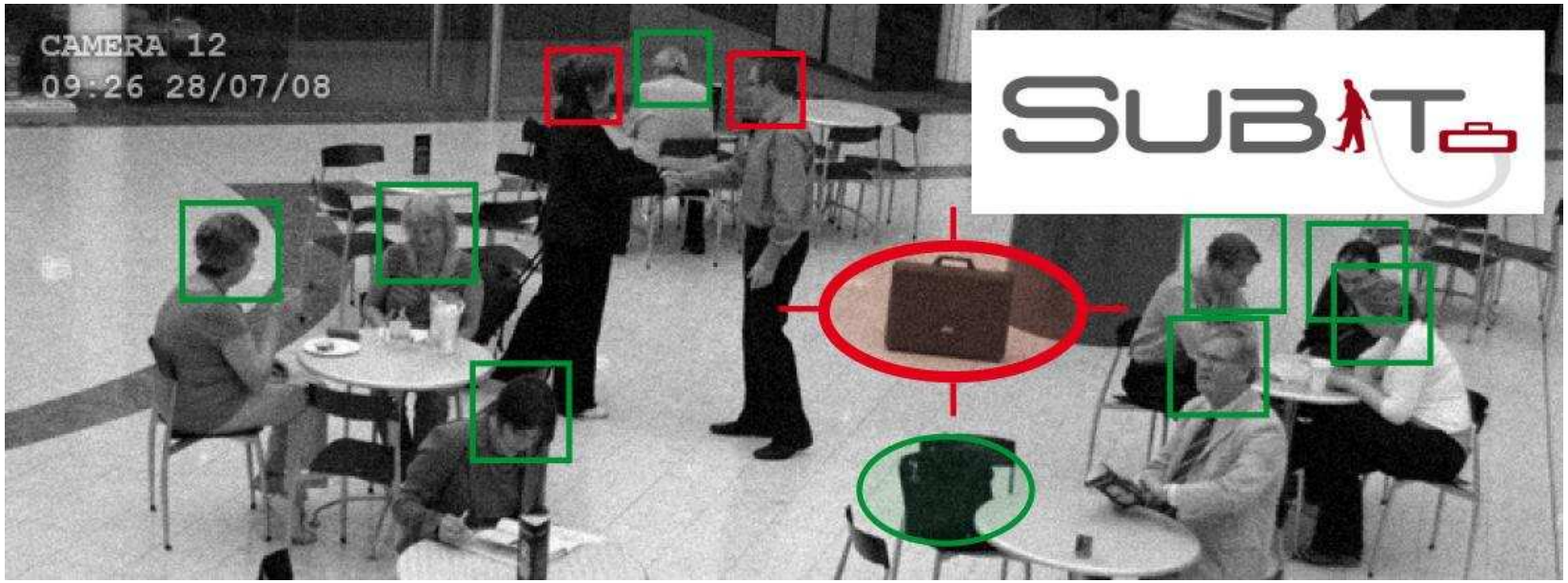


Use of PTZ in wide area surveillance

Maxime Taron, Constant Guillot, **Patrick SAYD**

Commissariat à l'Energie Atomique (CEA-LIST)
Lab Vision and Content Engineering
patrick.sayd@cea.fr

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2011) under grant agreement n° 218004.

Interest of PanTilt-Zoom cameras for Surveillance

Performance of object detection is limited in video-surveillance by the low resolution of images, because of the reduced number of sensors.

- **Pan-Tilt-Zoom devices show several advantages when compared with fixed camera devices :**
  - Monitoring very large area with a high spatial resolution.
  - Sufficient zoom capability to capture high resolution images of objects or people faces.
  - Reduce the cost of coverage for a given area in terms of number of devices.
  - May be manually controlled by operator to focus on a security incident.

High Resolution Views

But, use of PTZ cameras gives some constraints …

- **Low-cost equipment : limitation of the dynamic behaviour**
  - Uncertainty on motion (repeatability)
  - Uncertainty on zoom (repeatability)
  - Slow autofocus in case of zooming
- **Moving sensor implies constraints on processing methods**
  - Full calibration is complex
  - Approaches as background modelling for object detection developed in the static context are not easy to adapt
- **Strategy : Trade-off on zooming level**
  - Focus on an object of interest makes the PTZ blind for the rest of the scene
  - Enlarge the angle of view would reduce the high resolution capability

Which strategy to monitor large area ?

