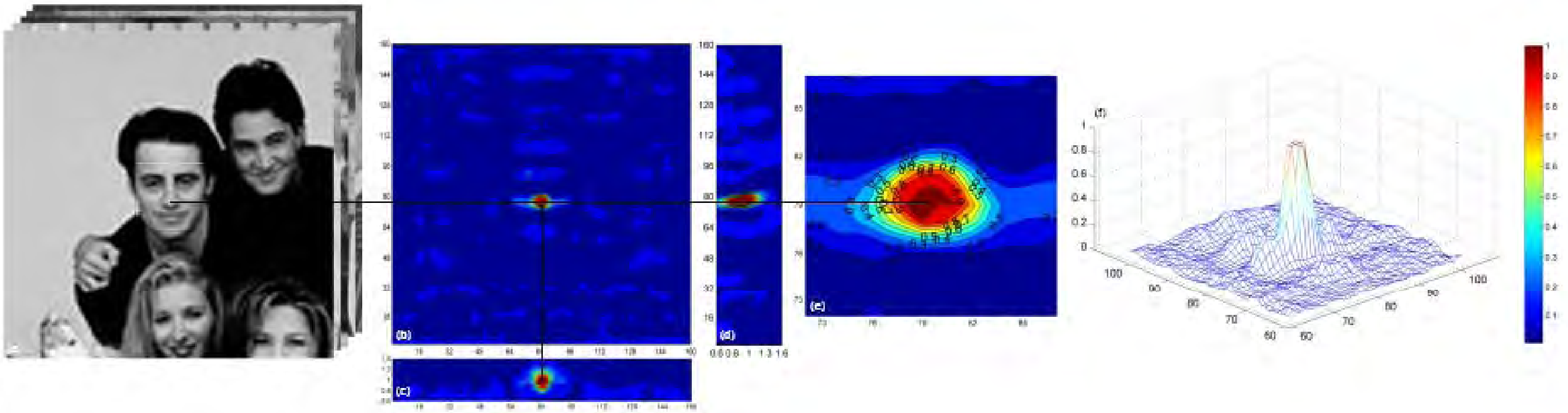
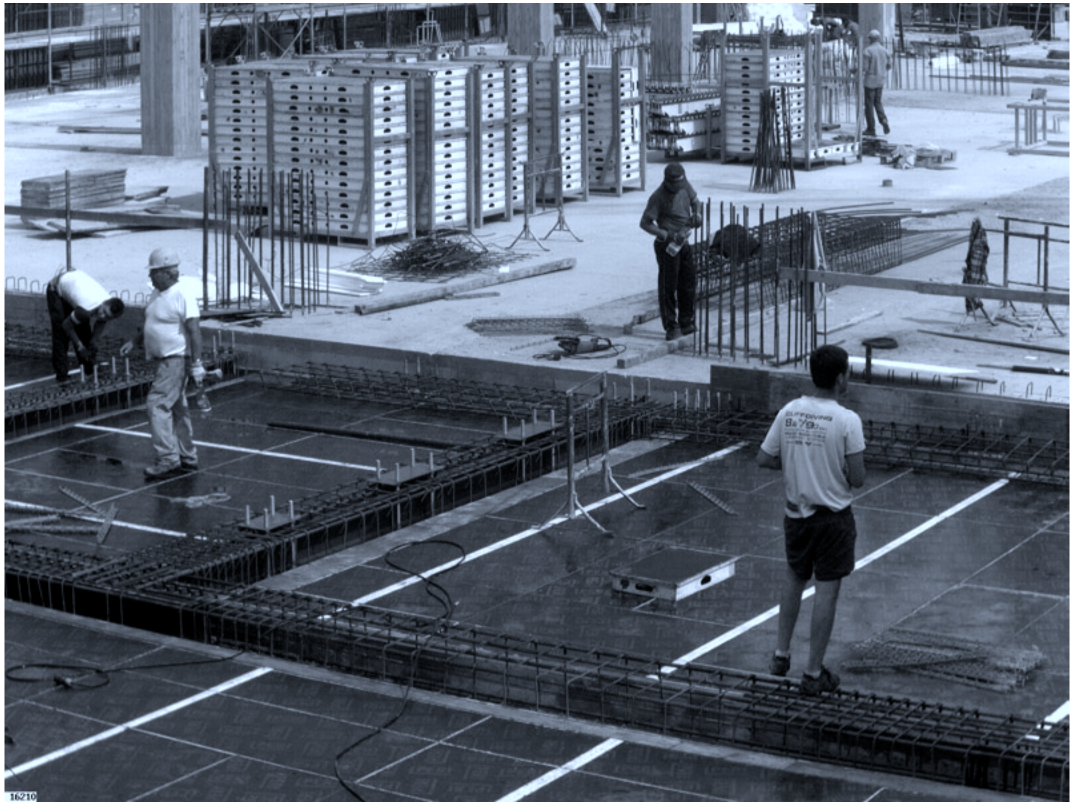


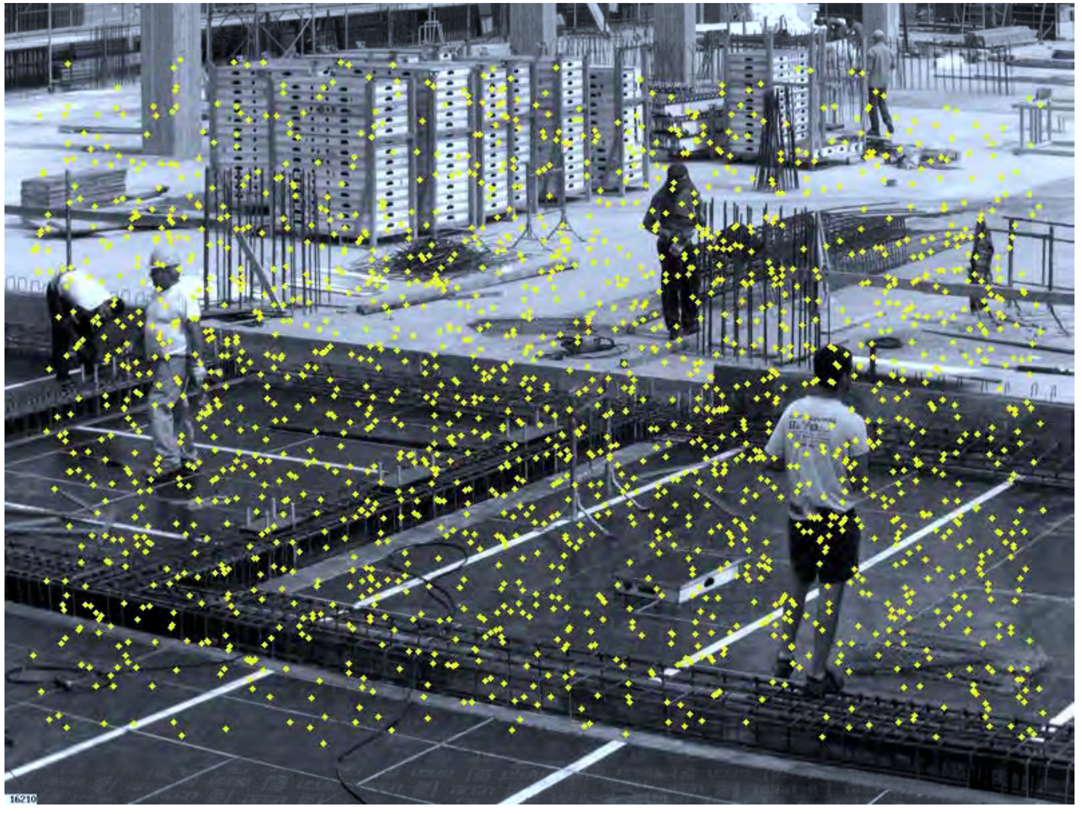
Fig. 2. Region of support and basin of attraction for the cascade of AdaBoost classifiers [33]. (a) a sample taken from a set of 100 patches from the MIT+CMU dataset: there is a centered face of 38x38 pixels and an average of 1.8 other faces at other positions and scales; (b-e) response of the classifier averaged on the 100 patches: (b) fixed scale w_s (equal to 38x38), sliding w_x, w_y; (c) fixed w_x (equal to x of patch center), sliding w_s and w_y; (d) fixed w_y (equal to y of patch center), sliding w_x and w_s; the central region of (b) is enlarged in (e) and plotted in 3D in (f).

Experiments with m=5 stages

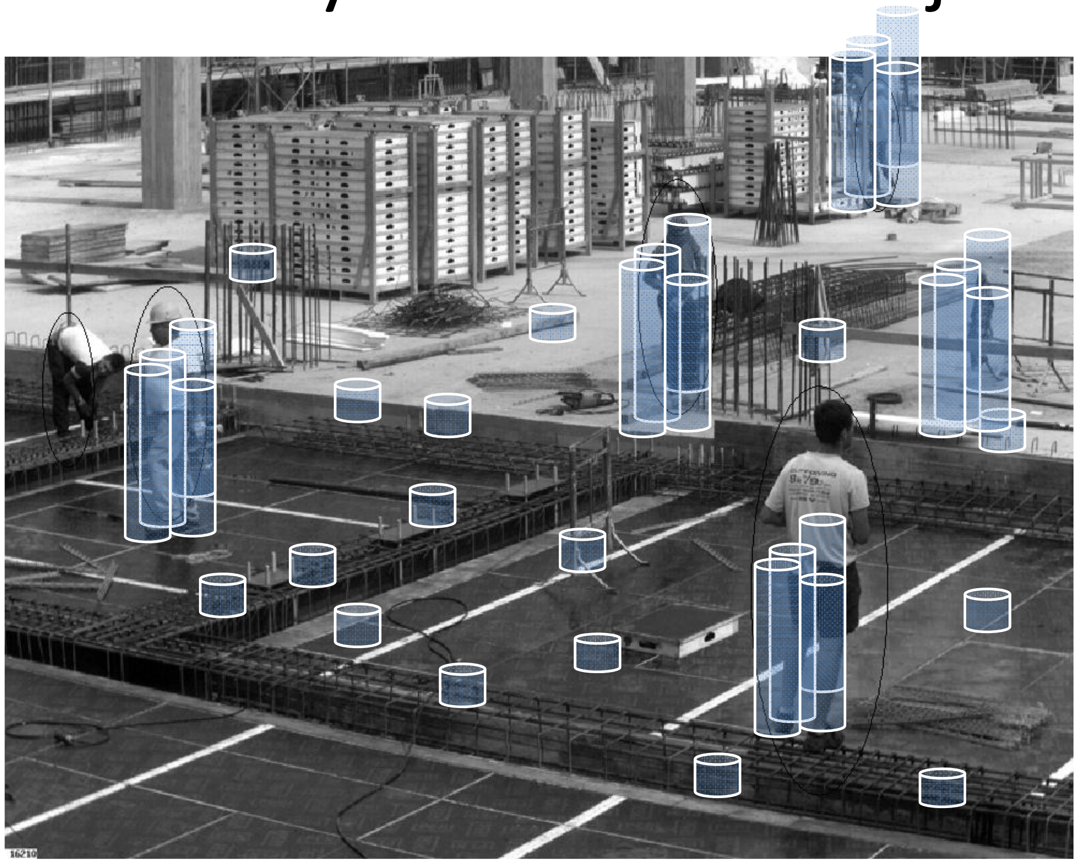
- Initial frame



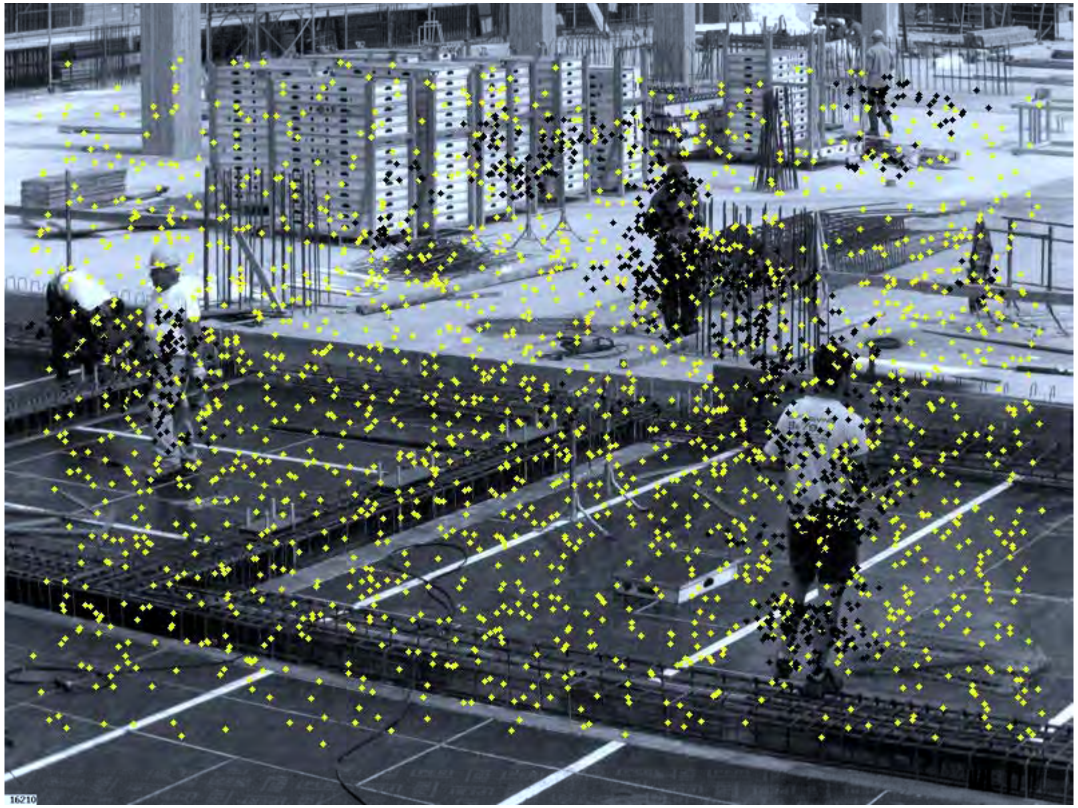
- Random initialization m=1



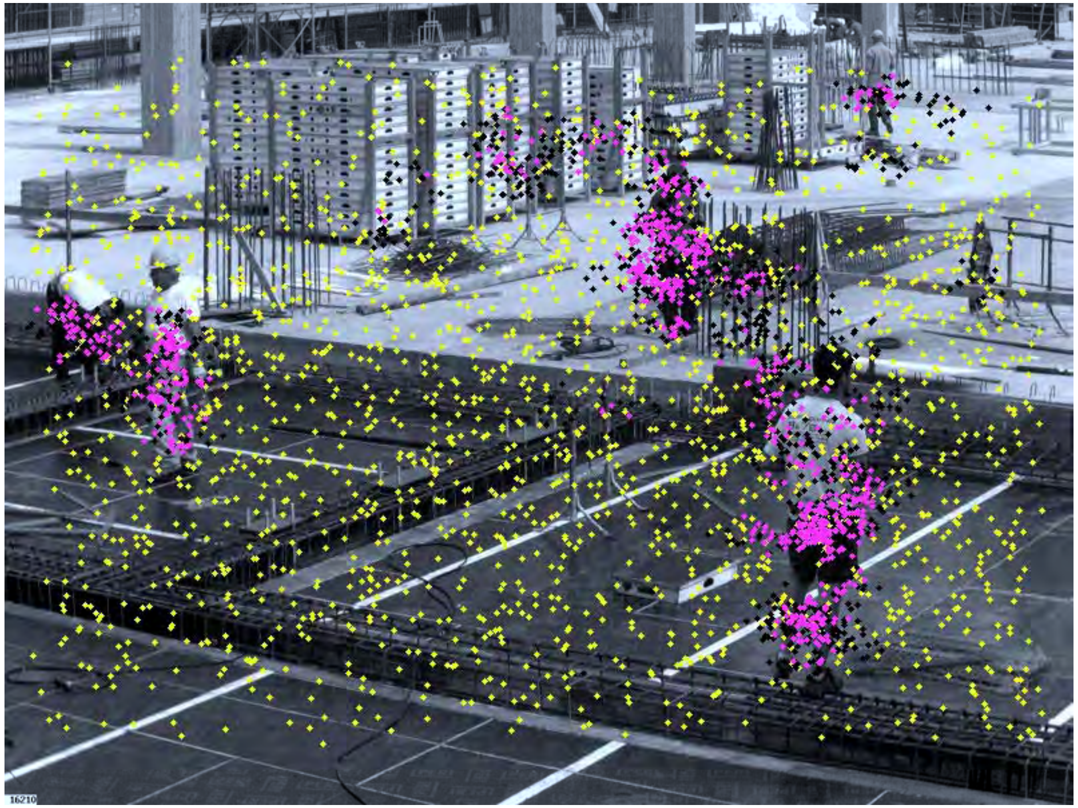
- Density function with rejection response R(x)



