Pedestrian Detection and Localization

Members:

Đặng Trương Khánh Linh 0612743

Bùi Huỳnh Lam Bửu 0612733

Advisor:

A.Professor Lê Hoài Bắc

UNIVERSITY OF SCIENCE

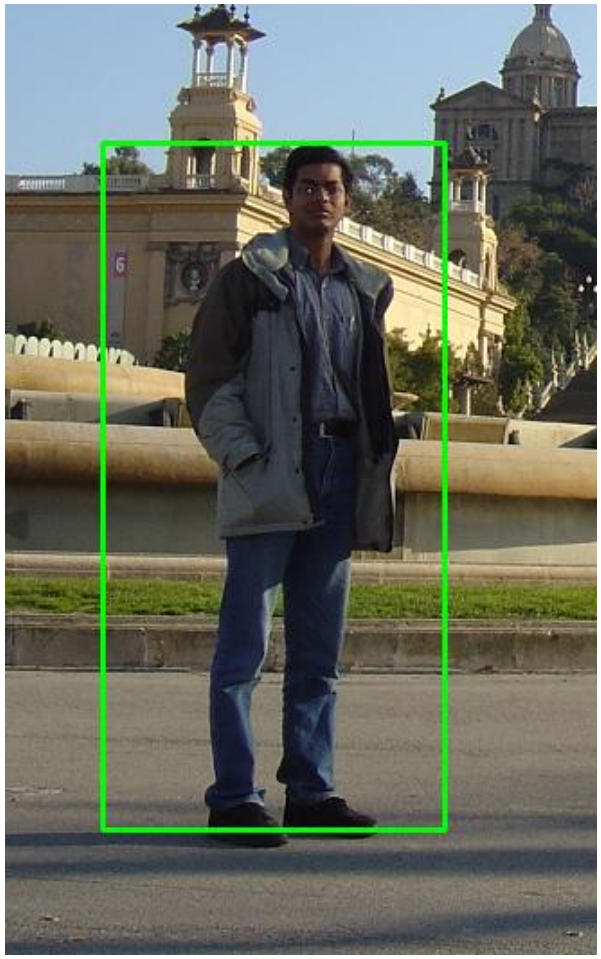
ADVANCED PROGRAM IN COMPUTER SCIENCE

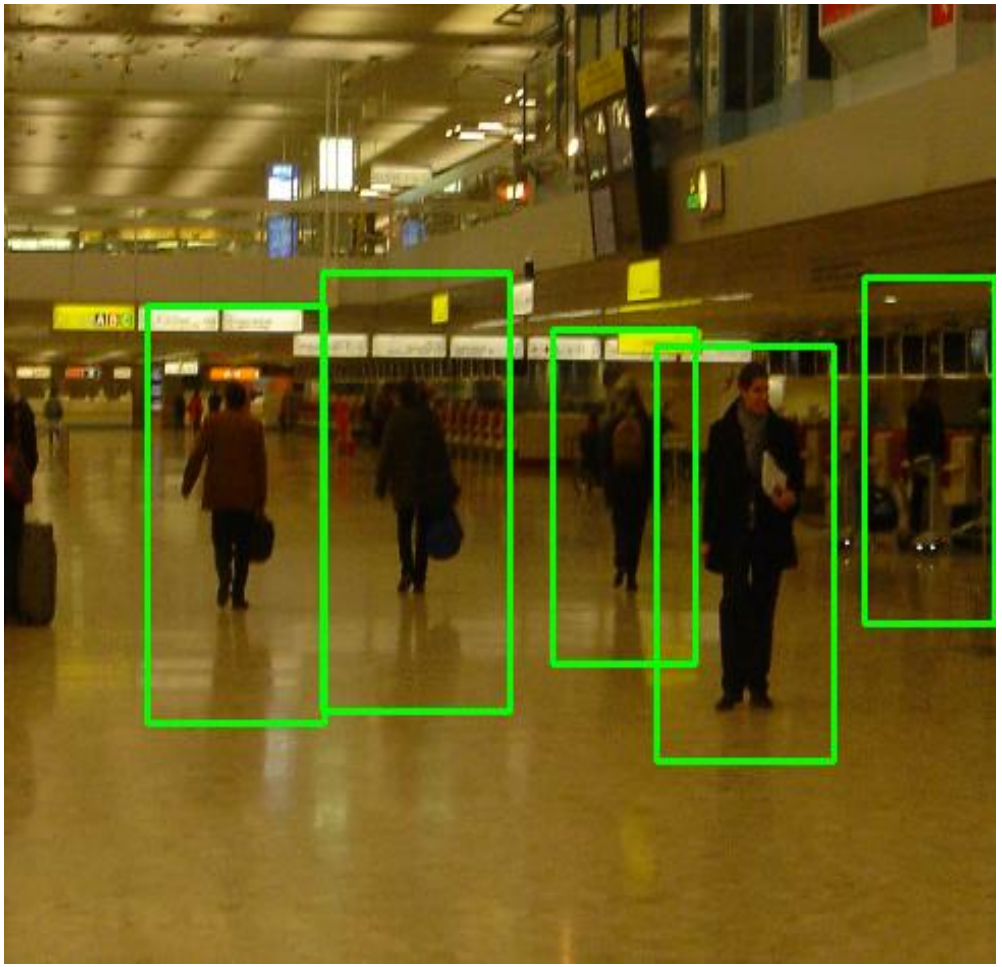
Year 2011

Problem statement

- Build up a system which automatically detects and localizes pedestrians in static image.
- Constraints:
  - Pedestrians stand up and fully visible people.
  - Size of pedestrian is not less than 64x128 pixels.

Some examples





Applications

- Using in smart car system, or smart camera in general.
- Build a software to categorize personal album images to proper catalogue.

Challenges

- Huge variation in intra-class.
- Variable appearance and clothing.
- Complex background.
- Non-constraints illumination.
- Occlusions, different scales.

Outline

- Existing approaches.
- Motivation.
- Overview of methodology.
  - Learning phase
  - Detection
- Some contributions:
  - Four regions based approach
  - Multi-level based approach
- Non-maxima Suppression
- Conclusions
- Future work
- Reference

Existing approaches

- Haar wavelets + SVM: Papageorgiou & Poggio, 2000; Mohan et al 2000
- Rectangular differential features + adaBoost: Viola & Jones, 2001
- Model based methods: Felzenszwalb & Huttenlocher, 2000; Loffe & Forsyth, 1999
- Lowe, 1999 (SIFT).
- LBP, HOG, …

Motivation of choosing HOG

- The blob structure based methods are false to object detection problem.
- Use the advantage of rigid shape of object.
- Low complexity and fast running time.
- Has a good performance.

Contributions

- Re-implement HOG description.
- Spatial Selective Method.
- Multi-level Method.

