Introduction to perception
Detection And Tracking of Moving Objects (DATMO)

Olivier Aycard

Professor

Grenoble INP - PHELMA

GIPSA Lab

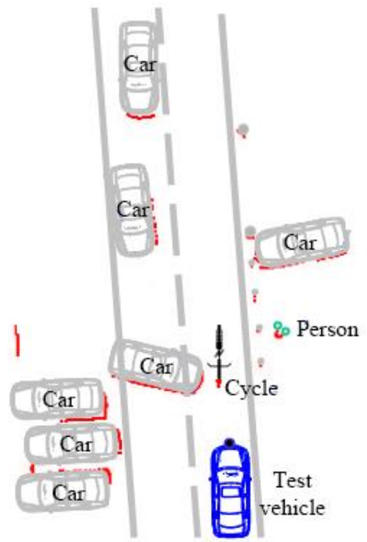
https://www.gipsa-lab.grenoble-inp.fr/user/olivier.aycard

olivier.aycard@grenoble-inp.fr

Perception

Goal

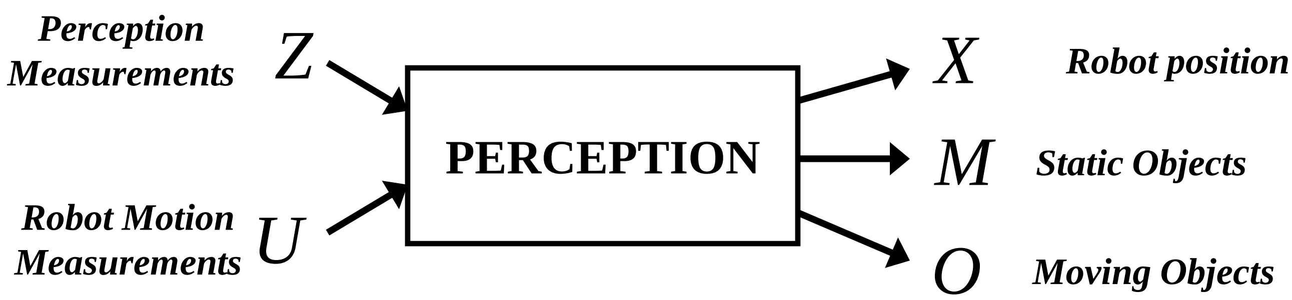
- Robot perception in dynamic environments
- **Laser scanner**
- Speed and robustness



Present Focus: interpretation of raw and noisy sensor data

- Identify static and dynamic part of sensor data
- Modeling dynamic part of the environment
  - Detection And Tracking of Moving Objects (DATMO)
- Modeling static part of the environment
  - Simultaneous Localization And Mapping (SLAM)

Problem statement



Static environments | Dynamic environments

SLAM → **SLAM + Moving object detection** → **SLAM + DATMO**

$$P(X,M \mid Z,U)$$

$$\begin{cases} Z = Z^{(s)} + Z^{(d)} \\ P(X,M \mid Z^{(s)},U) \end{cases}$$

$$P(X,M,O \mid Z,U)$$

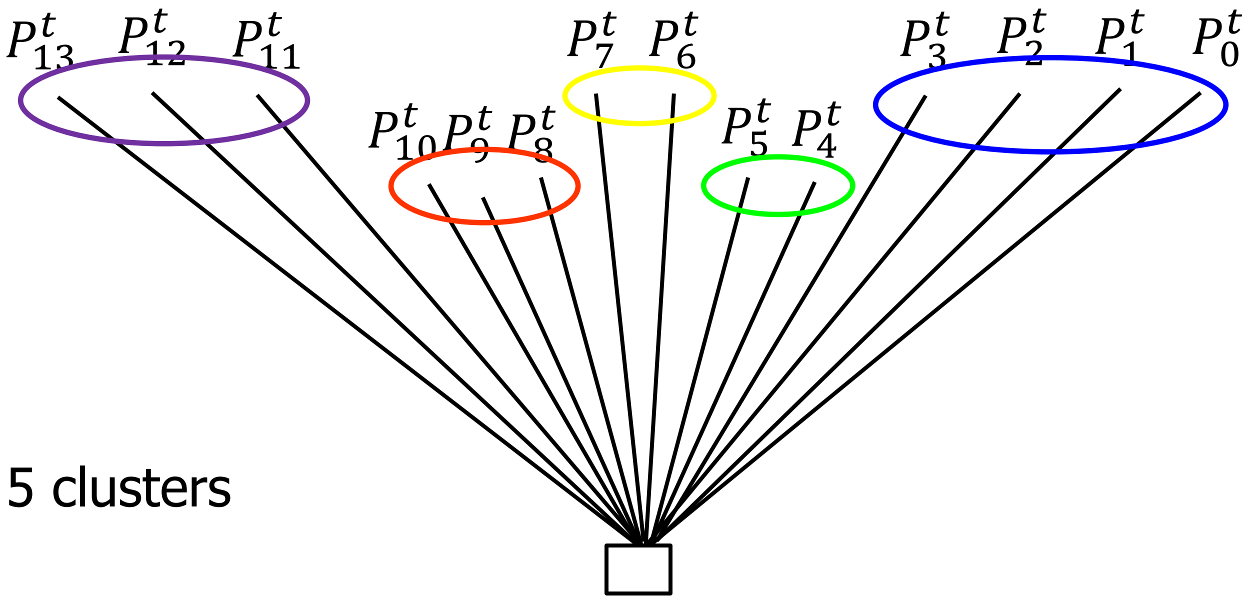
$$\begin{cases} P(X,M \mid Z^{(s)},U) \\ P(O \mid Z^{(d)}) \end{cases}$$

Outline

1. **Form objects**
   1. **Clustering hits of the laser**
2. **Detection of objects**
3. **Tracking of a moving object**
4. **Conclusion**

Cluster of raw data to form objects (1/2)

- **Hits of laser are not objects**
  - We need to cluster hits to form objects
  - Hits that are close in the **cartesian space** should belong to the same object: $P_i^t = (x_i^t, y_i^t)$ (i^th at time t)
- **Example of a person in front of a wall**



We have 5 clusters

Cluster of raw data to form objects (2/2)

- **Initialization of the first cluster**

  Create a first cluster with the first hit

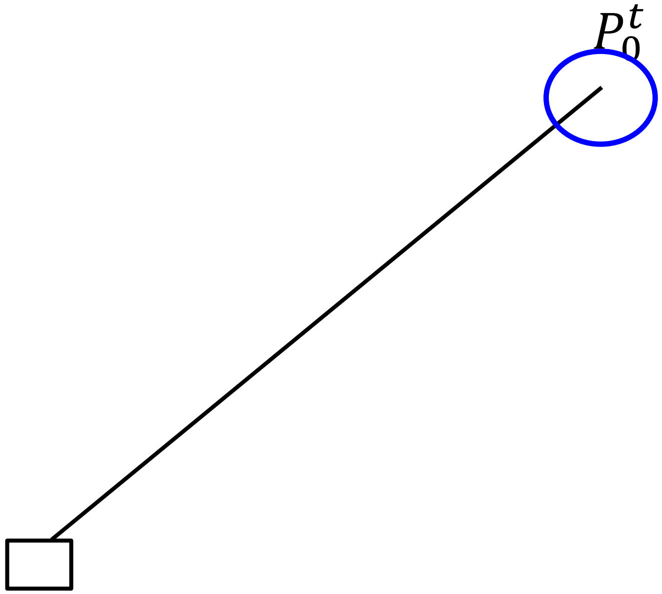
- **For all the hits except the first one**