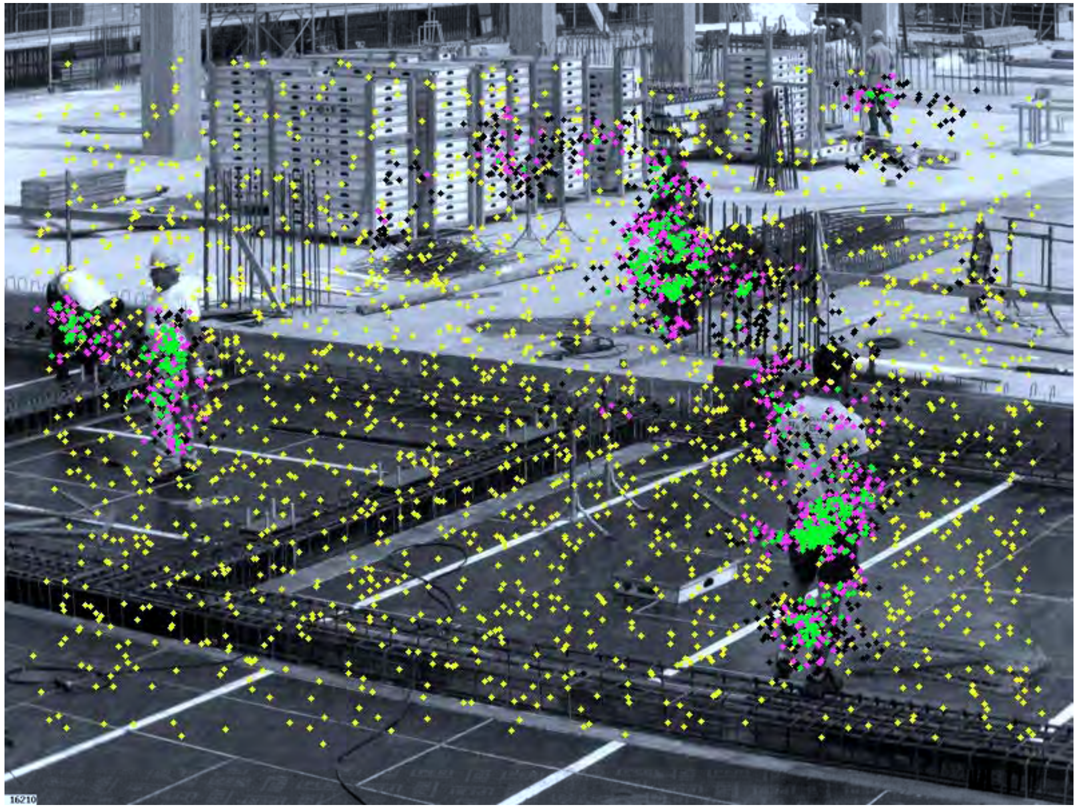
- Resampled particle windows m=2



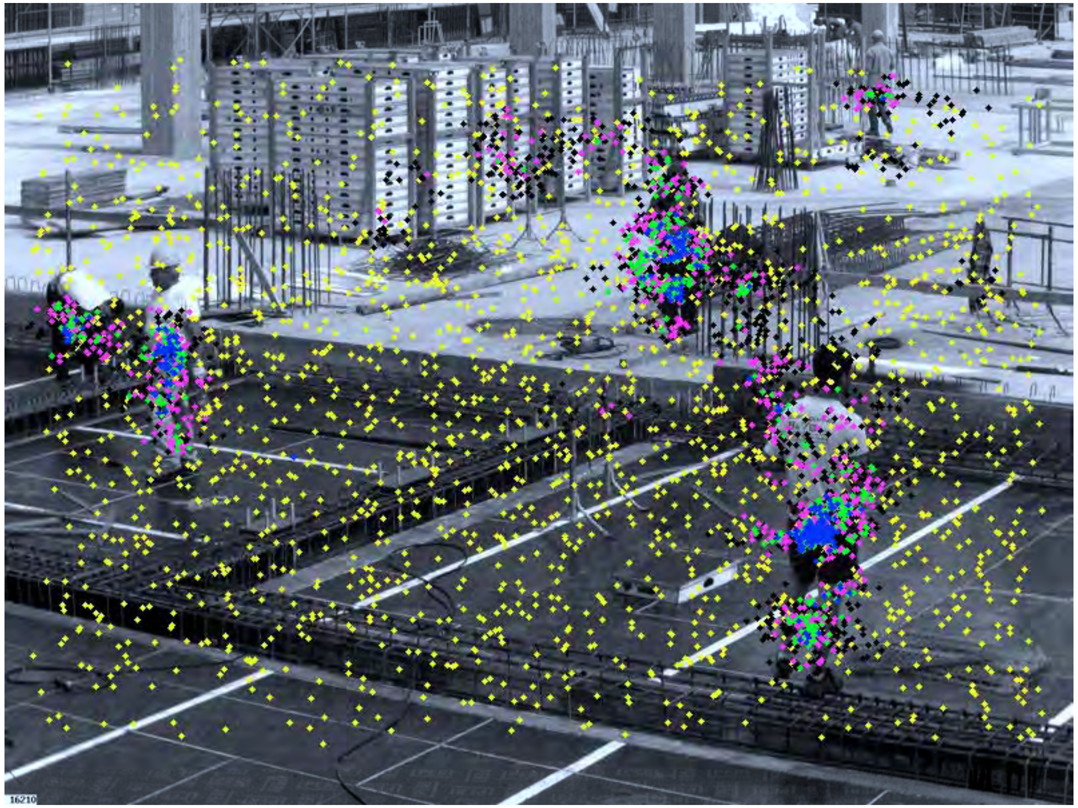
- Resampled particle windows m=3



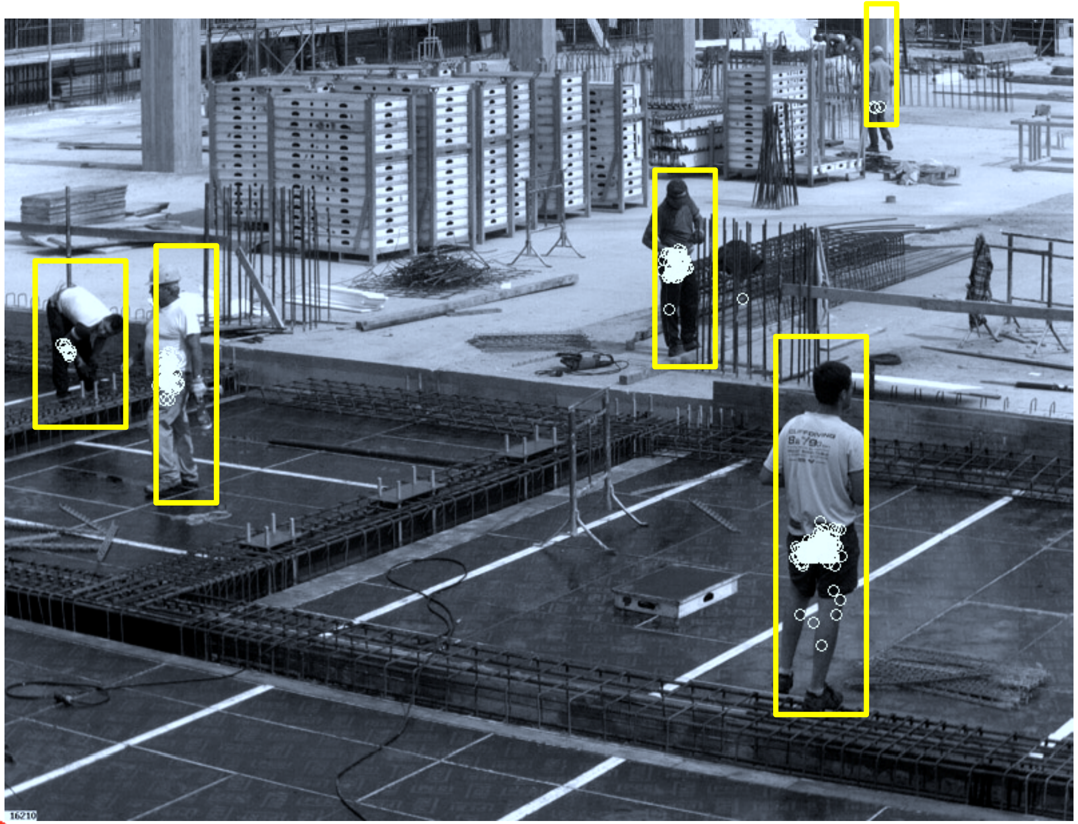
- Resampled particle windows m=4



- Resampled particle windows m=5



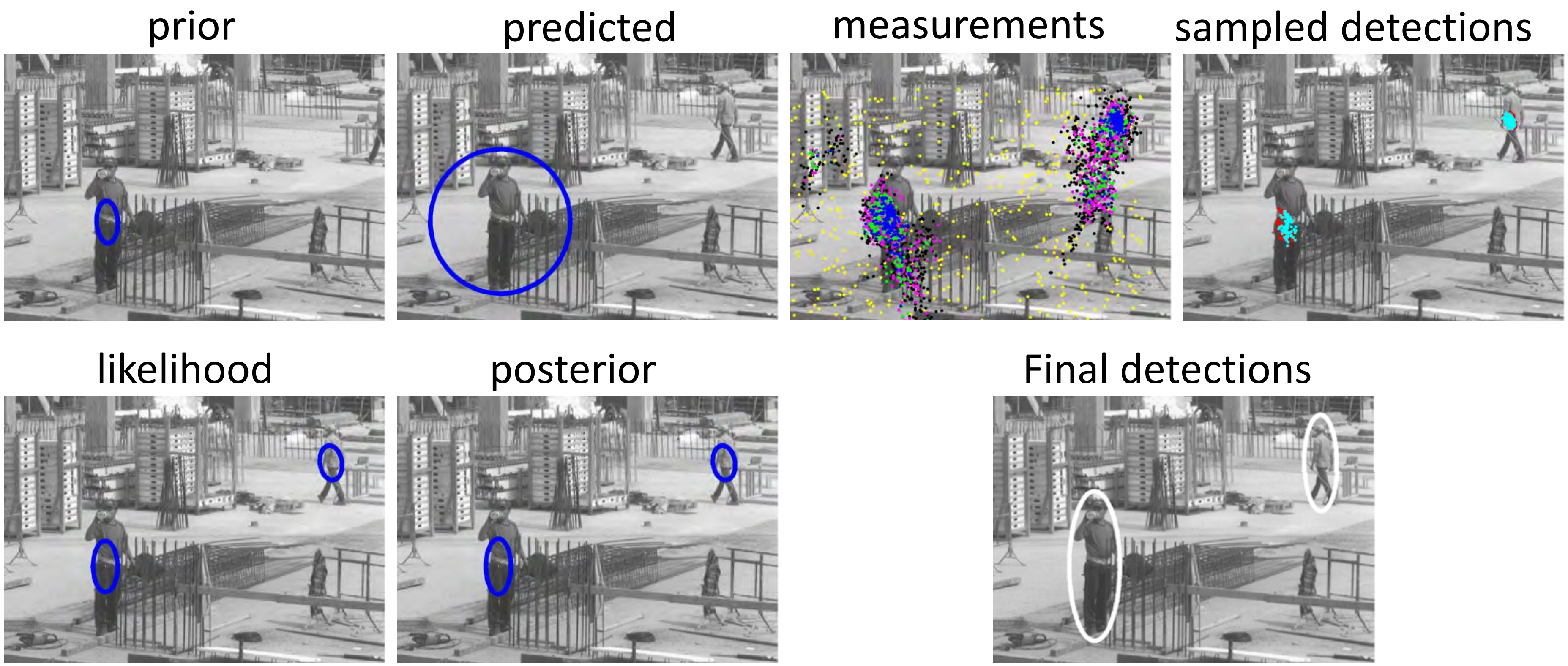
- Final people detection

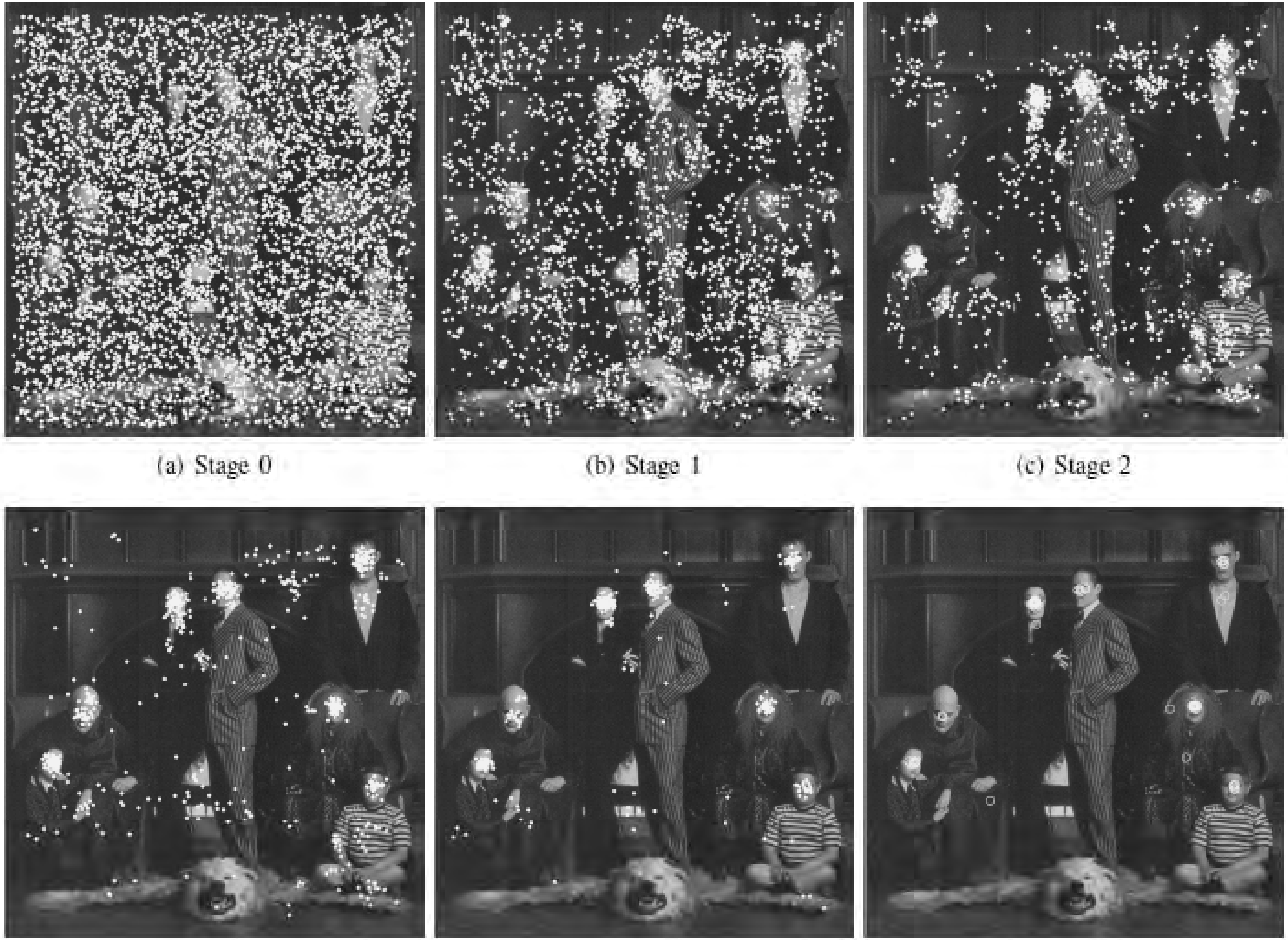


In Video: Exploit Bayesian Recursive Filter

$$p\left(\boldsymbol{X}_t|\boldsymbol{Z}_{1:t-1}\right) = \int p\left(\boldsymbol{X}_t|\boldsymbol{X}_{t-1}\right) p\left(\boldsymbol{X}_{t-1}|\boldsymbol{Z}_{1:t-1}\right) d\boldsymbol{X}_{t-1}$$

$$p\left(\boldsymbol{X}_t|\boldsymbol{Z}_{1:t}\right) = \frac{p\left(\boldsymbol{Z}_t|\boldsymbol{X}_t\right) p\left(\boldsymbol{X}_t|\boldsymbol{Z}_{1:t-1}\right)}{\int p\left(\boldsymbol{Z}_t|\boldsymbol{X}_t\right) p\left(\boldsymbol{X}_t|\boldsymbol{Z}_{1:t-1}\right) d\boldsymbol{X}_t}$$



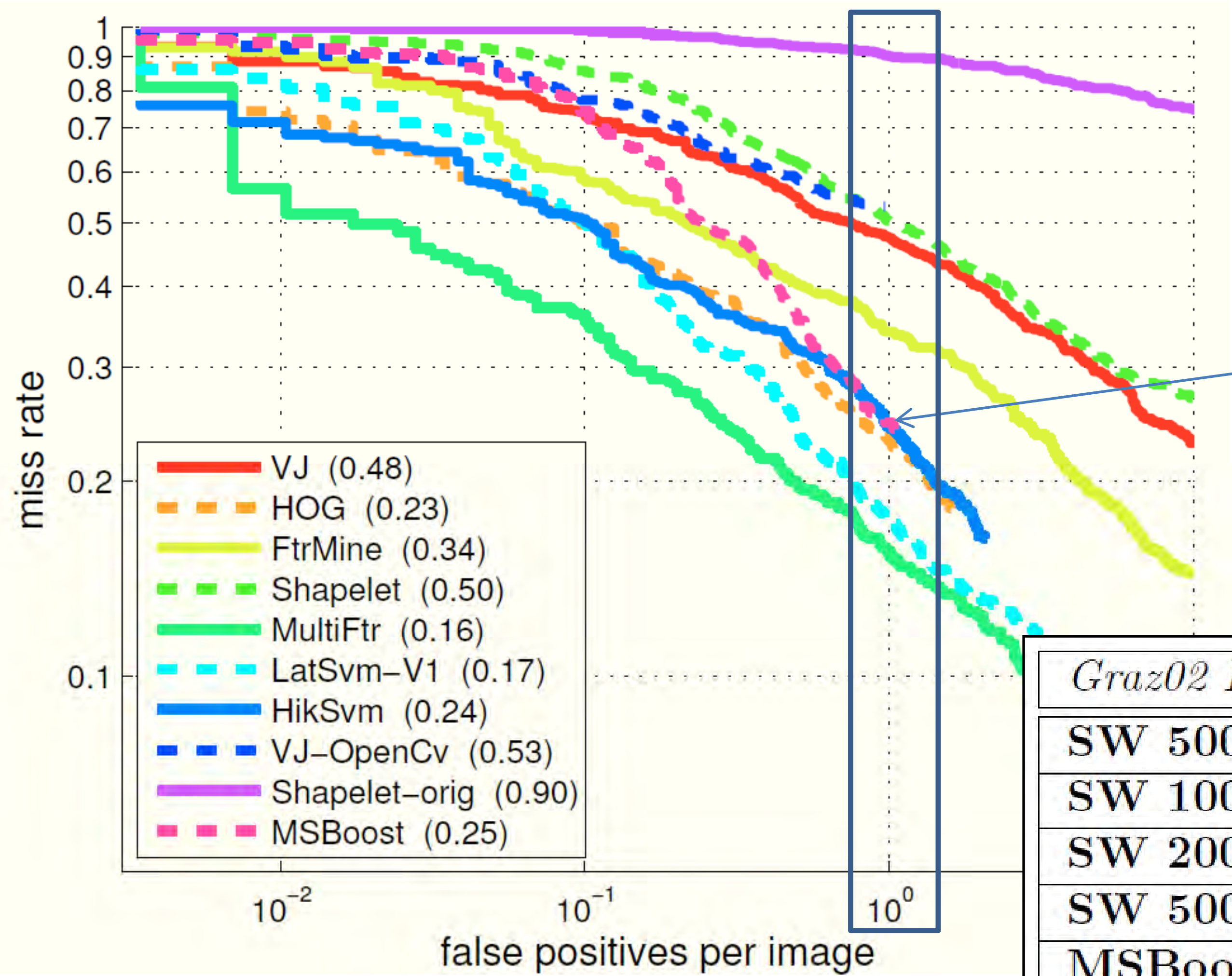


(a) Stage 0 (b) Stage 1 (c) Stage 2

(d) Stage 3 (e) Stage 4 (f) Detections

Experimental Results

Miss Rate vs False Positives Per Image



Dollar, P.; Wojek, C.; Schiele, B.; Perona, P.; **Pedestrian detection: A benchmark CVPR 2009**

