

MultiDrone



Introduction to Multidrone imaging

Presenter: Prof. Ioannis Pitas
Aristotle University of Thessaloniki
pitas@aiia.csd.auth.gr
www.multidrone.eu
Presentation version 1.11

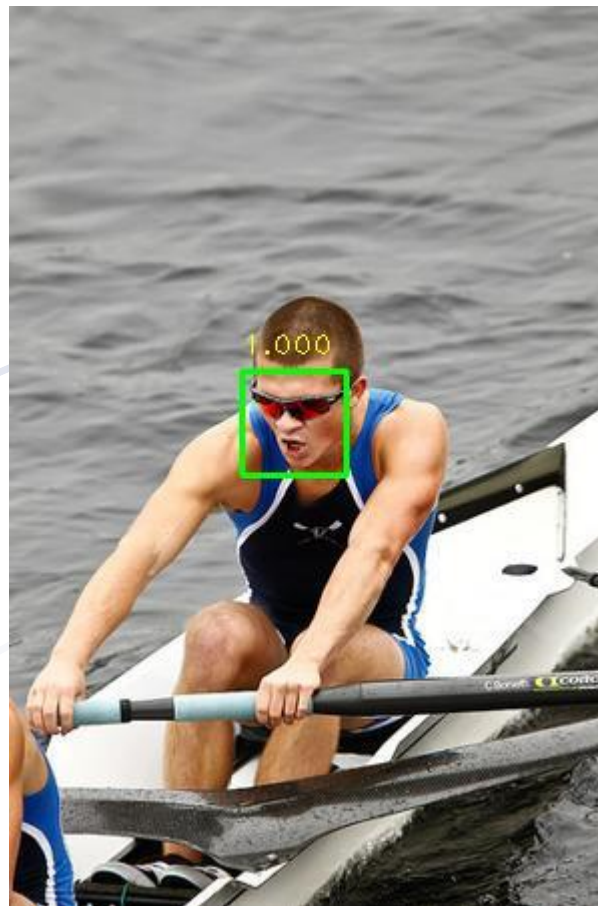
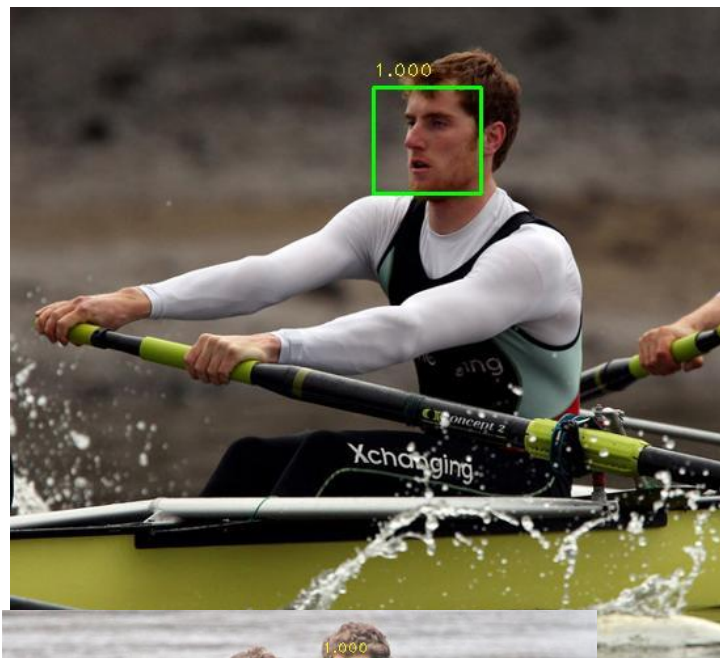