  If the euclidian distance between the current hit and the previous hit is lower than a given threshold

  Then add the current hit to the current cluster

  Else create a new cluster with the current hit

End for



Olivier.aycard@grenoble-inp.fr

Cluster of raw data to form objects (2/2)

- **Initialization of the first cluster**

  Create a first cluster with the first hit

- **For all the hits except the first one**

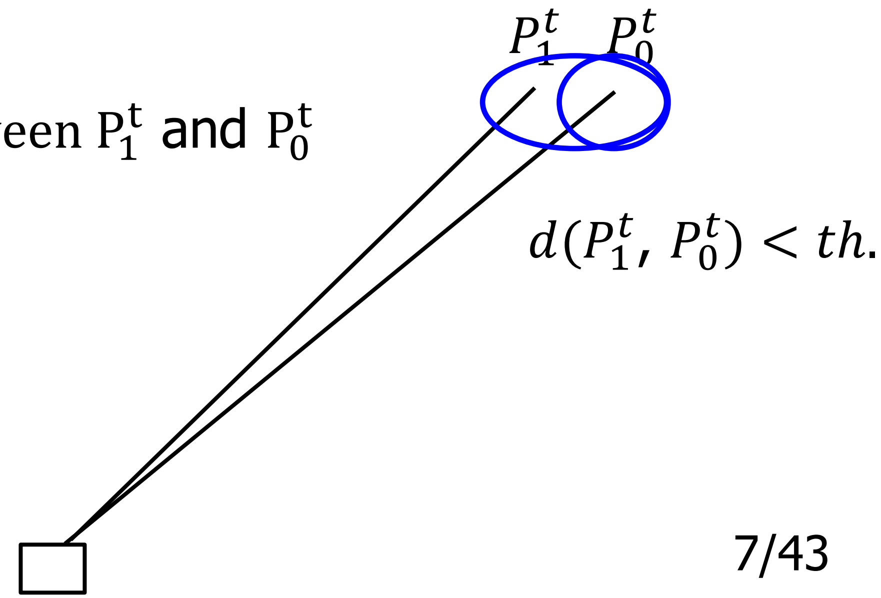
  If the euclidian distance between the current hit and the previous hit is lower than a given threshold

  Then add the current hit to the current cluster

  Else create a new cluster with the current hit

End for

$d(P_1^t, P_0^t)$ is the euclidean distance between P_1^t and P_0^t



Cluster of raw data to form objects (2/2)

- **Initialization of the first cluster**

  Create a first cluster with the first hit

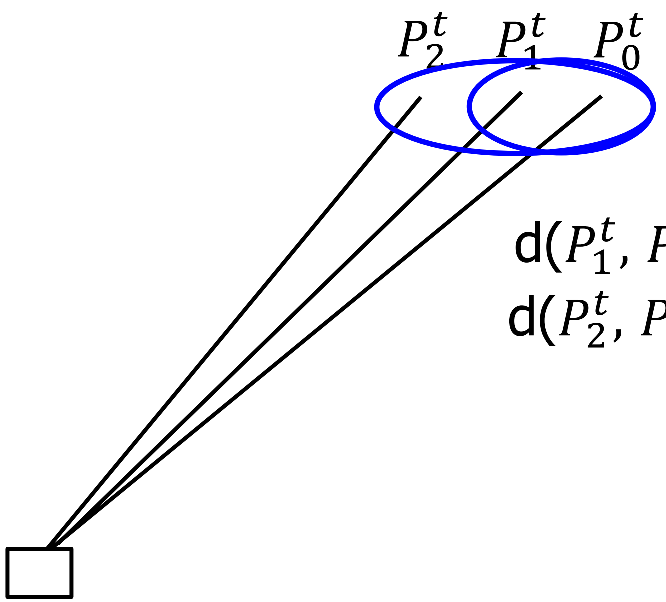
- **For all the hits except the first one**

  If the euclidian distance between the current hit and the previous hit is lower than a given threshold

  Then add the current hit to the current cluster

  Else create a new cluster with the current hit

End for



$\mathrm{d}(P_1^t, P_0^t) < th.$

$\mathrm{d}(P_2^t, P_1^t) < th.$

Olivier.aycard@grenoble-inp.fr

Cluster of raw data to form objects (2/2)

- **Initialization of the first cluster**

  Create a first cluster with the first hit

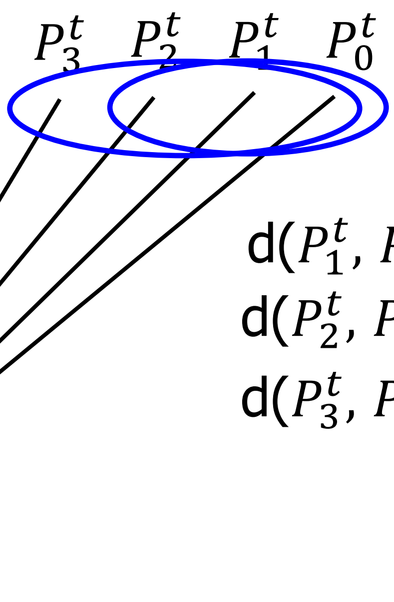
- **For all the hits except the first one**

  If the euclidian distance between the current hit and the previous hit is lower than a given threshold

  Then add the current hit to the current cluster

  Else create a new cluster with the current hit

End for



d$(P_1^t, P_0^t) < th.$

d$(P_2^t, P_1^t) < th.$

d$(P_3^t, P_2^t) < th.$

Cluster of raw data to form objects (2/2)

- **Initialization of the first cluster**

  Create a first cluster with the first hit

- **For all the hits except the first one**

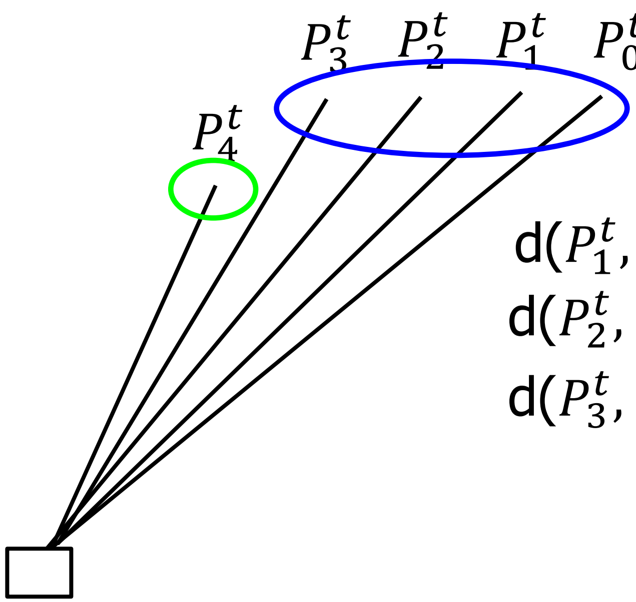
  If the euclidian distance between the current hit and the previous hit is lower than a given threshold

  Then add the current hit to the current cluster

  Else create a new cluster with the current hit

End for

d$(P_4^t, P_3^t) > th$.



d$(P_1^t, P_0^t) < th$.

d$(P_2^t, P_1^t) < th$.

d$(P_3^t, P_2^t) < th$.