Tuzel
And Our solution

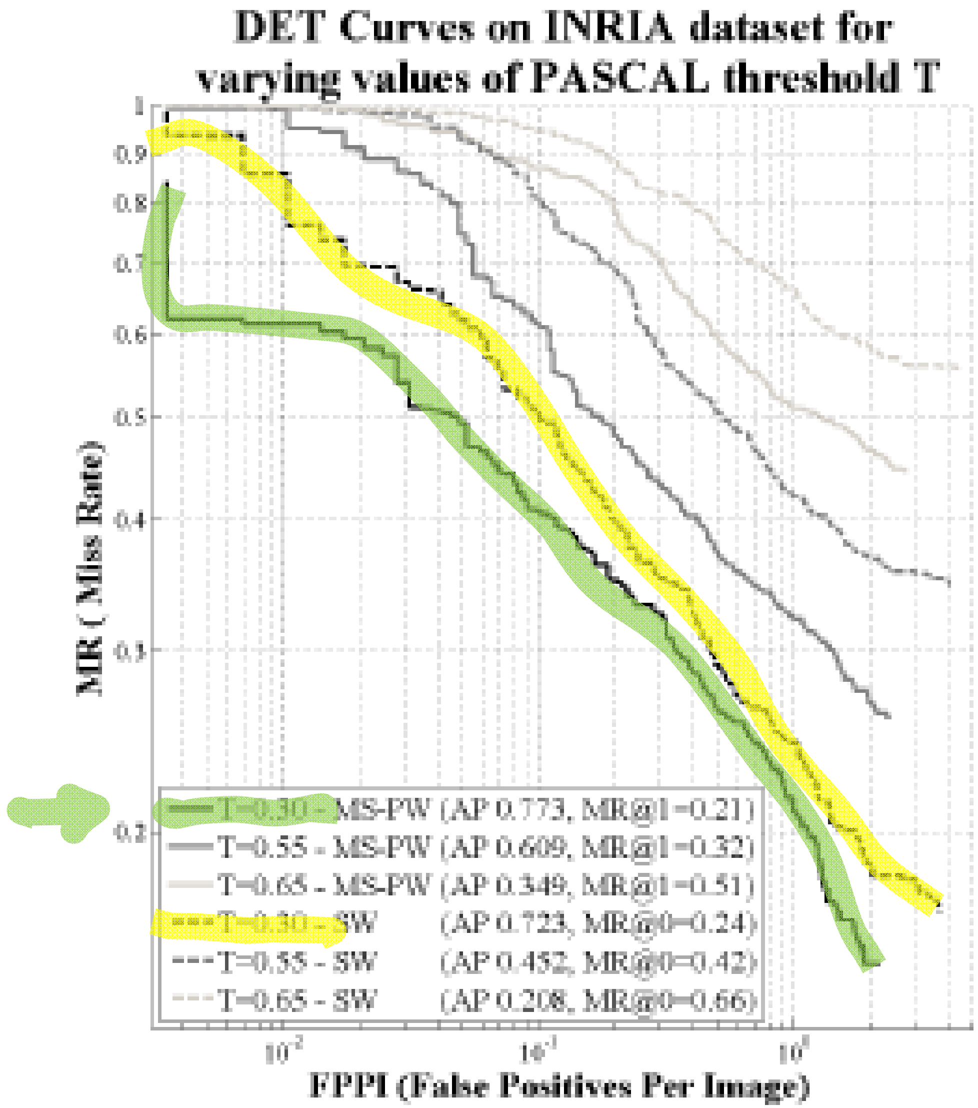
Graz02 DataSet	FPPI	MR
SW 5000 wnd	0.39	0.76
SW 10000 wnd	0.66	0.57
SW 20000 wnd	1.08	0.46
SW 50000 wnd	1.66	0.37
MSBoost 5000 p.	0.74	0.43

Experiments

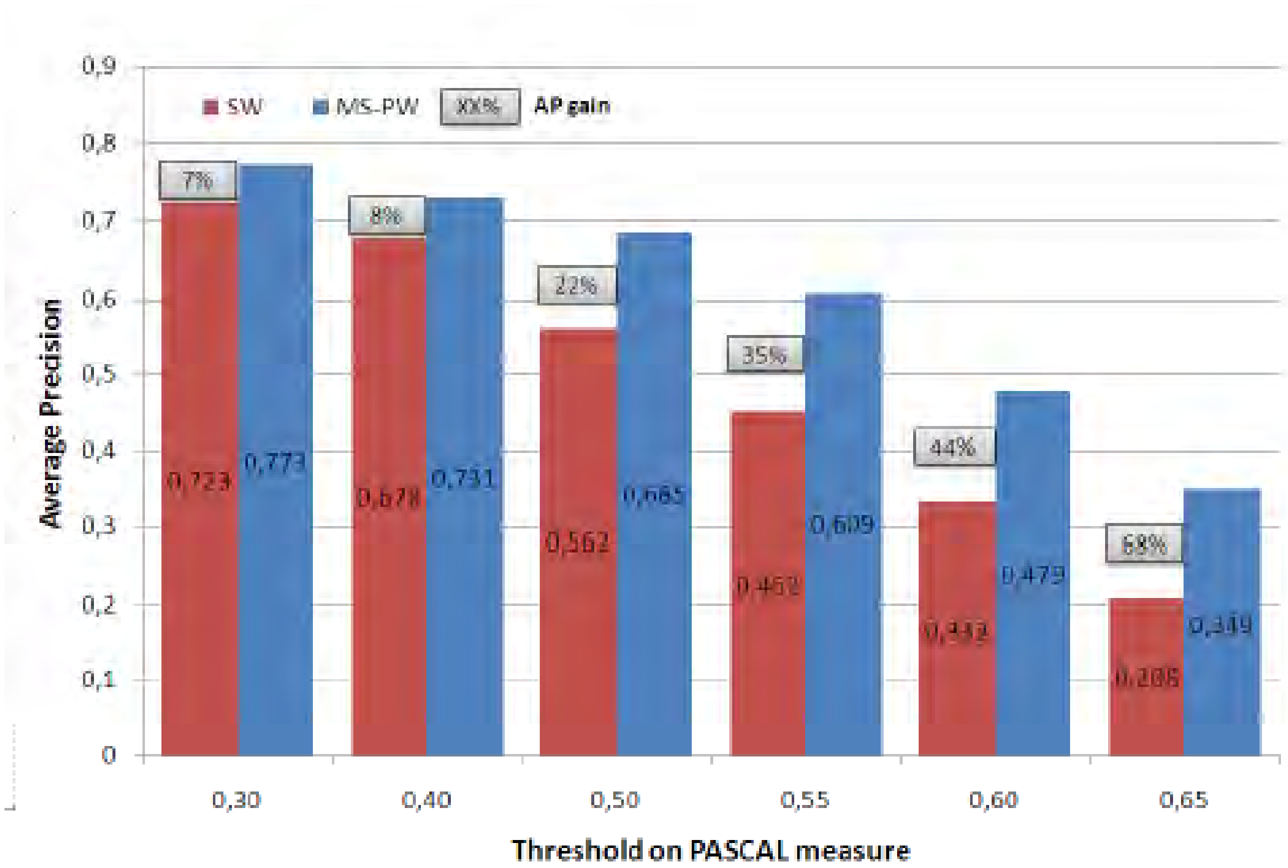
- Save time

- Or same time and

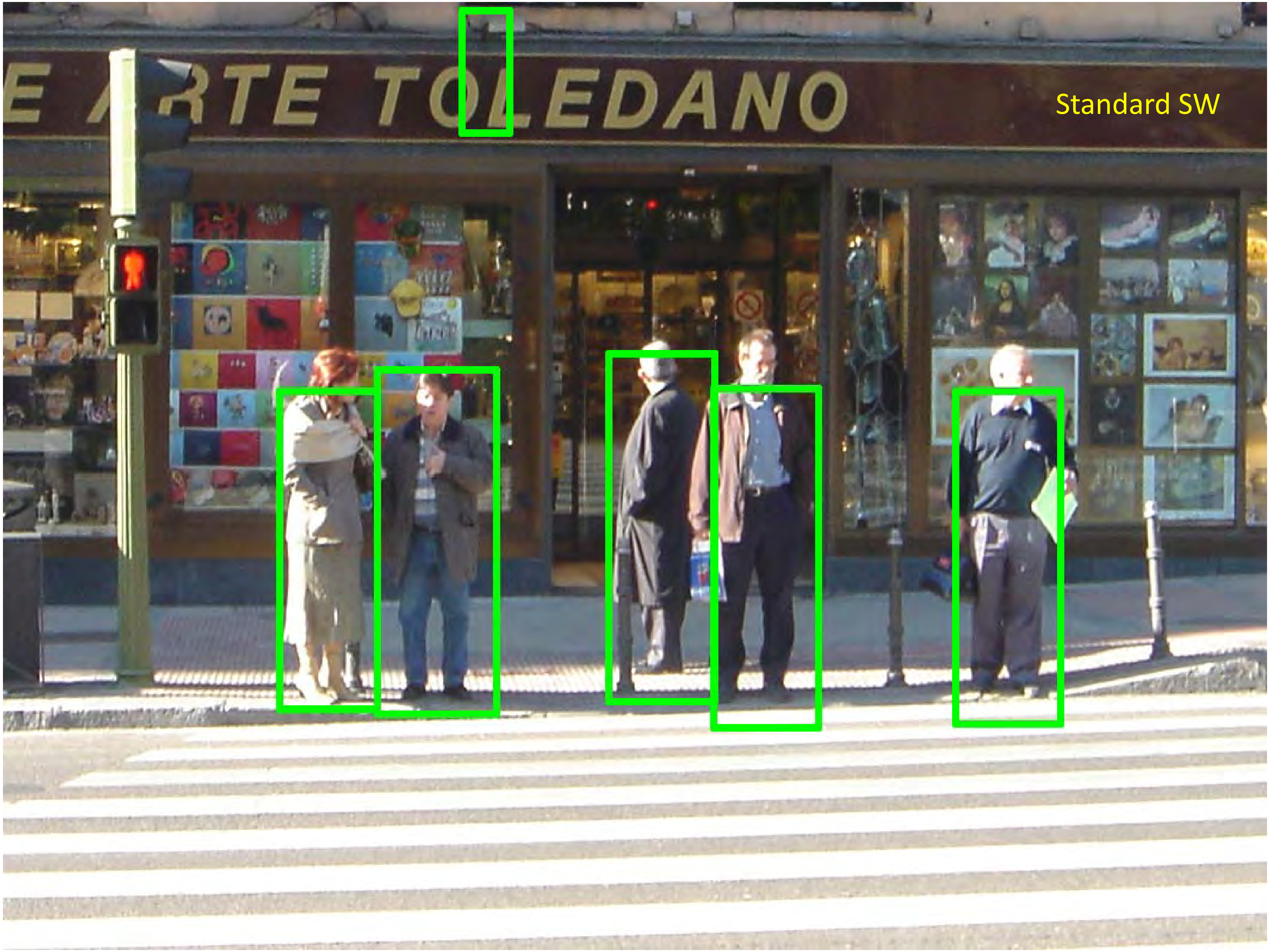
Better accuracy

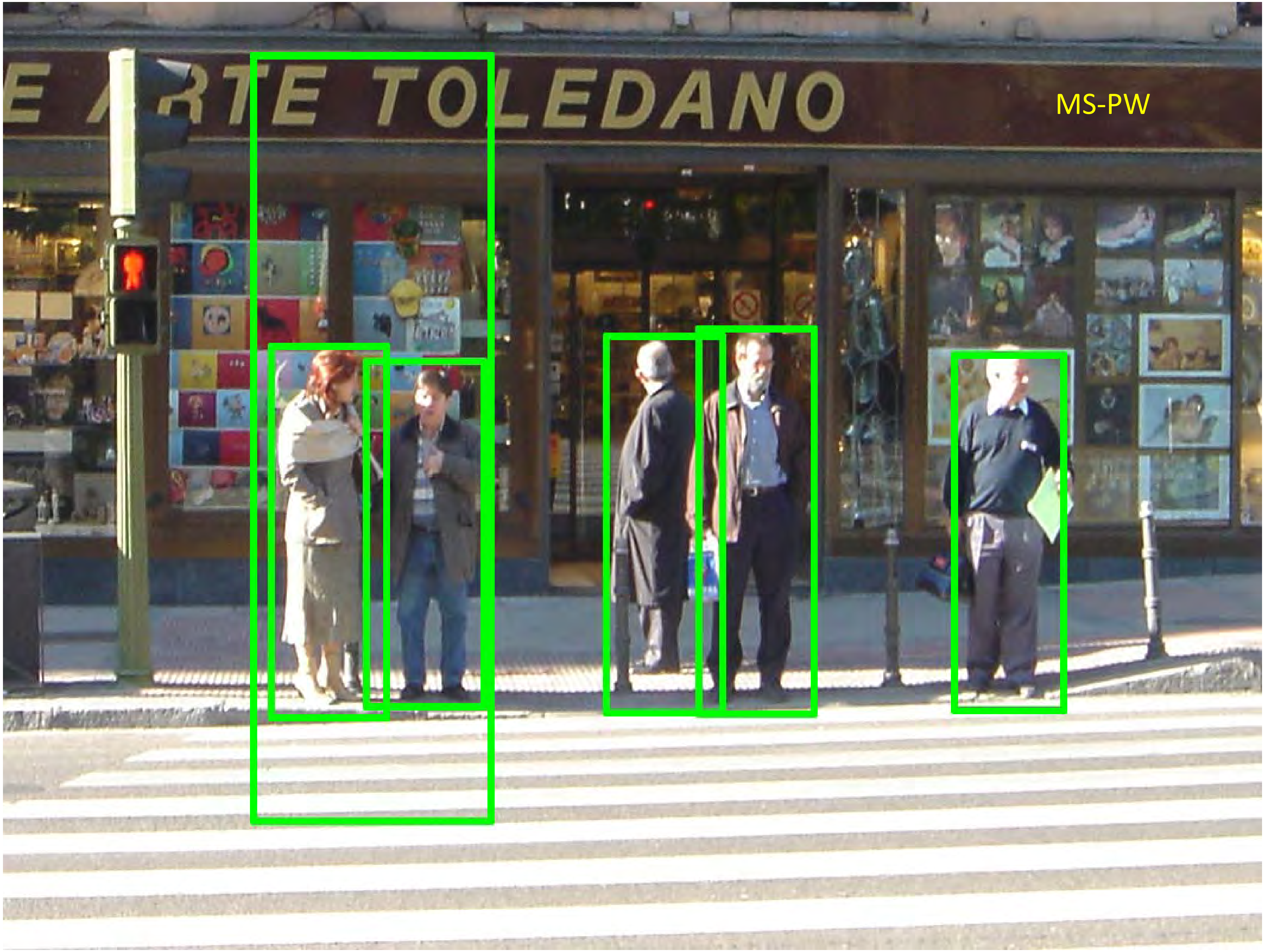


Gain in average precision



(b)


E ARTE TOLEDANO
Standard SW


E ARTE TOLEDANO
MS-PW

After people detection…
recognition

- Gait analysis
- Posture analysis
- Action/interaction/activity analysis
- Behavior analysis
- Motion in crowd
- Anomaly detection..