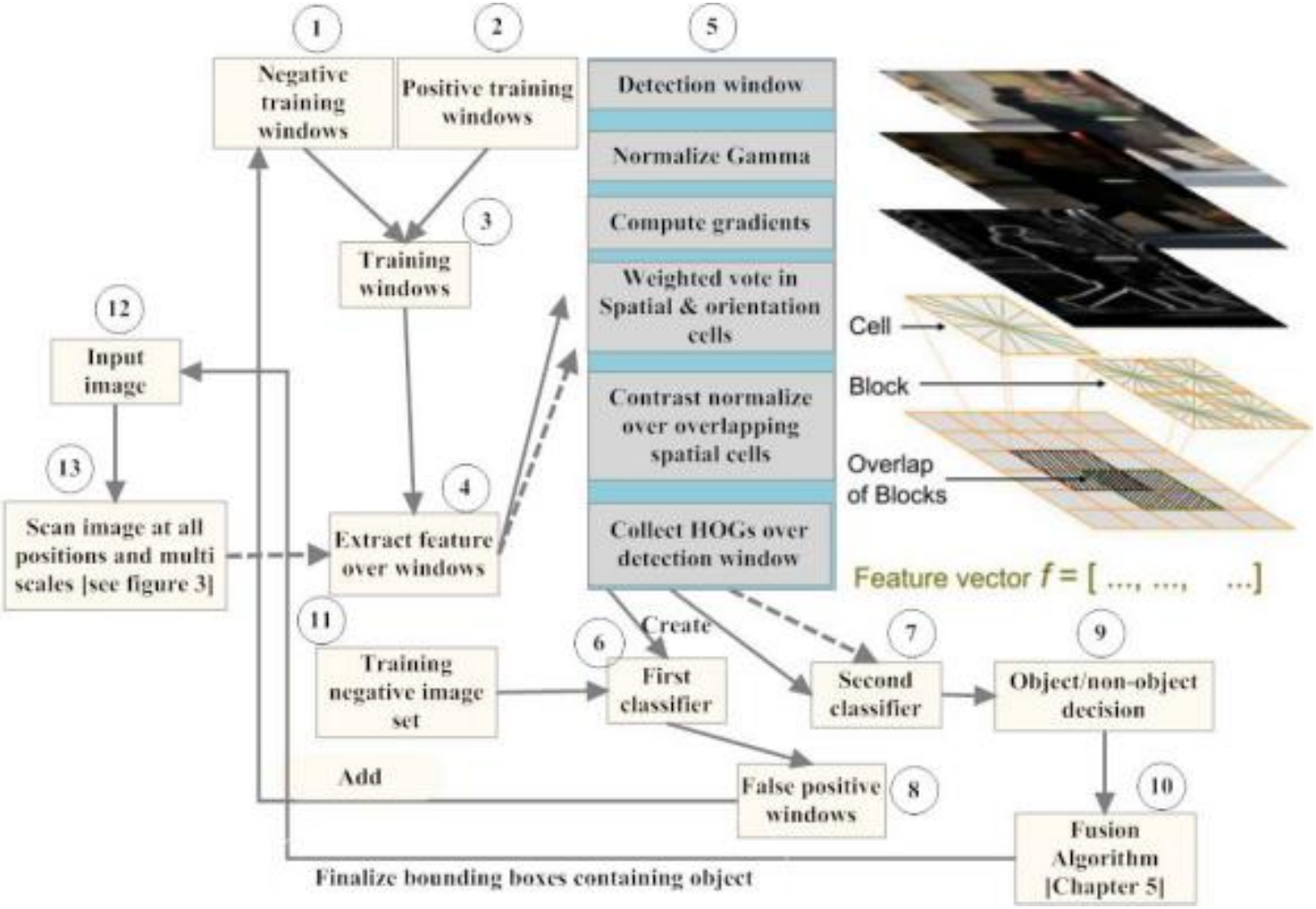
Dataset

INRIA pedestrian dataset

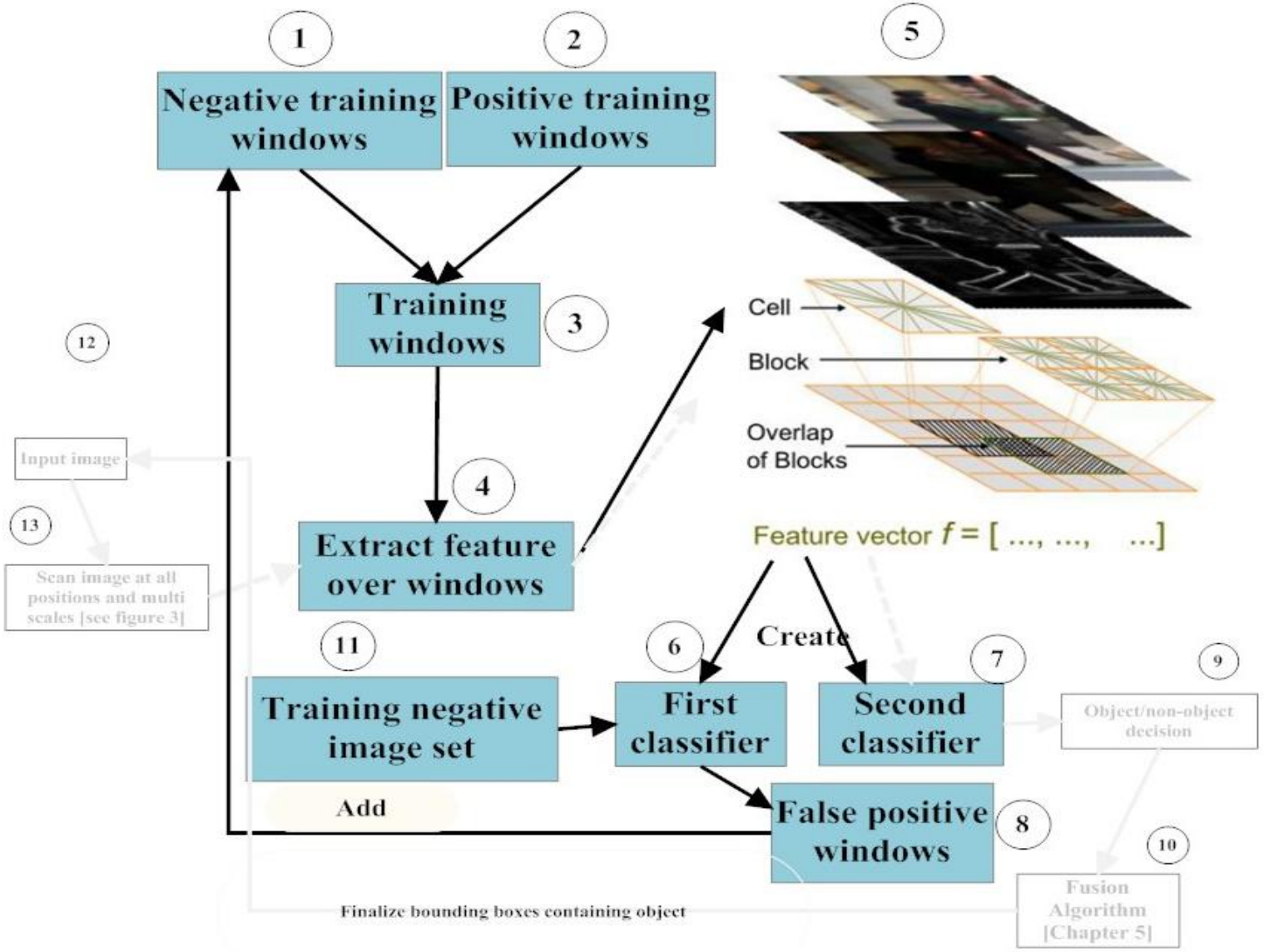
Train:
1208 positive windows
1218 negative images