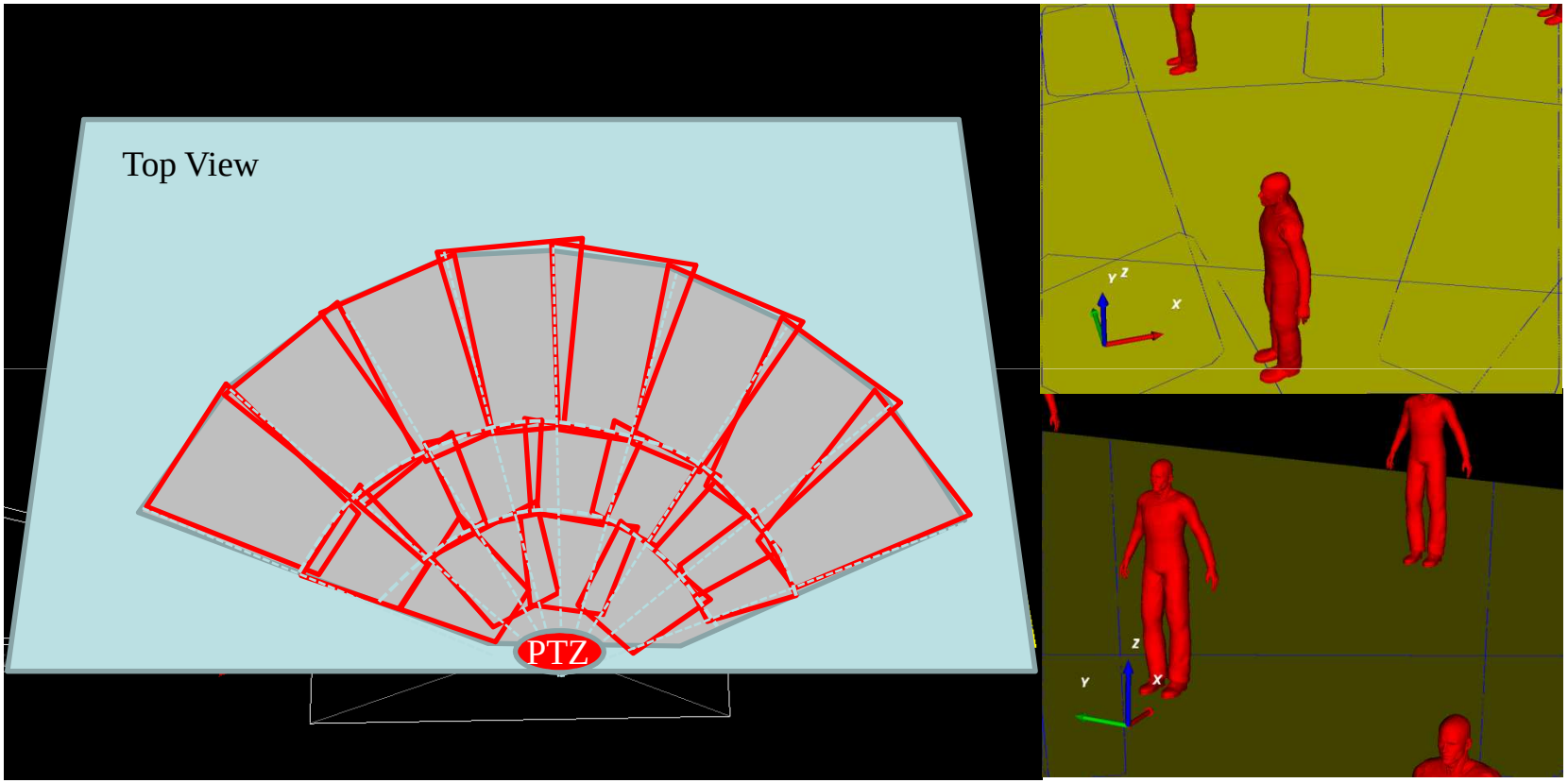
- **Classical approaches**
  - With a PTZ : the system covers the whole scene and focus on detection:
    → Detection on low resolution image and a blind system during focus
  - With a wide-angle camera and a PTZ : the wide angle camera monitors the whole scene and the PTZ focus on detection (master/slave):
    → Detection performance is depending of the low resolution sensors

- **Innovative approach in Subito: two PTZ performing guard tour for stationary object detection**
  - Guard tour : each PTZ monitors the whole space with a constant and high resolution. Repetition of a course defined by a set of orientation/zoom.
  - Classification to separate mobile and stationary objects.
  - Multi-view approach to manage occlusions and recover 3D information.