Multidrone case study: sports AV shooting



Rowing boat race

MultiDrone

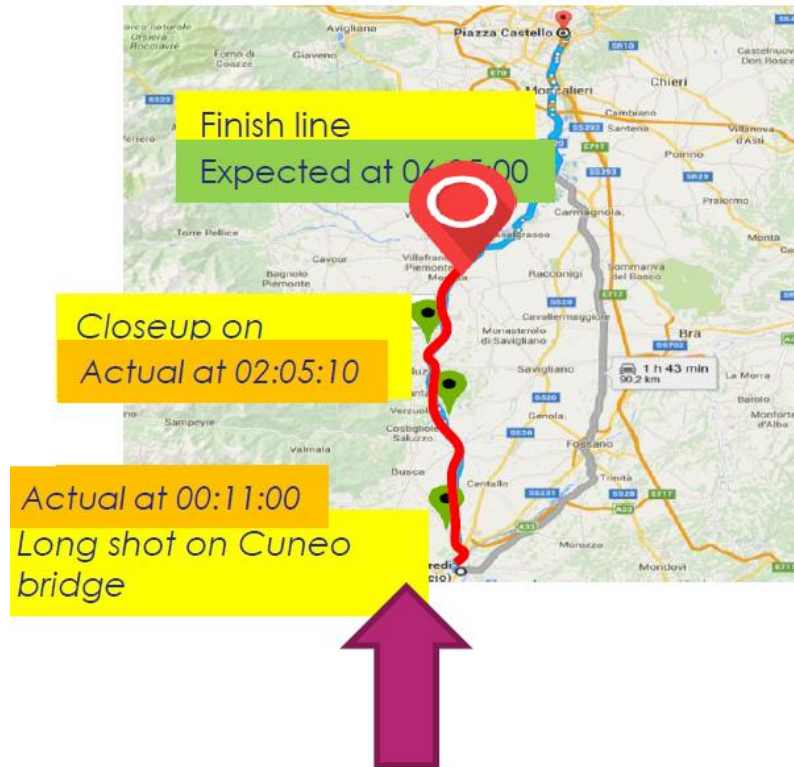


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)



