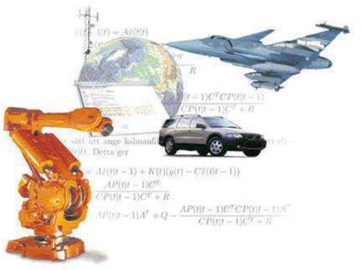
# Extended target tracking using PHD filters

*An application to pedestrian group tracking in video*



Karl Granström

Automatic Control
Department of Electrical Engineering
Linköping university

CADICS

Group tracking using publicly available video data:
“PETS 2012 dataset S1: Person count and density estimation”.



Frame 50 Frame 100 Frame 150

# Presentation outline



Group tracking using publicly available video data:
“PETS 2012 dataset S1: Person count and density estimation”.

Frame 50

Frame 100

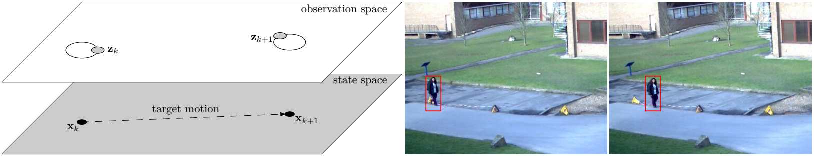
Frame 150

1. Multiple target tracking with PHD filters
   *Single target* $\rightarrow$ *Multiple targets* $\rightarrow$ *PHD filters*
2. Tracking of extended/group targets
3. Application: tracking groups of pedestrians in video data.

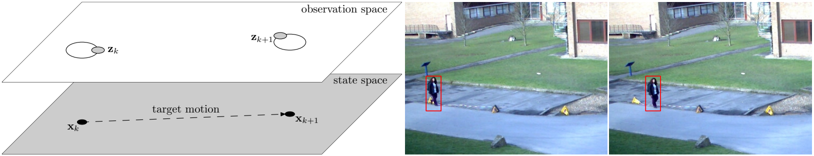


AUTOMATIC CONTROL
REGLERTEKNIK
LINKÖPINGS UNIVERSITET

- Process detections to maintain estimate of target state.



- Process detections to maintain estimate of target state.



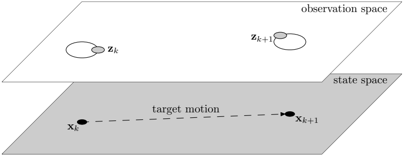

- Recursive Bayesian filter propagates full distribution:

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

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

# Single target tracking

- Process detections to maintain estimate of target state.

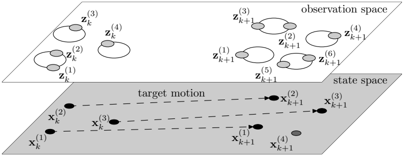



- E.g. constant gain Kalman filter propagates $1^{\text{st}}$ order moment:

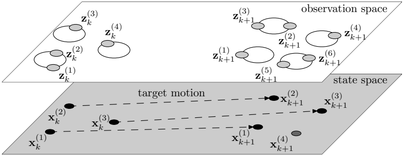
$$\cdots \quad \hat{\mathbf{x}}_{k|k} \quad \xrightarrow{\text{p}} \quad \hat{\mathbf{x}}_{k+1|k} \quad \xrightarrow{\text{c}} \quad \hat{\mathbf{x}}_{k+1|k+1} \quad \cdots$$

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 from multiple sources to maintain estimates of targets' states.

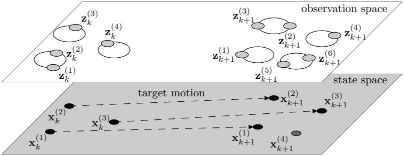



- Processing of multiple detections from multiple sources to maintain estimates of targets' states.



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

- Processing of multiple detections from multiple sources to maintain estimates of targets' states.




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

# Random Finite Set (RFS)



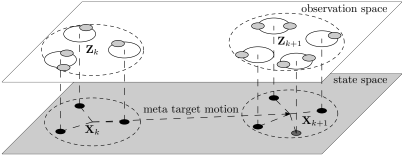
- A Random Finite Set $\mathbf{X}_k$ is a random variable
- Set cardinality $|\mathbf{X}_k| = N_{x,k}$ is a random variable.
- Each set memeber $\mathbf{x}_k^{(j)} \in \mathcal{X}_0$ is a random variable. Often $\mathcal{X}_0 = \mathbb{R}^{n_x}$.

- 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

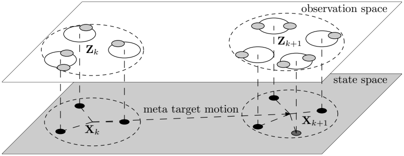
$$\begin{aligned} \mathbf{X}_k &= \emptyset \\ \mathbf{X}_k &= \left\{ \mathbf{x}_k^{(j)} \right\}_{j=1}^{N_{x,k}}, \ \mathbf{x}_k^{(j)} \in \mathbb{R}^{n_x} \text{ for } j = 1, \ldots, N_{x,k} \end{aligned}$$

- $\mathbf{X}_k$ 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 set of detections from set of sources to maintain estimates of targets' states.




- Processing of set of detections from set of sources to maintain estimates of targets' states.




- 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

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

- Multiple target equivalent to single target Bayes filter,

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

- Computationally intractable to implement due to set integrals.

- Use PHD filter instead.

# The Probability Hypothesis Density (PHD) filter

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

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

- An intensity function, not a pdf.

