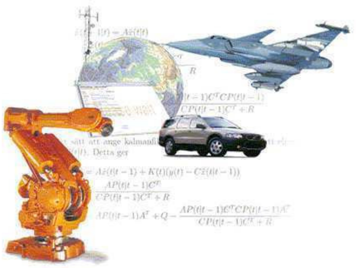
Extended target tracking using PHD filters

With applications to video data and laser range data



Karl Granström

Division of Automatic Control
Department of Electrical Engineering
Linköping university

Linköping, Sweden

Presentation outline

1. Introduction to target tracking using PHD filters
   - *Single target* $\rightarrow$ *Multiple targets* $\rightarrow$ *PHD filters*
   - Illustrative example: Pedestrian tracking in video data

2. Tracking of extended/group targets
   - Definition
   - Gamma-Gaussian-inverse Wishart-model
3. Applications
   - Video: group target tracking
   - Laser: pedestrian and bicycle tracking

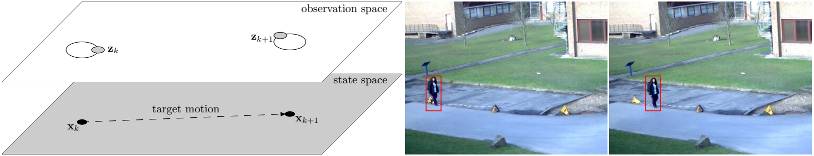
- *Single target* $\rightarrow$ *Multiple targets* $\rightarrow$ *PHD filters*
- Illustrative example: Pedestrian tracking in video data
- Publicly available data:
  "PETS 2012 dataset S1: Person count and density estimation".

Frame 50

Frame 100

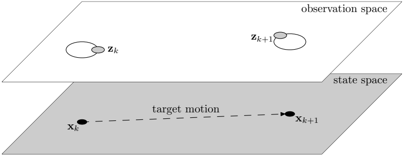
Frame 150

- Process sequence of detections $\mathbf{z}_k$ to maintain estimate of target state $\mathbf{x}_k$.



- Above: Target state $\mathbf{x}_k$ is, e.g., position and velocity in image.
- Above: Red rectangles are the detections $\mathbf{z}_k$.

- Process sequence of detections $\mathbf{z}_k$
  to maintain estimate of target state $\mathbf{x}_k$.

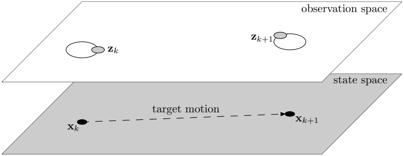


- Recursive Bayesian filter propagates full distribution:

$$\ldots \quad p\left(\mathbf{x}_k \middle| \mathbf{z}^k\right) \quad \xrightarrow{\text{p}} \quad p\left(\mathbf{x}_{k+1} \middle| \mathbf{z}^k\right) \quad \xrightarrow{\text{c}} \quad p\left(\mathbf{x}_{k+1} \middle| \mathbf{z}^{k+1}\right) \quad \ldots$$

where $\mathbf{z}^k = \{\mathbf{z}_0, \mathbf{z}_1, \ldots, \mathbf{z}_k\}$

- Process sequence of detections $\mathbf{z}_k$
  to maintain estimate of target state $\mathbf{x}_k$.

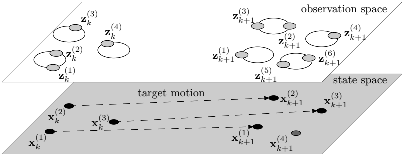


- E.g. constant gain Kalman filter propagates 1^{st} order moment:

$$\ldots \quad \hat{\mathbf{x}}_{k|k} \quad \stackrel{\mathsf{p}}{\rightarrow} \quad \hat{\mathbf{x}}_{k+1|k} \quad \stackrel{\mathsf{c}}{\rightarrow} \quad \hat{\mathbf{x}}_{k+1|k+1} \quad \ldots$$

where $\hat{\mathbf{x}}_{k|k} = \mathrm{E}\left[\mathbf{x}_k \,\middle|\, \mathbf{z}^k\right]$. Simpler, but computationally cheaper!

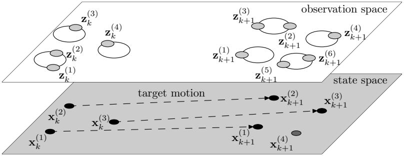
- Processing of multiple detections $\mathbf{z}_k^{(i)}$ from multiple sources to maintain estimates of the targets' states $\mathbf{x}_k^{(j)}$.




- Above: Target states $\mathbf{x}_k^{(j)}$ are, e.g., positions and velocities.
- Above: Red rectangles are the detections $\mathbf{z}_k^{(i)}$.