Giro d'Italia

MultiDrone



<<Accident Detected>>



Other
view 2

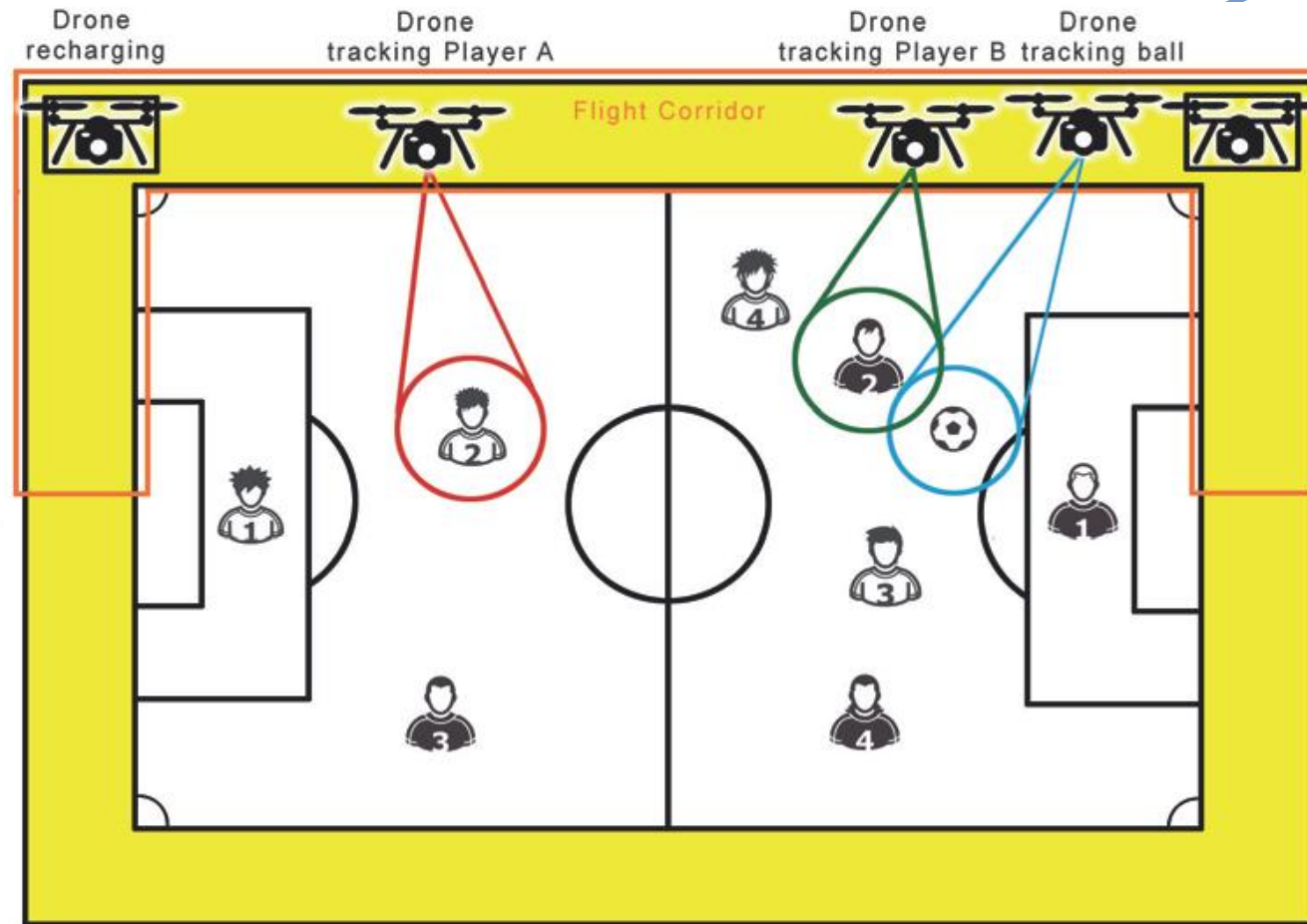


Panoramic (stitched)



Multi-angled shooting of a football match

MultiDrone



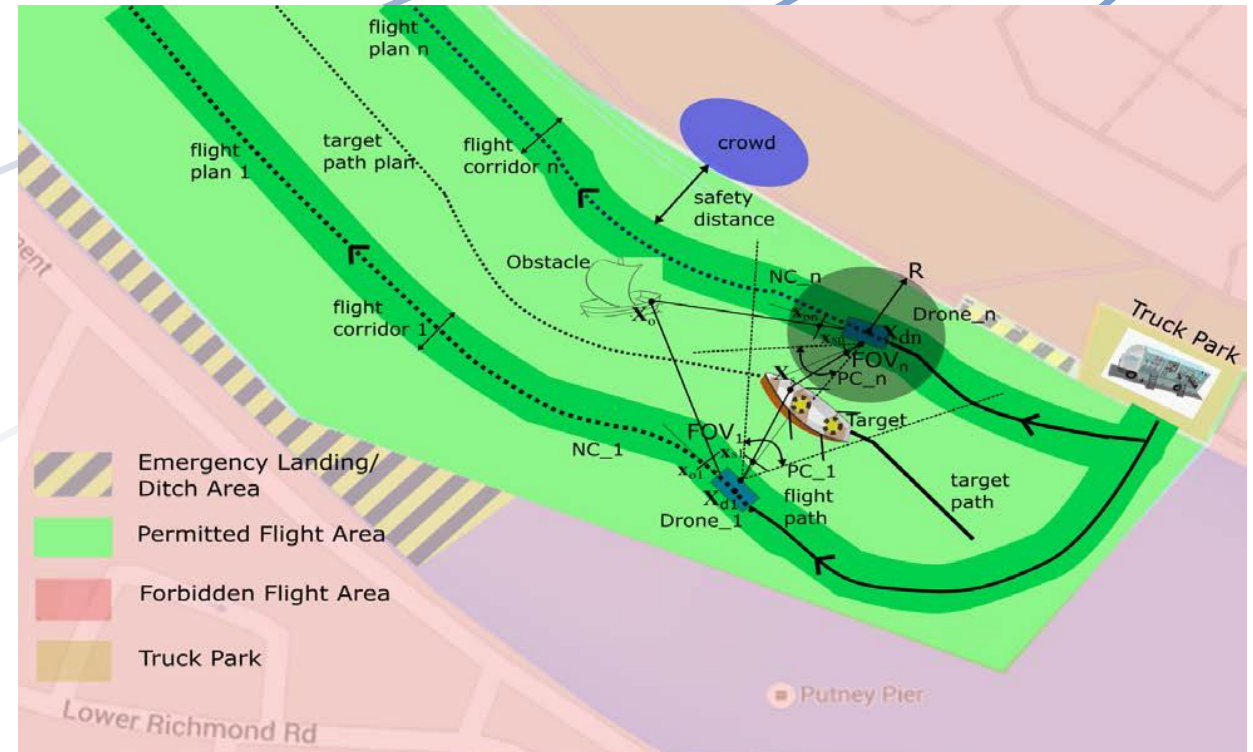
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)



Challenges in boat race shooting



- A) Drone decisional autonomy, robustness and safety:
 - Crowd, obstacle detection and avoidance
 - Emergency landing site detection.
- B) Multiple drone active perception and AV shooting:
 - Target tracking and following
 - Cinematographic shooting



Drone vision objectives and challenges



- **A) Improved multiple drone decisional autonomy, robustness and safety.**
- B) Innovative, safe and fast multiple drone active perception and AV shooting.