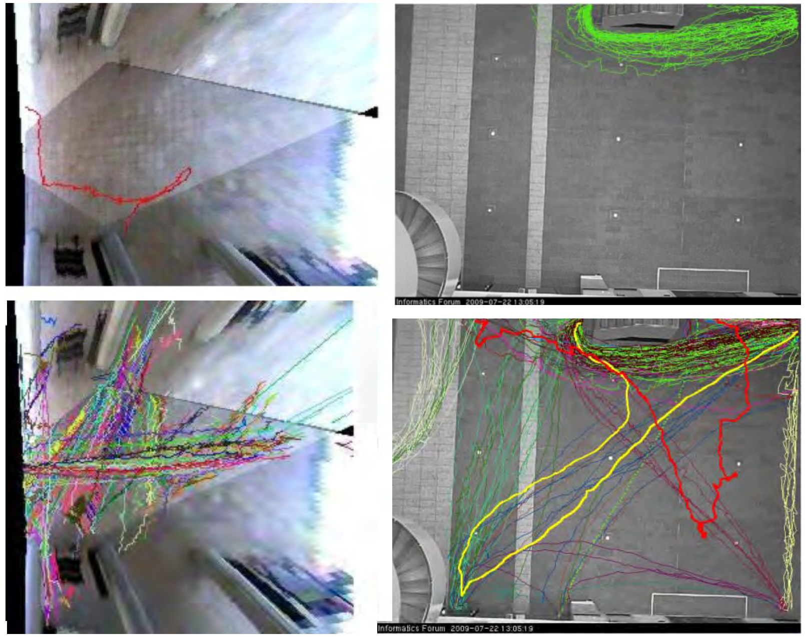
M. S. Ryoo, J. K. Aggarwal Semantic Representation and Recognition of Continued and Recursive Human Activities **Journal of Computer Vision** 2009

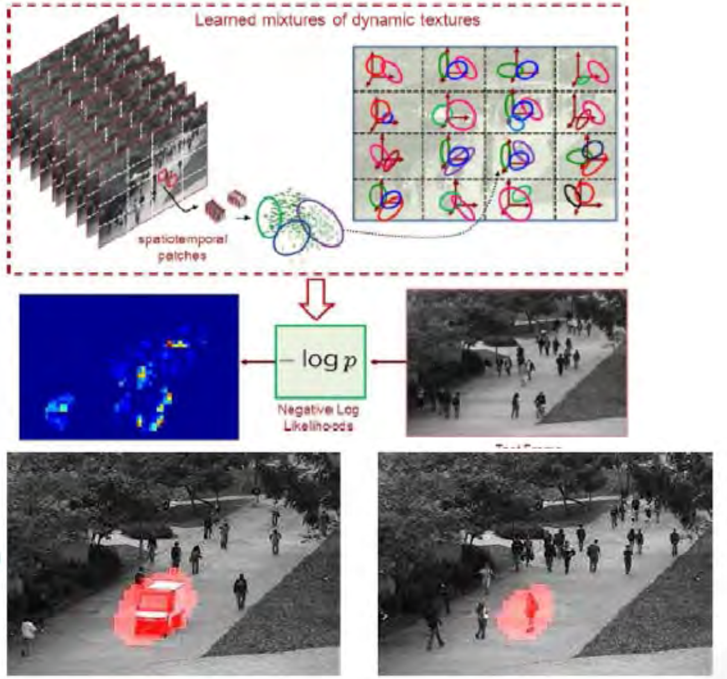
Kaiqi Huang; Dacheng Tao; Yuan Yuan; Xuelong Li; Tieniu Tan; View-Independent Behavior Analysis IEEE Trans SMC 2008

Cheriyadat, A.M.; Radke, R.J.; Detecting Dominant Motions in Dense Crowds, IEEE Journal of Selected Topics in Signal Processing 2008

Examples

Learn and Predict Trajectory anomaly detection using Circular Statistics[Calderara101





Holistic Anomaly detection using Mixtures of Dynamic Textures [Mahadevan10]

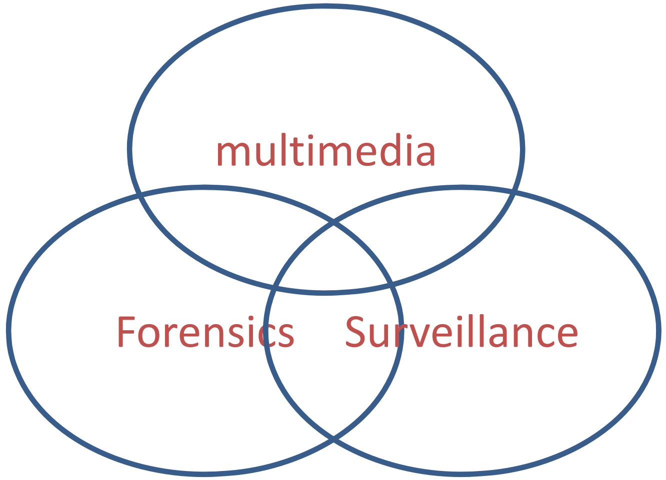
[Mahadevan10] Vijay Mahadevan, Weixin Li, Viral Bhalodia, Nuno VasconcelosAnomaly Detection in Videos Using Mixtures of Dynamic Textures in Proc of CVPR 2010
[Calderara11]Calderara S., Prati A. Cucchiara R. Mixtures of von Mises Distributions for People Trajectory Shape Analysis IEEE Trans. On Circuits and system for Video Technology 2011

Examples

Gait analysis in forensic medicine from
surveillanec video
[Larsen08]

Single View Metrology
[Criminisi02]

After people detection… re-identification



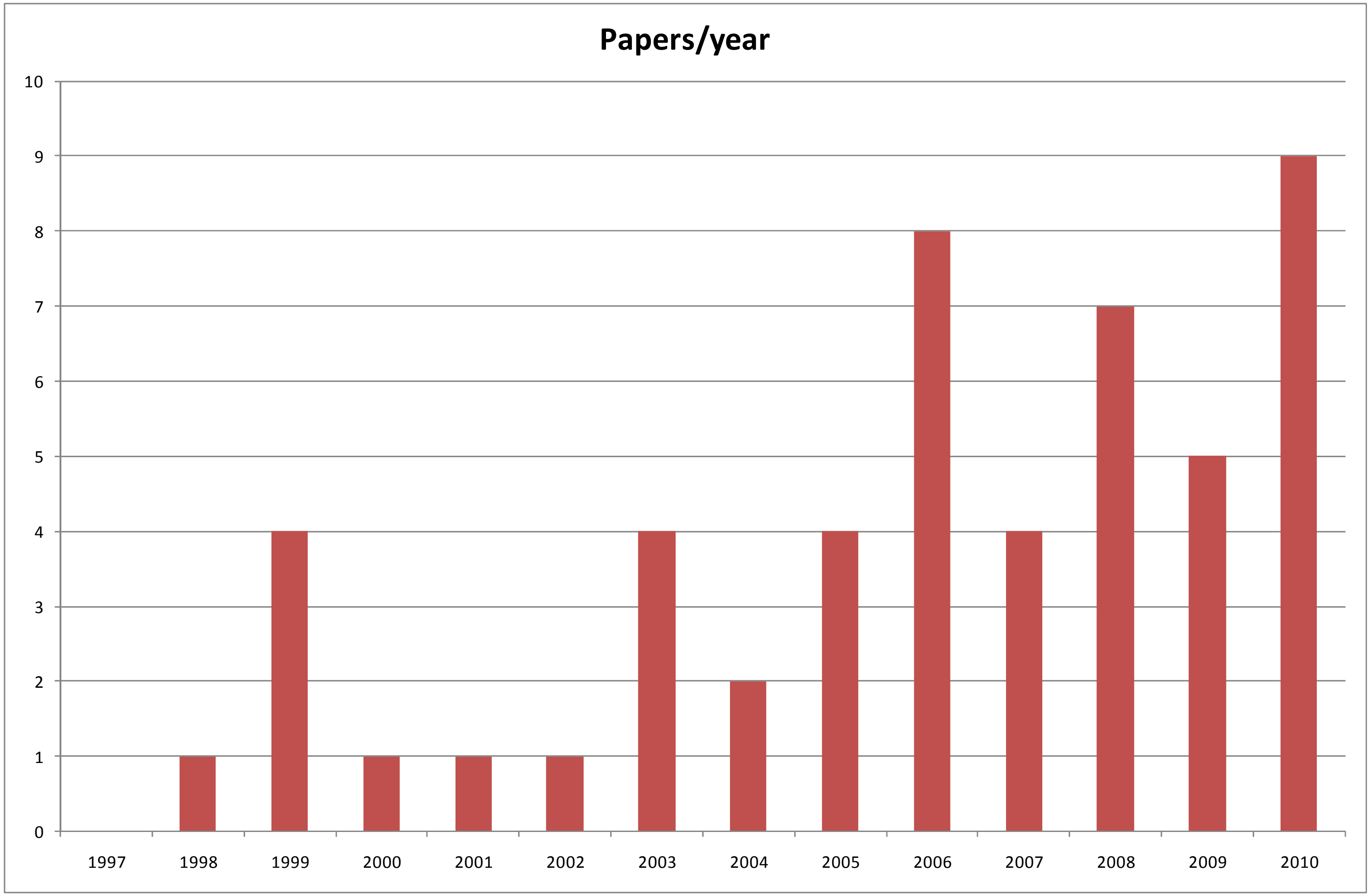
People Re-identification

Search for similarity (in a plain database of images and videos)

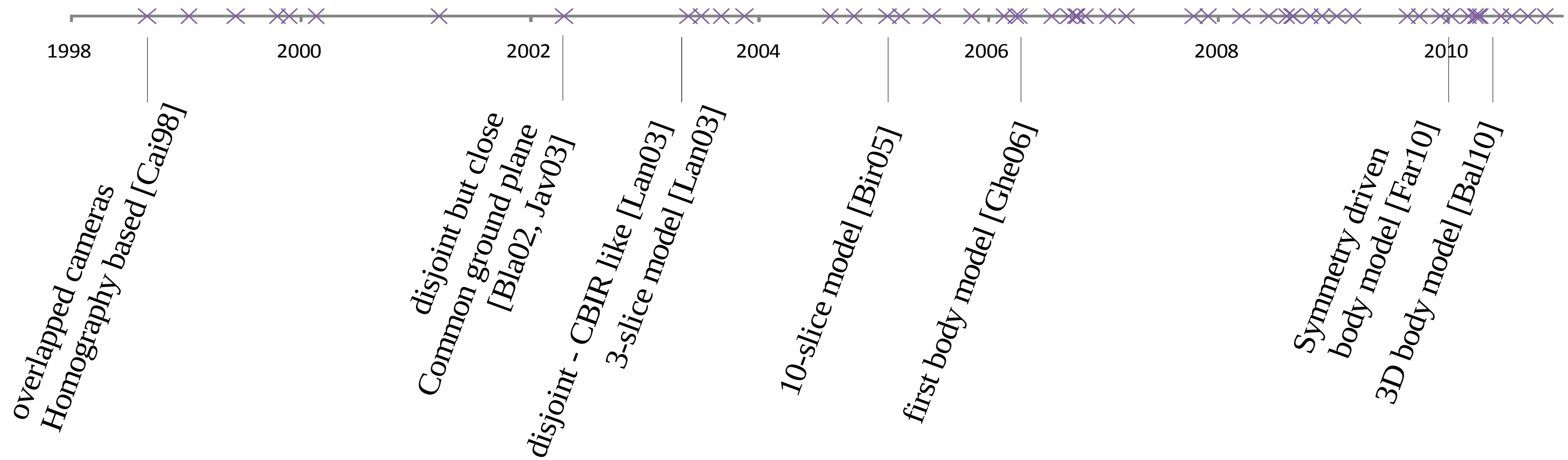
Extension of the **tracking** problem (in videos)

- The tracking problem, aims at finding an association between prediction and observation.
- Tracking matches a previously seen target if it appears again in the same camera, after a short time, in a position close to the previous one, and with a similar appearance.

Re-identification: a short survey

- 

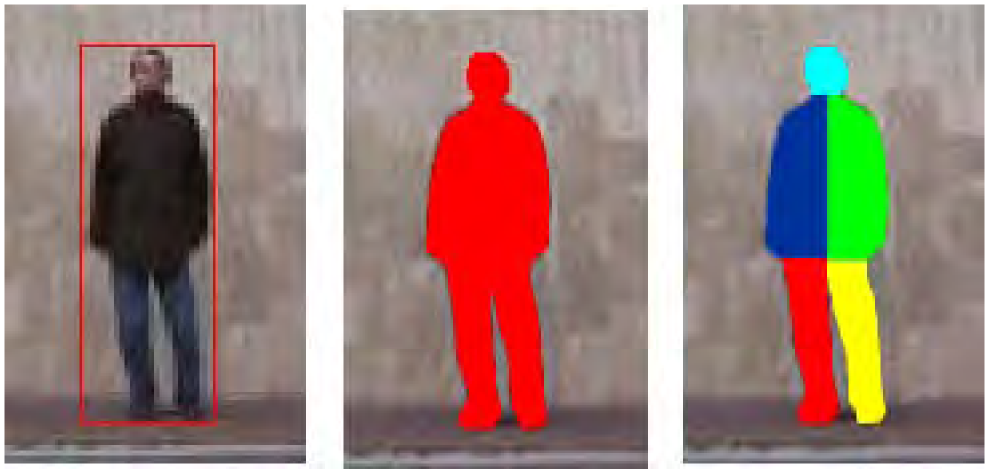
Many research works



- [Cai98] **Q. Cai and J.K. Aggarwal**, "Automatic tracking of human motion in indoor scenes across multiple synchronized video streams," Sixth International Conference on Computer Vision 1998, pp. 356-362.
- [Bla02**] J. Black, T. Ellis, and P. Rosin**, "Multi view image surveillance and tracking," Proceedings of Workshop on Motion and Video Computing, 2002., IEEE Comput. Soc, 2002, pp. 169-174.
- [Jav03] **O. Javed, Z. Rasheed, K. Shafique, and M. Shah**, "Tracking across multiple cameras with disjoint views," Proc. IEEE International Conference on Computer Vision,2003, pp. 952-957 vol.2.
- [Lan03] M. Lantagne, M. Parizeau, and R. Bergevin, "VIP: Vision tool for comparing Images of People," Vision Interface, 2003.
- [Bir05] N.D. Bird, O. Masoud, N.P. Papanikolopoulos, and A. Isaacs, "Detection of Loitering Individuals in Public Transportation Areas," IEEE Transactions on Intelligent Transportation Systems, vol. 6, Jun. 2005, pp. 167-177.
- [Ghe06] **N. Gheissari, T.B. Sebastian, and R. Hartley**, "Person Reidentification Using Spatiotemporal Appearance," Conference on Computer Vision and Pattern Recognition 2006, pp. 1528-1535.
- [Far10] **M. Farenzena, L. Bazzani, A. Perina, V. Murino, and M. Cristani**, "Person re-identification by symmetry-driven accumulation of local features," Conference on Computer Vision and Pattern Recognition, IEEE, 2010, pp. 2360-2367.
- [Bal10] **D. Baltieri, R. Vezzani, and R. Cucchiara,** "3D Body Model Construction and Matching for Real Time People Re-Identification," Proc. of Eurographics Italian Chapter Conference 2010 (EG-IT 2010), 2010.

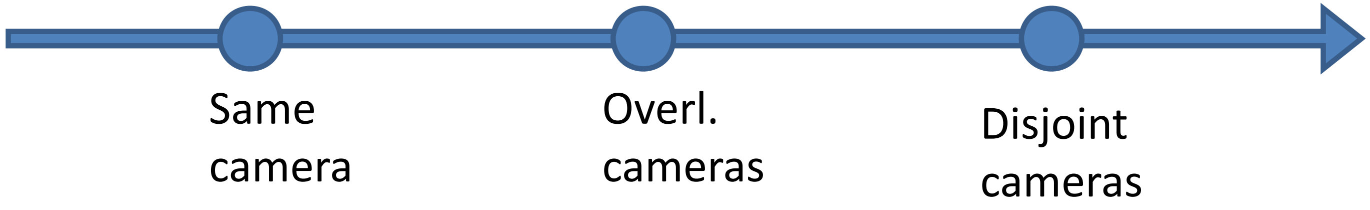
Requirements for re-identification

- Object/People detection (bounding box)
- Foreground detection (mask)
- Face/body-part detection (segmented regions)
- Single-camera tracking (temporal consistency and motion information)



A multi-dimensional problem

- Camera positioning



- **Same camera:** the system should be able to re-detect the same person whenever he appears again in the same camera. View point and color correction problems can be neglected [Yan99]
- **Overlapping cameras**: geometrical information can be exploited; main assumption: people to match are captured at the very instant [Cal08]
- **Disjoint cameras**: much complex case [Far10]

[Yan99] J. Yang, X. Zhu, R. Gross, J. Kominek, Y. Pan, and A. Waibel, "Multimodal people ID for a multimedia meeting browser," International **ACM Multimedia Conference**, 1999, p. 159.
[Cal08] S. Calderara, R. Cucchiara, and A. Prati, "Bayesian-competitive consistent labeling for people surveillance.," IEEE transactions on pattern analysis and machine intelligence, vol. 30, Feb. 2008, pp. 354-60.
[Far10] M. Farenzena, L. Bazzani, A. Perina, V. Murino, and M. Cristani, "Person re-identification by symmetry-driven accumulation of local features," 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, IEEE, 2010, pp. 2360-2367.