Multiple target tracking

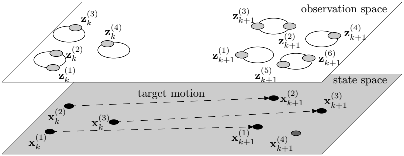
- Processing of multiple detections $\mathbf{z}_k^{(i)}$ from multiple sources to maintain estimates of the targets' states $\mathbf{x}_k^{(j)}$.



- Classic methods: Require solution to data association problem.
  1. Which detection belong to which target?
  2. Are there any false detections?

AUTOMATIC CONTROL
REGLERTEKNIK
LINKÖPINGS UNIVERSITET

- Processing of multiple detections $\mathbf{z}_k^{(i)}$ from multiple sources to maintain estimates of the targets' states $\mathbf{x}_k^{(j)}$.



- Alternative: Model with Random Finite Sets. No data association needed.

- A Random Finite Set $\mathbf{X}_k$ is a random variable
- Set cardinality $|\mathbf{X}_k| = N_{x,k}$ is a random variable.
- Each set memeber $\mathbf{x}_k^{(j)} \in \mathcal{X}_0$ is a random variable. Often $\mathcal{X}_0 = \mathbb{R}^{n_x}$.

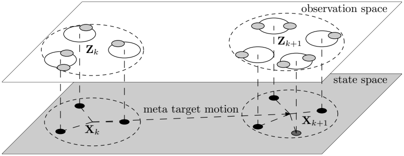
Random Finite Set (RFS)

- A Random Finite Set $\mathbf{X}_k$ is a random variable
- Set cardinality $|\mathbf{X}_k| = N_{x,k}$ is a random variable.
- Each set memeber $\mathbf{x}_k^{(j)} \in \mathcal{X}_0$ is a random variable. Often $\mathcal{X}_0 = \mathbb{R}^{n_x}$.
- Instantiations

$$
\begin{aligned}
\mathbf{X}_k &= \emptyset \\
\mathbf{X}_k &= \left\{ \mathbf{x}_k^{(j)} \right\}_{j=1}^{N_{x,k}}, \ \mathbf{x}_k^{(j)} \in \mathbb{R}^{n_x} \text{ for } j = 1, \ldots, N_{x,k}
\end{aligned}
$$

- $\mathbf{X}_k$ is without order, i.e. $\left\{ \mathbf{x}_k^{(1)}, \mathbf{x}_k^{(2)} \right\} = \left\{ \mathbf{x}_k^{(2)}, \mathbf{x}_k^{(1)} \right\}$

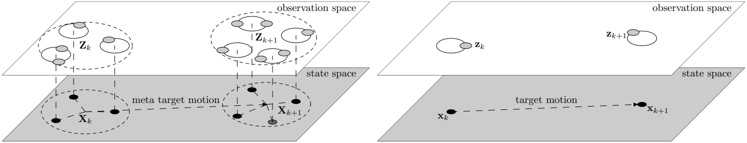
- Processing of sequence of RFS of detections $\mathbf{Z}_k$ from set of sources to maintain estimate of RFS of target states $\mathbf{X}_k$.



- RFS of targets $\mathbf{X}_k$. Above: the pedestrians
- RFS of detections $\mathbf{Z}_k$. Above: the detection windows
- Single target set, single detection set
  $\Rightarrow$ No data association needed

- Propagate the full set distribution

$$\ldots \quad p\left(\mathbf{X}_k \left| \mathbf{Z}^k \right.\right) \quad \xrightarrow{\text{p}} \quad p\left(\mathbf{X}_{k+1} \left| \mathbf{Z}^k \right.\right) \quad \xrightarrow{\text{c}} \quad p\left(\mathbf{X}_{k+1} \left| \mathbf{Z}^{k+1} \right.\right) \quad \ldots$$



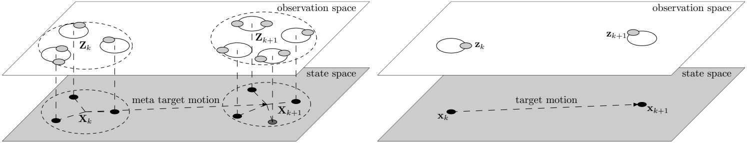
- Multiple target equivalent to single target Bayes filter,

$$\ldots \quad p\left(\mathbf{x}_k \left| \mathbf{z}^k \right.\right) \quad \xrightarrow{\text{p}} \quad p\left(\mathbf{x}_{k+1} \left| \mathbf{z}^k \right.\right) \quad \xrightarrow{\text{c}} \quad p\left(\mathbf{x}_{k+1} \left| \mathbf{z}^{k+1} \right.\right) \quad \ldots$$