- Expected value: $1^{\text{st}}$ order moment of single target pdf $p\left(\mathbf{x}_k \middle| \mathbf{z}^k\right)$
- PHD: $1^{\text{st}}$ 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

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

- An intensity function, not a pdf.
- PHD filter propagates the PHD in time

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

- PHD filter is analogous to constant gain Kalman filter,

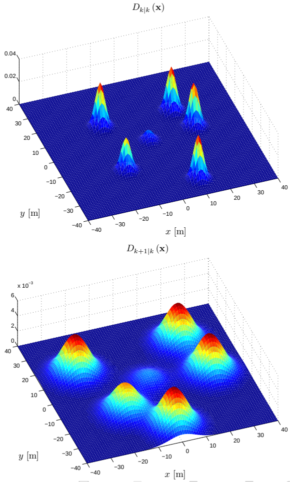
$$\ldots \quad \hat{\mathbf{x}}_{k|k} \quad \stackrel{\text{p}}{\rightarrow} \quad \hat{\mathbf{x}}_{k+1|k} \quad \stackrel{\text{c}}{\rightarrow} \quad \hat{\mathbf{x}}_{k+1|k+1} \quad \ldots$$

- Expected number of targets in any region $S \subseteq \mathcal{X}_0$ is

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

- The peaks in the intensity correspond to likely target locations.

- Implemented using SMC methods or distribution mixtures.



# Tracking of extended/group targets

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

- Point target assumption:
  $\leq 1$ 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.

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 $\mathbf{x}$,

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

# Group target state

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),

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



AUTOMATIC CONTROL
REGLERTEKNIK
LINKÖPINGS UNIVERSITET

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
- $\gamma_k$: number of detections, related to number of individuals in group


Gamma distributed random variable $\gamma$ (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 $\xi_k = (\mathbf{x}_k, X_k, \gamma_k)$

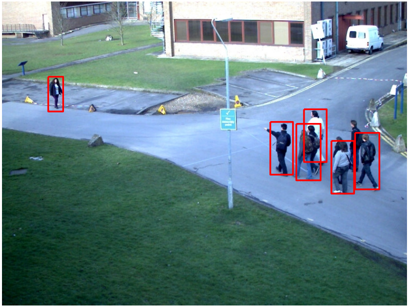
- $\mathbf{x}_k$: position and kinematics
- $X_k$: spatial extension, i.e. size and shape
- $\gamma_k$: number of detections, related to number of individuals in group

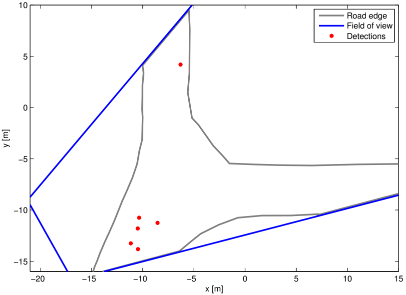

Group state $\xi_k$ is GGIW distributed,

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

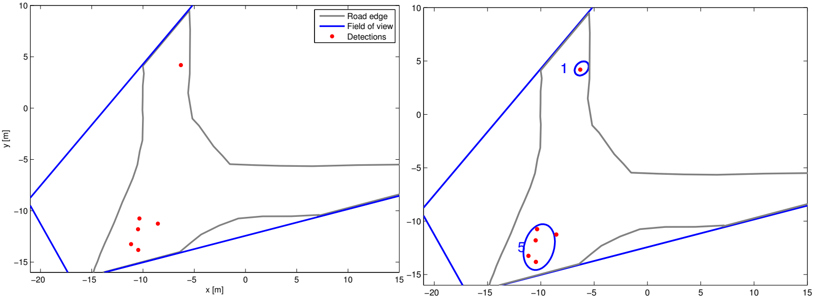
**1:** For each image, use pedestrian detector

**2:** Project each detection onto ground plane





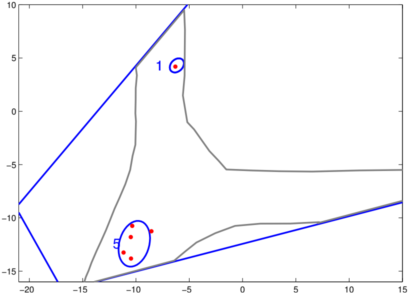
**3:** Use Extended/Group target PHD filter to track the groups





AUTOMATIC CONTROL
REGLERTEKNIK
LINKÖPINGS UNIVERSITET

**4:** Visualization: project filter output back into images

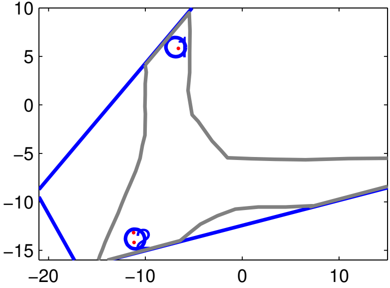




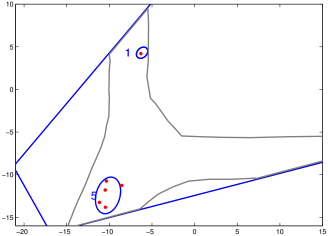
AUTOMATIC CONTROL
REGLERTEKNIK
LINKÖPINGS UNIVERSITET

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

# Group target tracking results



- 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.

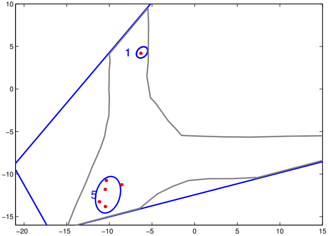



# Future work



Future work for group tracking in video includes

- Compare to alternative approach: track individuals, then form groups.
- Include methods for persons joining/leaving groups.
- PETS data has been used, e.g., for pedestrian counting. Initial results show that our methods compare well for pedestrian counting.




# Thank you for listening!

# Any questions?