Test:
566 positive windows
453 negative images

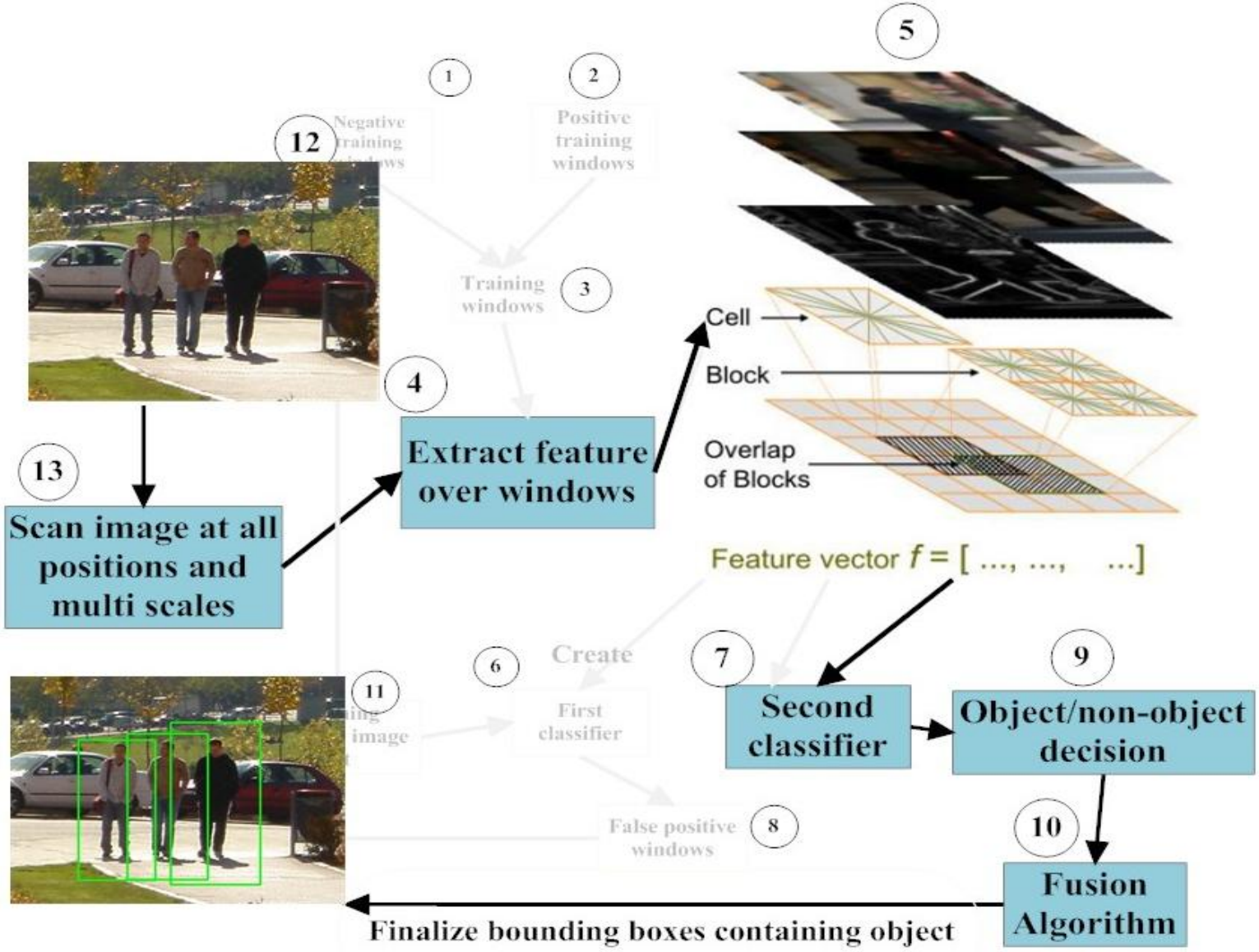
Overview of methodology



Learning Phase

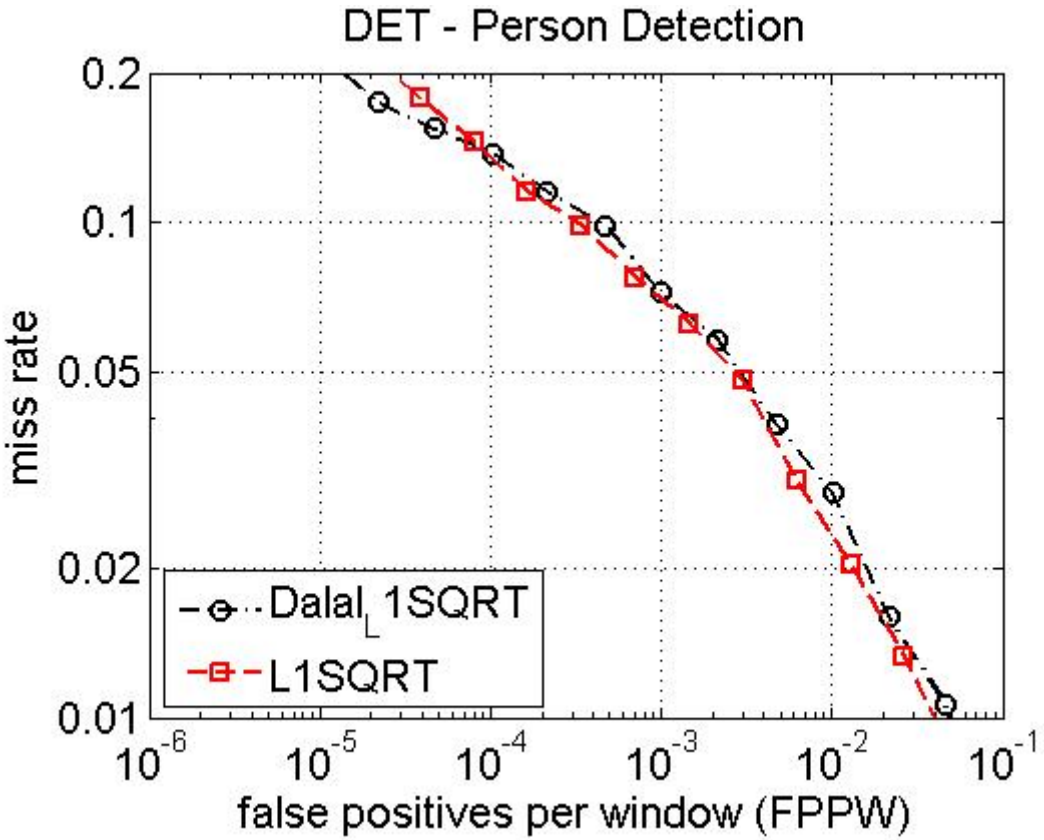


Detection Phase

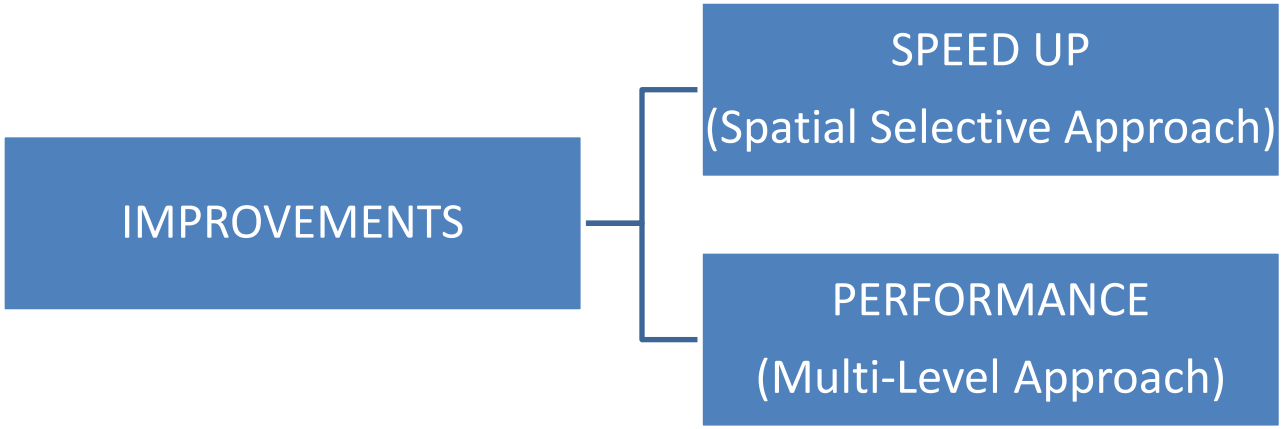


Scan image at all positions and scales

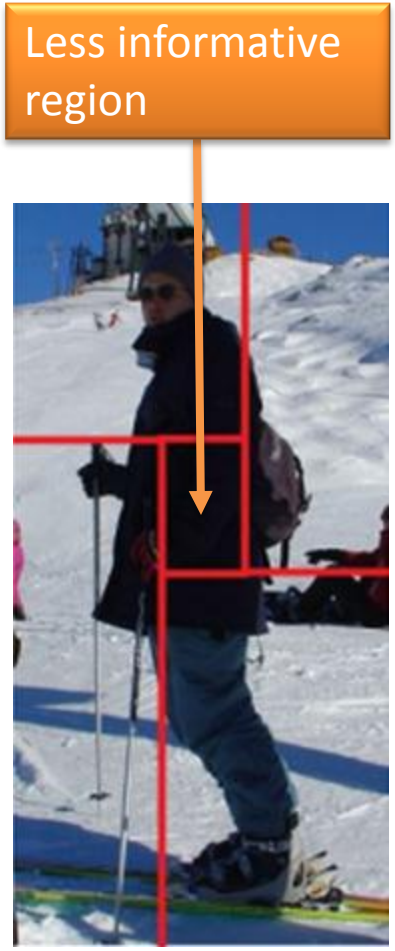
Result of experiment

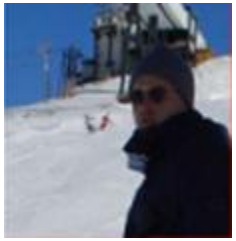


Some Contributions



Spatial Selective Approach



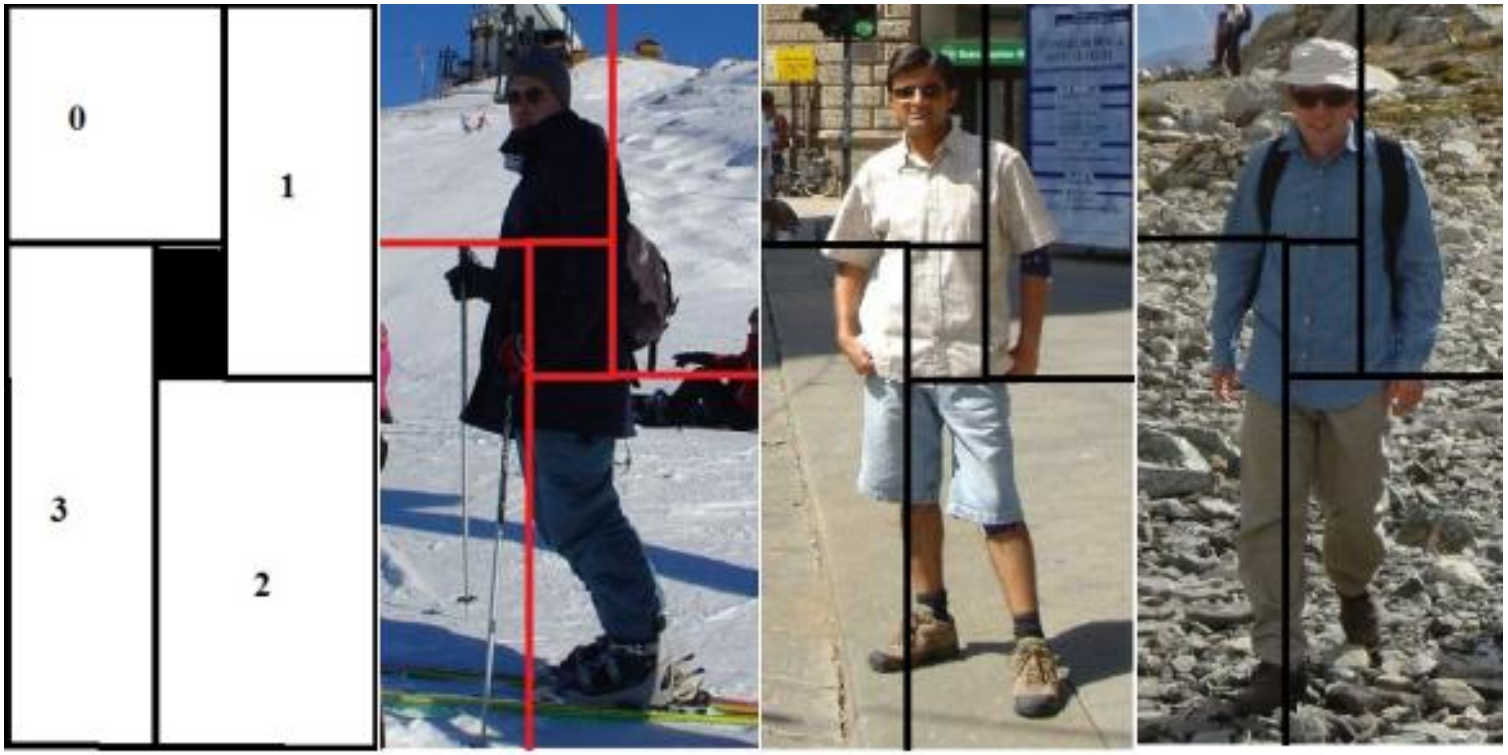


[A1,..,Z1] [A2,..,Z2] [A3,..,Z3] [A4,..,Z4]

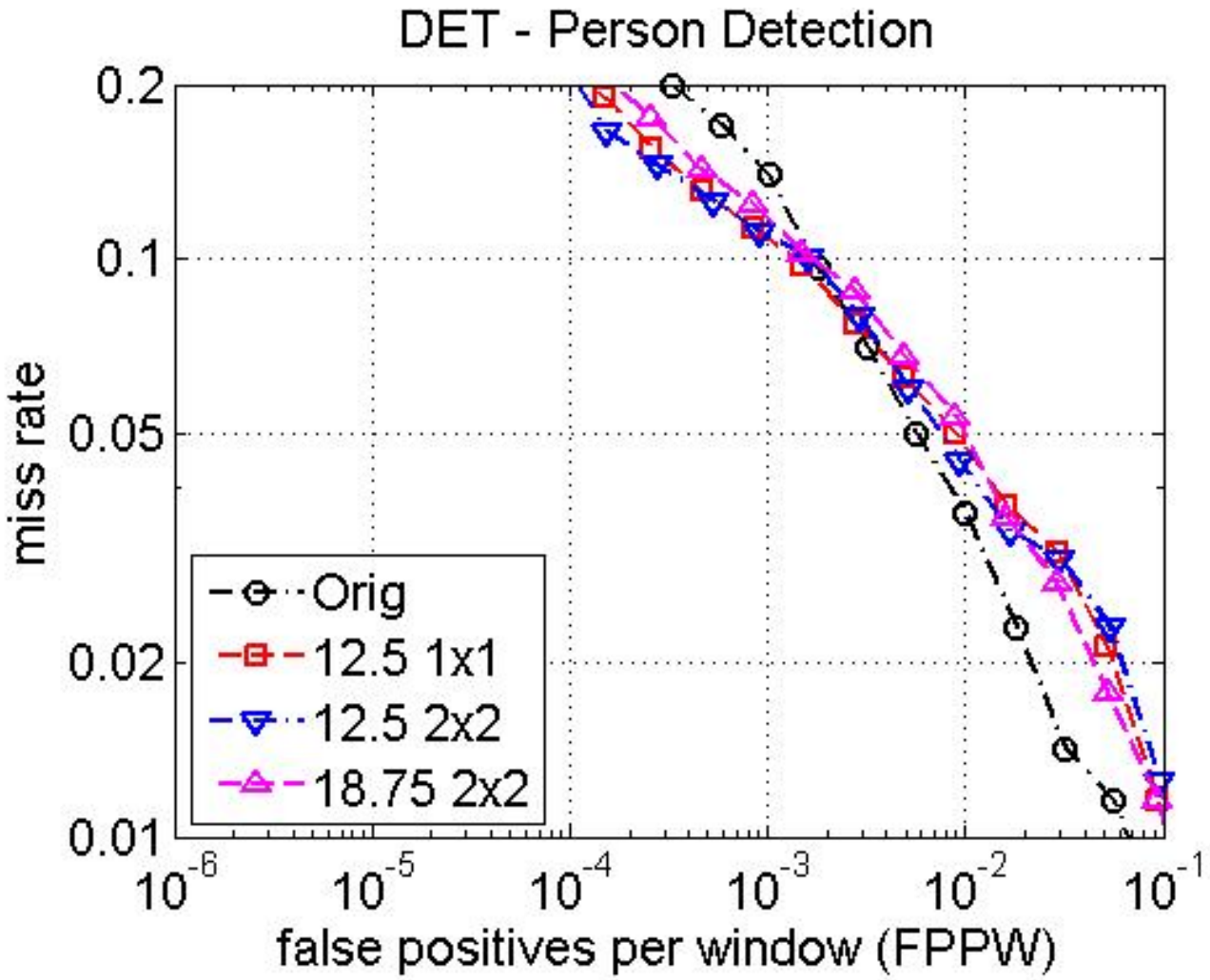
[A1,..,Z1, A2,..,Z2, A3,..,Z3, A4,..,Z4]