Recursive multi state Bayes filter

- Propagate the full set distribution

$$\ldots \quad p\left(\mathbf{X}_k \left| \mathbf{Z}^k \right.\right) \overset{\text{p}}{\rightarrow} p\left(\mathbf{X}_{k+1} \left| \mathbf{Z}^k \right.\right) \overset{\text{c}}{\rightarrow} p\left(\mathbf{X}_{k+1} \left| \mathbf{Z}^{k+1} \right.\right) \quad \ldots$$



- Multiple target equivalent to single target Bayes filter,

$$\ldots \quad p\left(\mathbf{x}_k \left| \mathbf{z}^k \right.\right) \overset{\text{p}}{\rightarrow} p\left(\mathbf{x}_{k+1} \left| \mathbf{z}^k \right.\right) \overset{\text{c}}{\rightarrow} p\left(\mathbf{x}_{k+1} \left| \mathbf{z}^{k+1} \right.\right) \quad \ldots$$

- Computationally intractable to implement due to set integrals.
- Alternative: Use PHD or CPHD filter instead.

- Expected value: 1st order moment of single target pdf $p\left(\mathbf{x}_k \middle| \mathbf{z}^k\right)$
- The probability hypothesis density (PHD):
  1st order moment of multi-target pdf $p\left(\mathbf{X}_k \middle| \mathbf{Z}^k\right)$.
- Defined on single target states $\mathbf{x} \in \mathcal{X}_0$, denoted

$$D_{k|k}\left(\mathbf{x}_k \middle| \mathbf{Z}^k\right)$$

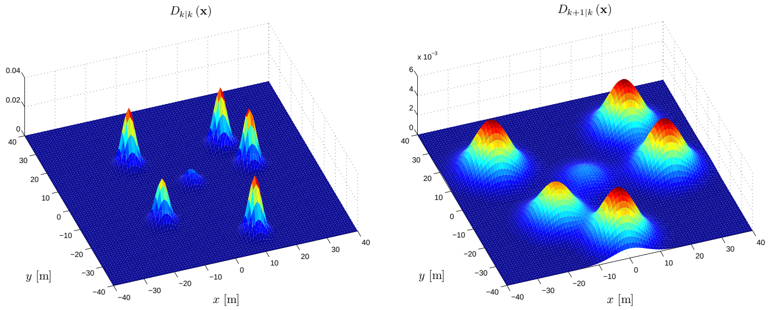
- The PHD is an intensity function, not a pdf.

The Probability Hypothesis Density

- The peaks in the intensity correspond to likely target locations.
- Expected number of targets in any region $S \subseteq \mathcal{X}_0$ is

$$\int_S D_{k|k}\left(\mathbf{x}_k \middle| \mathbf{Z}^k\right) \mathrm{d}\mathbf{x}_k$$

$S = \mathcal{X}_0$ gives xpected total number of targets



- The PHD filter propagates the PHD in time

$$\ldots \quad D_{k|k}\left(\mathbf{x}\,|\mathbf{Z}\right) \overset{\text{p}}{\rightarrow} D_{k+1|k}\left(\mathbf{x}\,|\mathbf{Z}\right) \overset{\text{c}}{\rightarrow} D_{k+1|k+1}\left(\mathbf{x}\,|\mathbf{Z}\right) \quad \ldots$$

- PHD filter is analogous to constant gain Kalman filter,

$$\ldots \quad \hat{\mathbf{x}}_{k|k} \overset{\text{p}}{\rightarrow} \hat{\mathbf{x}}_{k+1|k} \overset{\text{c}}{\rightarrow} \hat{\mathbf{x}}_{k+1|k+1} \quad \ldots$$

- Two types of implementations:
  1. Sequential Monte Carlo PHD approximations.
  2. Distribution Mixture PHD approximations,

$$D_{k|k}\left(\mathbf{x}\,|\mathbf{Z}\right) = \sum_{j=1}^{J_{k|k}} w_{k|k}^{(j)} p^{(j)}\left(\mathbf{x}\,\middle|\,\zeta_{k|k}^{(j)}\right)$$

Common choice: $p^{(j)}(\mathbf{x}|\mathbf{Z}) = \mathcal{N}\left(\mathbf{x};\zeta_{k|k}^{(j)}\right),\ \zeta_{k|k}^{(j)} = \left(m_{k|k}^{(j)}, P_{k|k}^{(j)}\right)$

1. Introduction to target tracking using PHD filters
   - *Single target* $\rightarrow$ *Multiple targets* $\rightarrow$ *PHD filters*
   - Illustrative example: Pedestrian tracking in video data

2. Tracking of extended/group targets
   - Definition
   - Gamma-Gaussian-inverse Wishart-model
3. Applications
   - Video: group target tracking
   - Laser: pedestrian and bicycle tracking

AUTOMATIC CONTROL
REGLERTEKNIK
LINKÖPINGS UNIVERSITET

- Point target assumption:
  ≤ 1 detection per target per time step.
- Typical in early RADAR-airplane-tracking.