  → Limitation : constant movement of the camera makes it impossible to track moving objects.

A strategy based on Automatic Guard tour

- **Cover the space with a constant resolution**



Stationary object detection during Guard Tour

- **At each position of the Guard Tour, object detection is performed based on robust background modelling**

- **Due to Guard Tour strategy : 20” between 2 views of a same place**
  - ➢ no statistic model of the background image
    - → only frame-to-frame differencing on local descriptors.
  - ➢ robust signature of image appearance
    - → SURF descriptors (texture) along with YUV color values are computed on image blocks as a background model.

- **Discriminate mobile and stationary object (potentially abandoned object)**
  - If background signature and current signature are different ➔ **object detection**. Background and Object signatures are saved.
  - Next view : the current signature is compared to the 2 previous ones:
    - if current signature = background signature ➔ **no more object**
    - if current signature is different of both signature ➔ **mobile object**
    - if current signature = previous object signature ➔ **stationary object**

Results of stationary blob detection

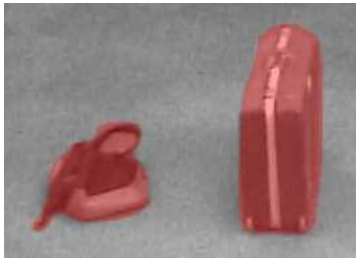
Example on one of the positions of the guard tour
Colour code : **mobile blob**, ..., stationary blob

Object labelling: single view object grouping

- **Background subtraction with re-identification.**
  - Produce local results → blobs
  - The concept of "object" is not directly present in this output.

- **Segmentation based on graph optimization form blobs in the image.**
  - Create a weighted graph with unary costs on nodes and binary costs on edge.
  - Input : background subtraction results (age of every detection)
  - Output : a map of labels corresponding to detected objects.

  - Unary cost :
    - include a minimal duration which defines an object as static.
    - several additional constraint for the labeling of blocks
  - Binary cost:
    - Include time difference between descriptor.



Results of object labelling

Multi-view : Object matching

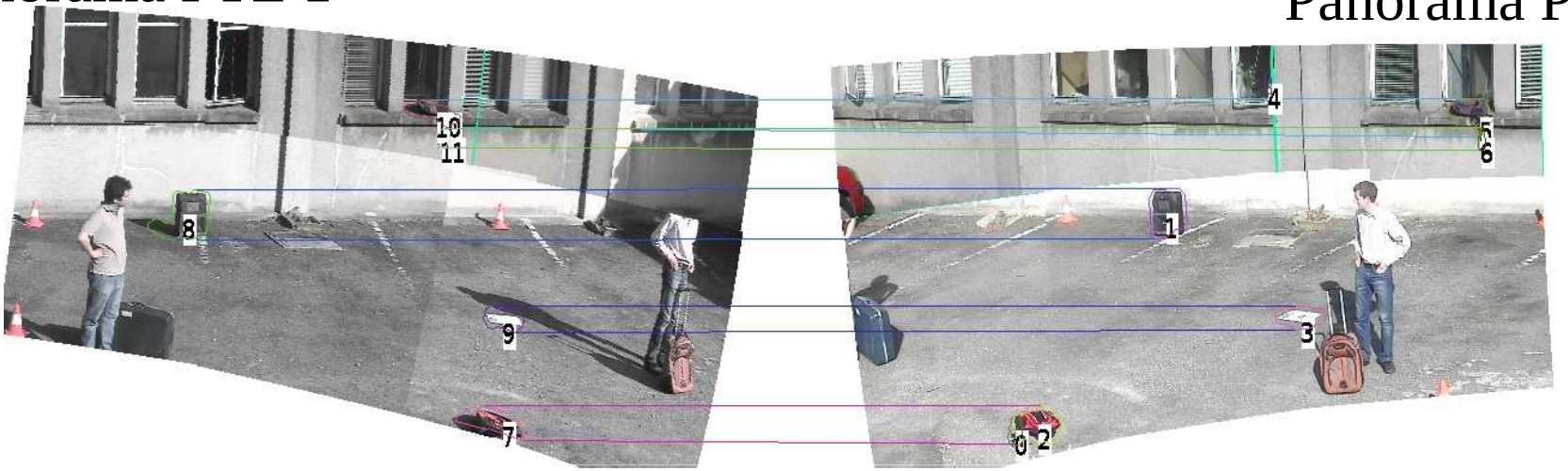
- **The use of two PTZ enables:**
  - the limitation of ambiguities due to occlusions
  - the recovery of 3D information