Spatial Selective Approach

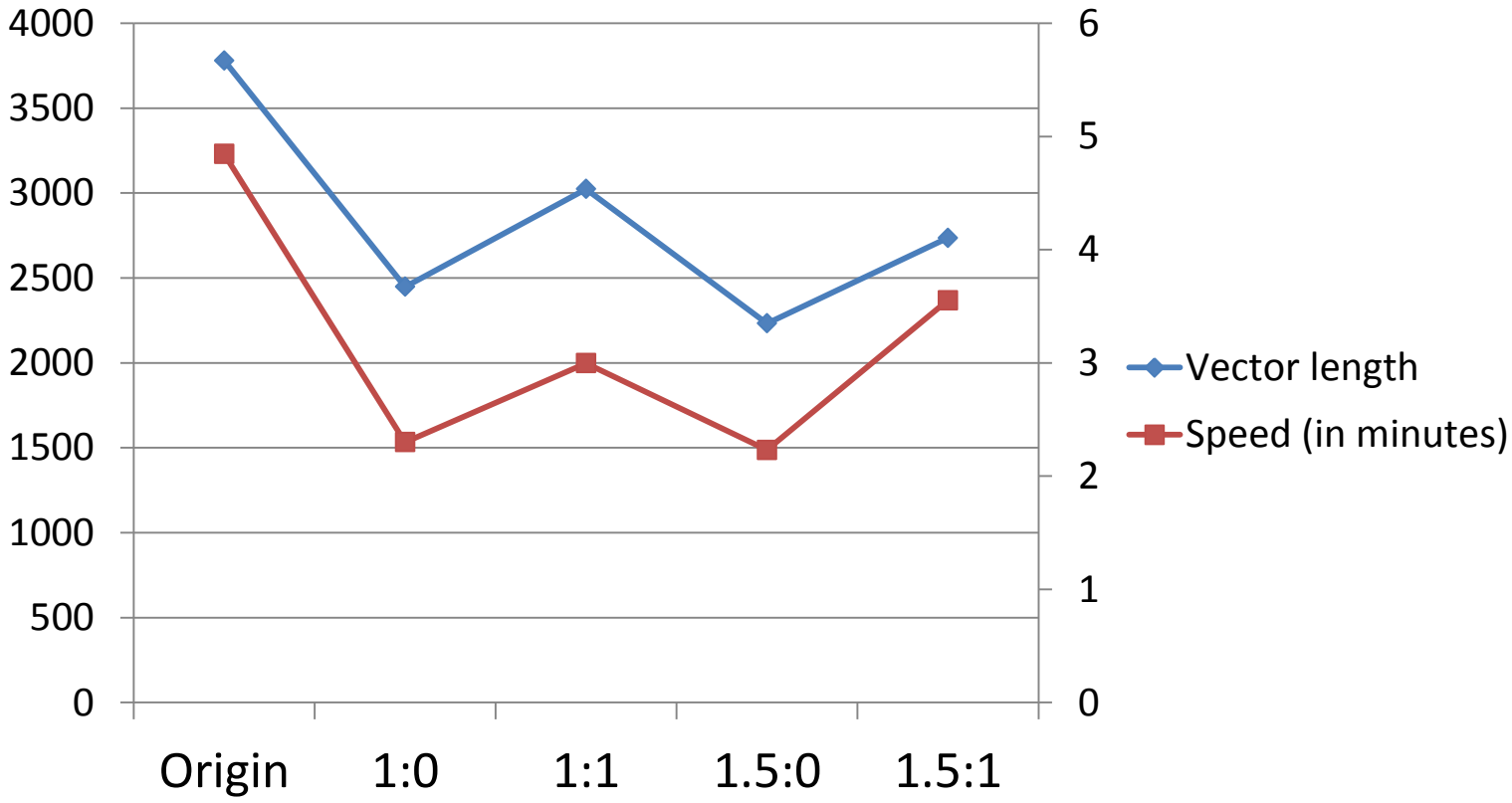
- Examples:



Result


DET - Person Detection
0.2
0.1
0.05
0.02
0.01
miss rate
Orig
12.5 1x1
12.5 2x2
18.75 2x2
10^-6
10^-5
10^-4
10^-3
10^-2
10^-1
false positives per window (FPPW)

Vector Length v.s Speed

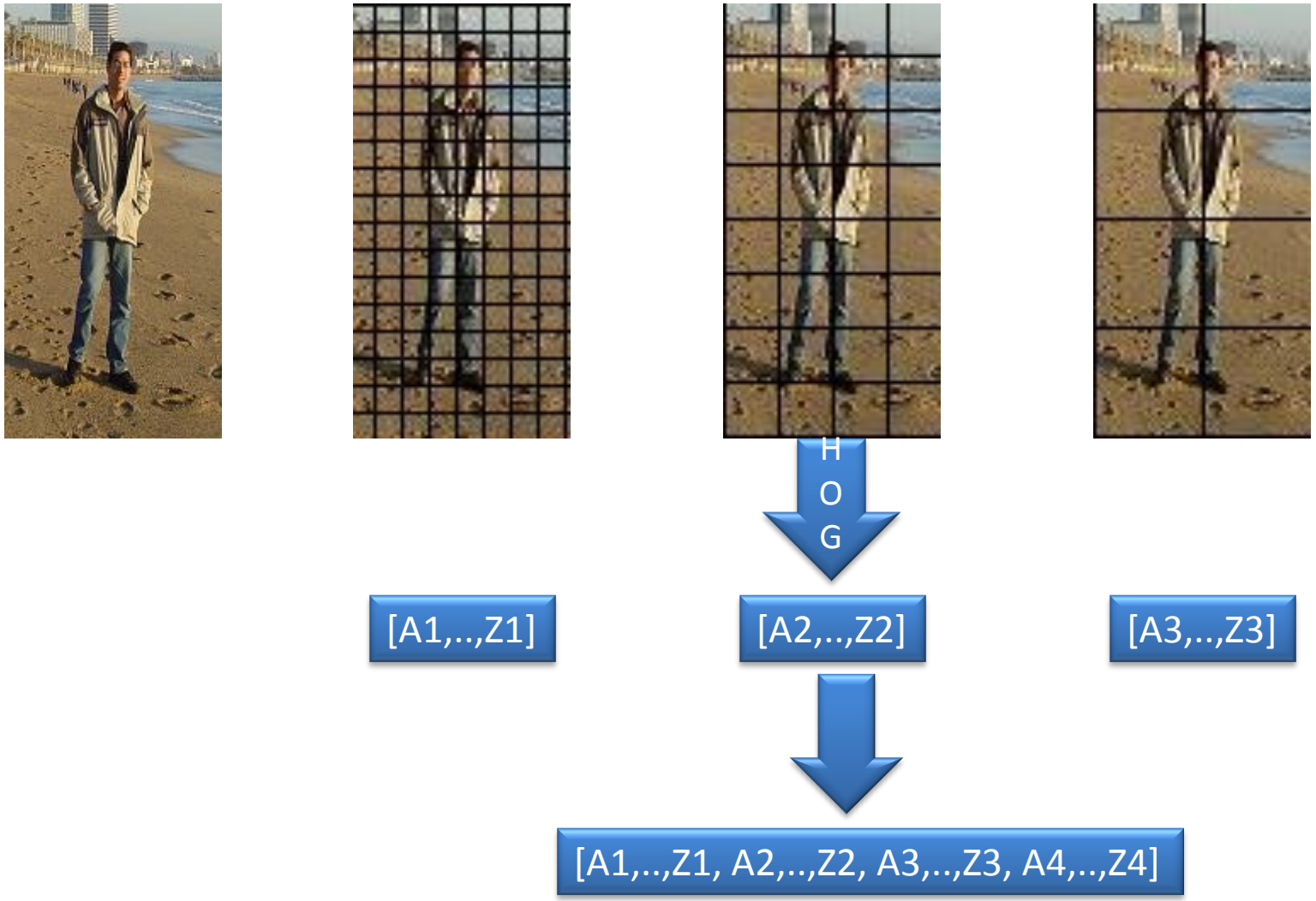


A:B → Deleted cell(s): Overlap cell(s)

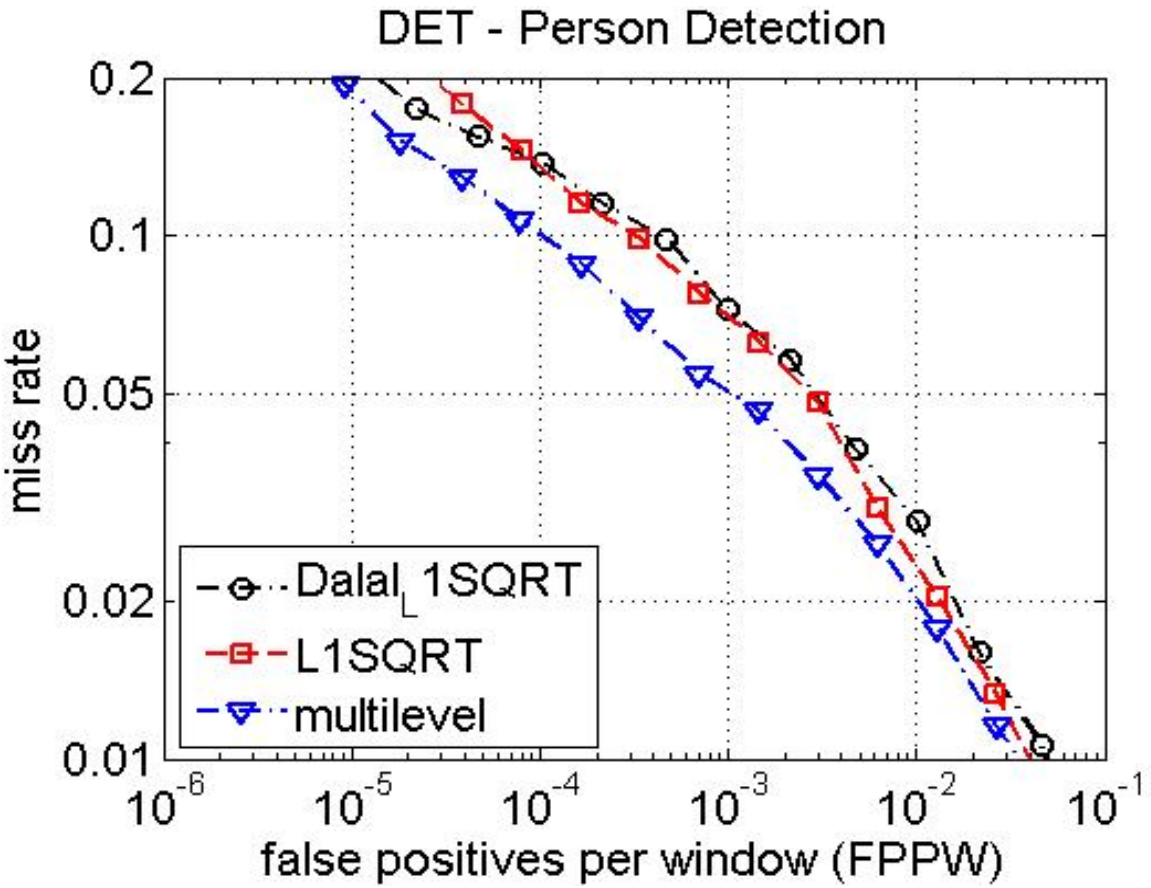
Multi-level Approach

- Purpose: enhance the performance by getting more information about shape and contour of object.