- Point target assumption:
  ≤ 1 detection per target per time step.
- Typical in early RADAR-airplane-tracking.

- Modern sensors typically have higher resolution, and can give multiple detections per target.
- A group of targets will give multiple detections per group.

Definition:
Extended/group targets are targets that potentially give rise to more than one detection per time step.

AUTOMATIC CONTROL
REGLERTEKNIK
LINKÖPINGS UNIVERSITET

For each group we estimate the group state $\xi_k = (\mathbf{x}_k, X_k, \gamma_k)$

For each group we estimate the group state $\xi_k = (\mathbf{x}_k, X_k, \gamma_k)$

- $\mathbf{x}_k$: position and kinematics

Gaussian distributed random vector $\mathbf{x}$,

$$p\left(\mathbf{x}_k \middle| \mathbf{Z}^k\right) = \mathcal{N}\left(\mathbf{x}_k\,;\, m_{k|k}, P_{k|k}\right)$$

For each group we estimate the group state $\xi_k = (\mathbf{x}_k, X_k, \gamma_k)$

- $\mathbf{x}_k$: position and kinematics
- X_k: spatial extension, i.e. size and shape

Inverse Wishart distributed random matrix X (group shape $=$ ellipse),

$$p\left(X_k \middle| \mathbf{Z}^k\right) = \mathcal{IW}_d\left(X_k\,;\, v_{k|k}, V_{k|k}\right)$$

Gamma-Gaussian-inverse Wishart-model

For each group we estimate the group state $\xi_k = (\mathbf{x}_k, X_k, \gamma_k)$

- $\mathbf{x}_k$: position and kinematics
- X_k: spatial extension, i.e. size and shape
- γ_k: number of detections, related to number of individuals in group

Gamma distributed random variable γ (Poisson rate),

$$p\left(\gamma_k \middle| \mathbf{Z}^k\right) = \mathcal{G}\left(\gamma_k\,;\,\alpha_{k|k}, \beta_{k|k}\right)$$

For each group we estimate the group state $\xi_k = (\mathbf{x}_k, X_k, \gamma_k)$

- $\mathbf{x}_k$: position and kinematics
- X_k: spatial extension, i.e. size and shape
- γ_k: number of detections, related to number of individuals in group

Group state ξ_k is GGIW distributed,

$$p\left(\xi_k \middle| \mathbf{Z}^k\right) = \mathcal{GGIW}\left(\xi_k\,;\,\zeta_{k|k}\right) \quad \zeta = (m, P, v, V, \alpha, \beta)$$

- Problem: Track pedestrian groups as they move along footpath.
- Estimate: position, size/shape, number of persons.

- **GGIW-PHD filter:** the PHD is approximated by a GGIW mixture,

$$D_{k|k}\left(\xi \,|\mathbf{Z}\right) = \sum_{j=1}^{J_{k|k}} w_{k|k}^{(j)} \mathcal{GGIW}\left(\xi\,;\,\zeta_{k|k}^{(j)}\right)$$

1. Introduction to target tracking using PHD filters
    - *Single target* → *Multiple targets* → *PHD filters*
    - Illustrative example: Pedestrian tracking in video data

2. Tracking of extended/group targets
    - Definition
    - Gamma-Gaussian-inverse Wishart-model
3. Applications
    - Video: group target tracking
    - Laser: pedestrian and bicycle tracking

Group target tracking approach

1: For each image, use pedestrian detector

2: Project each detection onto ground plane

3: Use GGIW-PHD filter to track the groups

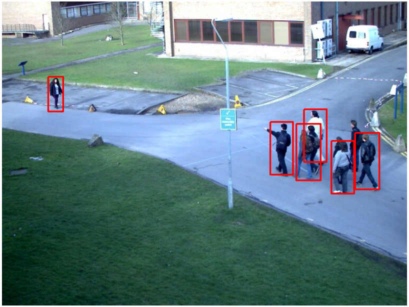
4: Results visualization: project filter output back into images

