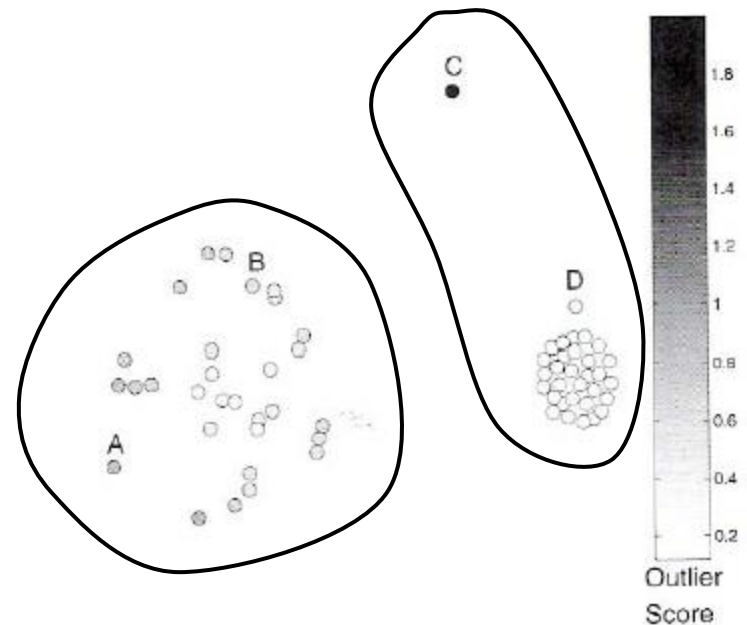


Anomaly Detection

- Motivation and Introduction
- Supervised Methods
- Semisupervised Methods
- Unsupervised Methods:
 - Graphical and Statistical approaches
 - Nearest neighbor based approaches
 - Clustering based approaches
- Evaluation

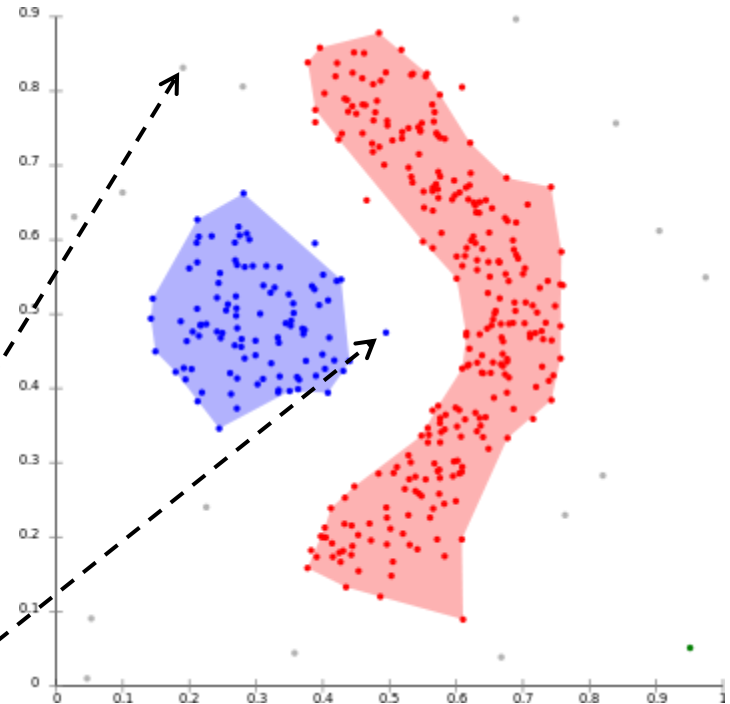


- Basic idea:
 - A set of clusters has already been constructed by any clustering method
 - An existing object or a new one is compared to these clusters in order to determine if it is an anomaly



