3D Object Localization summary

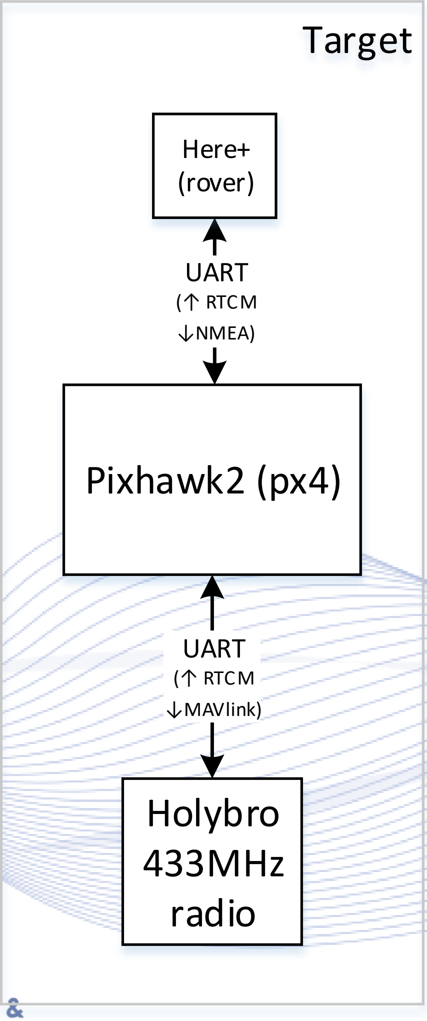
Contributors: R. Cunha (IST, Portugal)
J. Ramiro Martinez-de Dios (University of Seville)
E. Kakaletsis, C. Symeonidis Prof. Ioannis Pitas
Aristotle University of Thessaloniki
pitas@csd.auth.gr
www.aiia.csd.auth.gr
Version 2.6

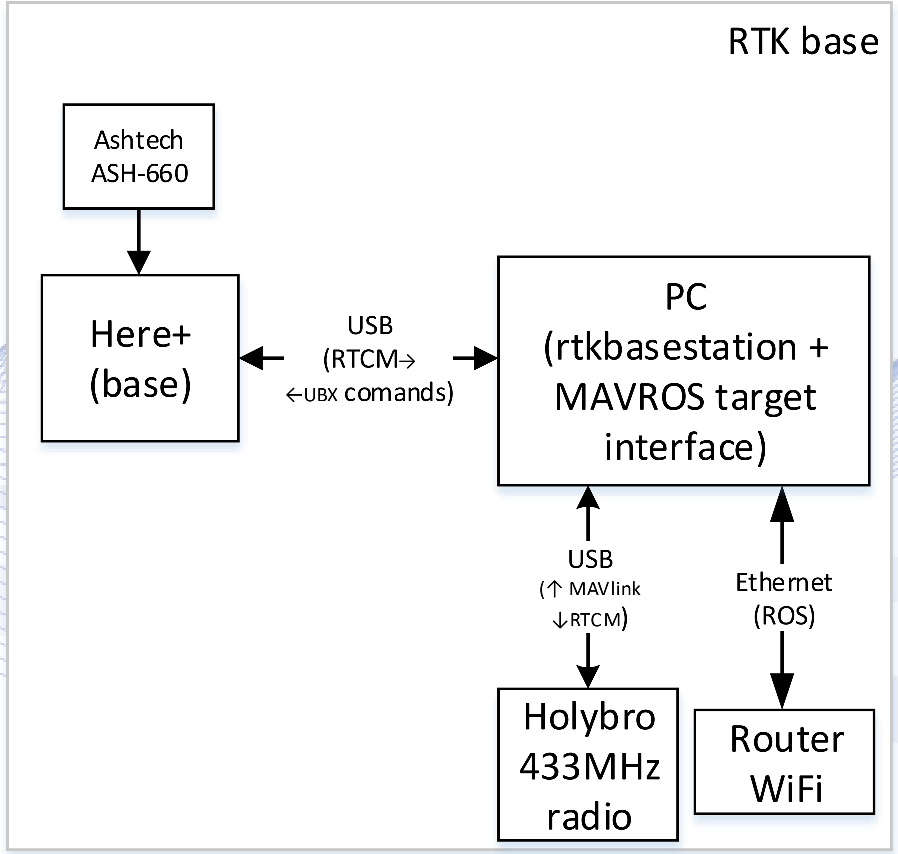
Artificial Intelligence & Information Analysis Lab

3D object localization

- **GPS object localization**
- Visual 3D object localization using 3D maps
- Multisensor object localization
- Multi-view object localization