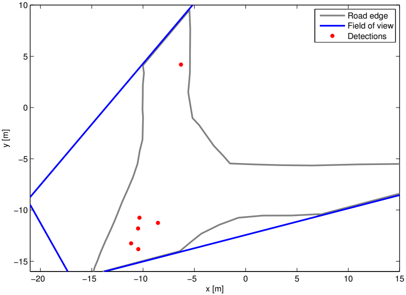
AUTOMATIC CONTROL
REGLERTEKNIK
LINKÖPINGS UNIVERSITET

1: For each image, use pedestrian detector

Group target tracking approach

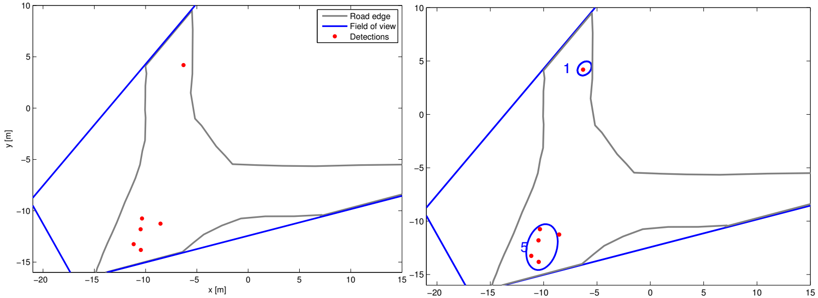
2: Project each detection onto ground plane



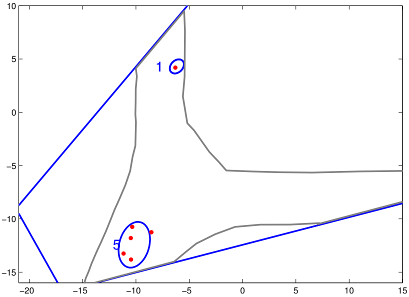


AUTOMATIC CONTROL
REGLERTEKNIK
LINKÖPINGS UNIVERSITET

3: Use Extended/Group target PHD filter to track the groups

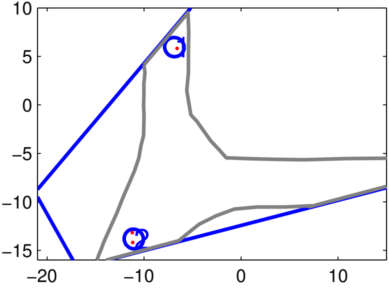


4: Visualization: project filter output back into images



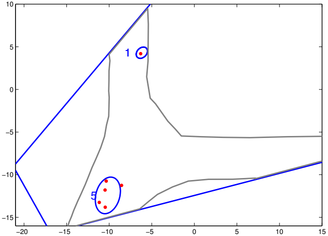
4: Visualization: projection is group "foot-print"

AUTOMATIC CONTROL
REGLERTEKNIK
LINKÖPINGS UNIVERSITET



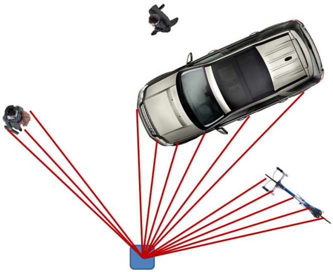
10
5
0
−5
−10
−15
−20
−10
0
10

- Ellipse is a good approximation of the group shape
- $\hat{\gamma}_{k|k}$ gives a lower bound for the number of pedestrians in the group.
- A model of occlusion from stationary objects, e.g. the lamp post, could improve results.

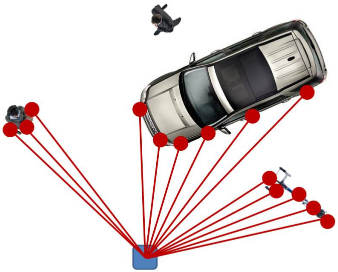


- Consider an urban scene:
  cars, bikes, pedestrians.

- Consider an urban scene: cars, bikes, pedestrians.
- Measures distance to closest object at given angles, relatively low sensor noise.

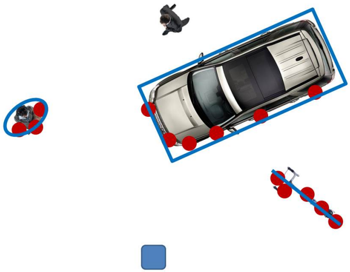


- Consider an urban scene: cars, bikes, pedestrians.
- Measures distance to closest object at given angles, relatively low sensor noise.
- Measurements $\mathbf{z}_k$ (red ●) along the parts of the target surface that are visible to sensor.

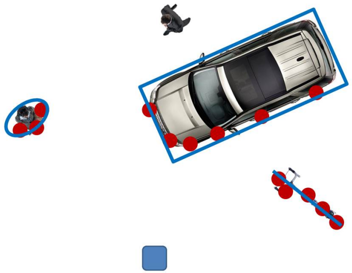


The laser range sensor

- Consider an urban scene: cars, bikes, pedestrians.
- Measures distance to closest object at given angles, relatively low sensor noise.
- Measurements $\mathbf{z}_k$ (red •) along the parts of the target surface that are visible to sensor.
- Car ≈ rectangle.
- Pedestrian ≈ ellipse.
- Bike ≈ line.