Cluster of raw data to form objects (2/2)

- **Initialization of the first cluster**

  Create a first cluster with the first hit

- **For all the hits except the first one**

  If the euclidian distance between the current hit and the previous hit is lower than a given threshold

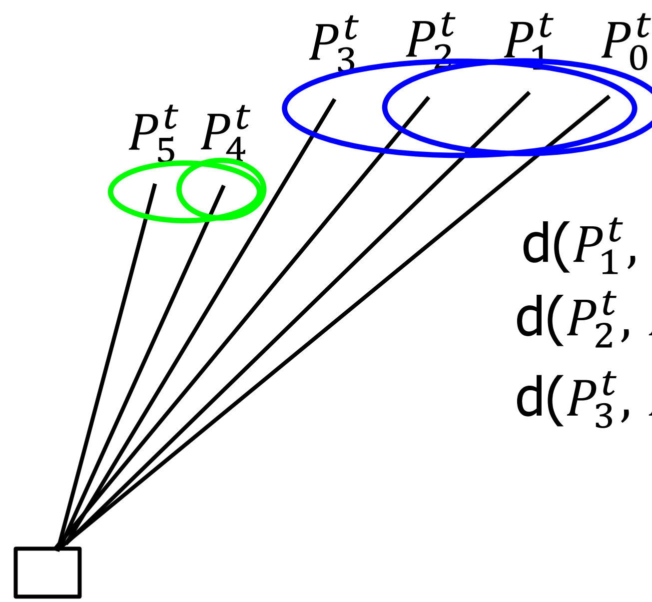
  Then add the current beam to the current cluster

  Else create a new cluster with the current beam

End for

$\mathrm{d}(P_4^t, P_3^t) > th.$

$\mathrm{d}(P_5^t, P_4^t) < th.$



$\mathrm{d}(P_1^t, P_0^t) < th.$

$\mathrm{d}(P_2^t, P_1^t) < th.$

$\mathrm{d}(P_3^t, P_2^t) < th.$

Olivier.aycard@grenoble-inp.fr

Cluster of raw data to form objects (2/2)

- **Initialization of the first cluster**

  Create a first cluster with the first hit

- **For all the beams except the first one**

  If the euclidian distance between the current hit and the previous hit is lower than a given threshold

  Then add the current beam to the current cluster

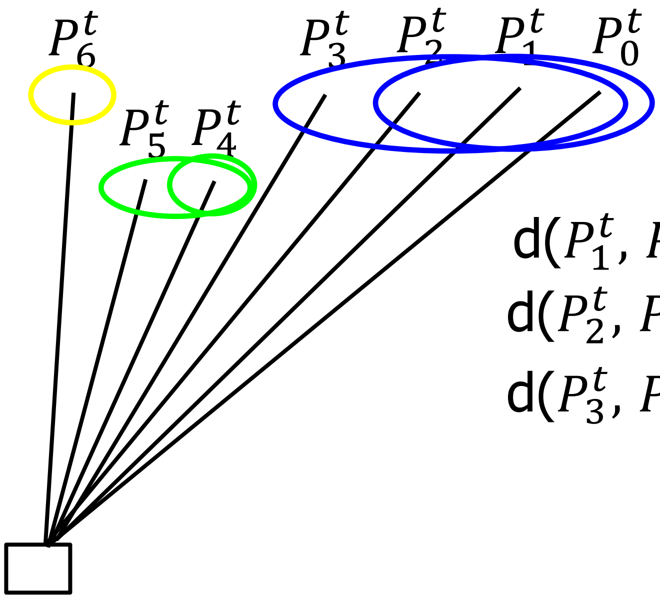
  Else create a new cluster with the current beam

End for

$d(P_4^t, P_3^t) > th.$

$d(P_5^t, P_4^t) < th.$

$d(P_6^t, P_5^t) > th.$



$d(P_1^t, P_0^t) < th.$

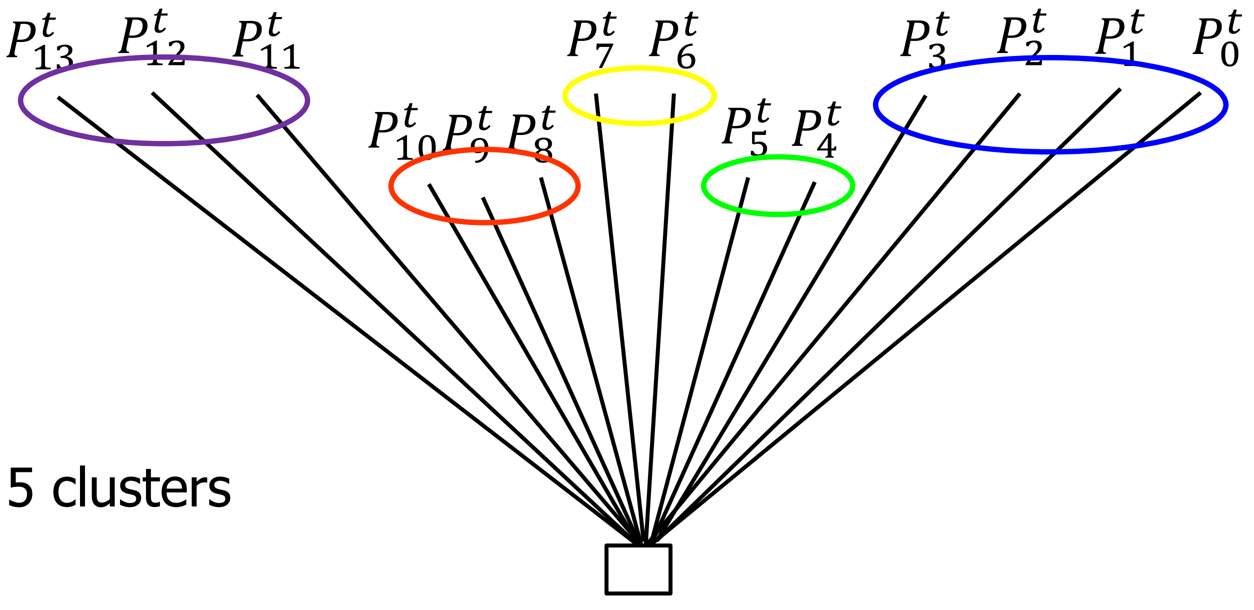
$d(P_2^t, P_1^t) < th.$

$d(P_3^t, P_2^t) < th.$

Olivier.aycard@grenoble-inp.fr

Cluster of raw data to form objects (2/2)

- **Hits of laser are not objects**
  - We need to cluster hits to form objects
  - Hits that are close in the **cartesian space** should belong to the same object: $P_i^t = (x_i^t, y_i^t)$ (i^th at time t)
- **Example of a person in front of a wall**



We have 5 clusters

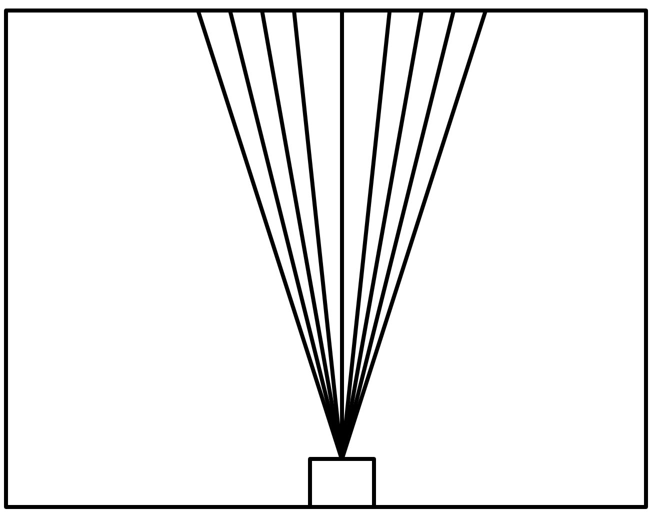
Outline

1. **Form objects**
2. **Detection of objects**
    1. **Motion based detection of objects**
    2. **Model based detection of objects**
    3. **Motion+model based detection of objects**
3. **Tracking of a moving object**
4. **Conclusion**