A1) Improved multiple drone decisional autonomy, robustness and safety: Objectives and results



- **OA1: Adaptive/cooperative/dynamic (online) multiple drone (re)planning:**
 - A novel dynamic/cooperative (re-)planner has been developed.
- **OA2: Improved, easy and transparent interaction with the production director and his/her crew.**
 - A novel drone cinematography mission definition and language has been developed.
 - An advanced director dashboard (interface) has been designed.



A2) Improved multiple drone decisional autonomy, robustness and safety: Objectives and results



- **OA3: Improved decisional / cognitive autonomy and robustness.**
 - Dynamic docking/recharge/emergency landing planning has been designed, including handling e.g.,
 - crowd detection, emergency landing site detection.
 - intelligent autonomous emergency handling, e.g.:
 - Flight replanning in case of emergency landing.
- **OA4: Improved safety during multiple drone mission execution:**
 - a) embedded flight regulation compliance is foreseen;
 - b) enhanced vision-based crowd avoidance is implemented;
 - c) a-priori emergency landing planning and autonomous multiple drone emergency landing re-planning has been designed.



A3) Improved multiple drone decisional autonomy, robustness and safety: Objectives and results.



- **OA5: Robust video streaming, communication and synchronization.**
 - Robust LTE/Wifi communication architecture has been implemented.
 - Full HD video compression and drone2ground streaming is developed.
 - Video frame timestamping and synchronization has been implemented.



MULTIDRONE challenges and objectives



- A) Improved multiple drone decisional autonomy, robustness and safety.
- **B) Innovative, safe and fast multiple drone active perception and AV shooting.**



B1) Innovative, safe/fast multiple drone active perception and AV shooting: Objectives and results.



- **OB1: Fast multiple drone semantic world modelling during pre-production, for e.g.:**
 - a) 3D world modelling procedures have been designed.
 - b) 3D map KML semantics have been designed to be populated.
- **OB2: Fast innovative multiple drone vision- and GPS-/RFID-based target tracking and shooting techniques:**
 - Fast novel real-time embedded target (e.g., boat, cyclist, football player) detection and tracking (visual/GPS) techniques have been developed.



B2) Innovative, safe/fast multiple drone active perception and AV shooting: Objectives and results.



- **OB3: Multiple drone AV shooting intelligence:**
 - Novel path/formation/gimbal/camera control techniques have been developed.
 - A novel drone cinematography has been developed.
- **OB4: Improved multiple drone human-centered visual information analysis both for individual persons and for crowds:**
 - Fast and improved crowd detection has been developed.
 - Fast novel real-time embedded cyclists and football player detection and tracking techniques have been developed.





Other issues-challenges

- **Security and privacy issues:**
 - Complete ethics review.
 - Development of privacy-by-design protection methods.
- **Overcoming barriers/obstacles due to regulations and other factors:**
- Plans on contribution to EBU standards on media production.
- **Boosting public awareness and dialog:**
 - 1 organized event (5/2018), 1 planned (8/2018).



Methodology



- **End user requirements.**
- **HW/SW system specs, design, implementation, integration.**
- **Strong interplay between:**
 - a) mission (AV shooting) planning, mission control/execution;
 - b) active perception
- **Pre-production:**
 - semantic world mapping
 - mission planning.