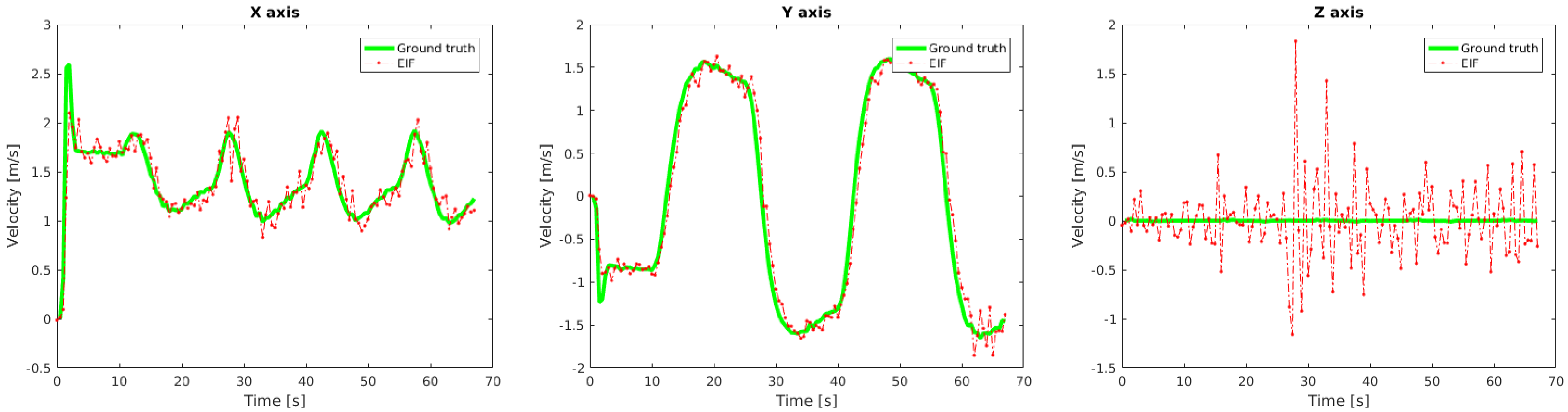
Target RTK GPS





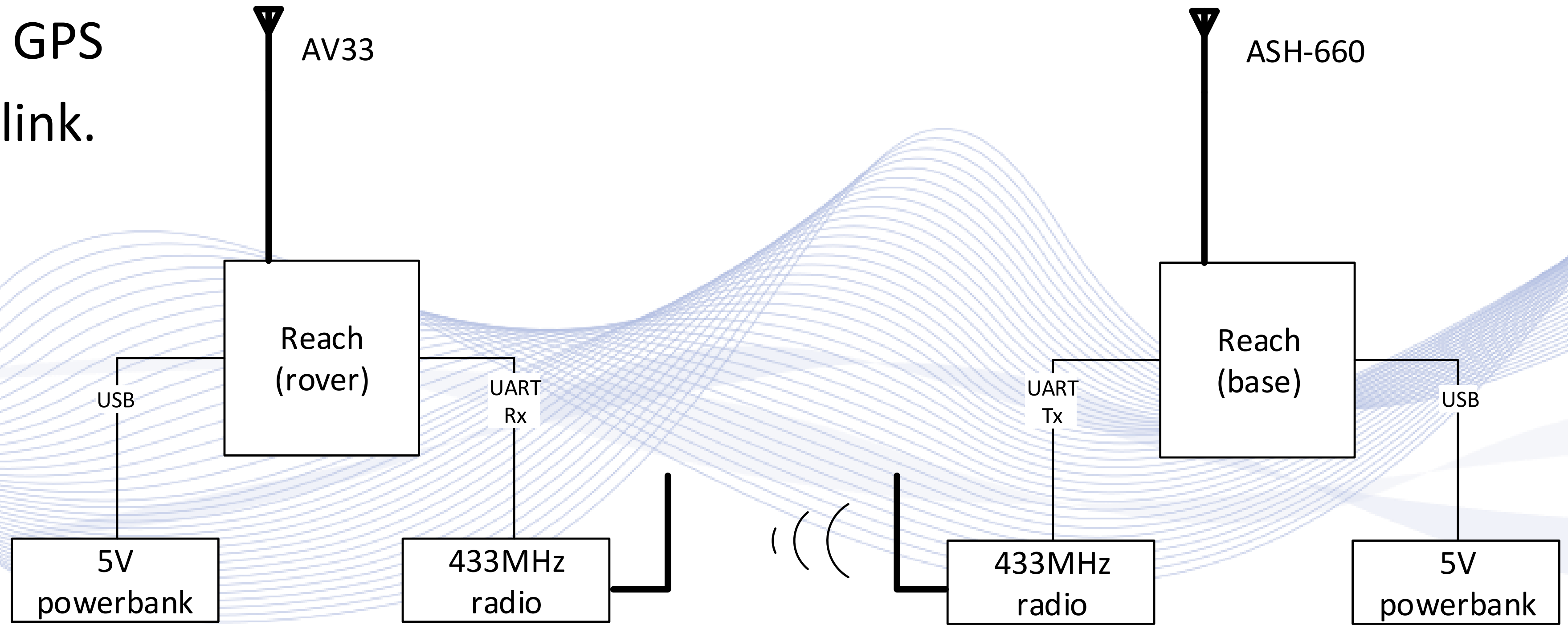
Optimal multi-sensor multi-drone 3D object localization & tracking