A multi-dimensional problem

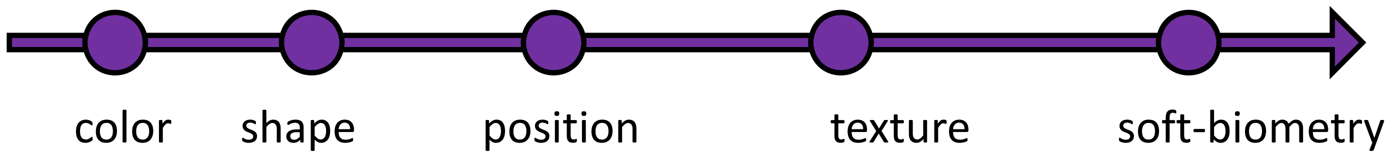
- Single/multiple shots

- **Single Shot** methods associate pairs of images, each containing one instance of an individual. These methods are mostly similar to those proposed for image retrieval with some particular specialization to people.
  - PROS: simple, fast. CONS: view dependent. Less stable with occlusions and noise

- **Multiple shot**: information coming from multiple frames (or images) containing the same person are used as training data.
  - PROS: more information gathered for the same person; CONS: alignment and increased data dimensionality

A multi-dimensional problem

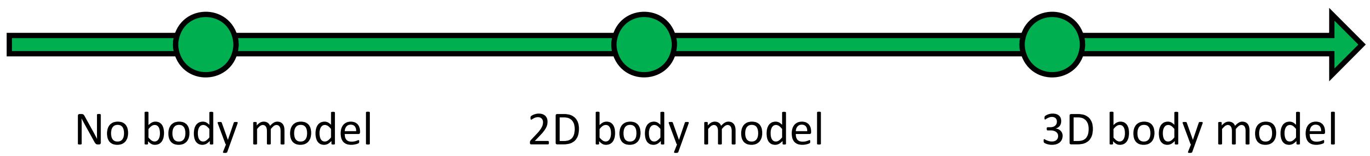
- Signature



- **color**
  - mean, histogram, Gaussian model, Mixture of Gaussians
  - color space RGB, rgb, HSV
- **shape:**
  - width, height, h/w ratio, contour
- **Spatial features (position/trajectory)** position in the image or in the ground plane
- **texture:** covariance matrix, SIFT/SURF
- **Soft-biometry:** face, gait

A multi-dimensional problem

- The body model

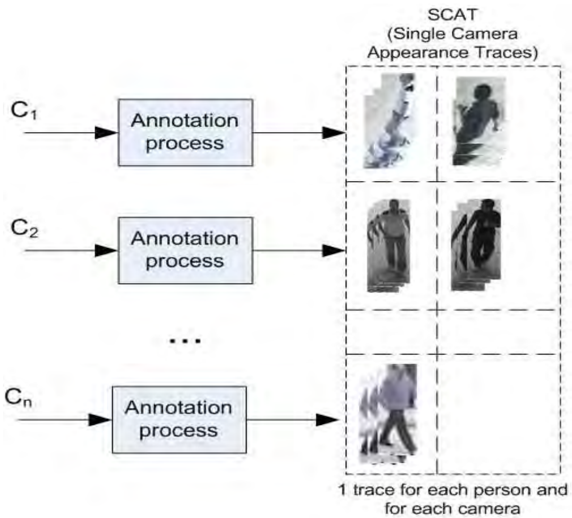


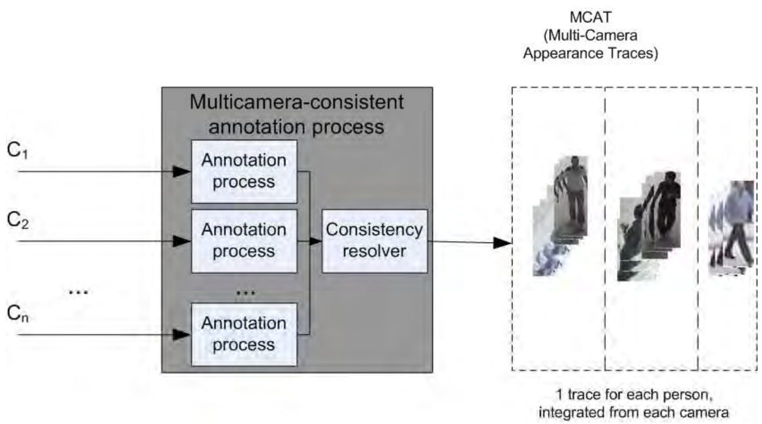
- No body model
- 2D body model
  - Cylindrical model
  - LTH Leg torso head
- 3D body model

No body model: search for similarity

Content based retrieval methods

Global descriptors: Histograms , texture, Medioni's
circular histograms, Mixture of gaussians, covariance matrix…

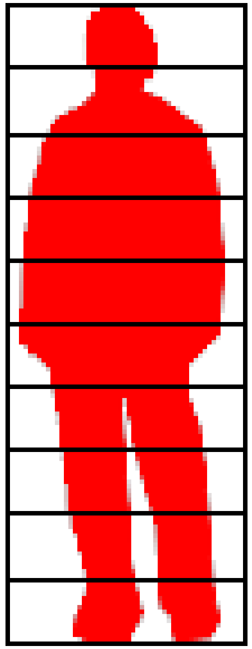




S. Calderara, R.Cucchiara, A. Prati **Multimedia Surveillance: Content based Retrieval with Multicamera People Tracking** Proc of workshop VSSN at acm multimedia 2006

Cylindrical body model

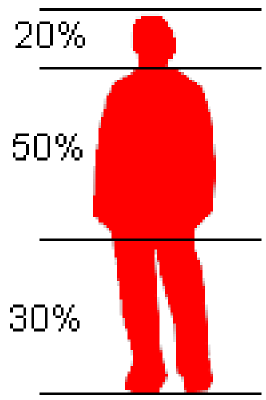
- Cylindrical shape (or more generally as a solid of revolution)
  - the horizontal variations of the people appearance are neglected, supposing that the color or texture
  - distribution along the vertical axis is the only important data.
  - [Bir05]: the person mask is divides into ten horizontal stripes and the mean color of each stripe is stored as representative feature.



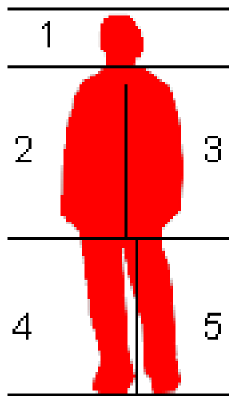
[Bir05] N.D. Bird, O. Masoud, N.P. Papanikolopoulos, and A. Isaacs, "Detection of Loitering Individuals in Public Transportation Areas," IEEE Transactions on Intelligent Transportation Systems, vol. 6, Jun. 2005, pp. 167-177.

Legs-torso-head model

- The reason of the legs-torso-head model, instead, is mainly due to the occidental traditional clothing.
- The target silhouette is divided into three horizontal parts, ideally corresponding to :
  - legs (and thus to the pants/skirt appearance)
  - torso (i.e., shirt or jacket)
  - head (i.e., hair).



Fixed [Lan03]



Estimated [Far10]

[Lan03] M. Lantagne, M. Parizeau, and R. Bergevin, “VIP: Vision tool for comparing Images of People,” Vision Interface, 2003.
[Far10] M. Farenzena, L. Bazzani, A. Perina, V. Murino, and M. Cristani, “Person re-identification by symmetry-driven accumulation of local features,” 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, IEEE, 2010, pp. 2360-2367.

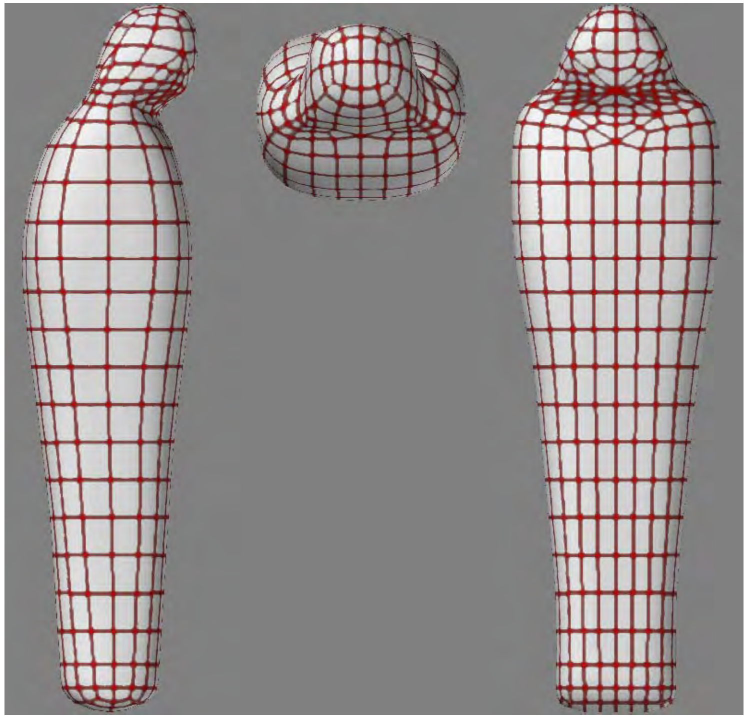
3d body models (1)

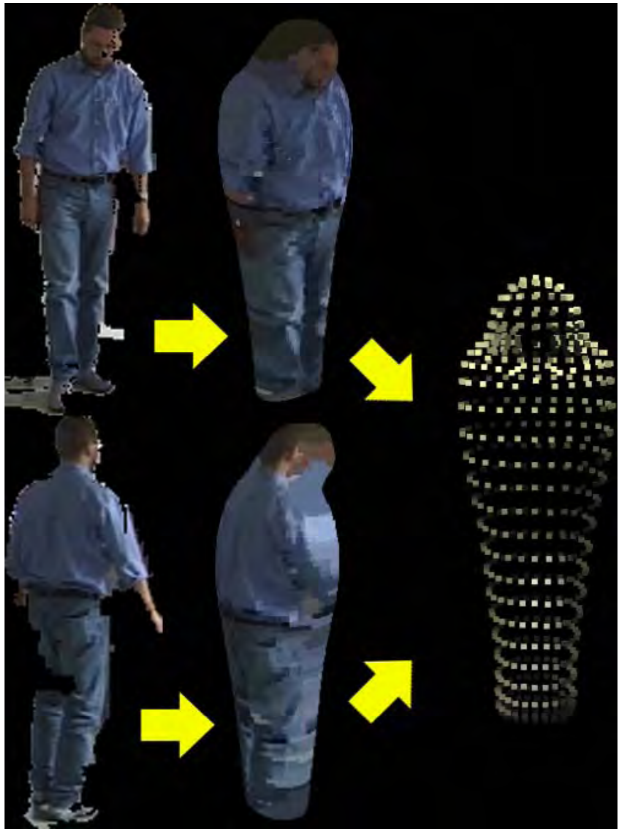
- Panoramic Appearance Map: surface of a 3D cylindrical model [GAN06].

[Gan06] T. Gandhi and M. Trivedi, “Panoramic Appearance Map (PAM) for Multi-camera Based Person Re-identification,” 2006 IEEE AVSS IEEE, 2006, pp. 78-78.

3D Body Models (2)

- 3D vertex model with features stored and related to each vertex [Bal10]

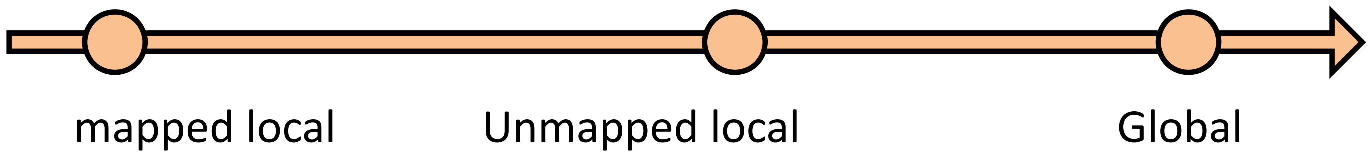




[Bal10] D. Baltieri, R. Vezzani, and R. Cucchiara, “3D Body Model Construction and Matching for Real Time People Re-Identification,” Proceedings of Eurographics Italian Chapter Conference 2010 (EG-IT 2010), Genova, Italy: 2010.

A multi-dimensional problem

- Spatial localization of features

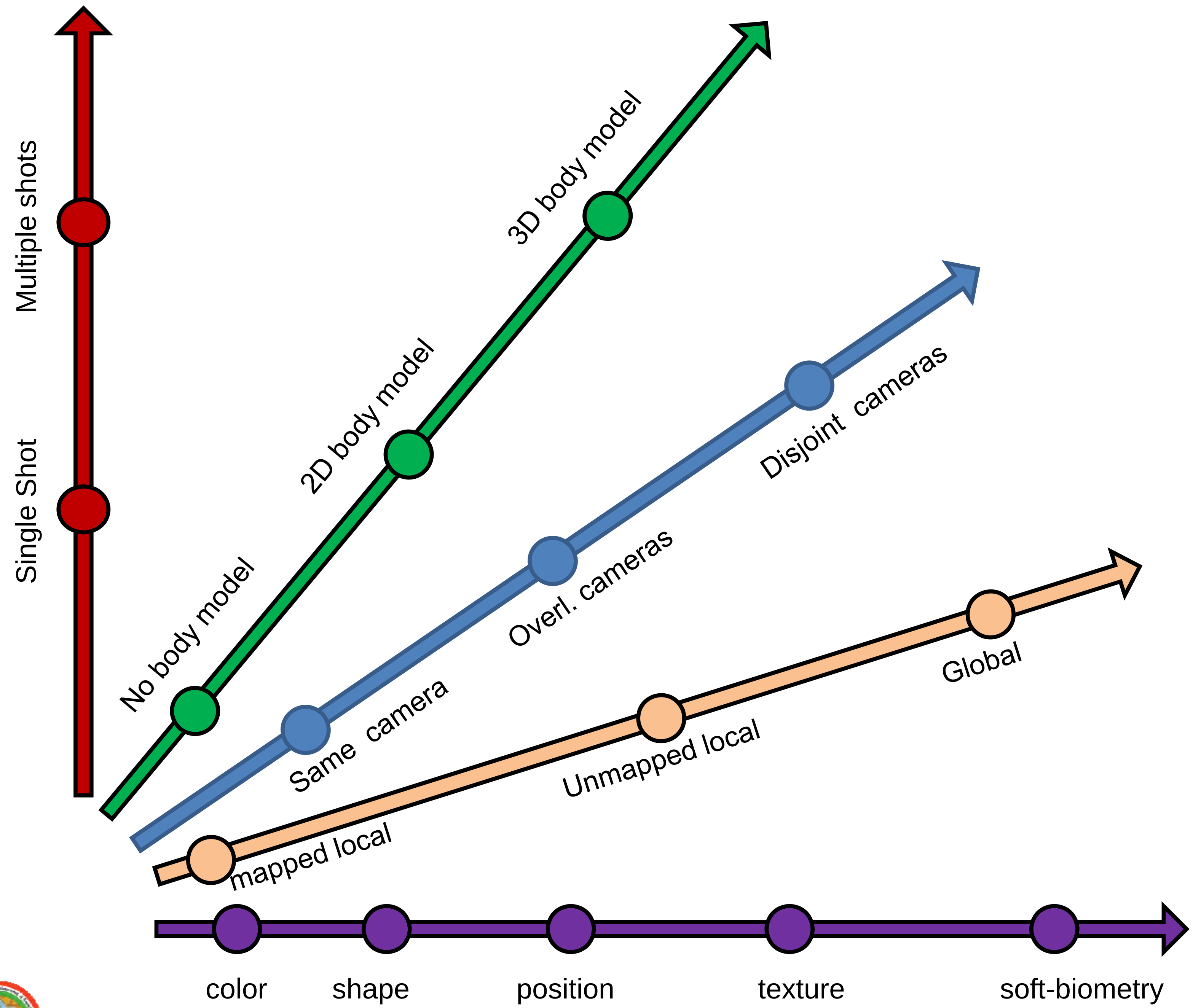


- **Global features**: global color histogram, shape descriptors [Orw99]
- **Unmapped local features**: features are computed on patches or blocks but are unmapped to a body model or a relative position. E.g., Bag-Of-Word with SIFT descriptors [Liu09]
- **Mapped local features**: features are referred to a human body model and to specific regions [Lan03,Met10]

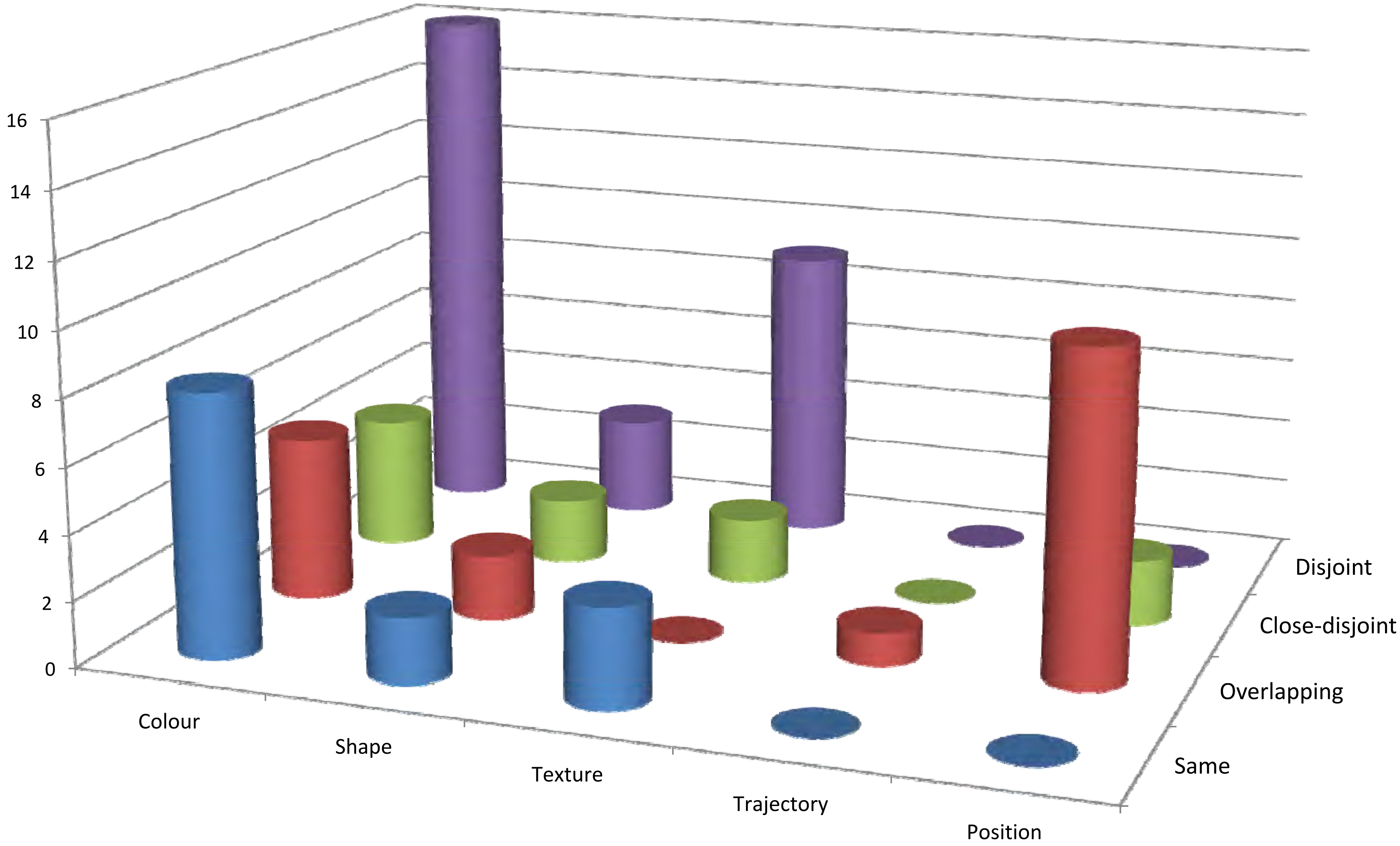
[Orw99] J. Orwell, P. Remagnino, and G.A. Jones, “Multi-camera colour tracking,” VS’99, pp. 14-21.