- **The challenge is the merge of information between the different views.**
  - A fast calibration of a couple of PTZ was developed, which gives a geometric relationship between images. Epipolar constraints reduce the matching possibility and enable the 3D reconstruction.
  - A strategy based on oriented graphs is proposed to match objects between views

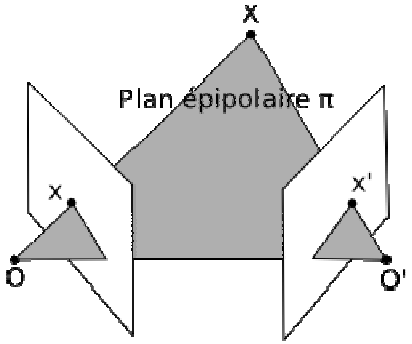
Object matching : Epipolar Gometry.

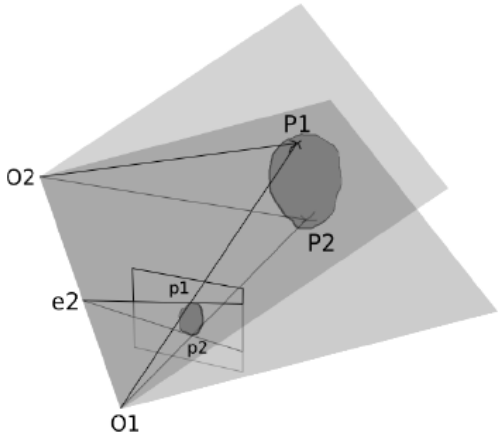
Panorama PTZ 1

Panorama PTZ 2



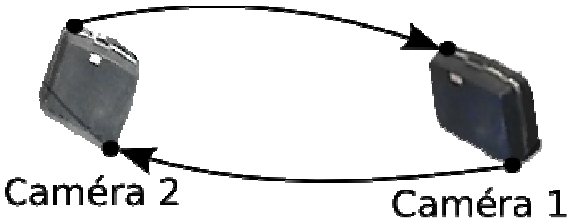
- Epipolar planes intersect each detected object of the 2 cameras.
- The epipolar planes tangent to the « top » and « bottom » tip of each object are named “frontier planes”. The corresponding points of each cameras the “frontier points”



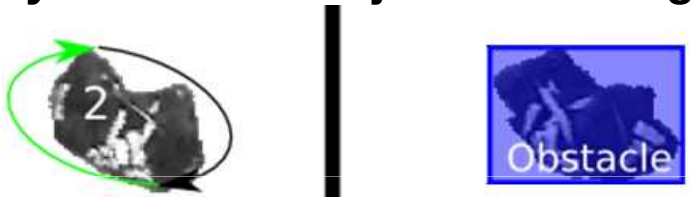


Graph construction for matching object.

- **Simple link 2 objects on 2 views**



- **Multiple scenarios are likely to complexify multi-view object matching.**
  - Object is being viewed in only one image (no possible match)

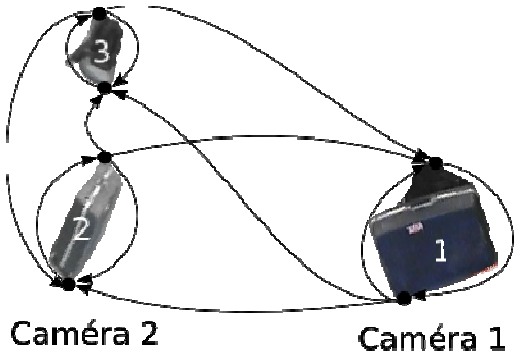


  - 2 Object are seen as one detection in one of the camera and two detection in the other.