➤ Approaches

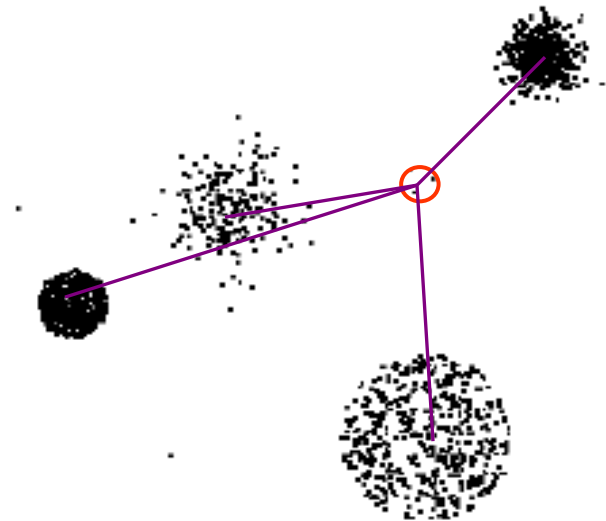
- One possibility: An object is an anomaly if it doesn't belong to any cluster
- Only for clustering methods which may not assign objects to clusters. Typically, density based algorithms as DBSCAN:
 - ➔ *Noise Points*



➤ Approaches

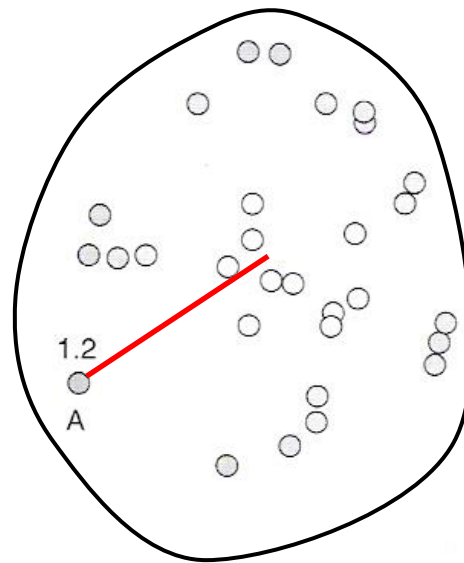
- Another possibility: The anomaly score is given by the distance to its nearest centroid.

How do we measure it?
→ *Noise Points*

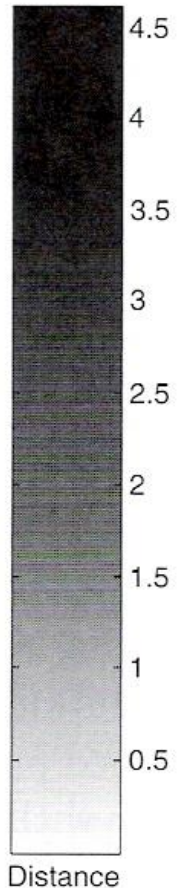
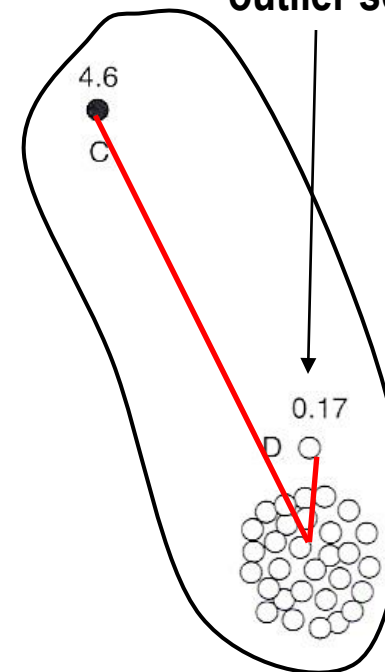


➤ Alternative a)