[Liu09] K. Liu and J. Yang, “Recognition of People Reoccurrences Using Bag-Of-Features Representation and Support Vector Machine,” Chinese Conference on Pattern Recognition, 2009, pp. 1-5.

[Lan03] M. Lantagne, M. Parizeau, and R. Bergevin, “VIP: Vision tool for comparing Images of People,” Vision Interface, 2003.
[Met10]M. Metternich, M. Worring, and A. Smeulders, “Color Based Tracing in Real-Life Surveillance Data,” Trans. on Data Hiding and Multimedia Security V, vol. 6010, 2010, pp. 18-33.


Multiple shots
Single Shot
3D body model
2D body model
No body model
Disjoint cameras
Overl. cameras
Same camera
Global
Unmapped local
mapped local
color
shape
position
texture
soft-biometry

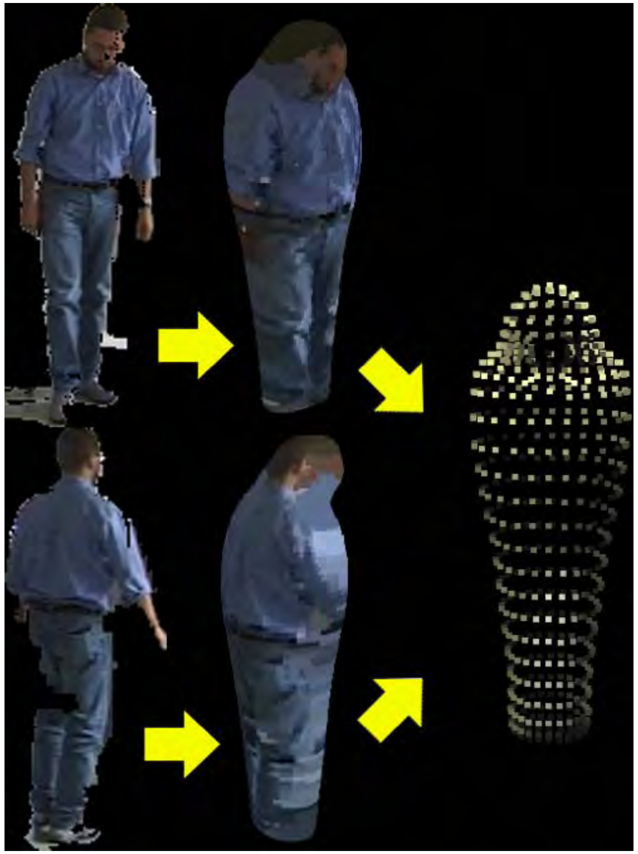
Papers for Camera positioning and Adopted Feature



Example 3D body model for ri-identification

- *Camera positioning:* disjoint
- *Body model*: 3D vertex based body model about 600 vertices with scale factor
- *Signature*: local color histograms
- *Requirements*: calibration

- Re-identification is provided comparing 3D models or view-specific projections of the model



ViSOR Re-Identification Dataset

- A new dataset designed for people re-identification

http://www.openvisor.org

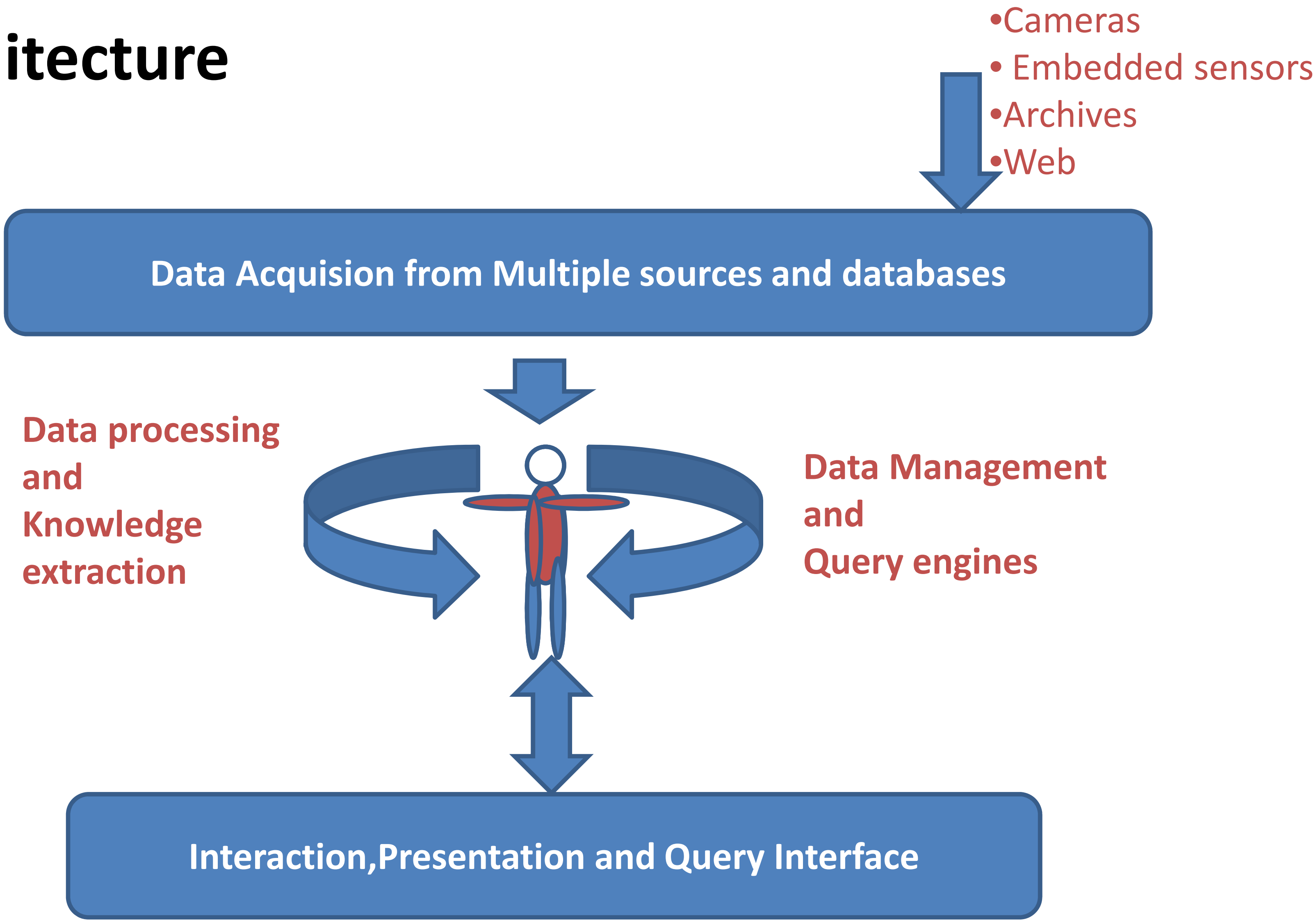
Re-Identification Dataset

- 50+ people
- At least 4 snapshot for each person from different angles
- Position and orientation of each person w.r.t. the camera for correct 2d/3d alignment
- Thanks to:
- A EU project in Prevention, Preparedness and Consequence Management of Terrorism and other Security-related Risks Programme European Commission – DG: JLS

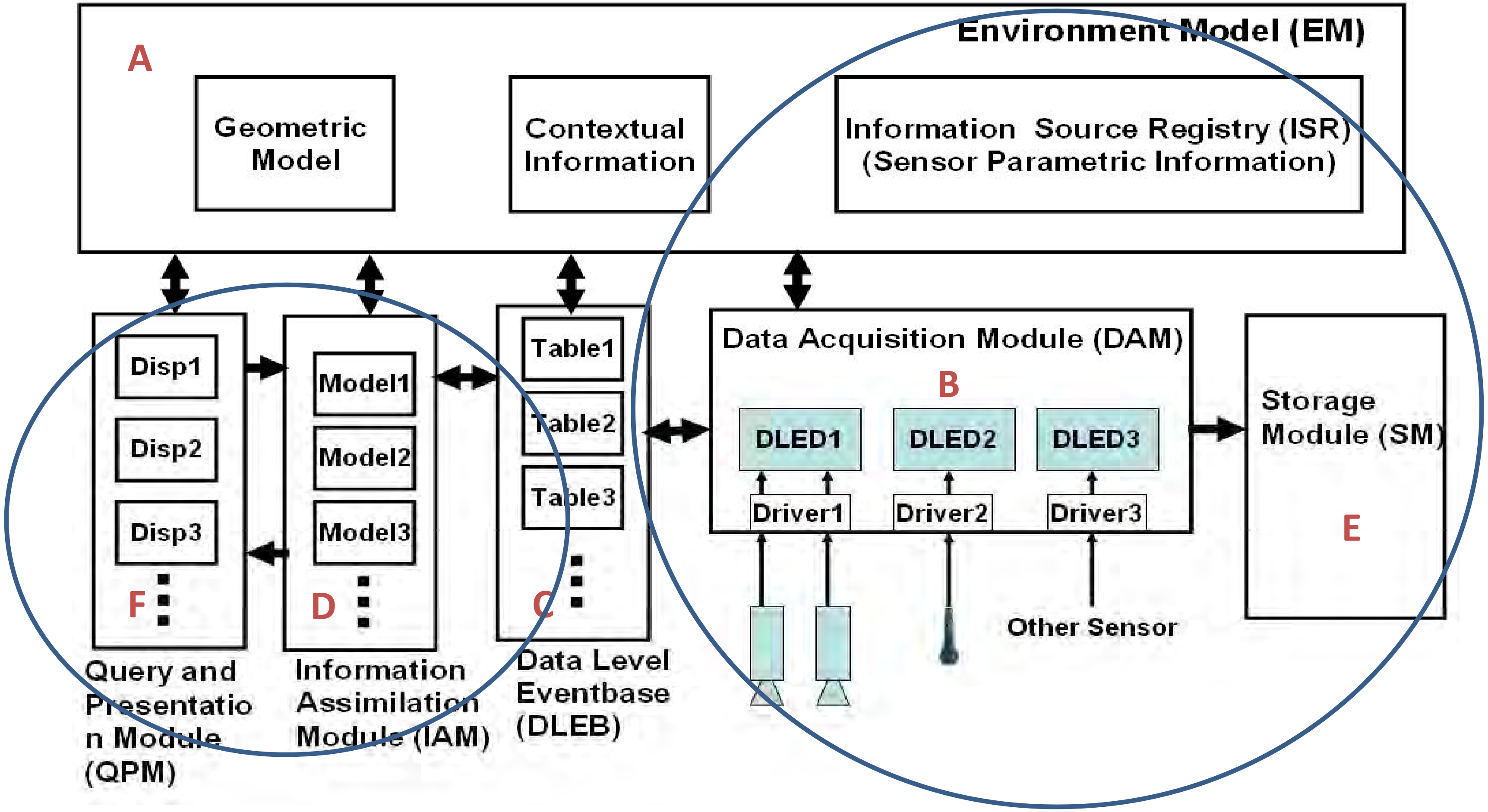
Conclusion…The Future of Multimedia Surveillance&Forensics Architecture

- Architecture for non IT experts.
- With software solutions for:
  - Combining different analysis (real-time knowledge extraction and data mining)
  - Allow 3D-4D world reconstruction
  - Presents data in innovative,intuitive and interactive way (touch, mobile..)
  - Allow traceability of operation (for legaly issues)
  - Deal with privacy
- Focus on performance when analysing very large databases of data

Multimedia Surveillance&Forensics Architecture



Video Surveillance Forensic [Saini09]

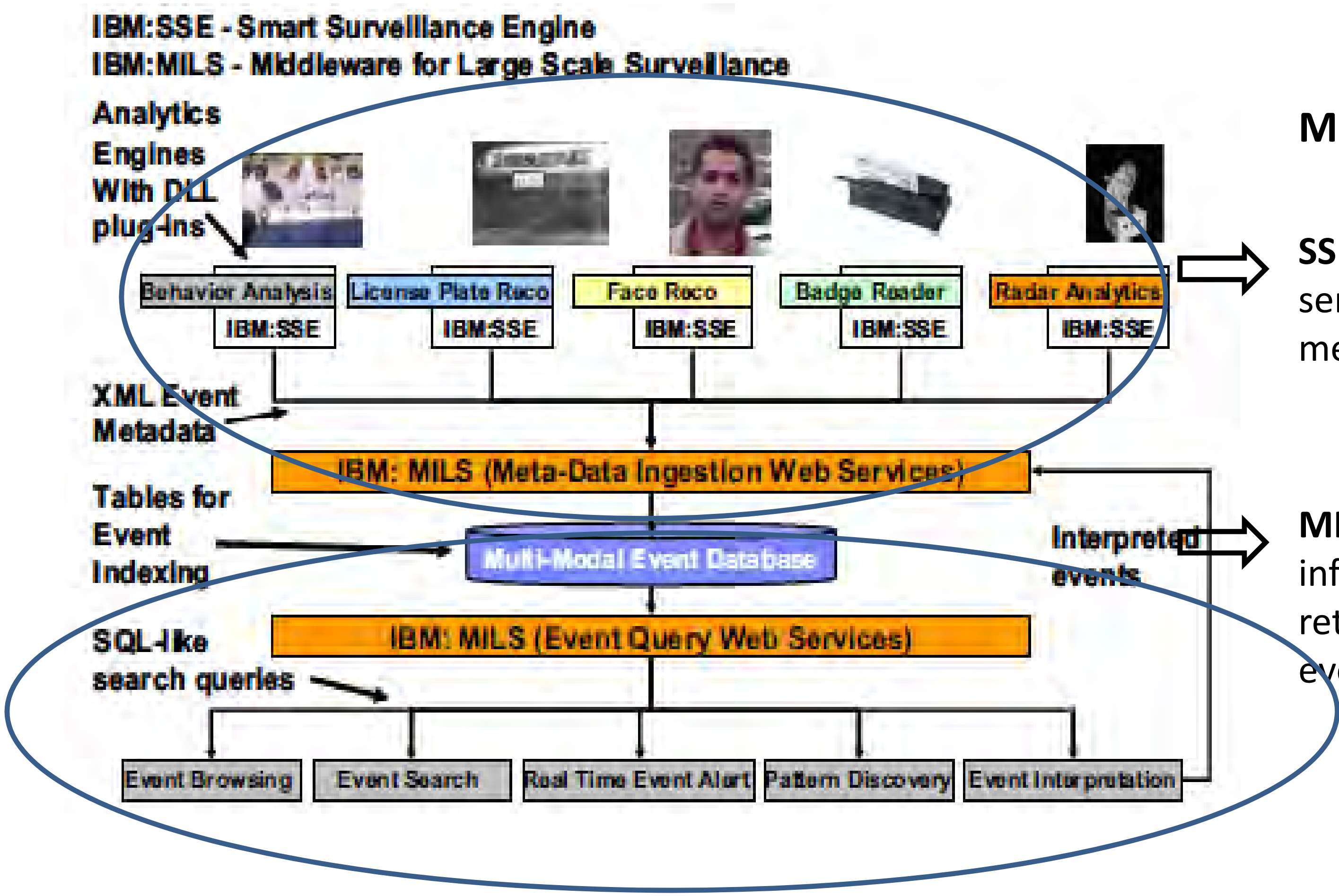


[Saini09] ", M. Saini, M. Kankanhalli, R. Jain "*A Flexible Surveillance System Architecture*, 2009 Advanced Video and Signal Based Surveillance AVSS09

Video Surveillance Forensic [Saini09] (2)

A. ***Environment Model*** **(EM):** abstracts environment dependent information into three categories: Geometric Information, Contextual Information and Sensor Parametric Information.
B. ***Data Acquisition Module*** **(DAM):** interacts with the hardware and identifies the type of information, choosing the relevant driver and the appropriate **DLED (*Data Level Event Detector*).**
C. ***Data Level Event Base*** **(DLEB):** an event repository having one event table for each sensor.
D. ***Information Assimilation Module*** **(IAM):** acceps the query from presentation module, evaluates it and returns the results back for presentation.
E. ***Storage Module*** **(SM):** stores compressed data streams for efficiency purposes.
F. ***Query and Presentation Module*** **(QPM)**: provides the interface between user and the system, with a simple GUI to make queries and a user friendly presentation of results.

IBM S3 Hybrid Surveillance Solution[Tian08]



Modules:

SSE: processes data from sensors and generates XML meta-data.