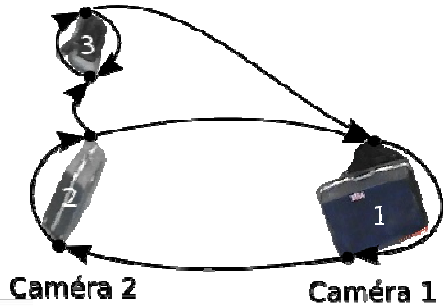
  - An object is seen as 2 detection in one camera because of an occlusion

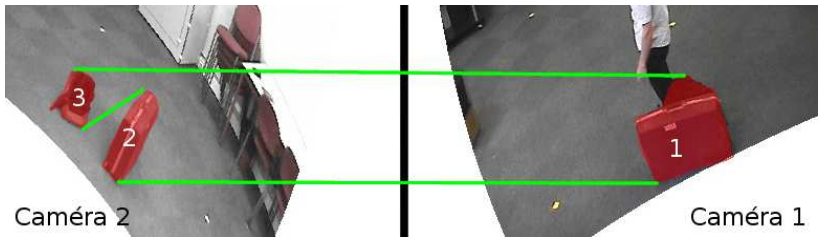
Optimization

- **An oriented graph links all objects of the two views is built over the single view detections.**



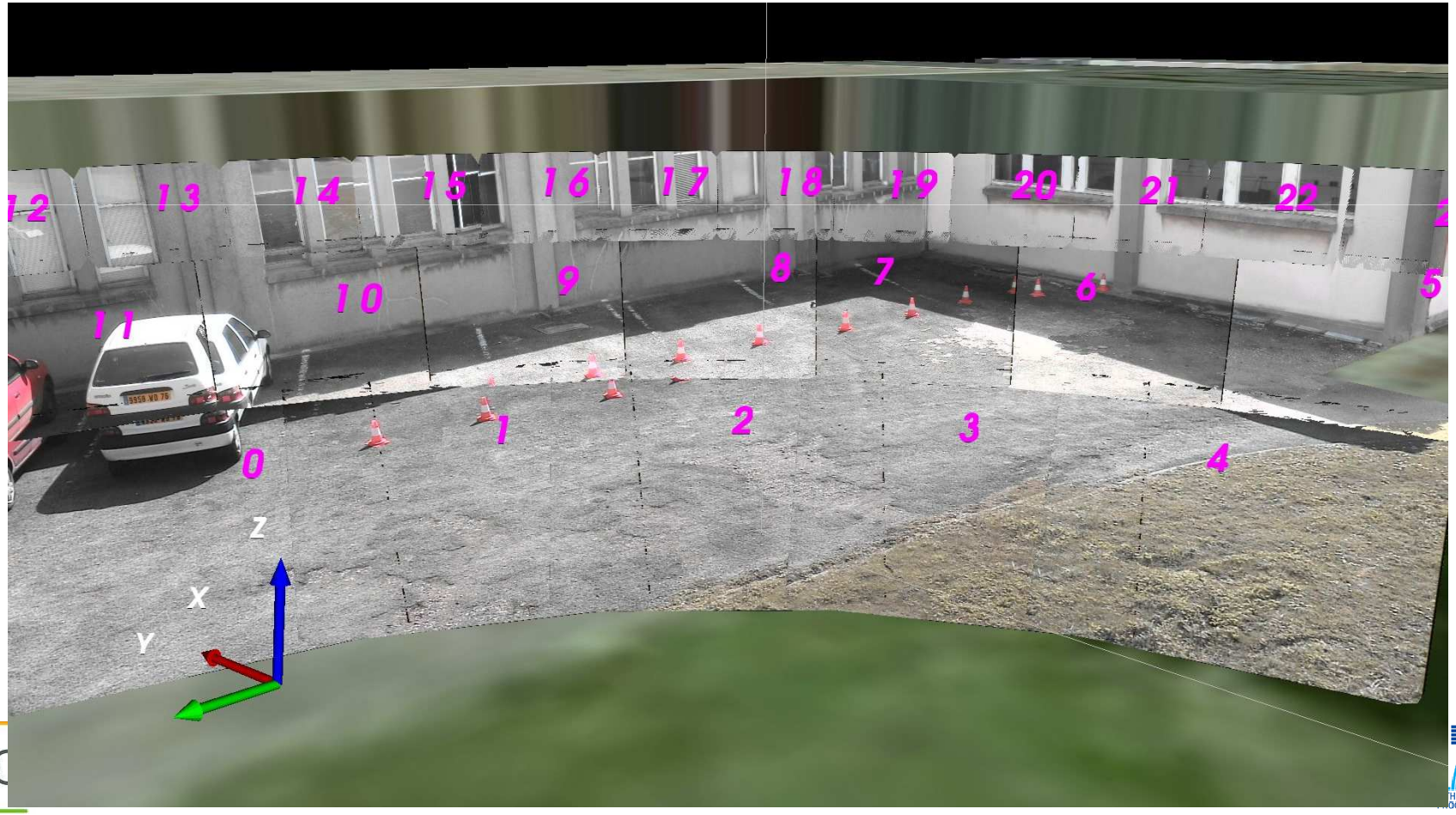
- **Every edge is given a cost as a function of the epipoloar constraint and object size.**
- **The “optimal object matching” solution is the set oriented cycle with minimal cost.**