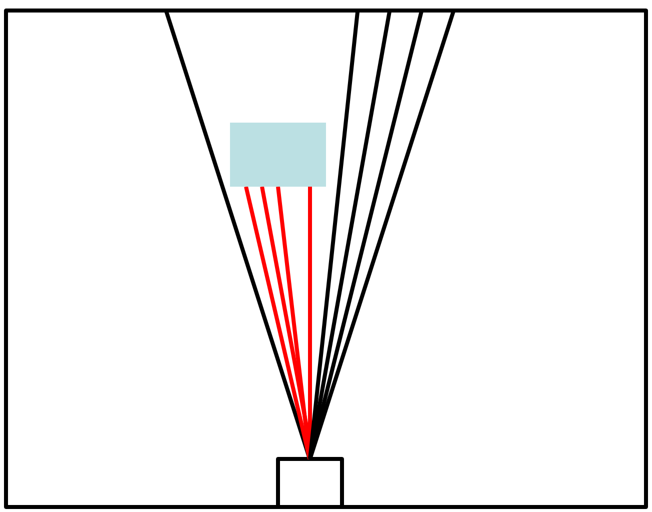
Motion based detection of moving objects(1/5)

A very simple idea (but it works):

- **A scan of the laser at time T composed of N observations:**
  - $r^T = \{r^T_0, r^T_1, \ldots, r^T_{N-1}\}$; (range in the polar space)
- **For each observation r^T_i at time T with i between 1 and N:**
  - we make the difference with the initial observation:
  - If $r^0_i - r^T_i >$ threshold then r^t_i should correspond to a moving object
  - We have to store r^0: background



Initial time t =0

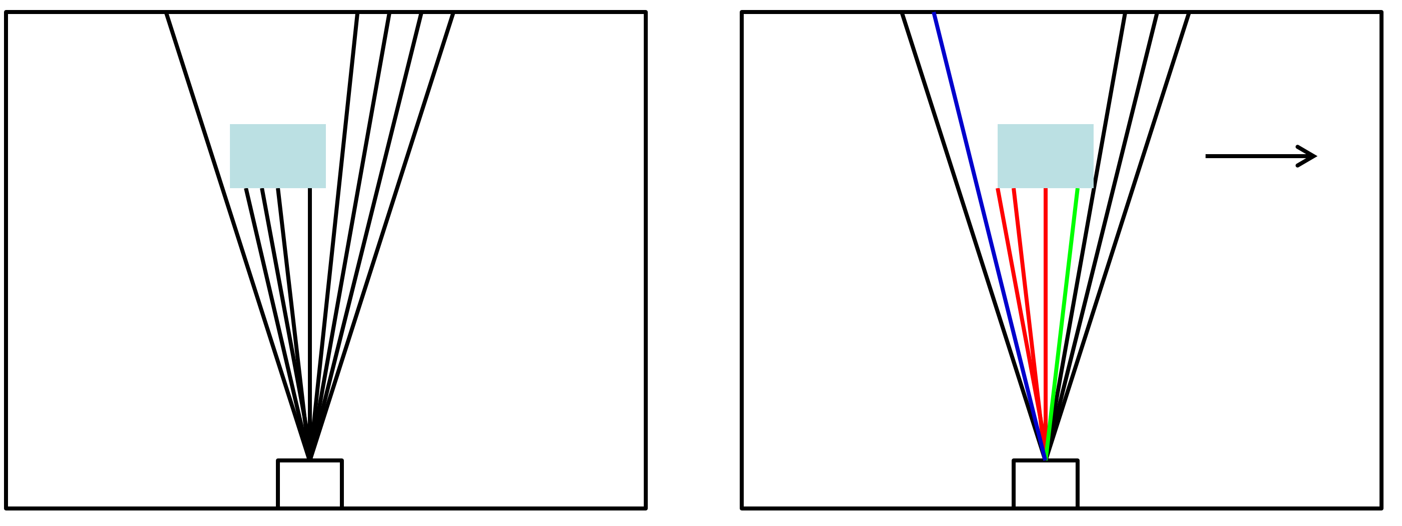


current time t = T

Motion based detection of moving objects(2/5)

An other idea (a bit more complex):

- For each observation r^T_i at time T with i between 1 and N:
  - we make the difference with the previous observation r^{T-1}_i :
  - If $r^T_i - r^{T-1}_i$ > threshold for one or several observations
  - If $r^T_j = r^{T-1}_j$ for one or several observations with j > i
  - If $r^{T-1}_k - r^T_k$ > threshold for one or several observations with k > j
  - an object should be moving from left to right

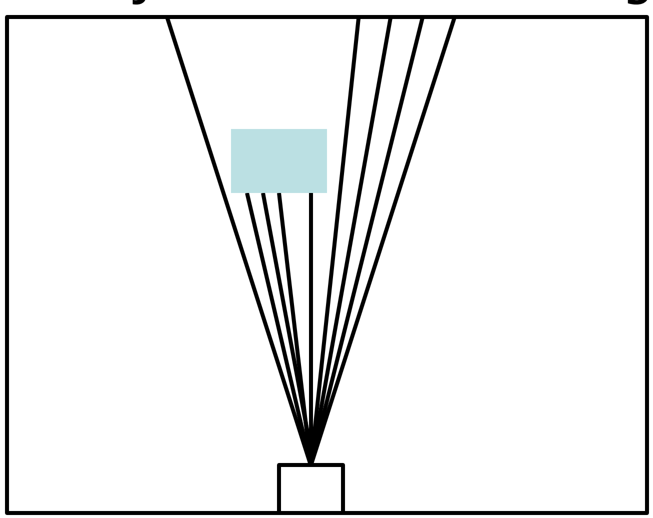


previous time current time

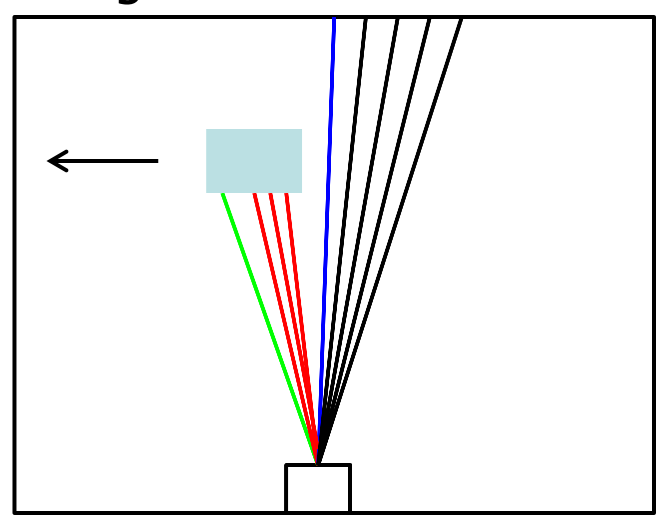
Motion based detection of moving objects(3/5)

An other idea (a bit more complex):