MILS: provides the infrastructure for indexing, retrieving and managing event meta-data

[Tian08]Tian, Y.L, Brown, L.M., Hampapur, A., Lu, M., Senior, A., Shu, C., IBM smart surveillance system (S3): event based video surveillance system with an open and extensibleframework, MVA(19), 2008,

IBM S3 Hybrid Surveillance Solution[Shu05] (2)

SSE provides support functionalities for the core analysis components

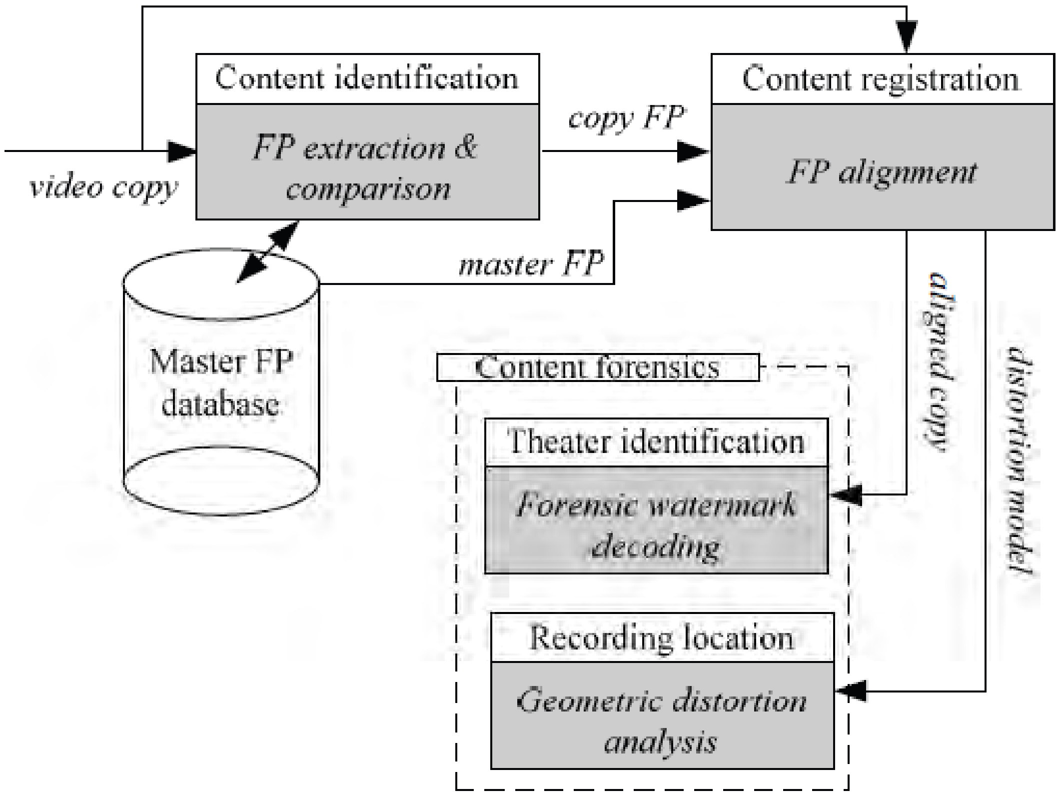
1. **Standard Plug-in Interfaces:** allows to plug into SSE any event analysis component complying with the defined interfaces.
2. **Extensible Meta-Data Interfaces:** provides a way to extend XML structures.
3. **Real-Time Alert Interfaces:** allows to plug-in application specific alerts.
4. **Compound Alers Interfaces:** compose multiple basic real-time events.
5. **Real-Time Actuation Interfaces:** allows to plug-in actuation modules driven from the defined event and alerts.
6. **Database communication Interfaces**.

MILS provides analysis engines via standard web services interfaces using XML documents

1. **Meta-Data Ingestion Services:** allow to manage both meta-data that can be retrieved through SQL-like queries and events detected by the SSE.
2. **Schema Management Services:** allow to menage the meta-data schema.
3. **Schema Management Services:** allow to manage the videosurveillance system, through Camera Management Services, Engine Management Services etc.

Mutimedia Forensic Video Forgery Detection [Baudry09]

Fingerprint: "a set of features automatically extracted from the video signal"



1. **Content identification** is performed by extracting a video fingerprint (FP) and matching it against a database of master content fingerprints.
2. **Alignment** of copy FP with master FP allows content registration.
3. The aligned video copy, together with the modeled distortions, is transmitted to **forensic modules**, such as **watermark decoding** or recording **location estimation**.

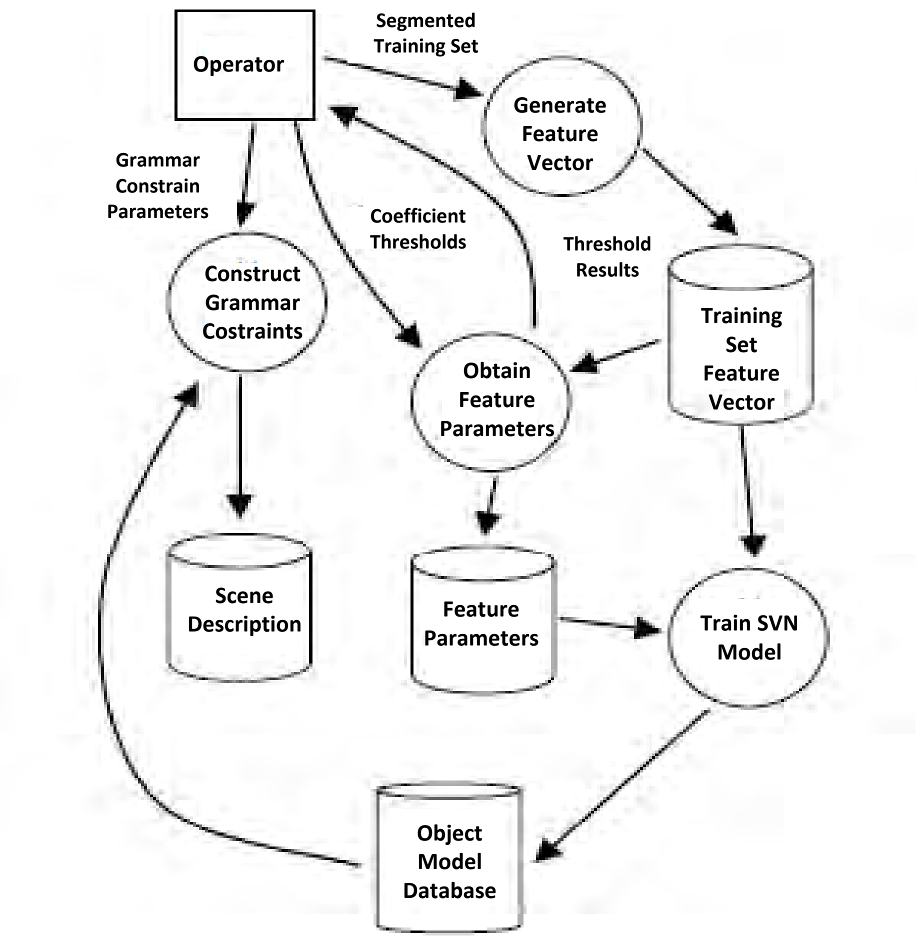
[Baudry09] ", S. Baudry, B. Chupeau, F. Lefèbvre *"A framework for Video Forensic Based on Local and Temporal Fingerprints*, ICIP 2009.

Forensic Image Mining [Brown05]

2. **Classification Module**

1. **Training Module**

G
Co
Pa



- Implements the **image retrieval process**
- Use of a **Bayesian Network** for query refinement with a set of relevance feedback parameters.

- execution of a chain of forensic tool for :
- filtering the data streams for:
  - Data reduction
  - Extraction of **meta-data**.

[Brown05] R. Brown, B. Pham, O. de Vel. *"Design of a Digital Forensic Image Mining System"*, Information and Engineering Systems, 2005

Natural interactivity for security forensics&surveillance

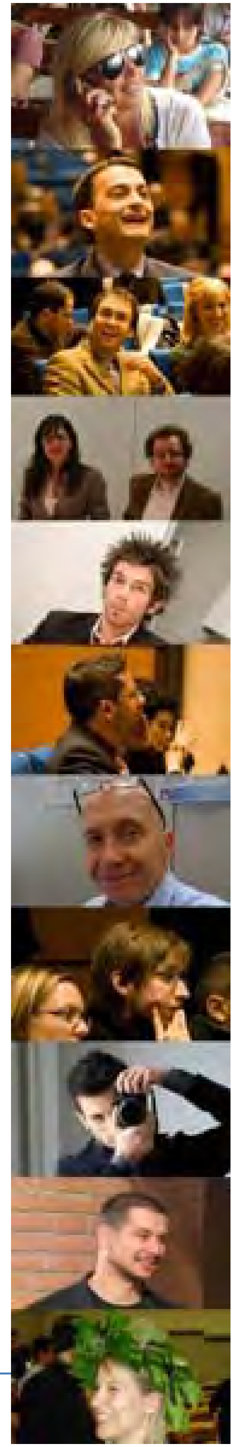
Virtual Autopsy Table (developed by Norrköping Visualization Centre and the Center for Medical Image Science and Visualization in Sweden)
http://www.youtube.com/watch?v=bws6vWM1v6g

Thanks.

For any details
http://Imagelab.ing.unimore.it

..And thanks to Imagelab

Andrea Prati, Roberto Vezzani, Costantino Grana, Simone Calderara, Giovanni Gualdi, Paolo Piccinini, Paolo Santinelli, Daniele Borghesani, Davide Baltieri, Sara Chiossi, Adnan Rashid, Michele Fornaciari, Manuel Aravecchia, Rudy Melli, Emanuele Perini, Giuliano Pistoni..

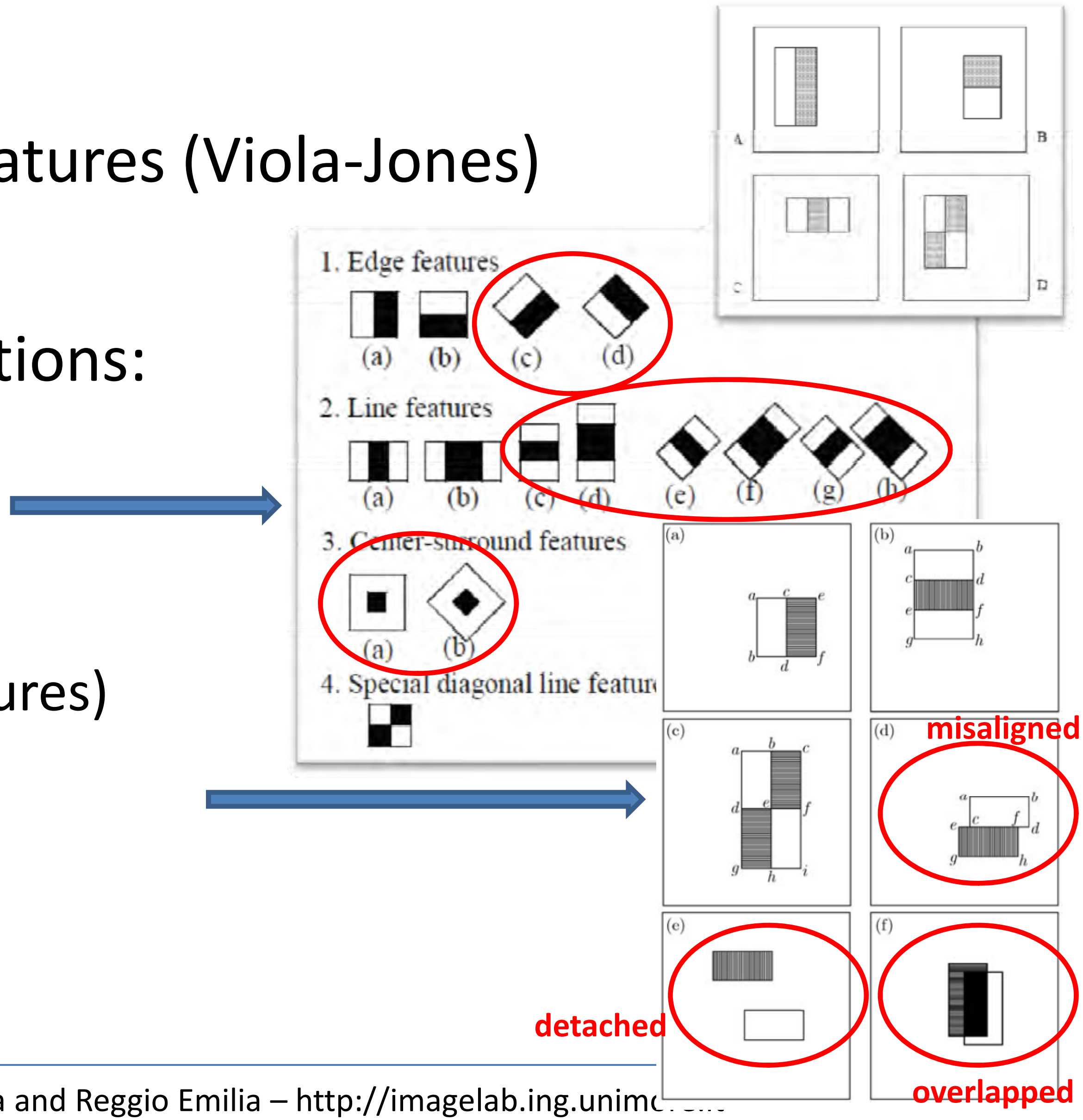


References

- R. Lienhart, A. Kuranov, V. Pisarevsky, “Empirical Analysis of Detection Cascades of Boosted Classifiers for Rapid Object Detection”, LNCS 2781:297-304, 2003
- P. Viola, M.J. Jones, “Robust Real-Time Face Detection”, IJCV, 57(2): 137-154, 2004
- A. Lehmann, B. Leibe, L. Van Gool, “Feature-Centric Efficient Subwindow Search”, in ICCV 2009
- B. Froba, A. Ernst, “Face detection with the modified census transform”, in Intl Conf on Automatic Face and Gesture Recognition, pp. 91-96, 2004
- J. Wu, M.D. Mullin, J.M. Rehg, “Linear Asymmetric Classifier for cascade detectors”, in ICML , pp. 988-995, 2005
- Q. Zhu, M-C. Yeh, K-T. Cheng, S. Avidan, “Fast Human Detection Using a Cascade of Histograms of Oriented Gradients”, in CVPR, 2:1491-1498, 2006
- W. Zhang, G. Zelinsky, D. Samaras, “Real-time Accurate Object Detection using Multiple Resolutions”, in ICCV, pp. 1-8, 2007
- R. Verschae, J. Ruiz-del-Solar, M. Correa, “A unified learning framework for object detection and classification using nested cascades of boosted classifiers”, Machine Vision and Applications, 19(2):85-103
- S. Paisitkriangkrai, S. Chunhua, Z. Jian}, “Fast Pedestrian Detection Using a Cascade of Boosted Covariance Features”, TCSVT, 18(8):1140-1151, 2008
- O. Tuzel, F. Porikli, P. Meer, “Pedestrian Detection via Classification on Riemannian Manifolds”, IEEE T-PAMI, 30(10):1713-1727, 2008
- W. Zhang, R. Tong, J. Dong, “Boosted cascade of scattered rectangle features for object detection”, Science in China Series F: Information Sciences, 52(2): 236-243, 2009
- R.N. Hota, K. Jonna, P.R. Krishna, “On-road vehicle detection by cascaded classifiers”, in COMPUTE, pp. 1-5, 2010.

Features used: HLF

- HLF=Haar-Like Features (Viola-Jones)
- Suitable modifications:
  - Enhanced HLF (Lienhart 2003)
  - SRF (Scattered Rectangular Features)



Weak Classifiers

$$h_j(x) = \begin{cases} 1 & \text{if } p_j f_j(x) < p_j \theta_j \\ 0 & \text{otherwise} \end{cases}$$

feature

parity sign

threshold

$$c(\mathbf{x}|j,\theta) := \begin{cases} 1 & x_j > \theta \\ 0 & \text{otherwise} \end{cases}$$

- Single perceptron
- Decision stumps
- CART (Classification And Regression Trees): a non-parametric technique; a form of binary recursive partitioning.
- Bayesian classifier
- Linear SVM
- Linear logistic regressor: use logistic function as objective function for linear regression

Strong Classifiers

- Mostly-used: AdaBoost

1. Given N examples $(x_1,y_1),\ldots,(x_N,y_N)$ with $x \in \Re^k, y_i \in \{-1,1\}$
2. Start with weights $w_i = 1/N$, $i = 1, \ldots, N$.
3. Repeat for $m = 1, \ldots, M$
   (a) Fit the classifier $f_m(x) \in \{-1,1\}$ using weights w_i on the training data $(x_1,y_1),\ldots,(x_N,y_N)$.
   (b) Compute $err_m = E_w[1_{(y \neq f_m(x))}]$, $c_m = \log((1-err_m)/err_m)$.
   (c) Set $w_i \leftarrow w_i \cdot \exp(c_m \cdot 1_{(y_i \neq f_m(x_i))})$, $i = 1, \ldots, N$, and renormalize weights so that $\sum_i w_i = 1$.
4. Output the classifier $sign\left[\sum_{m=1}^{M} c_m \cdot f_m(x)\right]$

Strong Classifiers

- Several modifications:

Gentle AdaBoost

1. Given N examples $(x_1,y_1),\ldots,(x_N,y_N)$ with $x \in \Re^k, y_i \in \{-1,1\}$
2. Start with weights $w_i = 1/N$, $i = 1, \ldots, N$.
3. Repeat for $m = 1, \ldots, M$

   (a) Fit the regression function $f_m(x)$ by weighted least-squares of y_i to x_i with weights w_i

   (c) Set $w_i \leftarrow w_i \cdot \exp(-y_i \cdot f_m(x_i))$, $i = 1, \ldots, N$, and renormalize weights so that $\sum_i w_i = 1$.
4. Output the classifier $sign\left[\sum_{m=1}^{M} f_m(x)\right]$

  – AsymBoost: the weight on positive examples is increased so that the minimum error criteria will also have very few false negatives.