Target velocities



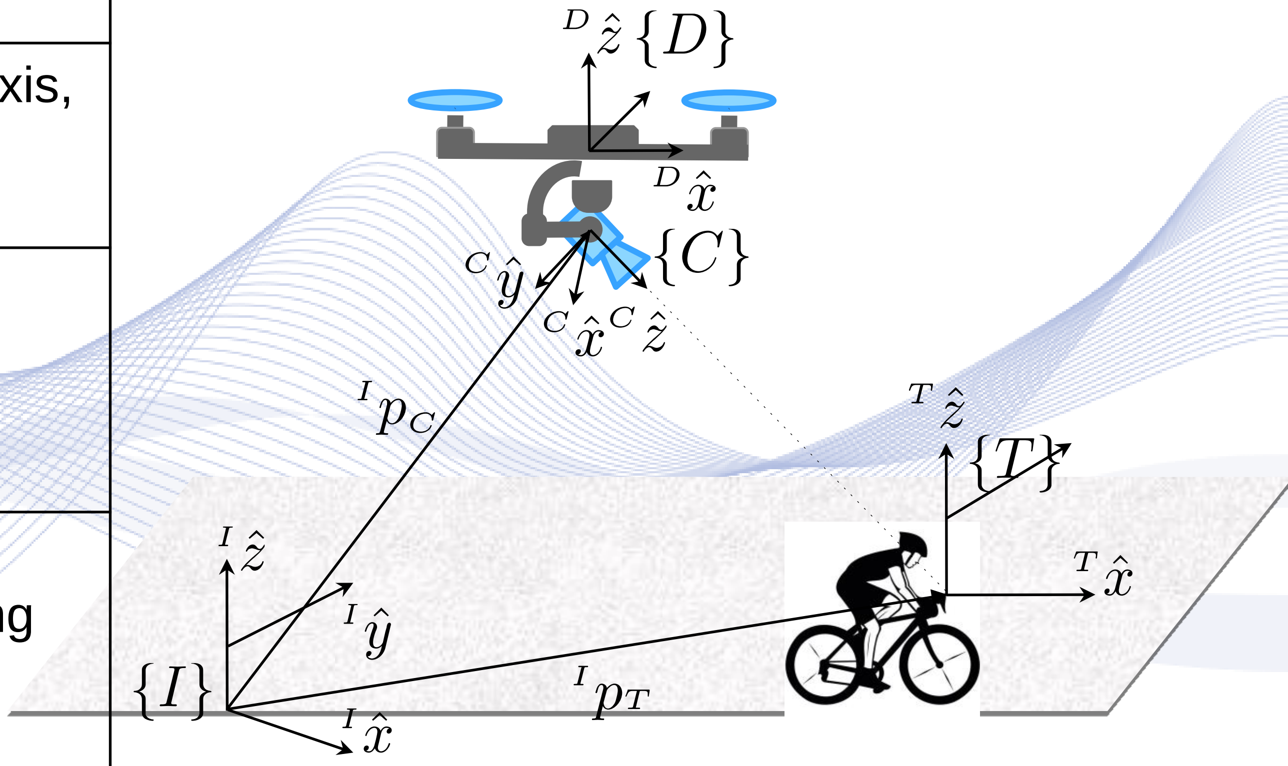
Object Localization using GPS

- Tests on Emlid Reach RTK GPS performance using:
  - GPS base station
  - Target GPS
  - Radio link.



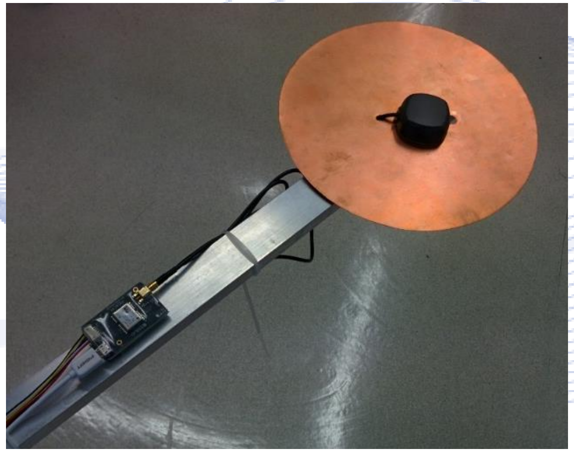
Coordinate systems

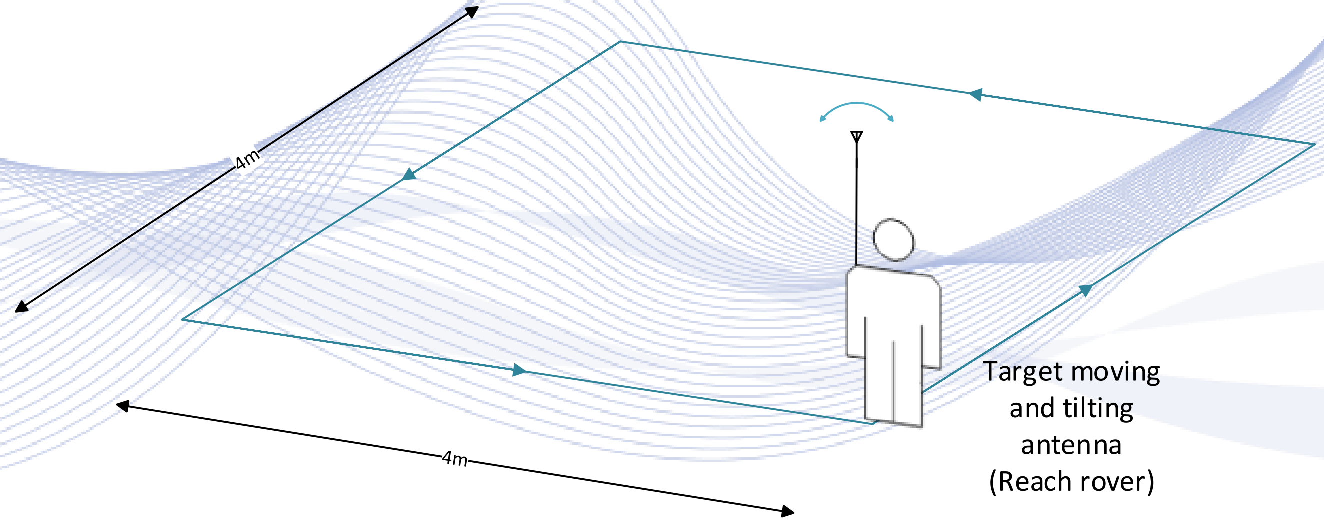
Name	Origin	Orientation
{I}Inertial	Mission parameter	X – Y – Z East – North – Up (ENU)
{D}Drone	Center of mass	x-axis aligned with front axis, z-axis aligned with rotors' axis
{C}Camera	Center of camera lens	z-axis aligned with optical axis, x-axis points right in image plane
{T}Target	(?) GPS antenna	z-axis up, x-axis aligned with heading