- Difference for each beam b between the current laser scanner acquisition and the previous one
  - we make the difference with the previous observation r^{T-1}_i :
  - If $r^{T-1}_i - r^T_i$ > threshold for one or several observations with i
  - If $r^T_j = r^{T-1}_j$ for one or several observations with j > i
  - If $r^T_k - r^{T-1}_k$ > threshold for one or several observations
  - an object should be moving from right to left



previous time



current time

Motion based detection of moving objects(4/5)

Goal:

- Detection based of moving objects using a lidar

Solution:

- Distinguish changes in the lidar scanner due to motion

Assumptions:

- Static (or known position) of the mobile robot
- Need of the initial observations r^0_i with 0 <= i <= N-1 (ie, background image) or the previous observations Z^{T-1}_i with 0 <= i <= N-1

Advantages:

- No a priori knowledge on object dynamic
- No a priori knowledge on object form

Motion based detection of moving objects(5/5)

Detection of moving objects with a mobile robot that could move:

- 3 important points should be taken into account

1. …
2. …
3. …

Outline

1. **Form objects**
2. **Detection of objects**
    1. **Motion based detection**
    2. **Model based detection**
        1. **Apriori model**
        2. **Learned model**
    3. **Motion+Model based detection**
3. **Tracking of a moving object**
4. **Conclusion**

Model based detection of objects(1/6)

We need to have a model of typical objects that we want to detect.

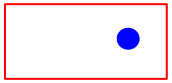
2 kinds of methods are used to detect typical objects:

1. **Define an apriori model[Dung'09];**
2. **Learn (using statistical learning) of the model [Arras'12];**

Model based detection objects(2/6)

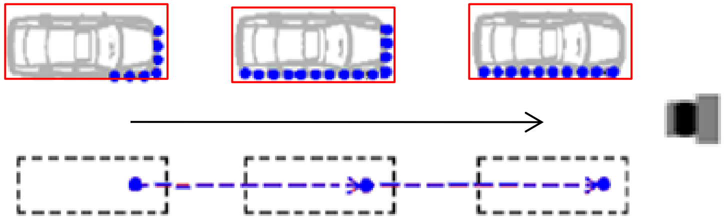
Use of an apriori model of typical objects[Dung'09]

⇒ Suppose that objects are a specific model of car.

⇒ We know the size of this specific model of car. 

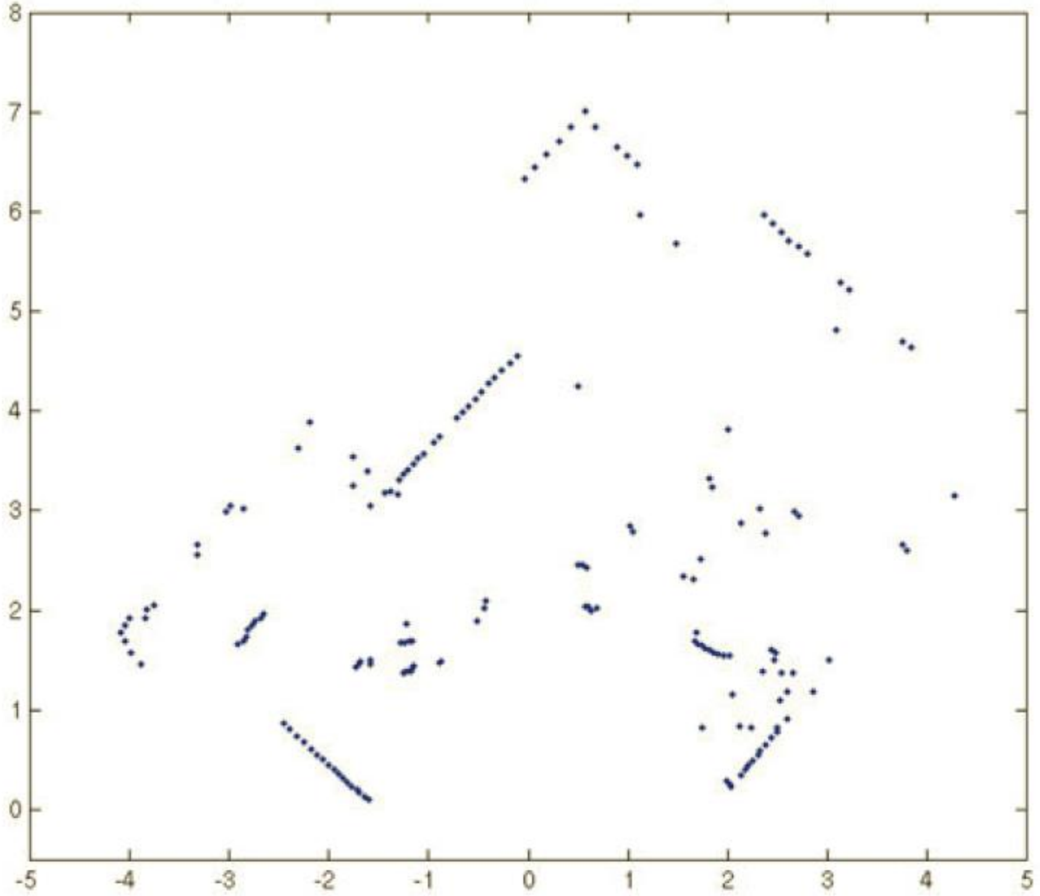
⇒ Most of the time only a part of objects is perceived

⇒ We should fit our data with the model



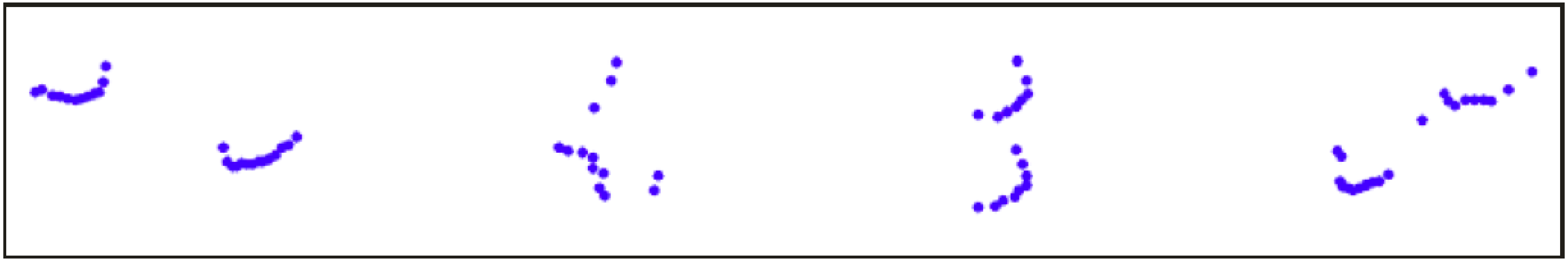
Model based detection of objects(3/6)

Example: scan of a typical office environment [Arras12]. Where are people ?