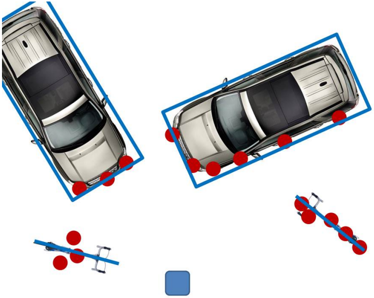


- Occlusion: objects behind other objects are invisible to the sensor.
- Fix: non-homogeneous detection probability.

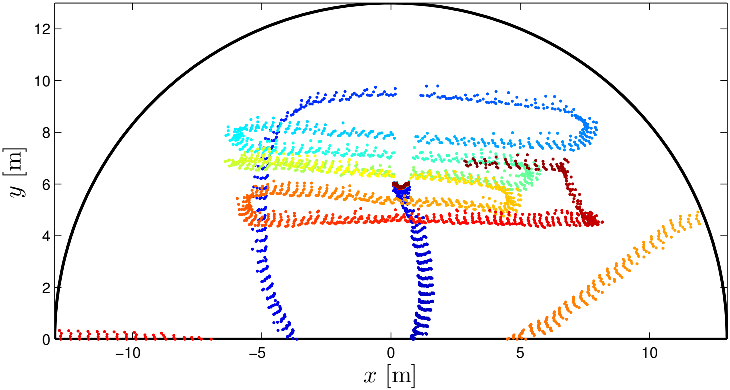


The laser range sensor: challenges

- Occlusion: objects behind other objects are invisible to the sensor.
- Fix: non-homogeneous detection probability.
- Time changing measurement appearance: depending on orientation measurement appearances is different.
- Fix: multiple measurement models.

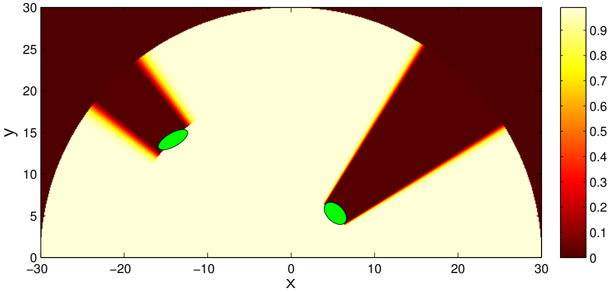


- Experiment data from SICK laser sensor
- Multiple human targets, at most 3 at the same time.
- Measurements of stationary objects removed beforehand.



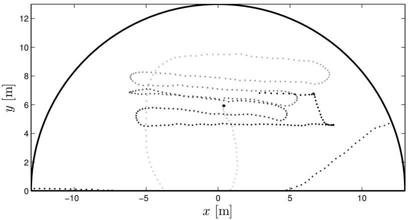
- Occlusion handled with non-homogeneous detection probability.

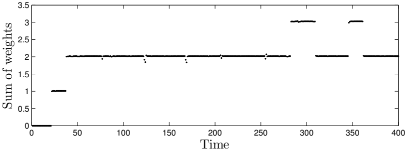
- Occlusion handled with non-homogeneous detection probability.
- Low p_D behind predictions. Gaussian smoothing on edges.



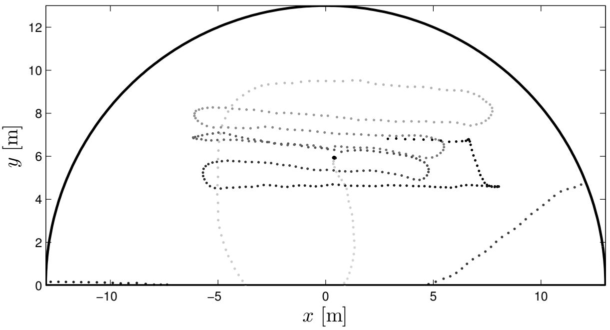
Similar idea used in

Reuter, Dietmayer, *"Pedestrian tracking using random finite sets"*, FUSION 2011



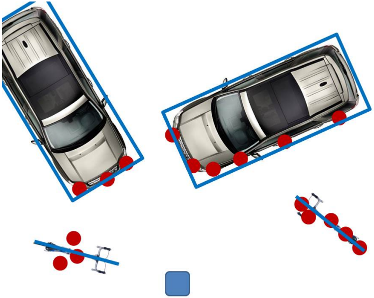


- Occlusion handled by non-homogeneous detection probability.
- Extension estimate makes it possible to handle partial occlusion.