Object Localization using GPS

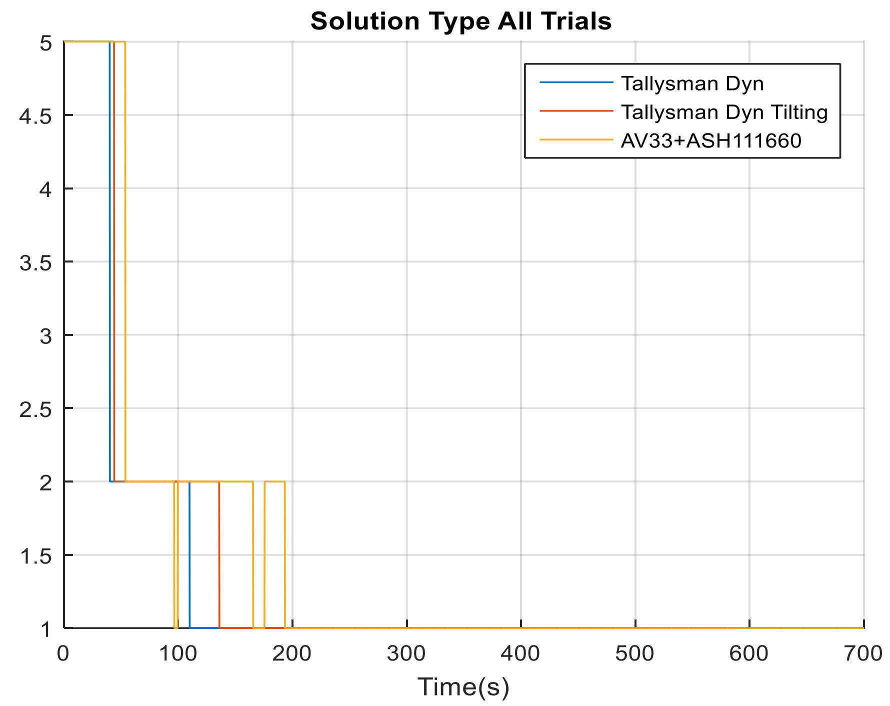
- Two different pairs of antennas:
  - 1 Ashtech ASH-660 on base + 1 Trimble AV33 on rover.
  - 2 Tallysman TW4721 (supplied with Reach) with ≥ 10cm ground plane.
- Target moving and antenna tilting.

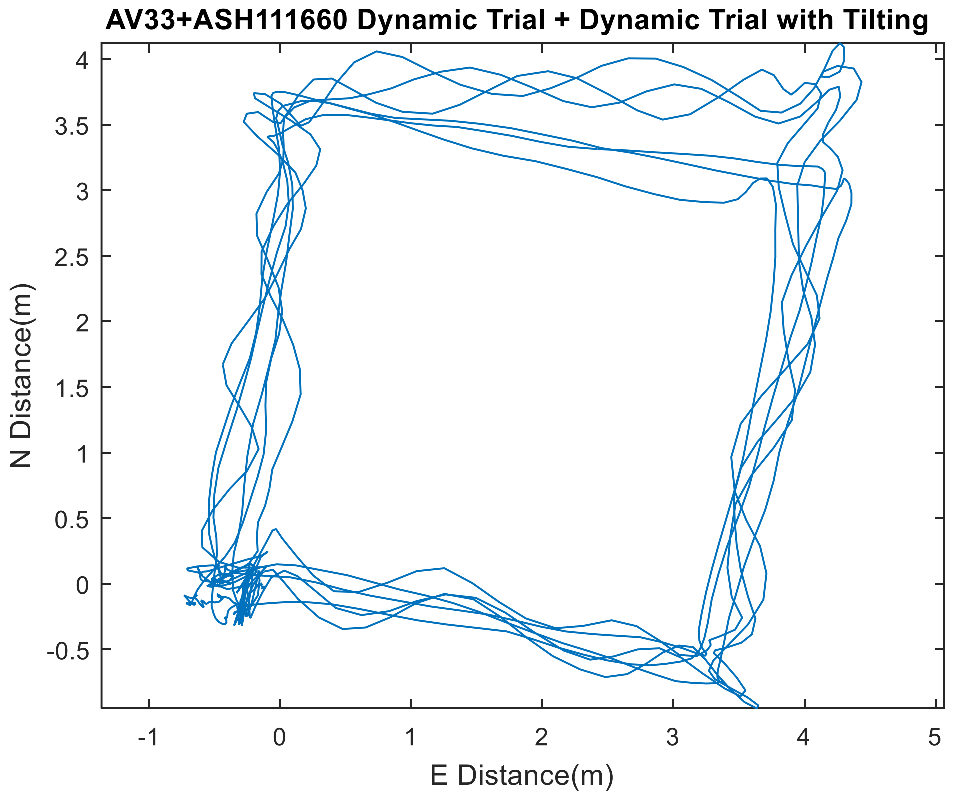




Object Localization using GPS

- Trial results:





Object Localization using GPS

- Conclusions
  - Rover with RTK fix solution for both antenna testings.
  - GPS RAW data (UBX) and position (LLH, NMEA or ENU) logged at 5Hz on Reach.
  - Better performance with ASH-660 + AV33 antennas.