By measuring the Euclidean distance to its closest centroid



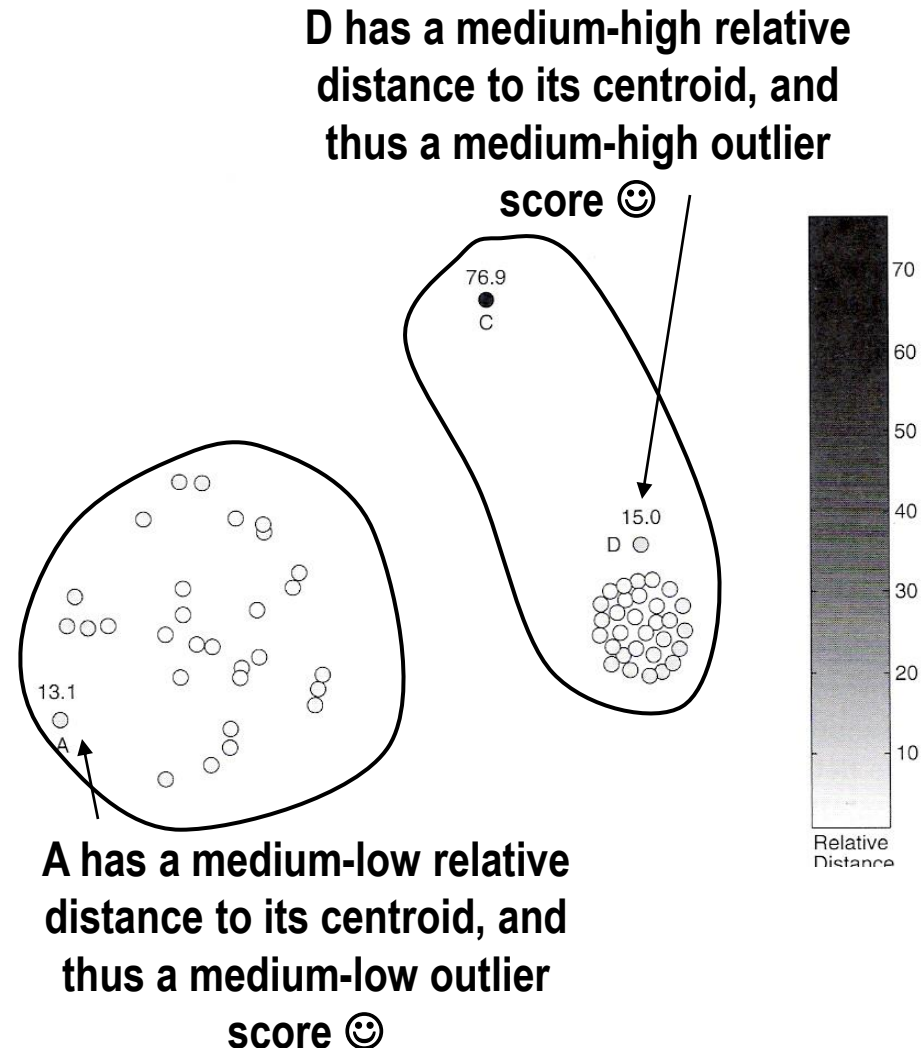
D is near to its centroid,
and thus it has a low
outlier score ☹



➤ Alternative b)

By measuring the relative distance to its closest centroid.

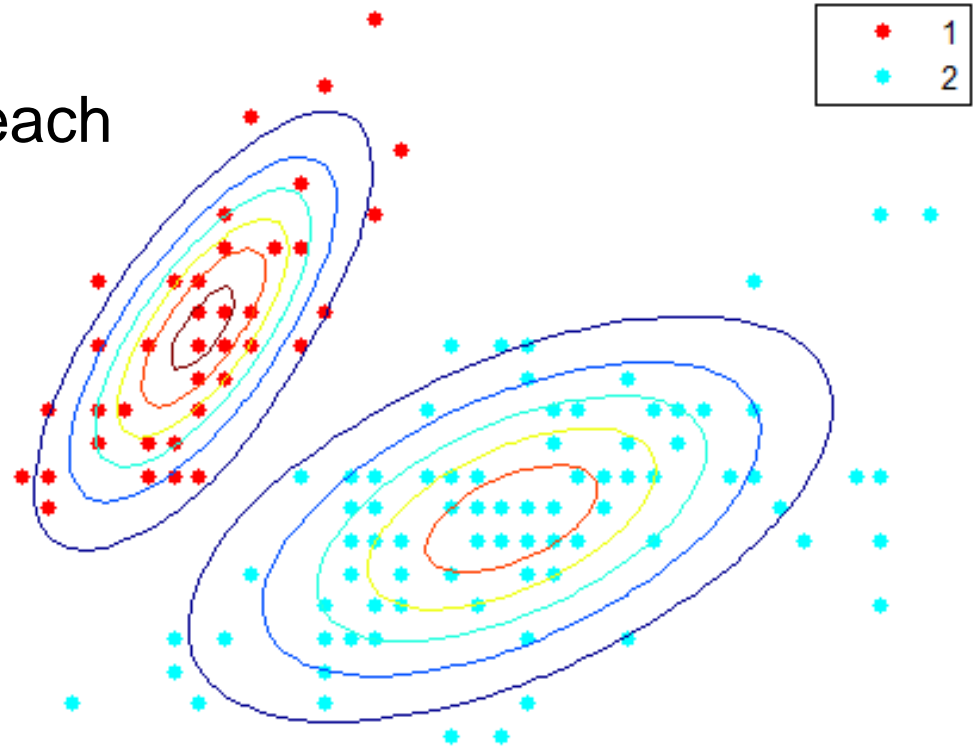
Relative distance is the ratio of the point's distance from the centroid to the median distance of all the points in the cluster from the centroid