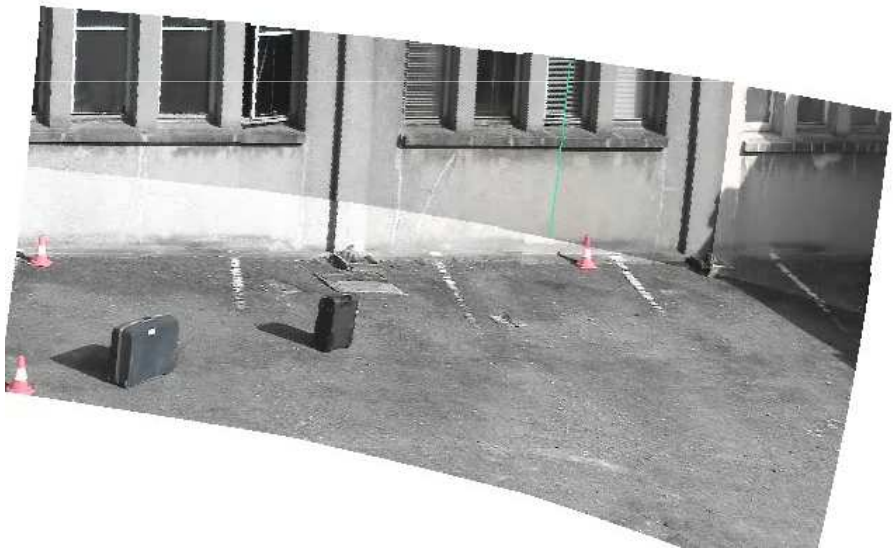
Test site in CEA faciities

- **2 PTZ Cameras at 5m high**
- **Guard tours with 22 preset views.**
- **1 frame of the same area every 30s**
- **The overlapping area covered is 10x20x2m**



Results of object association.

- Object detection tested on 8 preset views
- Sequence about 20’ = 5000 images
- 20 stationnary baggage → 6000 occurences



Evaluation of the Multiview approach

Detection rate

	Single camera	Couple of Camera
Recall	0.95	0.95
Precision	0.86	0.92

False alarms rate = mainly due to system latency (duration of the guard tour)

Reconstruction accuracy

Association	Real Size (cm)	Estimated Size (cm)
8 - 0	53	57
7 – 1	51	51
9 – 2	1	9
10 – 3	20	21
12 – 4	15	15
11 - 5	13	7

Conclusion

In SUBITO, we show that PTZ is relevant for large area surveillance:

- The Guard tour strategy is a way to monitor the area with a high resolution
- Stationary object detection is made possible thanks to:
  - a robust descriptor for background/foreground
  - an object labelling based on a graph approach, weighted by age information
- Multi-view-Approach reduces the level of false alarm even if easy improvements need to be implemented.

Future works:

- Robustness of the PTZ system must be improved to cope with real surveillance environment
- The PTZ system must be integrated into a global framework for unattended baggage detection, in order to collaborate with owner identification and tracking functions.

Thank your for your attention