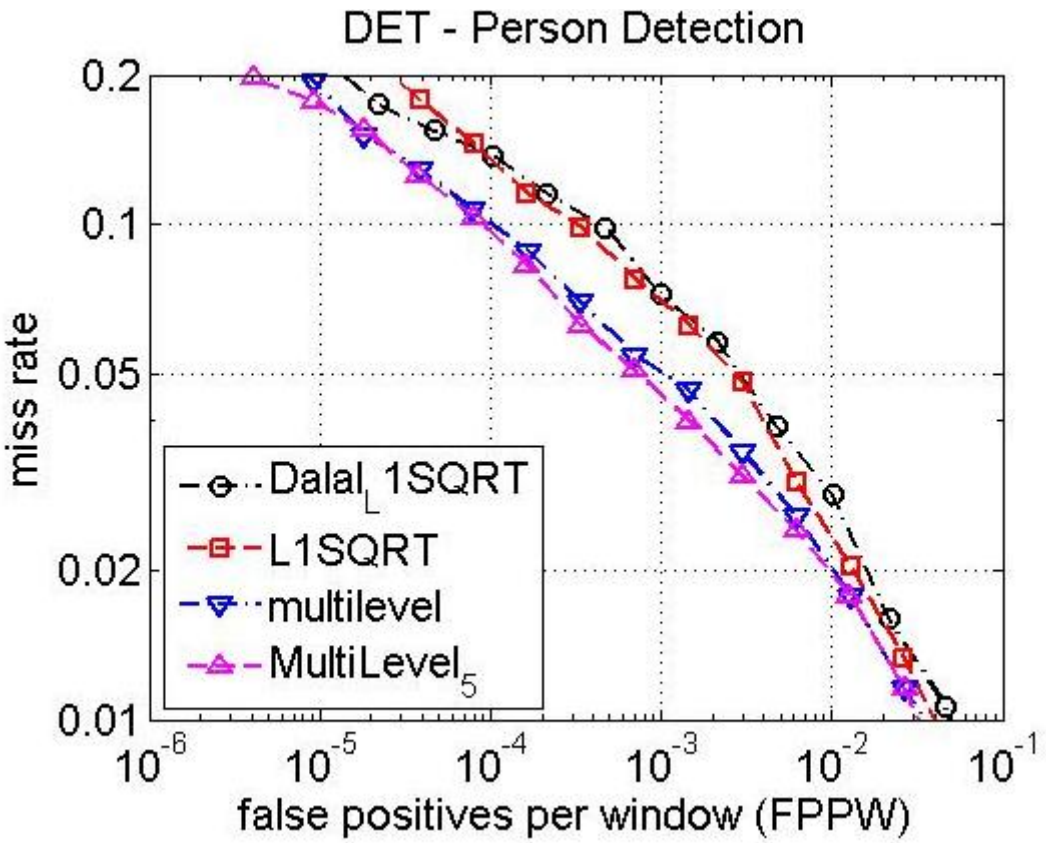
Multi-level Approach



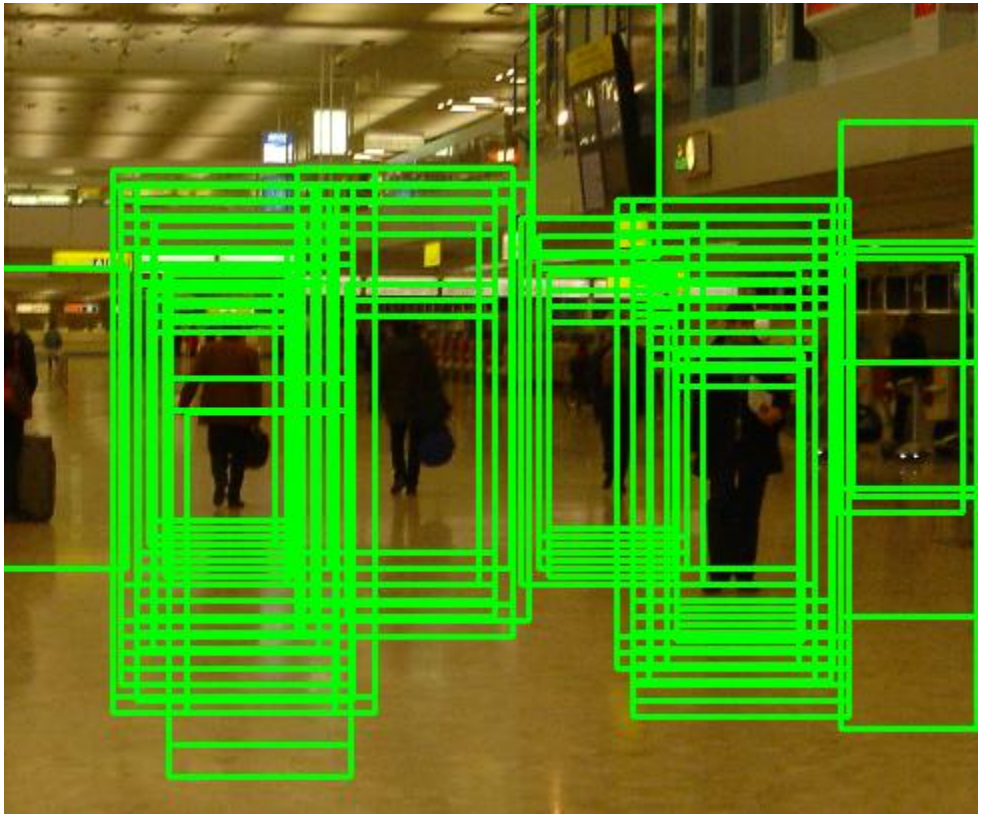
Result


DET - Person Detection
miss rate
0.2
0.1
0.05
0.02
0.01
10^{-6}
10^{-5}
10^{-4}
10^{-3}
10^{-2}
10^{-1}
false positives per window (FPPW)
Dalal_L 1SQRT
L1SQRT
multilevel

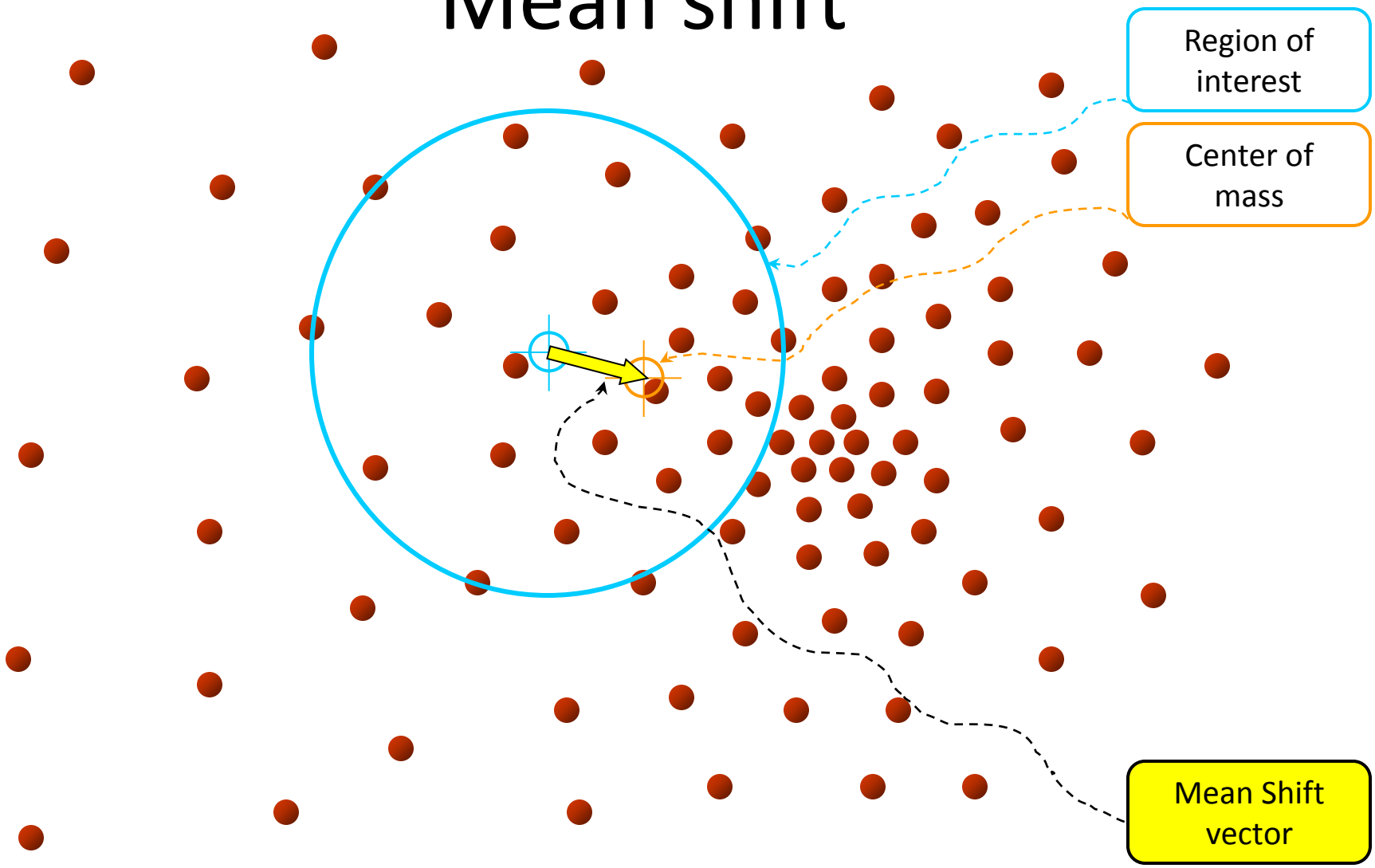
Result(cont…)