➤ Alternative c)

By measuring the Mahalanobis distance of each point to the cluster distribution

Mahalanobis distance involves computation of the covariance matrix of each cluster, which is quite expensive

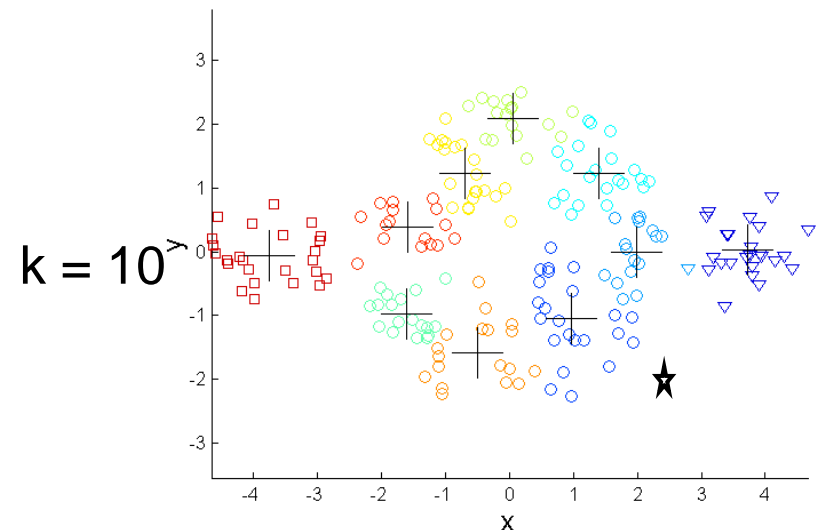
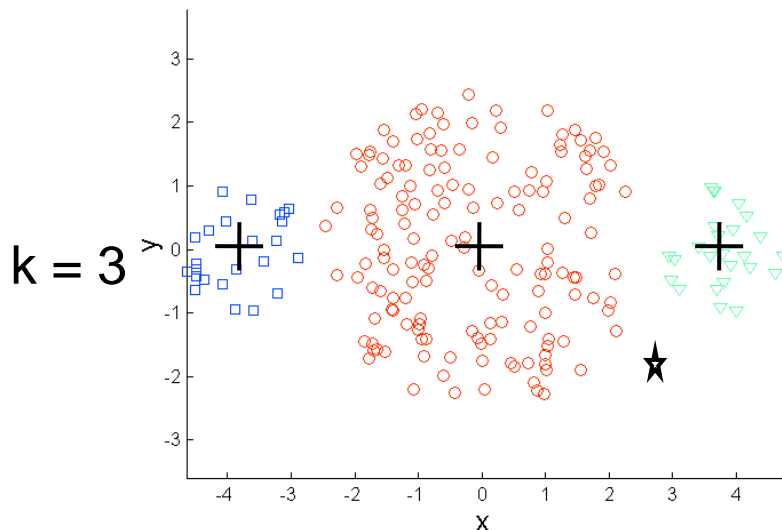


Choice of k is problematic

(k is now the number of clusters)

Usually, it's better to work with a large number of small clusters.

An object identified as outlier when there is a large number of small clusters, it's likely to be a true outlier.