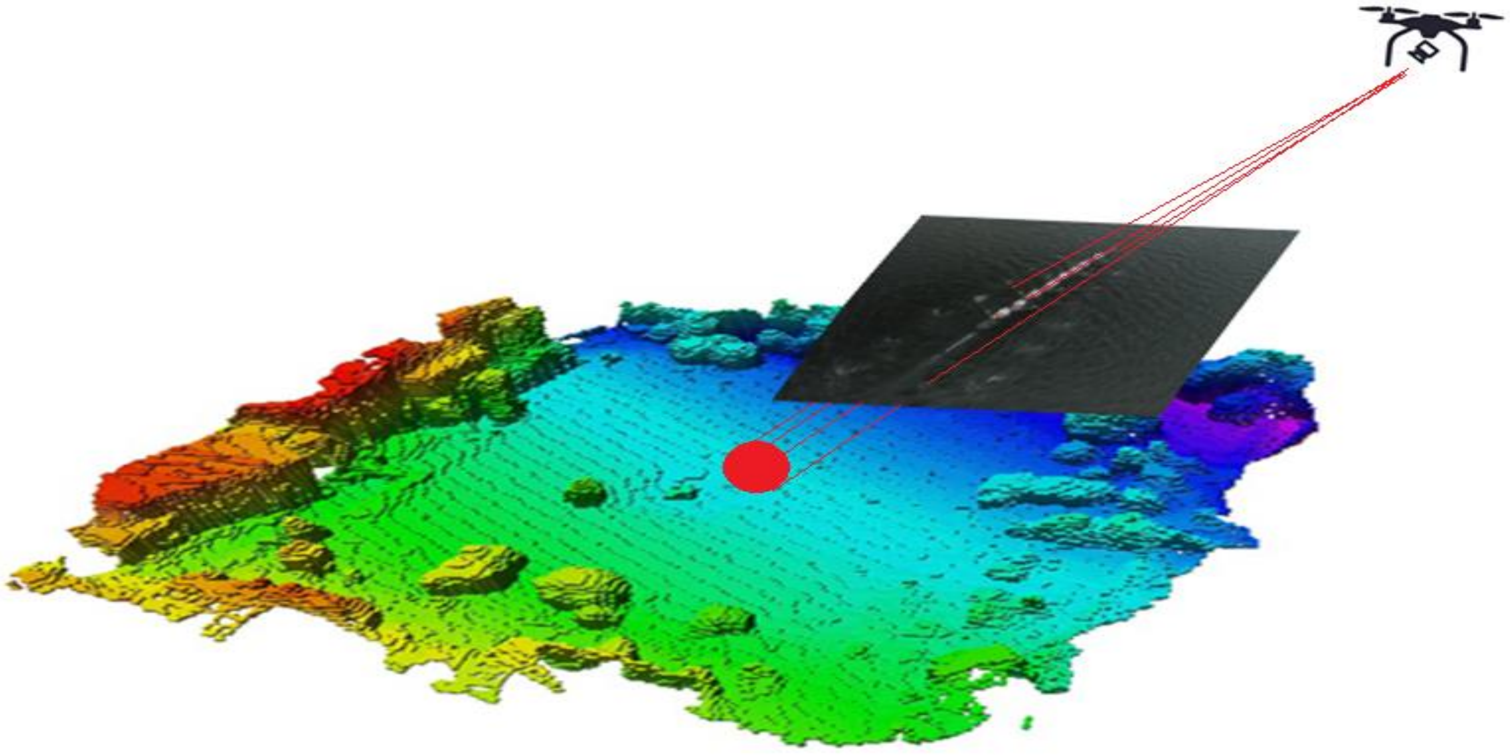
3D object localization

- GPS object localization
- **Visual 3D object localization using 3D maps**
- Multisensor object localization
- Multi-view object localization

3D object localization using 3D maps

- Specify target on the frame focal plane of the drone video camera by the object detection procedure.
- Raycasting from the object’s focal plane onto the 3D octomap according to camera drone COP.
- The rays from the COP (camera center of projection) hit the 3D octomap to specific voxels which is the location of the object in the 3D environment.
- Provided that we know the intrinsic and extrinsic parameters of the drone camera.