Model based detection of objects(4/6)

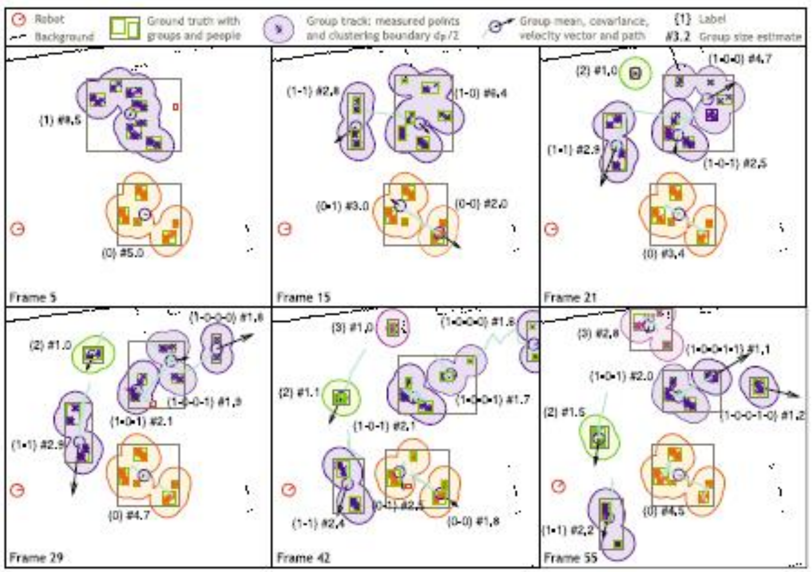
This is not easy to recognize people: the appearance can change drastically**[Arras12]**



1. **Learn a statistical model of the the typical object**
   1. Extract some interesting features of objects and perform statistical learning on these features;
      - For recognition of legs (14 features): number of hits, standard deviation, circularity, radius…
   2. Combine these features using statistical learning.
      - Adaboost, svm…
2. **Use this model to perform detect/recognition of the typical objects**

Recognition of typical objects(5/6)

Fig. 15 Space where we have recorded the datasets for our experiments.



[Arras12]

Model based detection of objects(6/6)

Suppose that objects are persons

1. Define an apriori model of a person

$\Rightarrow$ A person has two legs located at less than 70cms one from the other

$\Rightarrow$ A leg has a size between 5cms and 25cms

Outline

1. **Form objects**
2. **Detection of objects**
    1. **Motion based detection**
    2. **Model based detection**
    3. **Model+Motion based model**
3. **Tracking of a moving object**
4. **Conclusion**

Model+motion based detection of objects

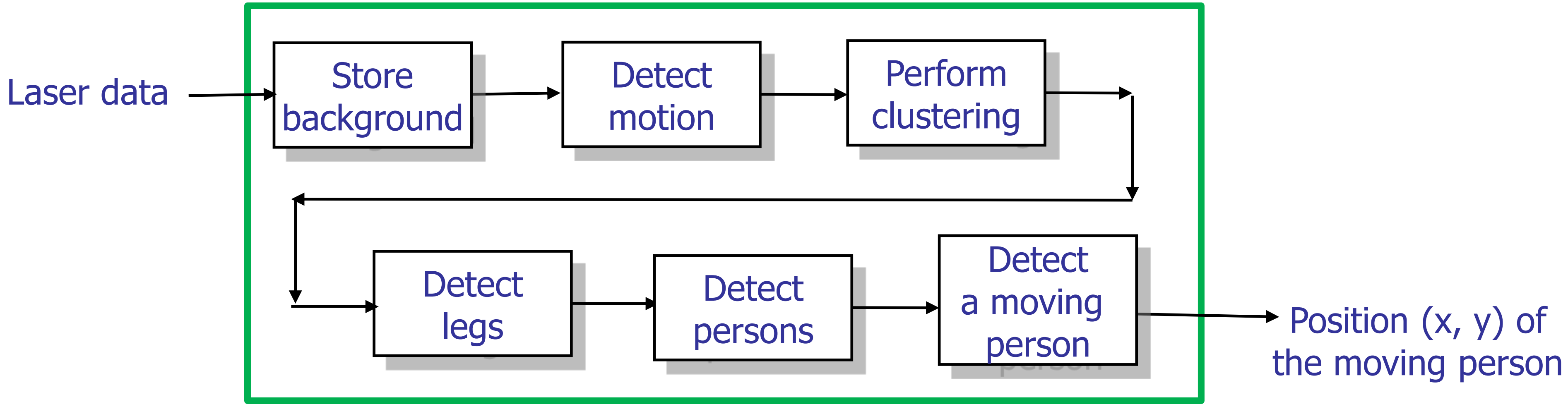
Suppose that objects are moving persons

1. Define an apriori model of a person

⇒ A person has two legs located at less than 70cms one from the other

⇒ A leg has a size between 5cms and 25cms

2. A person is supposed to move



Lab of week 2

Outline

1. **Form objects**
2. **Detection of objects**
    1. **Motion based detection**
    2. **Model based detection**
    3. **Model+Motion based model**
3. **Tracking of a moving object**
    1. **2 examples**
    2. **Tracking of a moving object**
4. **Conclusion**

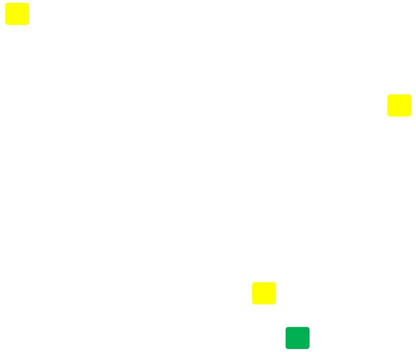
Olivier.aycard@grenoble-inp.fr

Example 1 (1/3)

- We track a moving person at time t;
- Robair starts to move to this tracked moving person. The moving tracked person moves as well;

Example 1 (2/3)

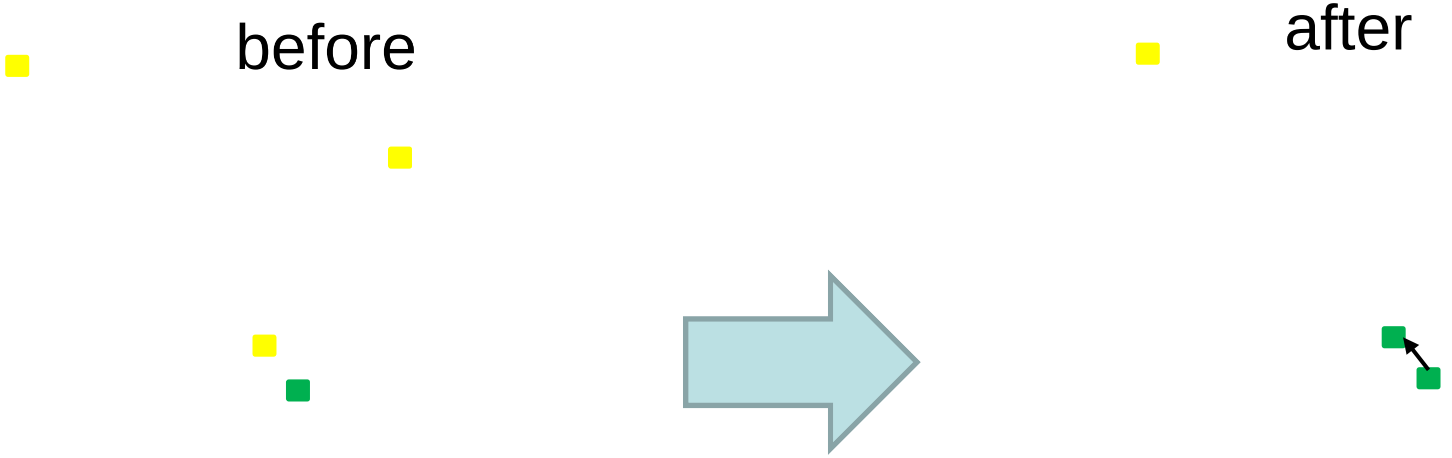
- We track a moving person at time t;
- Robair starts to move to this tracked person. The tracked person moves as well;
- At time t+1, we detect 3 static/moving persons;