Unsupervised Methods

Cluster based

Mark Schwabacher , Nikunj Oza, Bryan Matthews, **2007**. Unsupervised Anomaly Detection for Liquid-Fueled Rocket. Propulsion Health Monitoring,” AIAA Infotech@Aerospace Conference, American Institute of Aeronautics and Astronautics

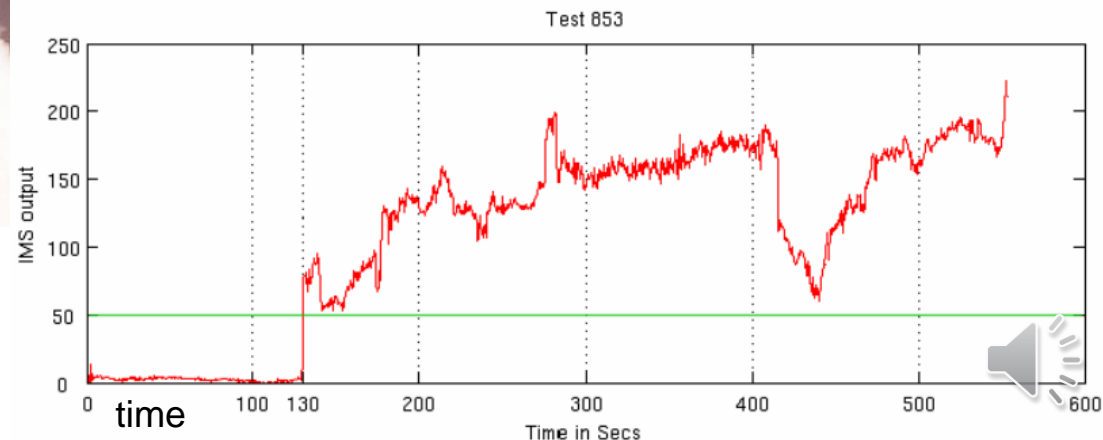
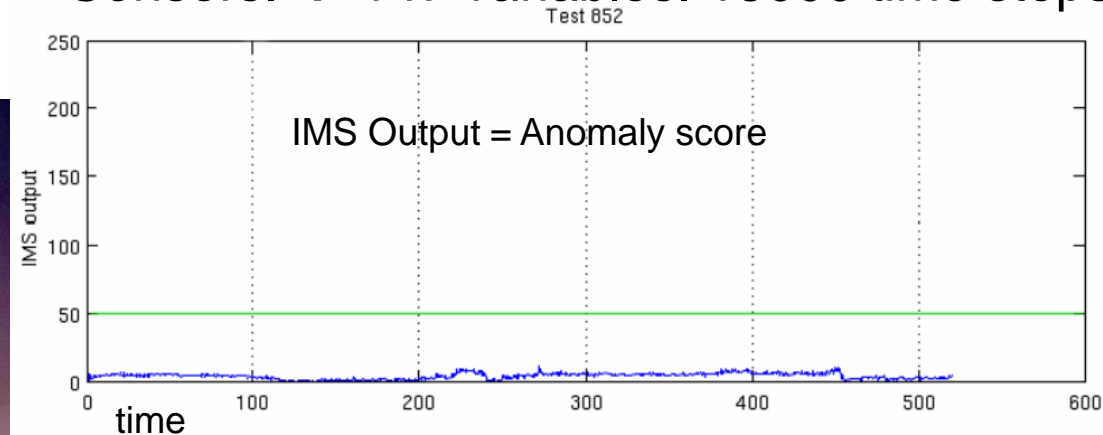
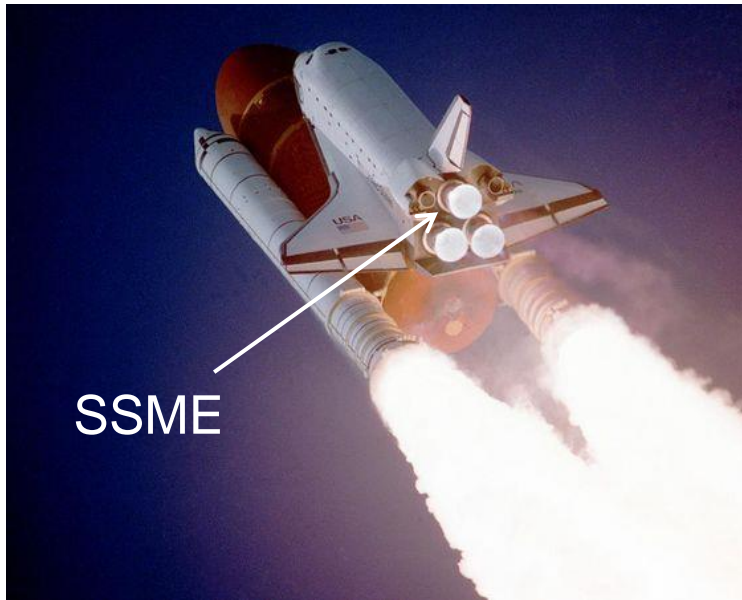
Inductive Monitoring System(IMS)