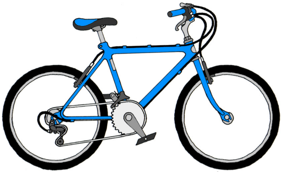


12
10
8
6
4
2
0
–10
–5
0
5
10
y [m]
x [m]

- Most extended targets have constant extensions (orientation may change over time).
- We consider constant extension and time changing appearance, especially abrupt changes.
- Observability of state variables may change with appearance.

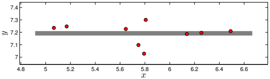


- Most extended targets have constant extensions (orientation may change over time).
- We consider constant extension and time changing appearance, especially abrupt changes.
- Observability of state variables may change with appearance.

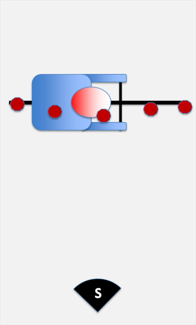


- Example: Bicycles measured by a laser mounted at pedal height

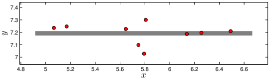
- Measurement appearance varies significantly

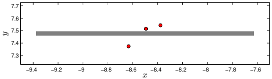


Along bike length

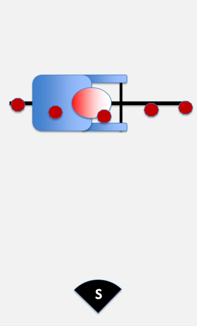


- Measurement appearance varies significantly

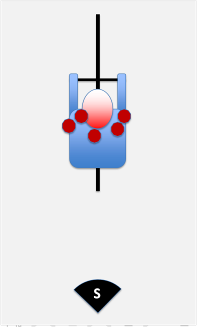




Along bike length

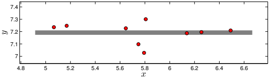


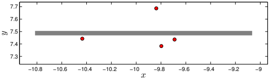
Around pedals

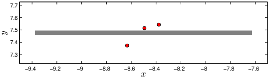


AUTOMATIC CONTROL
REGLERTEKNIK
LINKÖPINGS UNIVERSITET

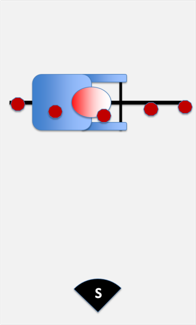
- Measurement appearance varies significantly



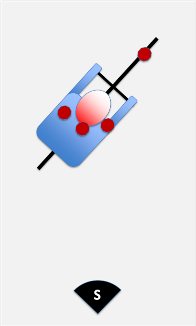




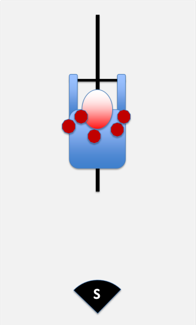
Along bike length



In-between

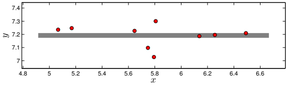


Around pedals

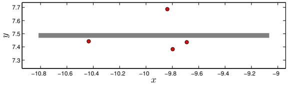


AUTOMATIC CONTROL
REGLERTEKNIK
LINKÖPINGS UNIVERSITET

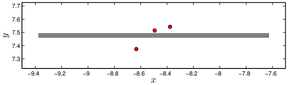
- Measurement appearance varies significantly



Along bike length



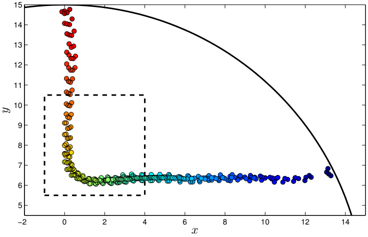
In-between



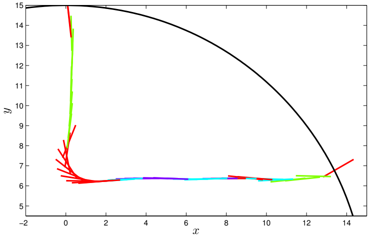
Around pedals

- Two “simple” cases can be identified,
  1. line shaped measurements,
  2. point clusters,

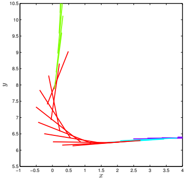
  and ambiguous cases in-between the two.
- Assume geometric shape: bike $\approx$ line with constant length
- Use MM-PHD filter with a point model and a line model.