3D object localization using 3D maps

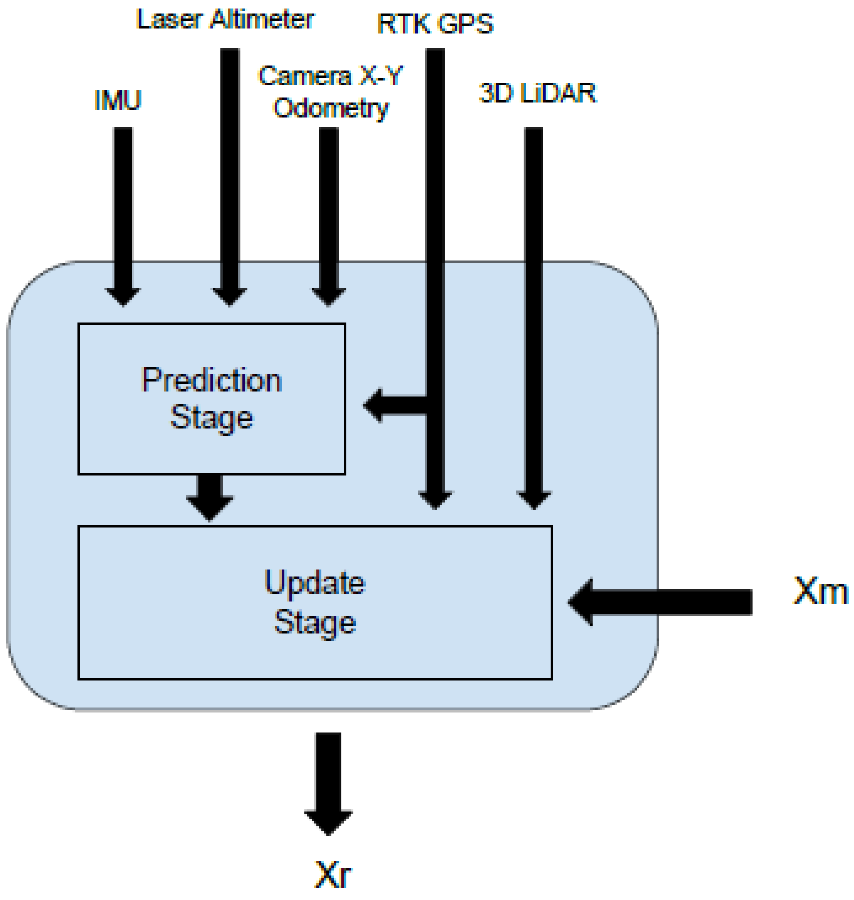


3D object localization

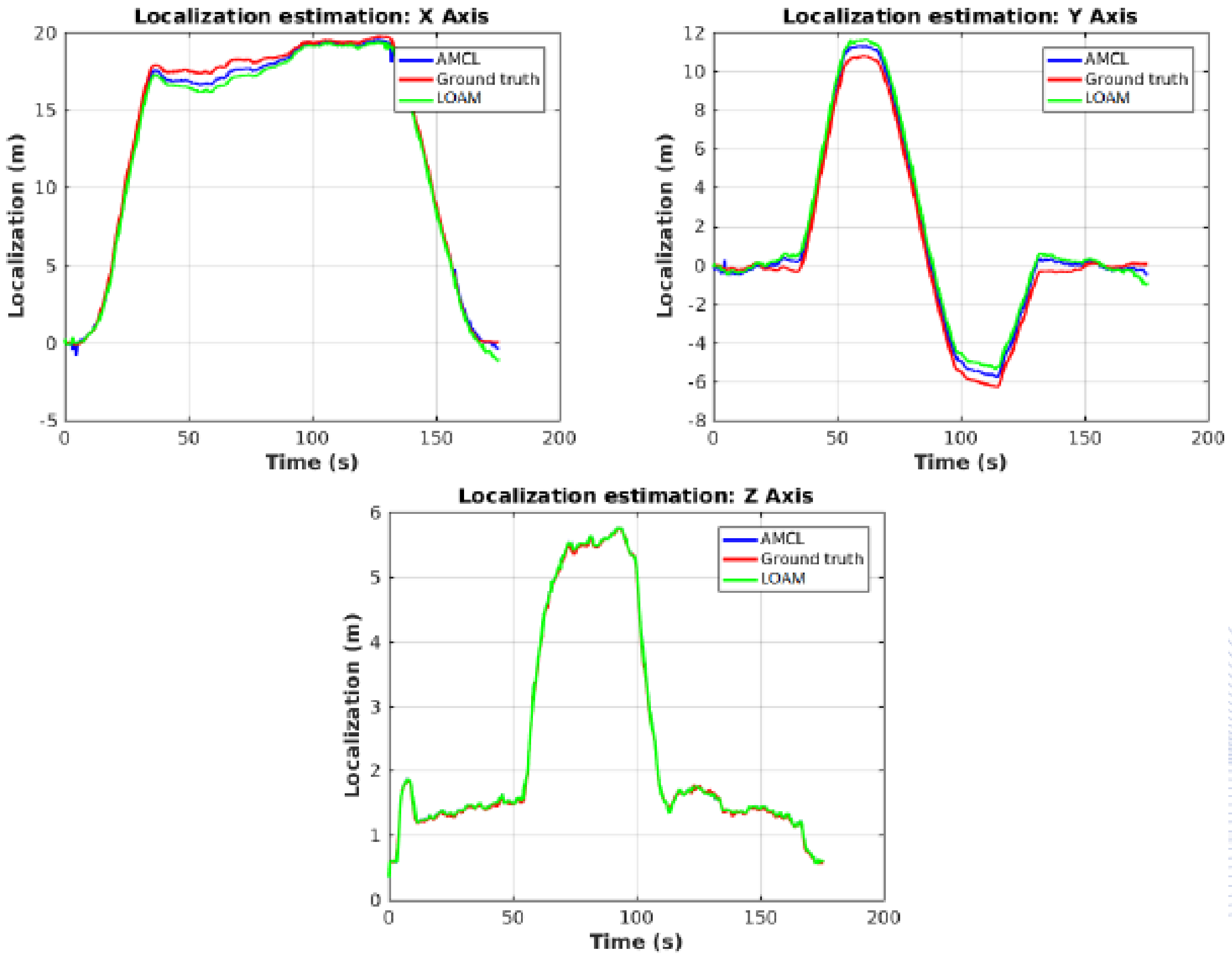
- GPS object localization
- Visual 3D object localization using 3D maps
- **Multisensor object localization**
- Multi-view object localization

6 DoF localization

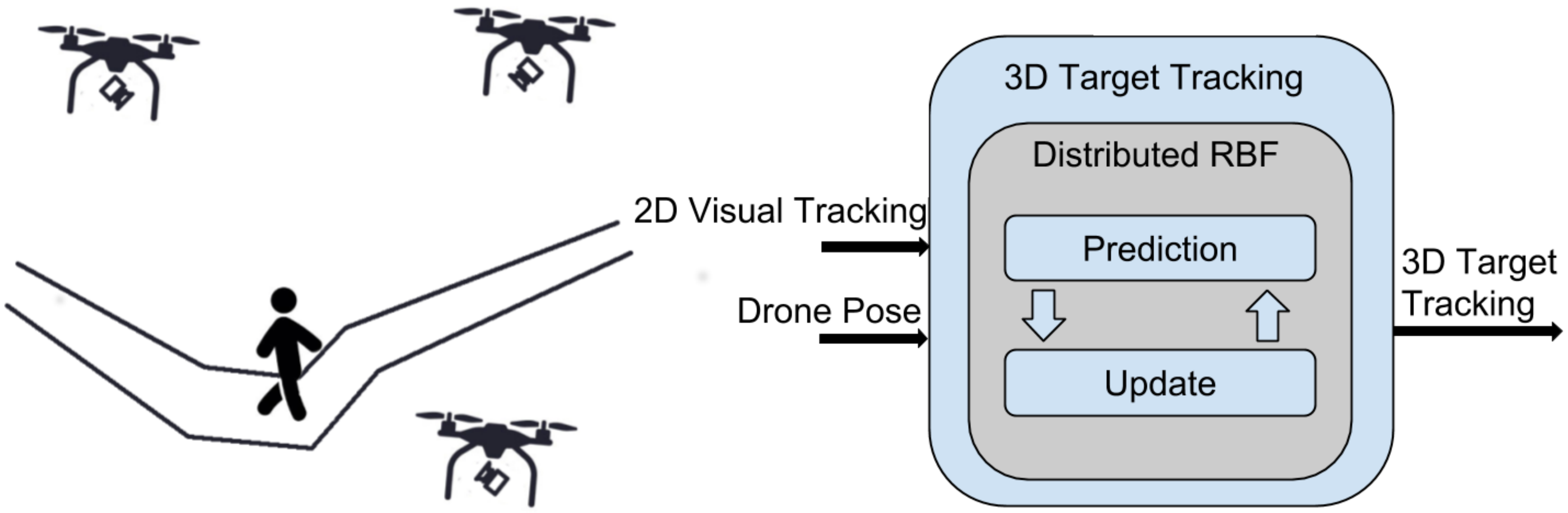
- Multi-sensor MCL for real-time 6DoF localization:
  - MCL Prediction: LIDAR odometry
  - Update of particles X, Y, Yaw: LIDAR point-clouds + camera features
  - Update of particles Z, pitch, roll: altimeter + IMU
  - MCL Update using the consistency of LIDAR point clouds with the map
- SLAM-based localization
  - SLAM that uses a previous map
  - Rely on previous maps but at the same time incorporates map changes