Pressure, temperature, vibration, etc.

Sensors. → 147 variables. 13000 time steps

SSME:

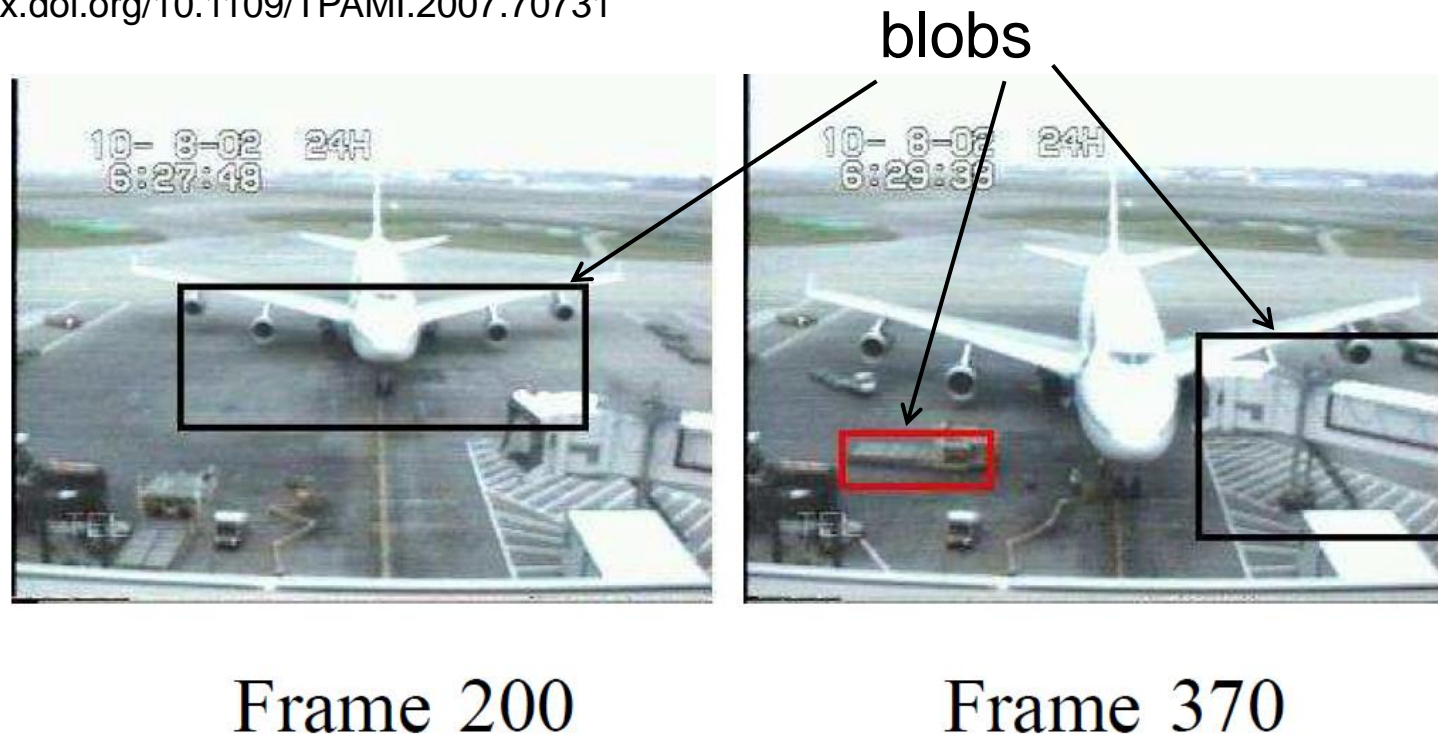
Space Shuttle Main Engines
(reusable)



In complex scenarios, several techniques work together.