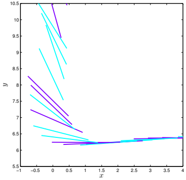
Measurements



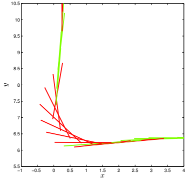
Proposed MM approach



MM approach



Only L model



Only P model

Legend:

CT , P model

CV , P model

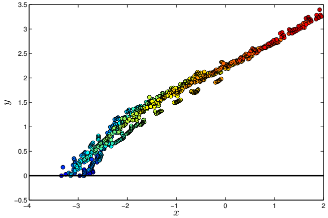
CT , L model

CV , L model

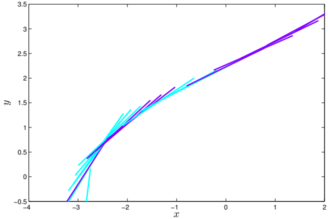
AUTOMATIC CONTROL

REGLERTEKNIK

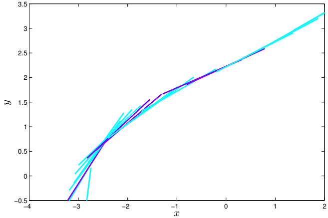
LINKÖPINGS UNIVERSITET



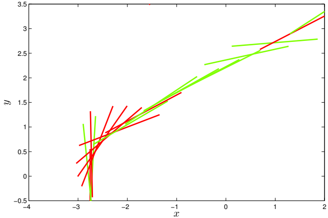
Measurements



MM approach



Only L model



Only P model

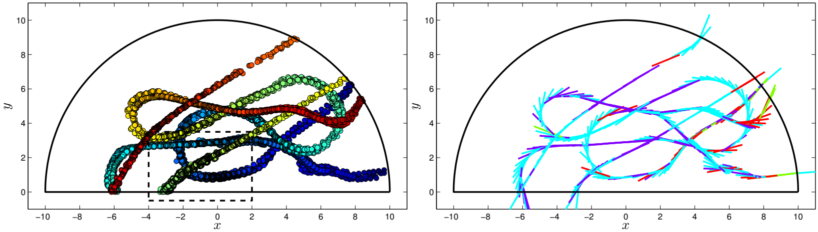
Legend:

CT , P model

CV , P model

CT , L model

CV , L model



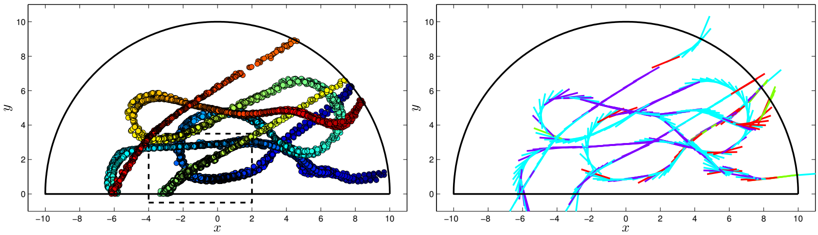
Legend:

CT , P model

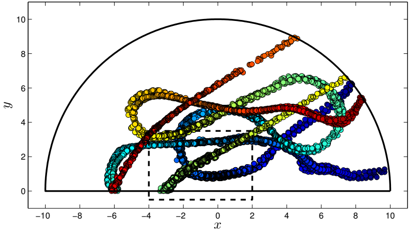
CV , P model

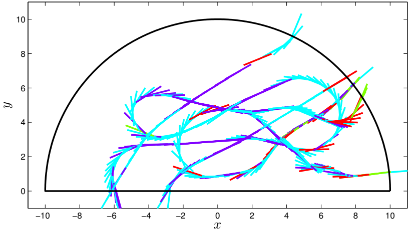
CT , L model

CV , L model



- In general: estimated mode corresponds to expectation/intuition.
- Only P model works in most cases,
  however heading/orientation is more uncertain.
- Only L model sometimes fails during turns.





- In general: estimated mode corresponds to expectation/intuition.
- Only P model works in most cases, however heading/orientation is more uncertain.
- Only L model sometimes fails during turns.
- Using both models generally superior.
- In MM case L model can aid in detecting CT maneuvers.

Summary

- Brief introduction to PHD filters.
- Extended/group objects appear in many modern sensors.
- PHD filter adaptations:
  - Non-homogeneous detection probability
  - Multiple models
- Experimental results using video and laser
  - Groups of pedestrians
  - Pedestrians
  - Bicycles
- List of publications, with pdf:s, can be found at `http://users.isy.liu.se/en/rt/karl/`

AUTOMATIC CONTROL
REGLERTEKNIK
LINKÖPINGS UNIVERSITET

Thank you for listening!

Any questions?

- List of publications, with pdf:s, can be found at `http://users.isy.liu.se/en/rt/karl/`