Olivier.aycard@grenoble-inp.fr

Example 1 (3/3)

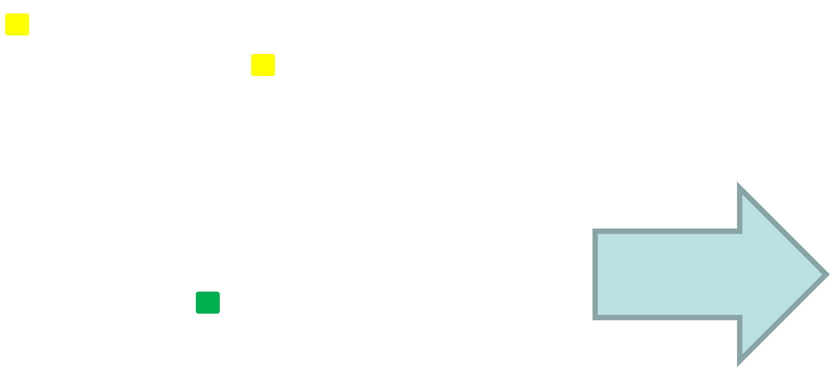
- We track a moving person at time t;
- Robair starts to move closer to this tracked person. The tracked person moves as well;
- At time t+1, we detects 3 static/moving persons;
- We associate the tracked person at time t with the closest moving person detected at t+1;
- We estimate the position of the tracked person at time t+1.



Example 2 (1/2)

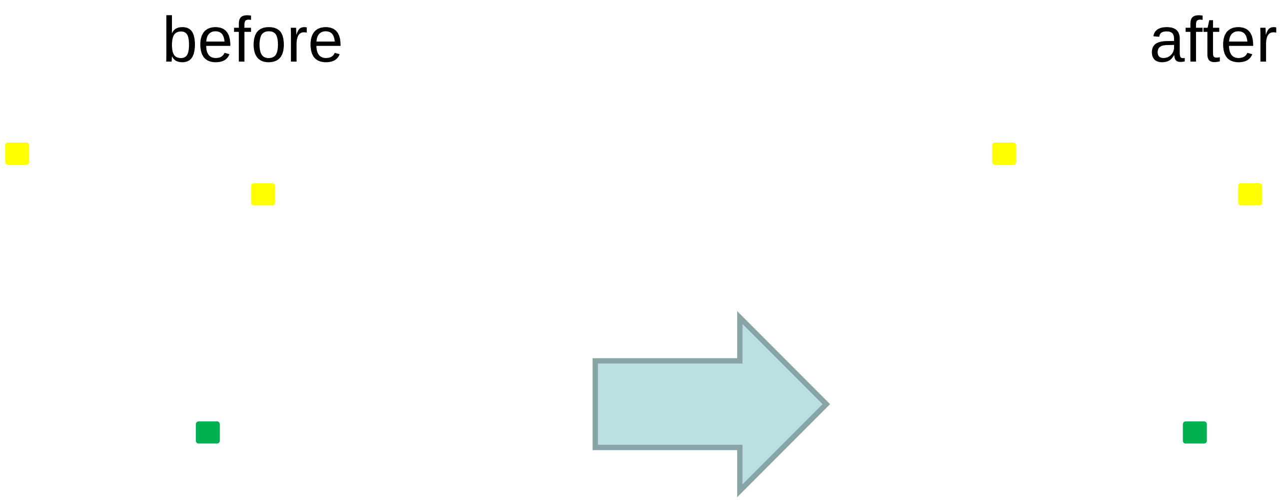
- We track a moving person at time t;
- Robair starts to move to this tracked person. The tracked person moves as well;
- At time t+1, we detect 2 static/moving persons;

before



Example 2 (2/2)

- We track a moving person at time t;
- Robair starts to move to this tracked person. The tracked person moves as well;
- At time t+1, we detect 2 static/moving persons;
- We dont associate the tracked person at time t with the closest person detected at t+1
- The tracked person is still at the same position at time t+1



Olivier.aycard@grenoble-inp.fr

Tracking of a Moving Object (1/2)

- The goal of tracking is to integrate the sequence of detections in time of a moving object to know its position while this object moves and the robot moves as well;
- Several problems could appear:
  - Several moving objects are present in the environment
  - The tracked objects is sometimes not detected
- We generally manage the tracking of an object with 2 variables:
  - Uncertainty to know which detected objects could be associated with the tracked object;
  - Frequency to know the number of times the tracked object has been detected in the last times.

Tracking of Moving Objects (2/2)

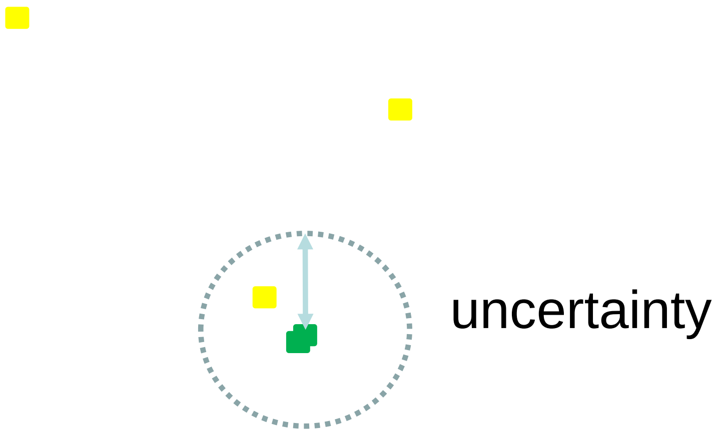
1. Using the information about the motion of the tracked object at time t, we predict its position at time t+1;
   => prediction phase
2. Using the different detected moving objects at time t+1, we associate one (or several) of these detections with the tracked moving object;
   => association phase
3. Using the information about the associated objects detected, we estimate the position of the tracked object at time t+1;
   => estimation phase

Example 1 revisited: prediction phase

- A moving person is detected at time t
- Robair start to move to this moving person
- To simplify, we dont take into account the information, we have about the motion of the tracked object
  => no prediction

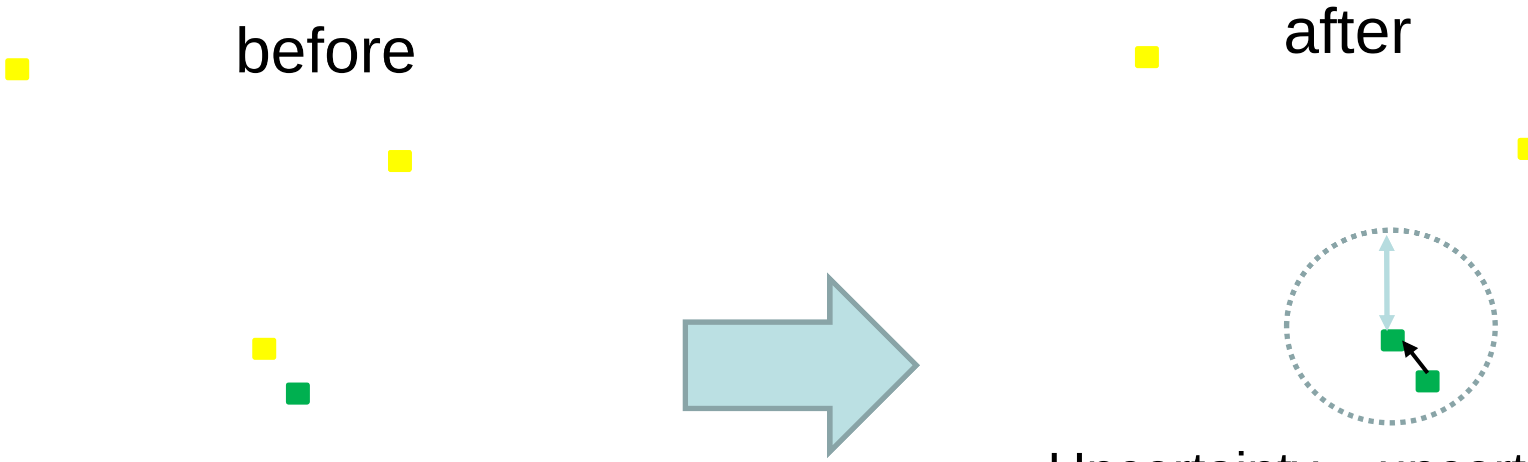
Example 1 revisited: association phase

- At time t+1, we detects 3 static/moving persons;
- We associate the tracked person at time t with the closest moving person detected at t+1 and taking into account the uncertainty;
  => association phase