Anomaly detection sequences in video

Tao Xiang and Shaogang Gong. **2008**. Video Behavior Profiling for Anomaly Detection. *IEEE Trans. Pattern Anal. Mach. Intell.* 30, 5 (May 2008), 893-908. DOI=10.1109/TPAMI.2007.70731
<http://dx.doi.org/10.1109/TPAMI.2007.70731>



- A foreground detection method is applied to detect changes in image sequences
- Foreground pixels in a vicinity are grouped into a blob
- Instead of working with pixels, the method uses:

$$f = (\overline{x}, \overline{y}, w, h, Rf, M_px, M_py)$$

$(\overline{x}, \overline{y})$ are location features.

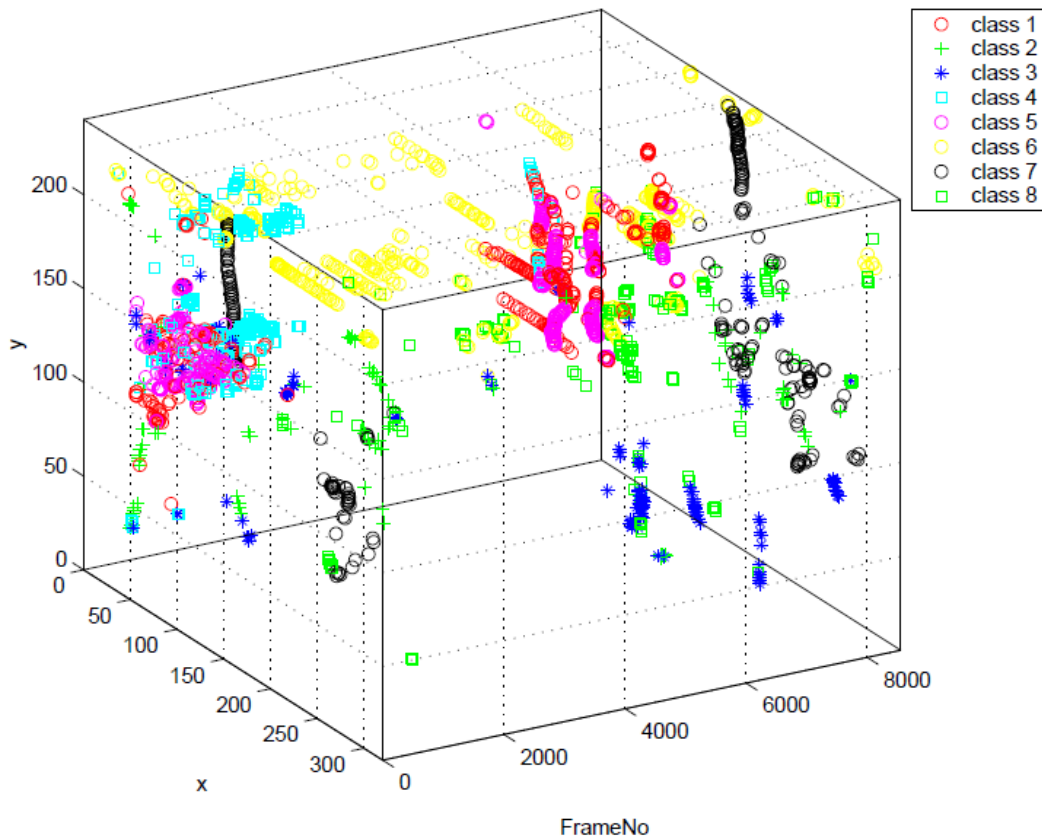
(w, h) and Rf are mainly shape features but also contain some indirect motion information

(M_px, M_py) are motion features capturing the direction of object motion.



A clustering is performed in the 7-dimension feature space

$$f = (\bar{x}, \bar{y}, w, h, Rf, M_px, M_py)$$



Each cluster (class) represents an scene-event:

- aircraft moving and stopping
- airbridge moving
- rear catering vehicles
-

A Bayesian net to model the scene-events interaction is constructed