DET - Person Detection
miss rate
false positives per window (FPPW)
0.2
0.1
0.05
0.02
0.01
10^-6
10^-5
10^-4
10^-3
10^-2
10^-1
Dalal_L 1SQRT
L1SQRT
multilevel
MultiLevel_5

Mean Shift as Non-maxima Suppression

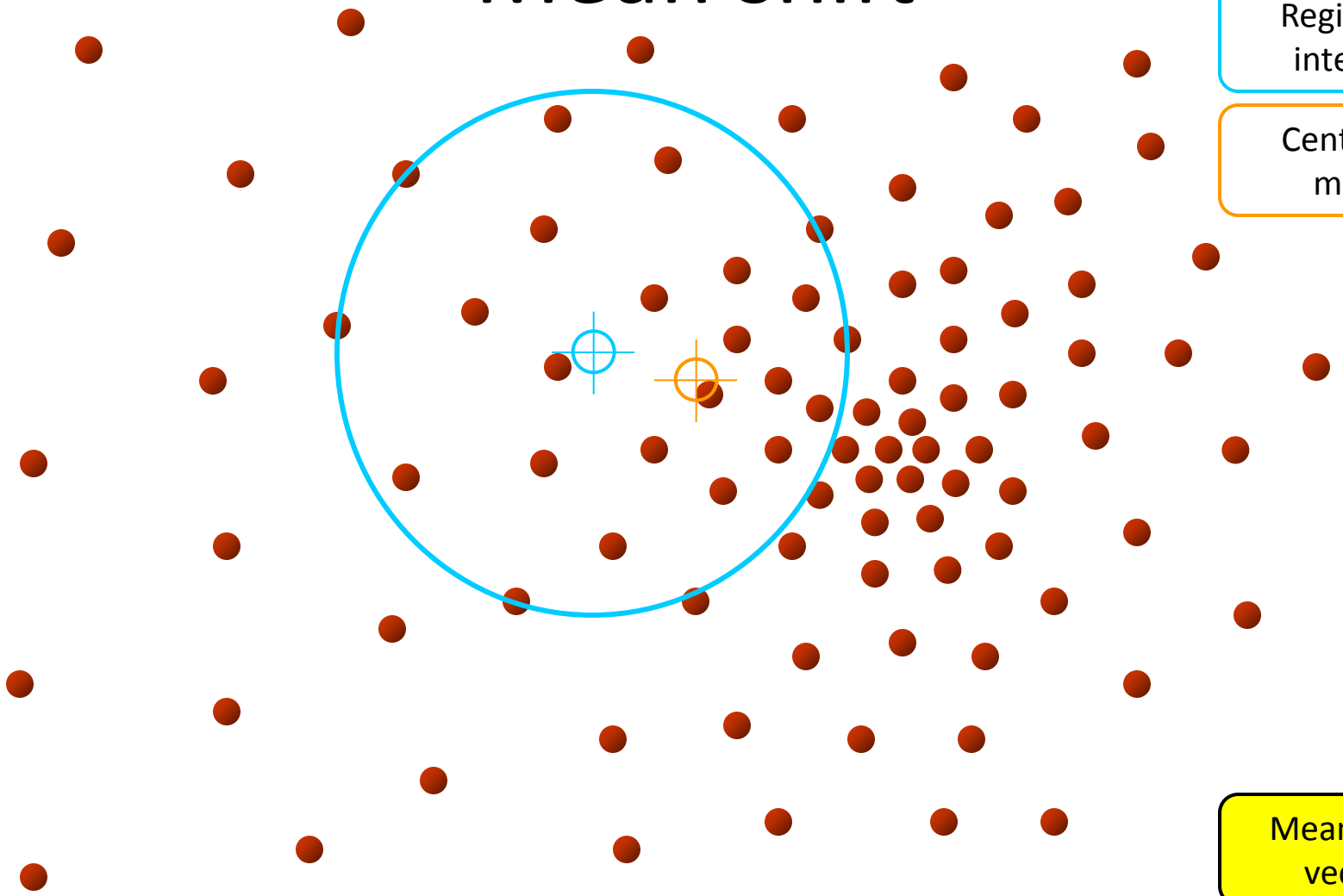


Mean shift



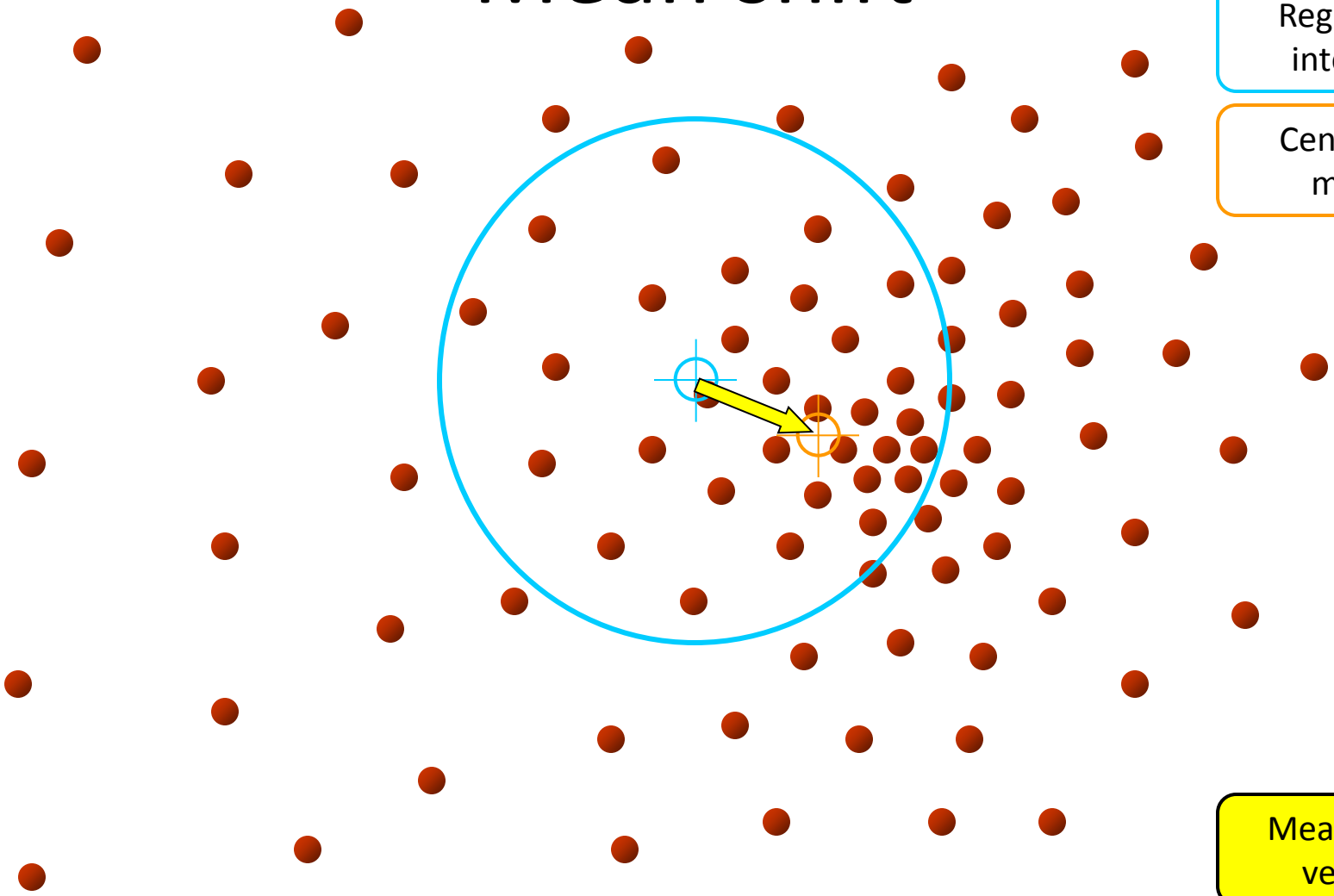
Slide by Y. Ukrainitz & B. Sarel

Mean shift



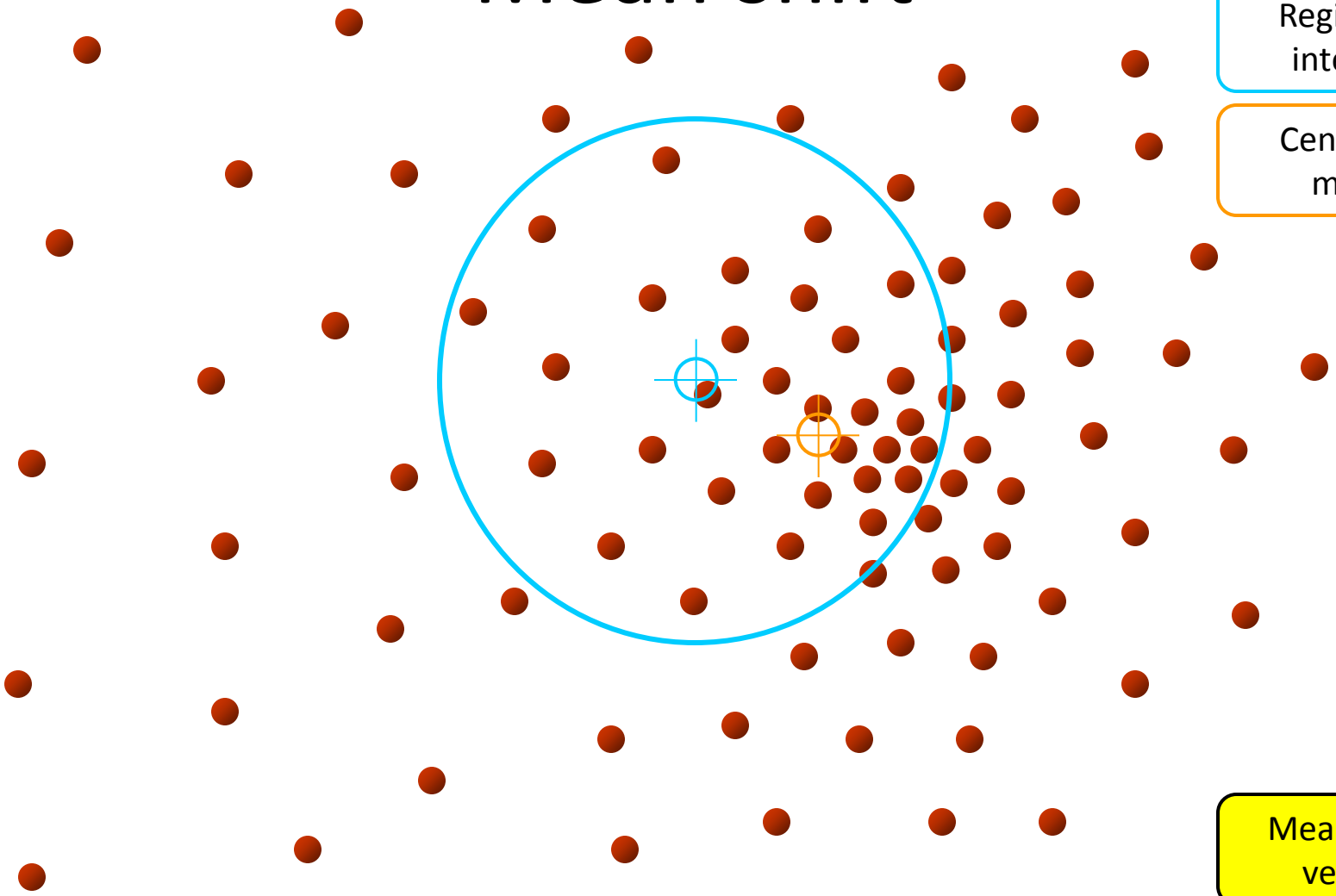
Slide by Y. Ukrainitz & B. Sarel

Mean shift



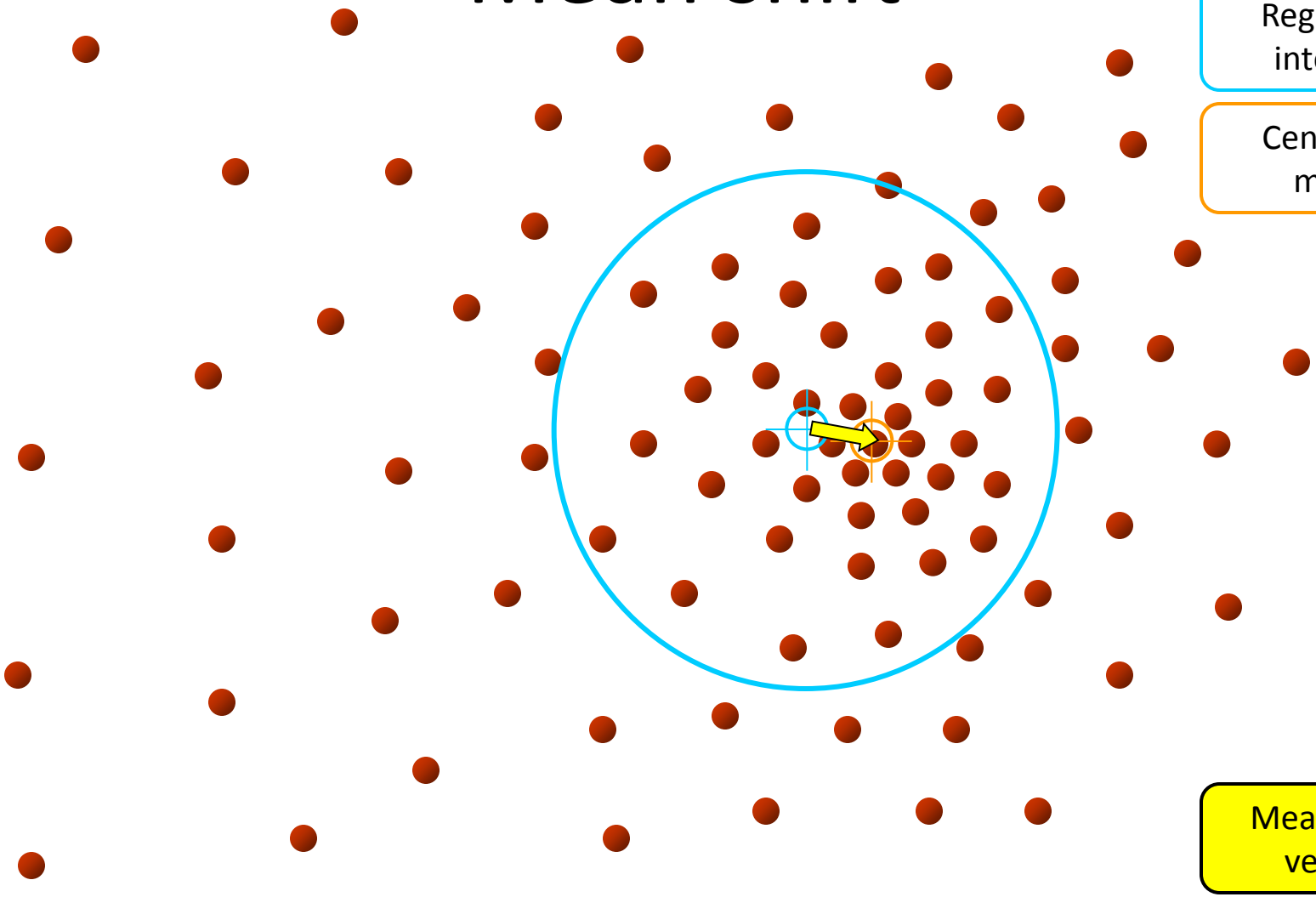
Slide by Y. Ukrainitz & B. Sarel

Mean shift



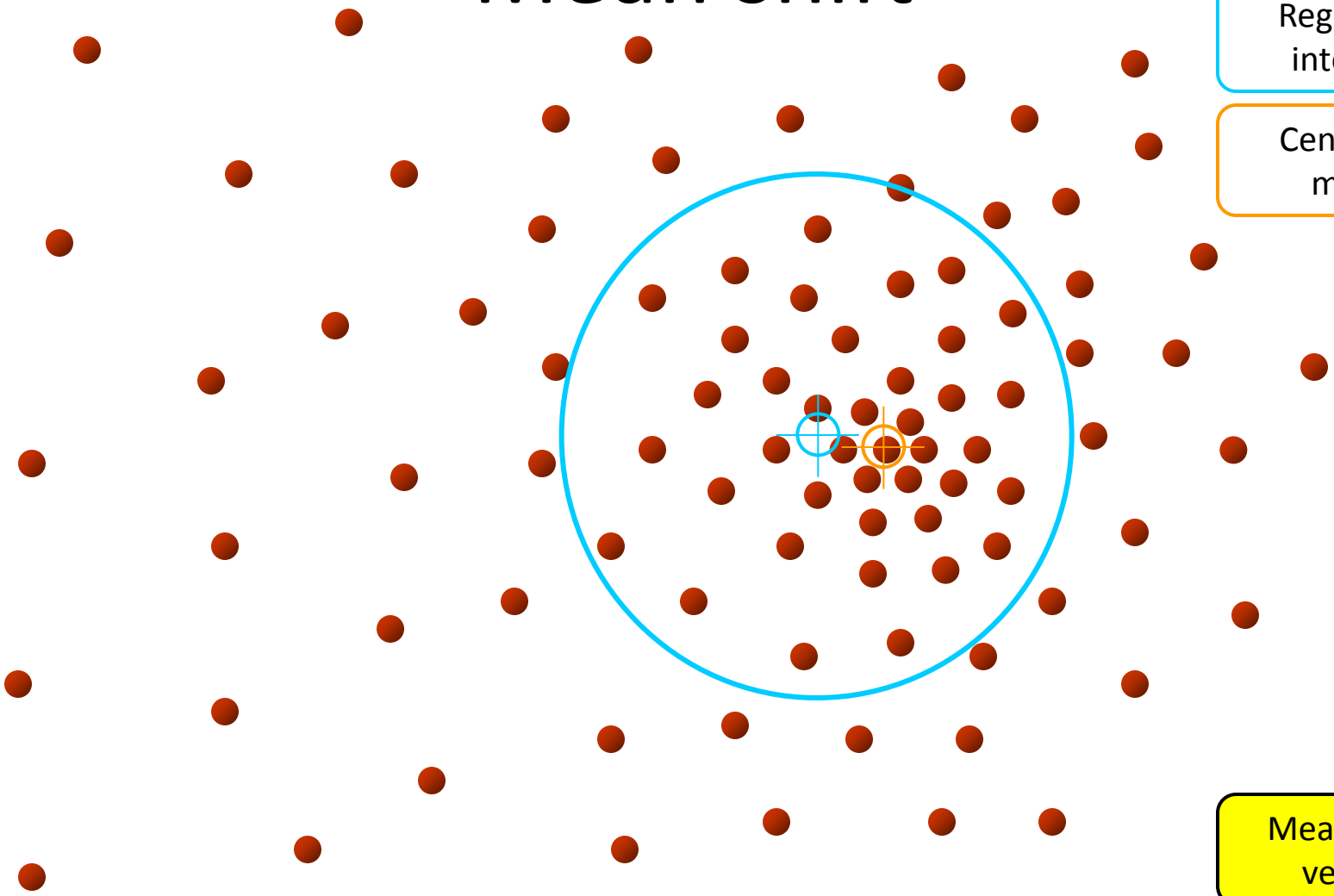
Slide by Y. Ukrainitz & B. Sarel

Mean shift



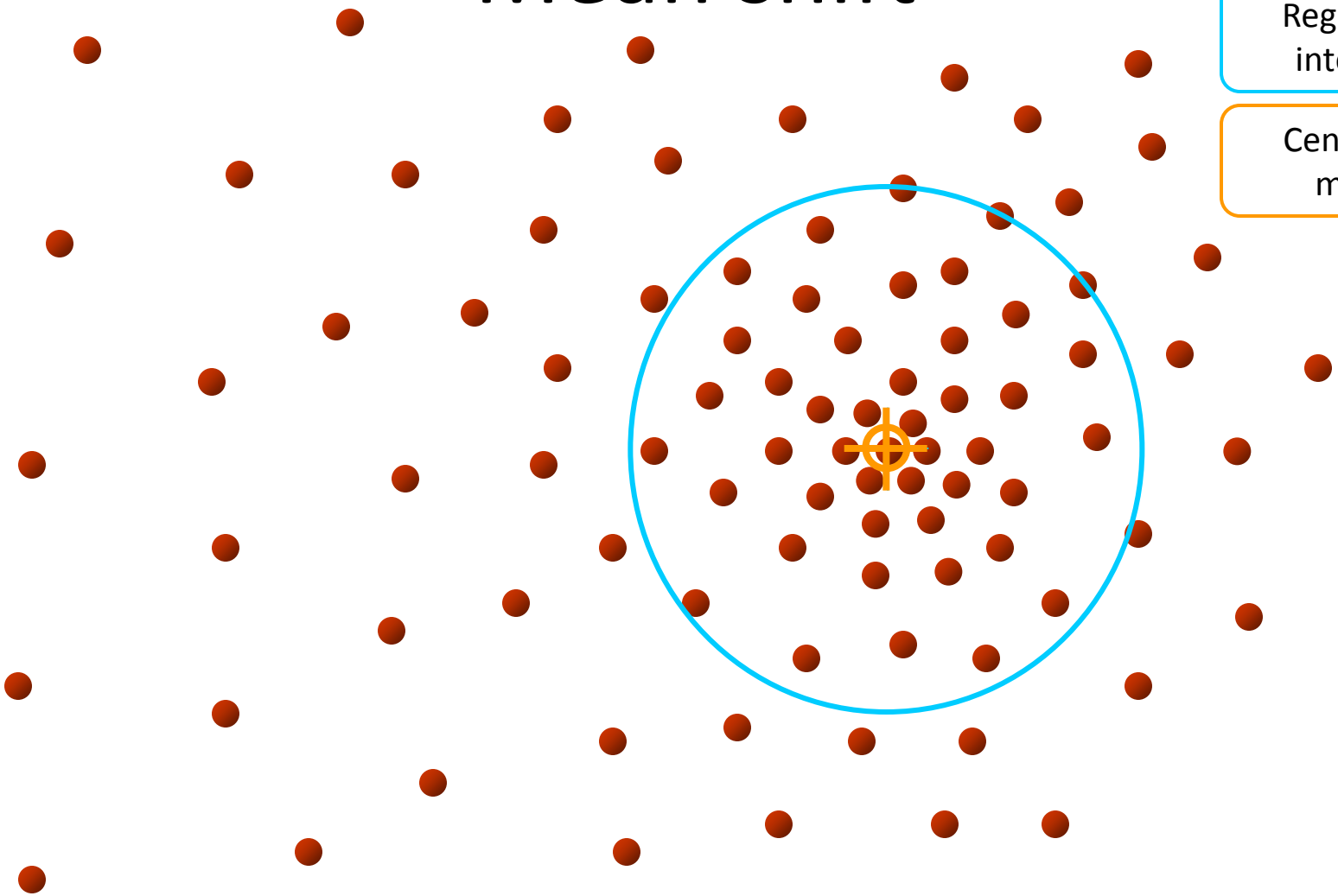
Slide by Y. Ukrainitz & B. Sarel