Strong Classifiers

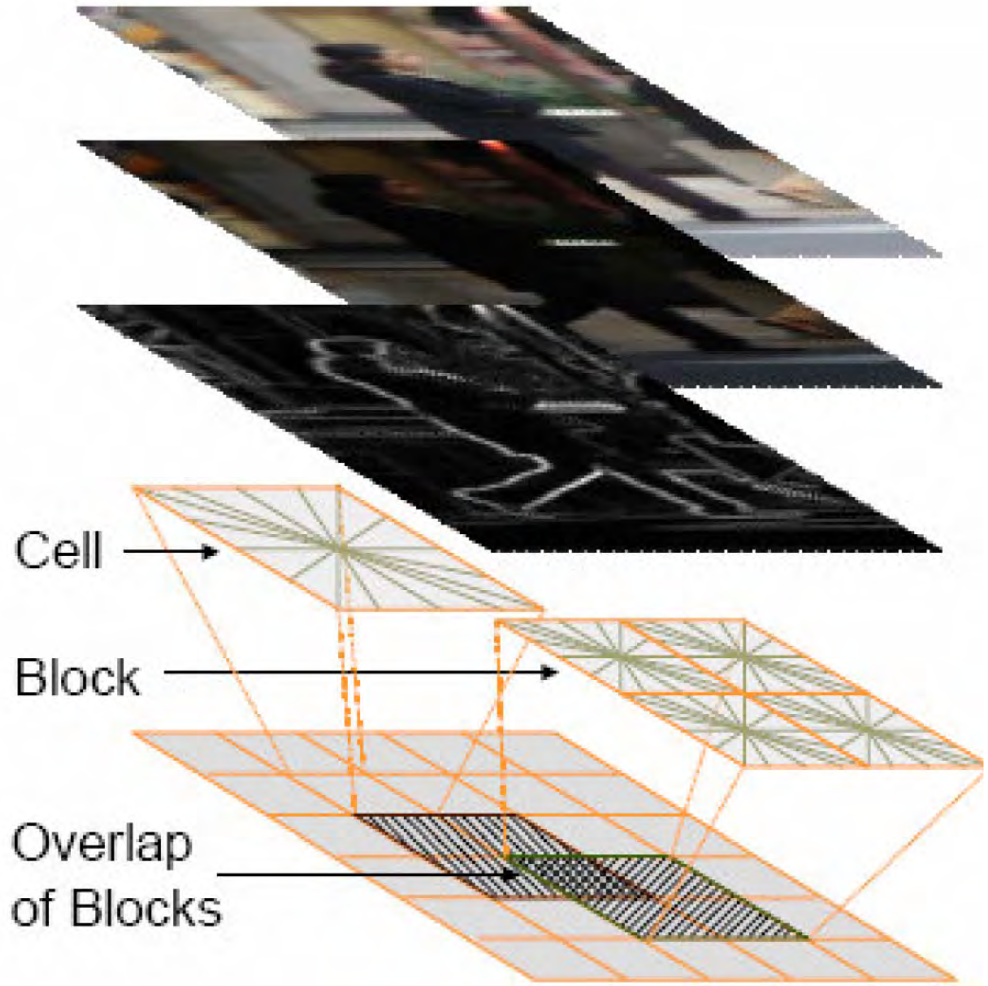
- LogitBoost: it casts the AdaBoost algorithm into a statistics framework. Obtained by considering AdaBoost as a generalized additive model and then applying the cost functional of logistic regression. It solves:

$$\min_{f(x)} E_{w(x)} \left(F(x) + \frac{1}{2} \frac{y^* - p(x)}{p(x)(1-p(x))} - (F(x) + f(x)) \right)^2$$

Histograms of Oriented Gradients 1/2

*N. Dalal and B. Triggs "**Histograms of Oriented Gradients for Human Detection,**" CVPR 2005*

- Compute the gradients of the input image in a detection window
- Compute histograms of the gradients over 16 directions in overlapping blocks
- Combine all the histograms into a single feature vector ***f***
- Apply a classifier (linear SVM) trained off-line on a large training set of ***f***



Anomaly Detection for Forensics	
Training Strategy	
Pre learnt set of Normal Activities	Explicit Modelling of the desired Anomaly
[Adam07] [Cui08] [Zhong04]…	[Chan07] [Gryn05]…

[Adam07] A. Adam, E. Rivlin, I. Shimshoni, D. Reinitz, **Robust Real-Time Unusual Event Detection using Multiple Fixed-Location Monitors** in IEEE Trans. on PAMI 30 (2007)
[Cui08] P. Cui, L. Sun, Z.-Q. Liu, S. Yang, **A Sequential Monte Carlo Approach to Anomaly Detection in Tracking Visual Events** in: Proc. of IEEE Int'l CVPR 2008
[Zhong04] H. Zhong, J. Shi, M. Visontai, **Detecting Unusual Activity in Video** in: Proc. of IEEE CVPR , 2004

[Chan07] M. T. Chan, A. Hoogs, J. Schmiederer, M. Petersen, **Detecting Rare Events in Video Using Semantic Primitives with HMM** in: Proc. of ICPR, 2007

[Gryn05] J. M. Gryn, R. P. Wildes, J. K. Tsotsos, **Detecting Motion Patterns via Direction Maps with Application to Surveillance** in: WACV-MOTION , 2005

Anomaly Detection for Forensics

Training Strategy

Pre learnt set of Normal Activities	Explicit Modelling of the desired Anomaly
[Adam07] [Cui08] [Zhong04]…	[Chan07] [Gryn05]…

Feature Selection

Holistics Approaches (Scene Motion)	
Anomaly Detected using Temporal Bayesian Networks	
Parametric PdF modelling of recurrent motion vectors	[Saleemi 08] [Mahadevan10]
Generative Model of common motion behaviors	[Xiang08]
Non parametric PdF modelling of motion patterns	[Basharat08]

[Saleemi08] I. Saleemi, K. Shaque, M. Shah**, Probabilistic Modeling of Scene Dynamics for Applications in Visual Surveillance** IEEE Trans. on PAMI 31 (2008)

[Mahadevan10] Vijay Mahadevan, Weixin Li, Viral Bhalodia, Nuno Vasconcelos**Anomaly Detection in Videos Using Mixtures of Dynamic Textures** in Proc of CVPR 2010

[Xiang08] T. Xiang, S. Gong, **Video Behavior Profiling for Anomaly Detection** IEEE Trans. on PAMI 30 (2007)

[Basharat08]A. Basharat, A. Gritai, M. Shah, **Learning object motion patterns for anomaly detection and improved object detection** in: Proc. of IEEE Int'l CVPR 2008

Anomaly Detection for Forensics

Training Strategy

Pre learnt set of Normal Activities	Explicit Modelling of the desired Anomaly
[Adam07] [Cui08] [Zhong04]…	[Chan07] [Gryn05]…

Feature Selection

[Picciarelli08]Piciarelli, C.; Micheloni, C.; Foresti, G.L., "**Trajectory-Based Anomalous Event Detection**," *Circuits and Systems for Video Technology, IEEE Transactions on* , vol.18, no.11, pp.1544-1554, Nov. 2008

[Junejo04] Junejo, O. Javed, and M. Shah, "**Multi feature path modeling for video surveillance**," *in Proc. of Int'l Conference on Pattern Recognition, vol. 2, Aug. 2004, pp. 716– 719*.

[calderara11] Calderara S., Prati A. Cucchiara R. **Mixtures of von Mises Distributions for People Trajectory Shape Analysis IEEE** Trans. On Circuits and system for Video Technology 2011

Feature Based (e.g. People Trajectories)	
Anomaly Detected using Trained Classifiers	
[Picciarelli08]	SVM training on normal people trajectories
[Junejo04]	Graph cut clustering of motion paths to detect normal cluster
[Calderara11]	Evolutionary learn and predict K-medoids on people trajectory

Application of People Soft-Identification analysis in Forensic		
Search-by-example		
	Feature	Method
People Off-line Video Tracking [Koppen09]	People Extraction using HOG. Color Similarities	Searching for optimal graph path using optimization techniques
Video sequences of people appearance Similarity Search [Calderara06]	Gaussian Color GMM over a temporal sequence	KNN similarity search using KL Distance
Single query Image Mining [Vaquero09]	Multiple Techniques exists based on: • Local feature(e.g Sift) [Olivera09] •Color (e.g Histogram, Correlogram Color Segmentation) [Cong10] •Mixed local and Global approaches [Farenzena10]	
Single View Metrology and 3D reconstruction		
Height Estimation from Single Image [Benabde08] [Viswanath09]		
3D semi-automatic Scene Modelling And Rendering [Mohan09]		
Forensic Application of Gait Analysis		
	Feature	Classification
Age Identification [Lu10]	Gabor Phase and Magnitude computed over a gait sequence	Multi label K-NN
Low Resolution Gait Identification [Jiwen10]	Spectral Embedding of Gait characteristics in superresolution images.	Tensor voting scheme for classification

Soft Identification analysis

Soft identification refers to tools and techniques for find similar istances of a query person based on visual features:

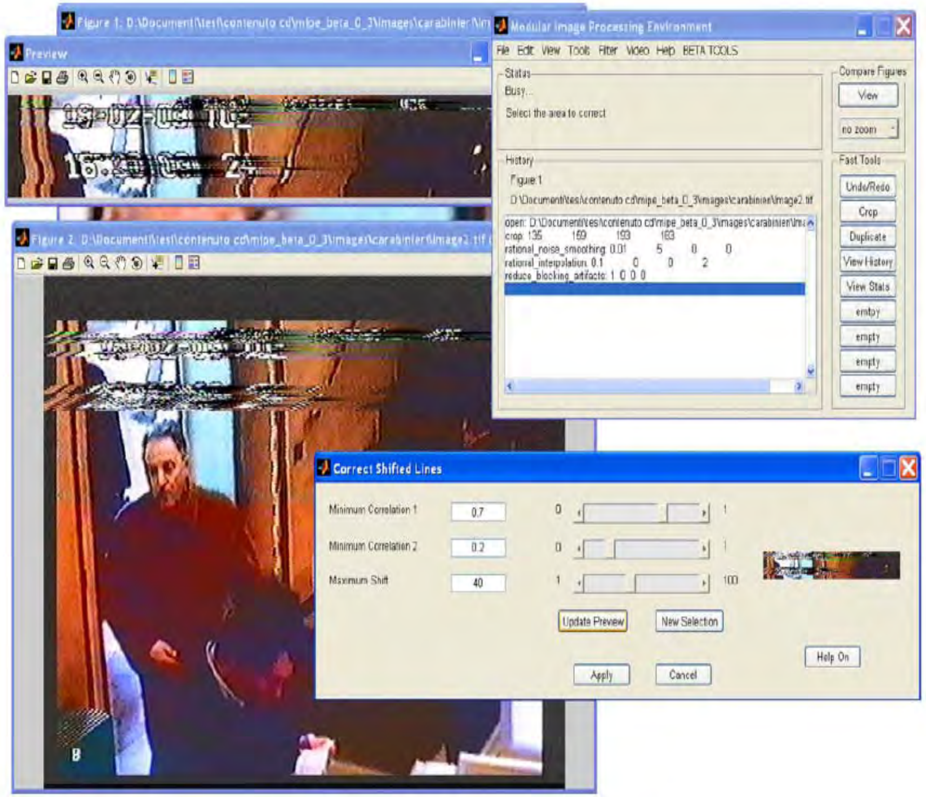
- Image mining techniques aims at :
  - Recognizing multiple temporal correlated istances of the same person in video sequences ("off-line people tracking") [Koppen09]
  - Finding similar images of a desired subject in videos from etherogeneus sources (Appearance similarity search) [Vaquero09]
  - Finding multiple instances of the same subject in different videos from disjoint cameras (people-reidentification) [Calderara06]
- Single View Metrology aims at finding measurement of biometric indicators from a single image (i.e. people height)
- Soft Biometry aims at analysing distinctive (but not primary) elements of a person that can help identifying the subject [Jain06]:
  - Gait analyis
  - Posture analysis
  - Age,gender height

Multimedia Forensic Image Enhancement [Jerian07]

•Image Enhancement Techniques For Image Restoration.

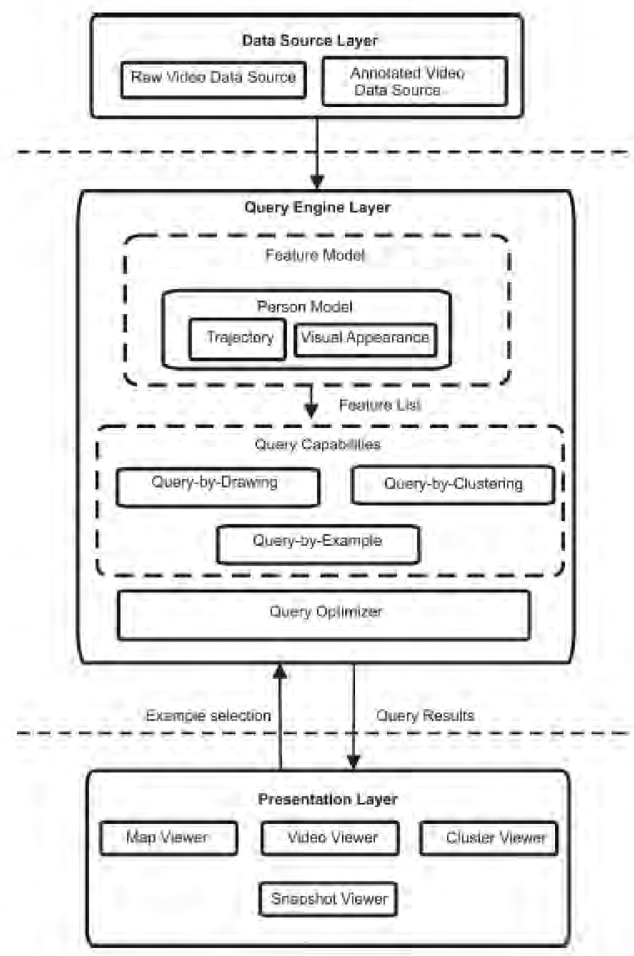
- A complete **transparency** is achieved :
  - All algorithms applied to an image are listed in the right sequence on a log file automatically saved with the image
  - For each employed algorithm all the involved parameters are listed
  - For each algorithm the reference article is provided
  - For each algorithm the source code is provided

▪Main Objective is the traceability of operation



[Jerian07] M. Jerian, S. Paolino, F. Cervelli, S. Carrato, A. Mattei, L. Garofano “*A forensic image processing environment for investigation of surveillance video*”, *For. Science Int.* **167**, p. 207 (2007)

Video Surveillance Forensic Data Browsing [Aravecchia10]



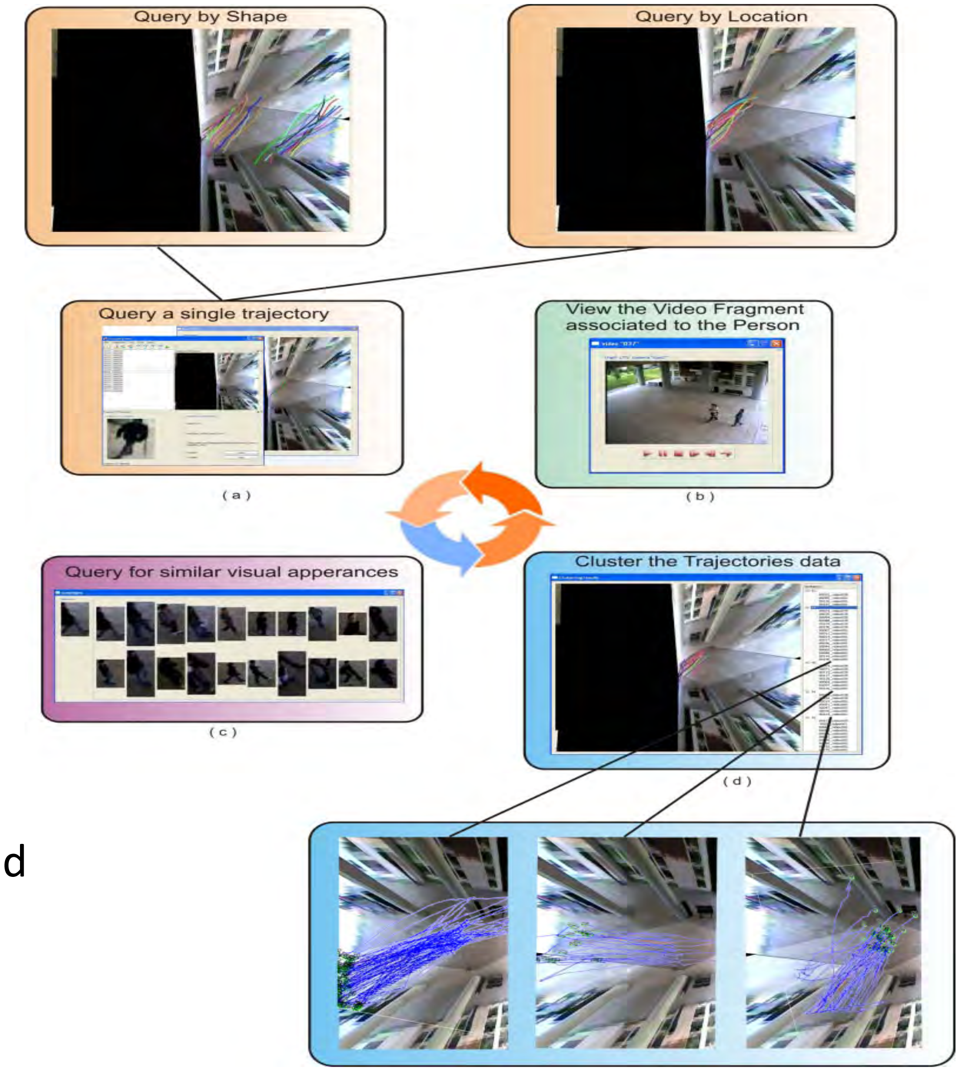
The Source Data Layer ensure to manage different types of data, from raw to annotated videos.

The Query Engine Layer is the core of the application, based on Person Model and allowing to browse relevant information inferred from large video dataset.

The Presentation Layer allows user-friendly access to data in order to support fast e simple forensic investigation.

[Aravecchia10] M. Aravecchia, S. Calderara, S. Chiossi, R. Cucchiara, "*A Videosurveillance Data Browsing Software Architecture for Forensics: from Trajectories Similarities to Video Fragments*", MiFor 2010.

Video Surveillance Forensic Data Browsing Example [Aravecchia10] (2)



<u>Person Model:</u>

- **Trajectories feature comparison:**
  - Shape
  - Location
- **People Appearance Comparison**

<u>Query Engine:</u>

browse video data according to the adopted Person Model:

- ***Query-by-Drawing***: (Trajectories)
- ***Query-by-Clustering****:* (Trajectories)
- ***Query-by-Example***: (Trajectories and people appearances