Example 1 revisited: estimation phase

- We estimate the position of the tracked person at time t+1;
  => estimation phase

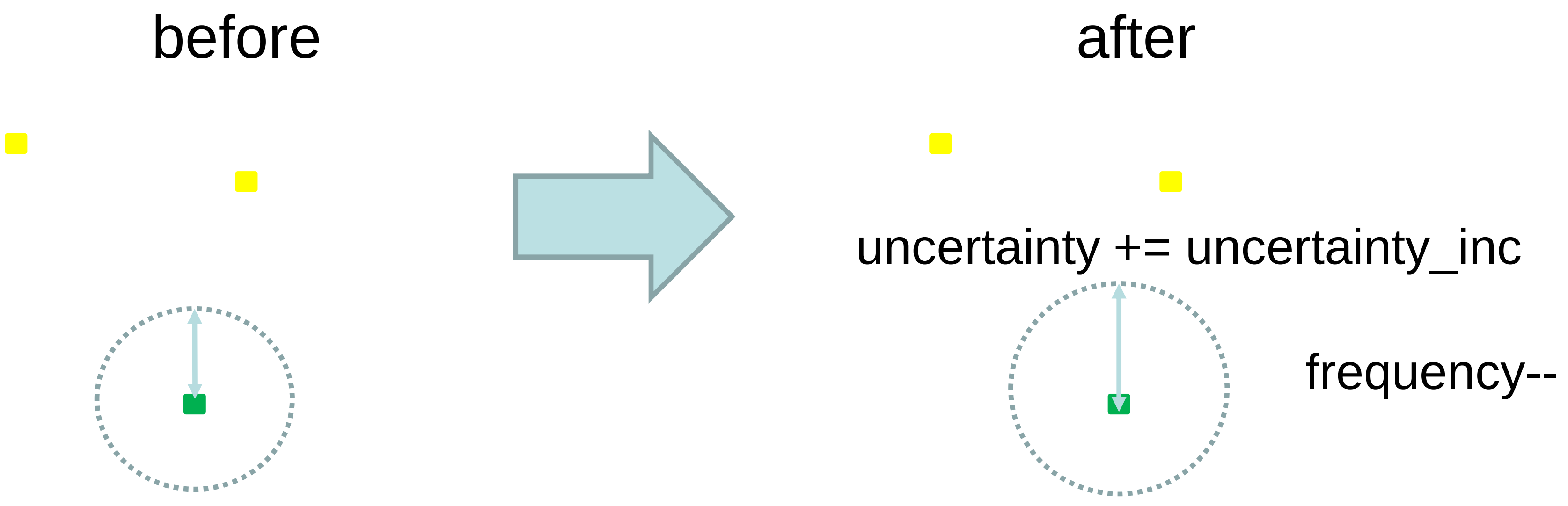


Uncertainty = uncertainty_min
Frequency++

Olivier.aycard@grenoble-inp.fr

Example 2 revisited

- We track a moving person at time t;
- Robair start to move to this tracked person. The tracked person moves as well;
- At time t+1, we detect 2 static/moving persons;
- We dont associate the tracked person at time t with the closest person detected at t+1;
- The tracked person is still at the same position at time t+1

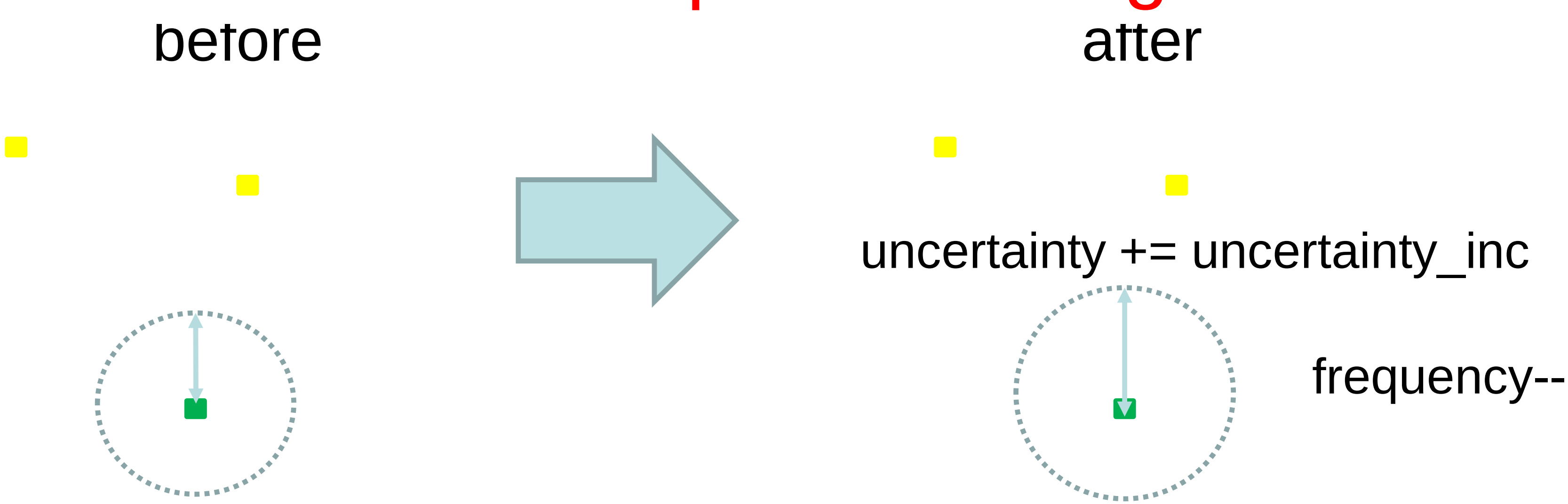


Olivier.aycard@grenoble-inp.fr

Example 2 revisited

- We track a moving person at time t;
- Robair start to move to this tracked person. The tracked person moves as well;
-
-
- :+1

If frequency == 0 then we consider that the tracked person has been lost and we stop the tracking



Outline

1. **Form objects**
2. **Detection of objects**
3. **Tracking of a moving object**
4. **Conclusion**

Conclusion

- Hits of the laser are not objects: need to cluster hits to form objects
- Need to detect objects and know their type and if there are moving or not
- First lab, You will implement a moving person detector

- Detection is not enough: tracking is a must
- There exists a lot of implementation of a tracker:
  - Kalman filter, Mean shift, particle filter…
- You will implement a tracker of a moving person

Olivier.aycard@grenoble-inp.fr