Mean shift



Slide by Y. Ukrainitz & B. Sarel

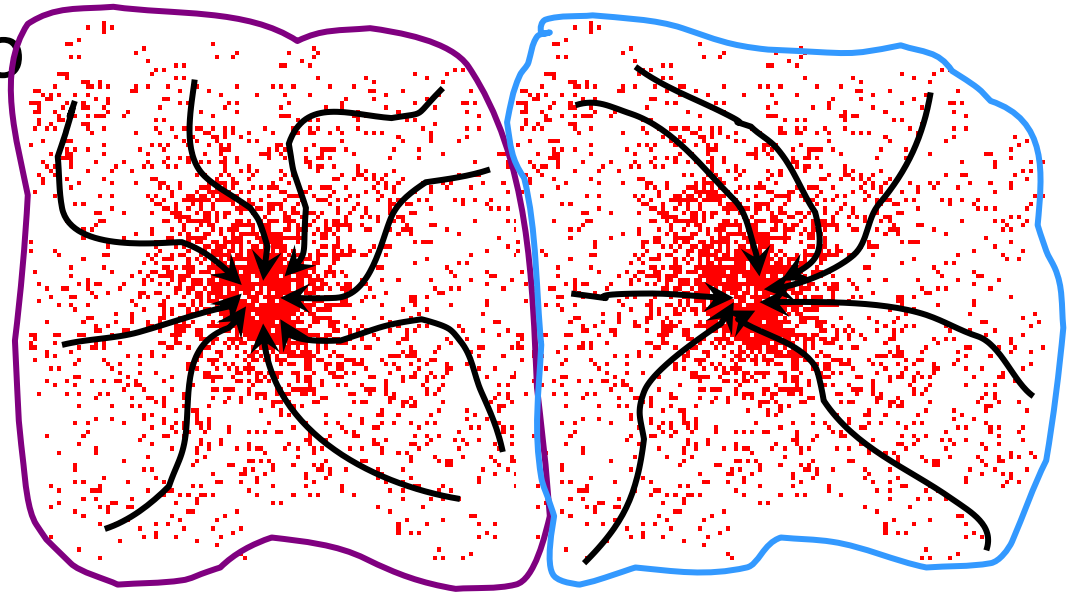
Mean shift



Slide by Y. Ukrainitz & B. Sarel

Mean shift clustering

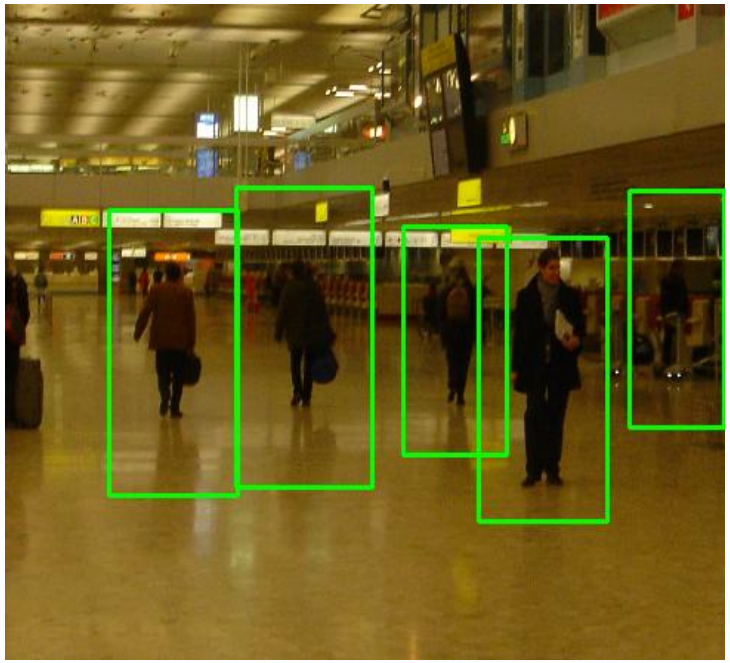
- Cluster: all data points in the attraction basin of a mode
- Attraction basin: the region for which all trajecto



Slide by Y. Ukrainitz & B. Sarel

Non-maximum suppression

- Using non-maximum suppression such as mean shift to find the modes.



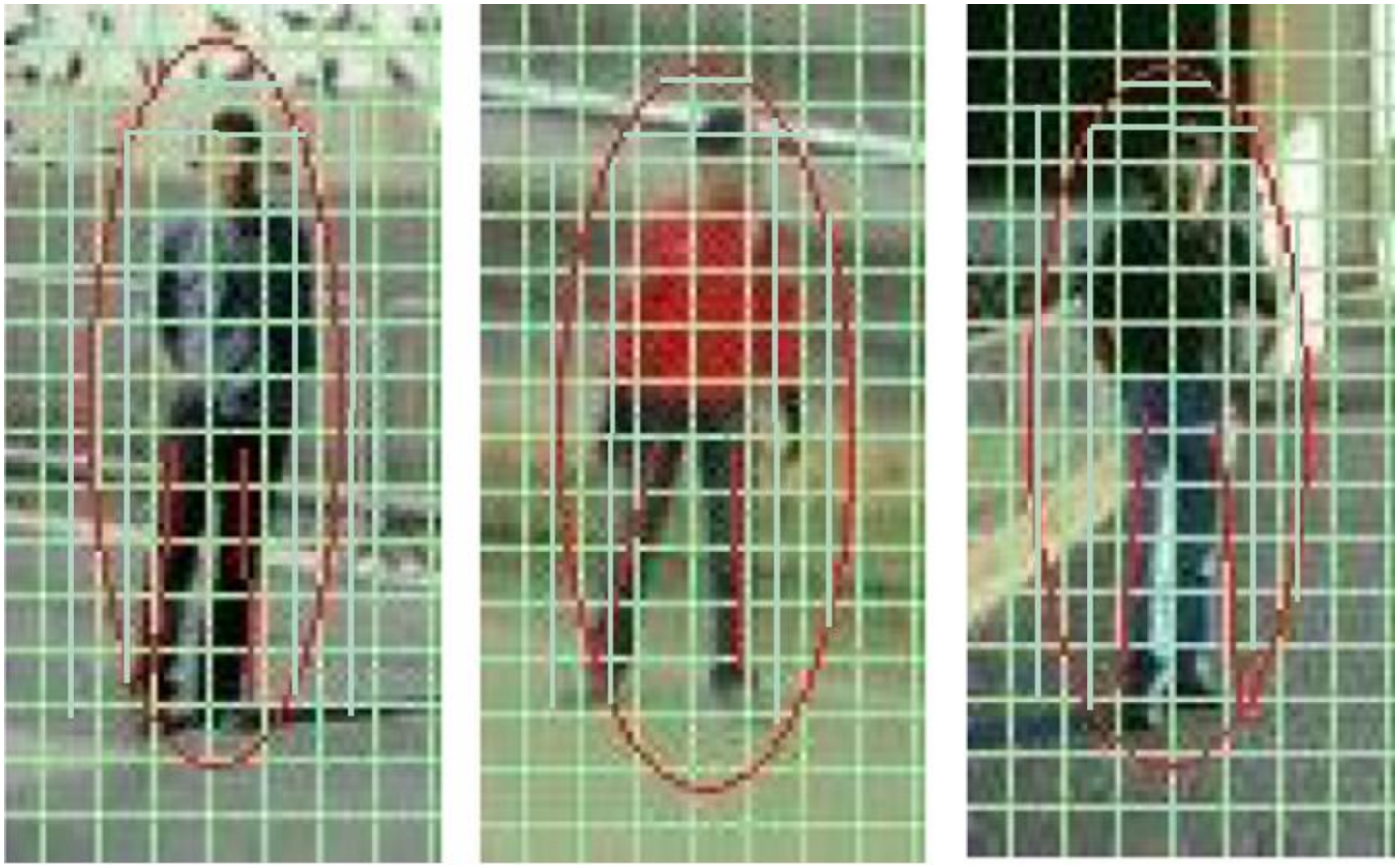
Conclusions

- Successfully re-implement HOG descriptor.
- Propose the Spatial Selective Approach which take advantages of less informative center region of image window.
- Multi-level has more information about shape and contour of object.

Future work

- Non-uniform grid of points.
- Combination of Spatial Selective and Multi-level approach.

Non-uniform grid of points



References

- N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection,” in IEEE Conference on Computer Vision and Pattern Recognition, 2005.
- Subhransu Maji et al. Classification using Intersection Kernel Support Vector Machines is Efficient. IEEE Computer Vision and Pattern Recognition 2008
- C. Harris and M. Stephens. A combined corner and edge detector. In Alvey Vision Conference, pages 147–151, 1988.
- D. G. Lowe. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 60(2):91–110, 2004.