6 DoF localization



Optimal multi-sensor multi-drone 3D object localization & tracking

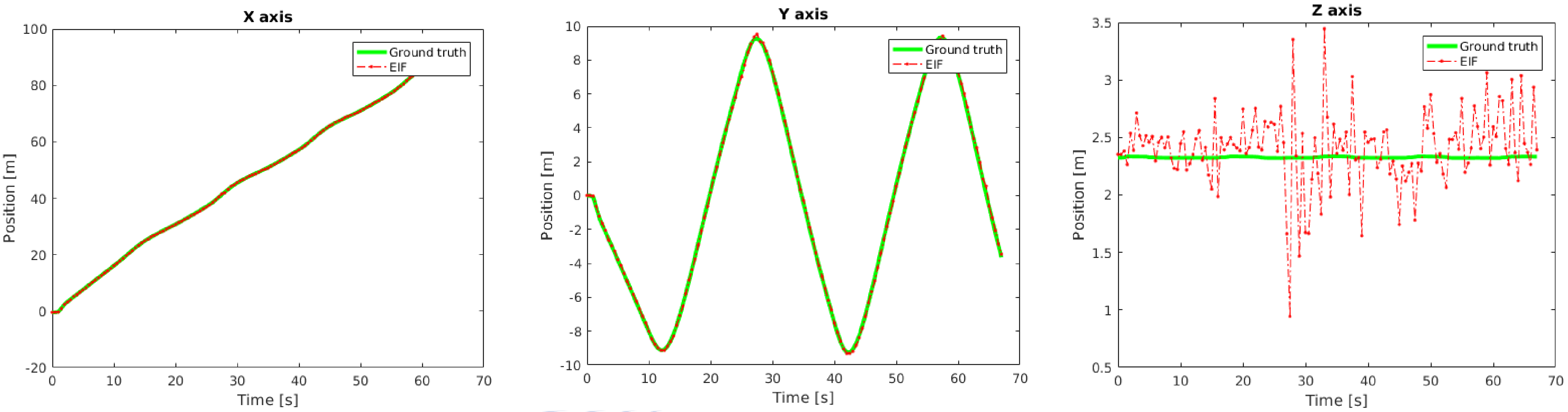


Optimal multi-sensor multi-drone 3D object localization & tracking

Artificial Intelligence &
Information Analysis Lab

Optimal multi-sensor multi-drone 3D object localization & tracking

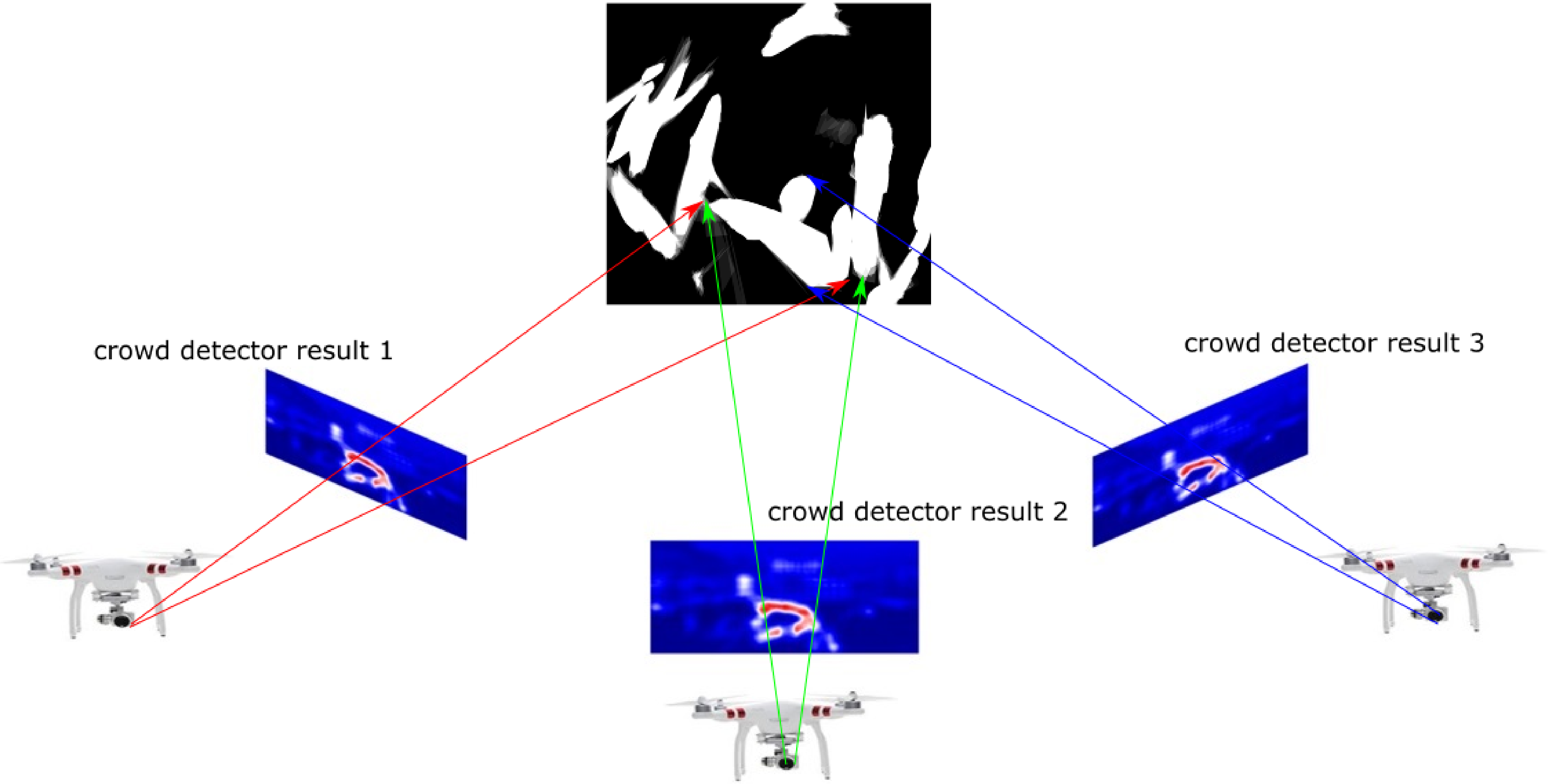
Target localization



3D object localization

- GPS object localization
- Visual 3D object localization using 3D maps
- Multisensor object localization
- **Multi-view object localization**

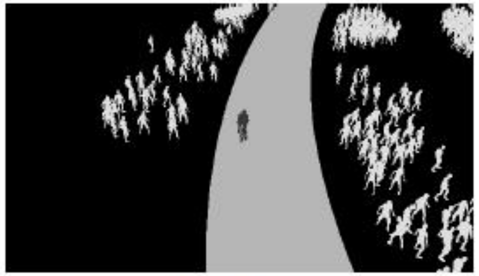
Multi-view Crowd Heatmap Fusion



Multi-view Crowd Heatmap Fusion

- *AirSim Synthetic Dataset:*
A mountainous terrain model with moving crowds set on the sides of a road


(a) RGB image


(b) Ground truth segmentation

- Three (3) simulated drones were deployed to follow three cyclists (one drone per cyclist).

- *Performance metrics:* mean IoU for the *N* number of frames

$$IoU_{mean} = \frac{1}{N}\sum_{i=1}^{N}\frac{Overlap_i}{Union_i}$$

Multi-view Crowd Heatmap Fusion

- ***Real world Dataset:*** Multiple drones (2) are seeing a crowd existing in a flat terrain

- Multi-view video sequences are derived.

Multi-view Crowd Heatmap Fusion

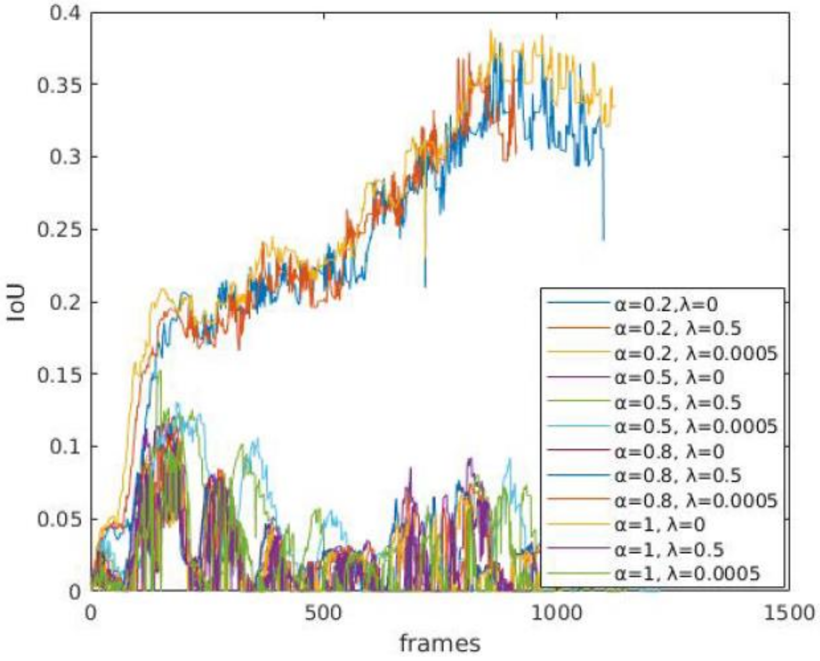


Table 1: Experimental Results

method	m IoU ($\pm std$)
single-view [4]	0.2048
multi-view $a = 0.2, \lambda = 0$	0.2330($\pm$0.0780)
multi-view $a = 0.2, \lambda = 0.5$	0.2195($\pm$0.0875)
multi-view $a = 0.2, \lambda = 0.0005$	**0.2565**($\pm$0.0815)
multi-view $a = 0.5, \lambda = 0$	0.0366 ($\pm$0.0316)
multi-view $a = 0.5, \lambda = 0.5$	0.0385($\pm$0.0343)
multi-view $a = 0.5, \lambda = 0.0005$	0.0400 ($\pm$0.0323)
multi-view $a = 0.8, \lambda = 0$	0.0294 ($\pm$0.0271)
multi-view $a = 0.8, \lambda = 0.5$	0.0258($\pm$0.0265)
multi-view $a = 0.8, \lambda = 0.0005$	0.0284 ($\pm$0.0277)
multi-view $a = 1, \lambda = 0$	0.0198 ($\pm$0.0234)
multi-view $a = 1, \lambda = 0.5$	0.0176($\pm$0.0231)
multi-view $a = 1, \lambda = 0.0005$	0.0210 ($\pm$0.0241)

- The proposed method outperforms the single-view method
- When the exponential decay of the blended probability is conducted with smoother way (for λ smaller), the proposed method introduces the forgetting policy to the previous detections with more significance.
- The maximum of mean IoU value for probability blending with parameter α= 0.2 and forgetting policy namely for λ=0.0005

Q & A

Thank you very much for your attention!

More material in
http://icarus.csd.auth.gr/cvml-web-lecture-series/

Contact: Prof. I. Pitas
pitas@csd.auth.gr