Methodology



- **Production:**

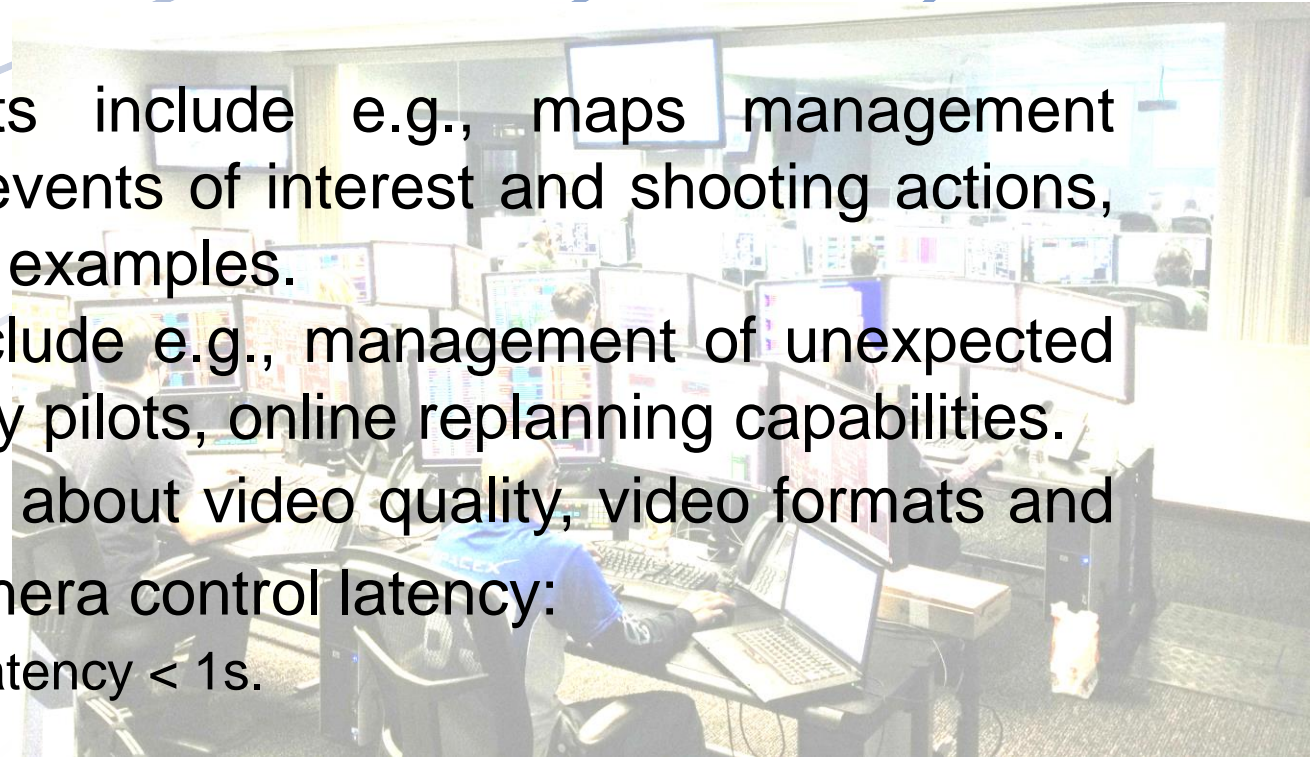
- multiple drone flight/formation control
- active perception (multiple drone and target localization tracking),
- cinematographic AV shooting.
- safety/emergency monitoring/sensing
- emergency handling at the production phase.



Media production requirements



- Distinction between pre-production (planning) and production (live) requirements:
 - Pre-production requirements include e.g., maps management functionalities, definition of events of interest and shooting actions, visual target specification by examples.
 - Production requirements include e.g., management of unexpected events, emergency control by pilots, online replanning capabilities.
- Other general requirements are about video quality, video formats and codecs, camera control and camera control latency:
 - E.g., we want camera command latency $< 1\text{s}$.



Media production requirements - shooting



- Important part of media production requirements are about shooting capabilities of the system ...

Central framing shot selection

Reverse angles shot transition selection

Rule of thirds shot selection

180 degree rule shot transition selection

Avoid same camera perspectives

Leave Looking room for subjects



Media production requirements – basic multi-drone shooting

MultiDrone



System Platform requirements



- These requirements are from the point of view of media production, therefore they are mainly about:
 - Drone physical parameters (weight, maximum speed, etc.)
 - Autonomy in flight and in perception
 - Storage, communication, control
 - Logistics
 - Interfaces to studio
 - Director and flight supervisor dashboards.





Personnel and Roles

- **Director.** Person in charge of the media production. Specify the shots to be taken by the drone team. He will interface with the system through the Dashboard.
- **Supervisor Operator.** Person in charge of the security of the system. Throughout the Supervisor module, this person will validate plans as safe, and will give a green light to the Director.
- **Drone Pilots(?).** For security reasons, each drone will have a human pilot in charge to take over in case of emergency.
- **Cameramen (?).** There will be a cameraman in charge of each camera on board the drones to take manual control if required by Director.



Drone vision for cinematography: HW issues



- 1. Drone platform:**
 1. Flight machine
 2. AV and visual perception payload
- 2. Ground station platform**
- 3. Drone-ground station communications**
- 4. Human centered interfaces:**
 1. Director, (photographers?)
 2. Flight supervisor, (pilots?).



Overview of the drone hardware



Drone core

Flight Control Unit with main sensors, RTK GPS, Thales LTE & Wi-Fi module, back-up radio for commands

Batteries

At least 2 batteries on the drone

Drone platform

Frame, arms, landing gears, propulsion systems, ESCs

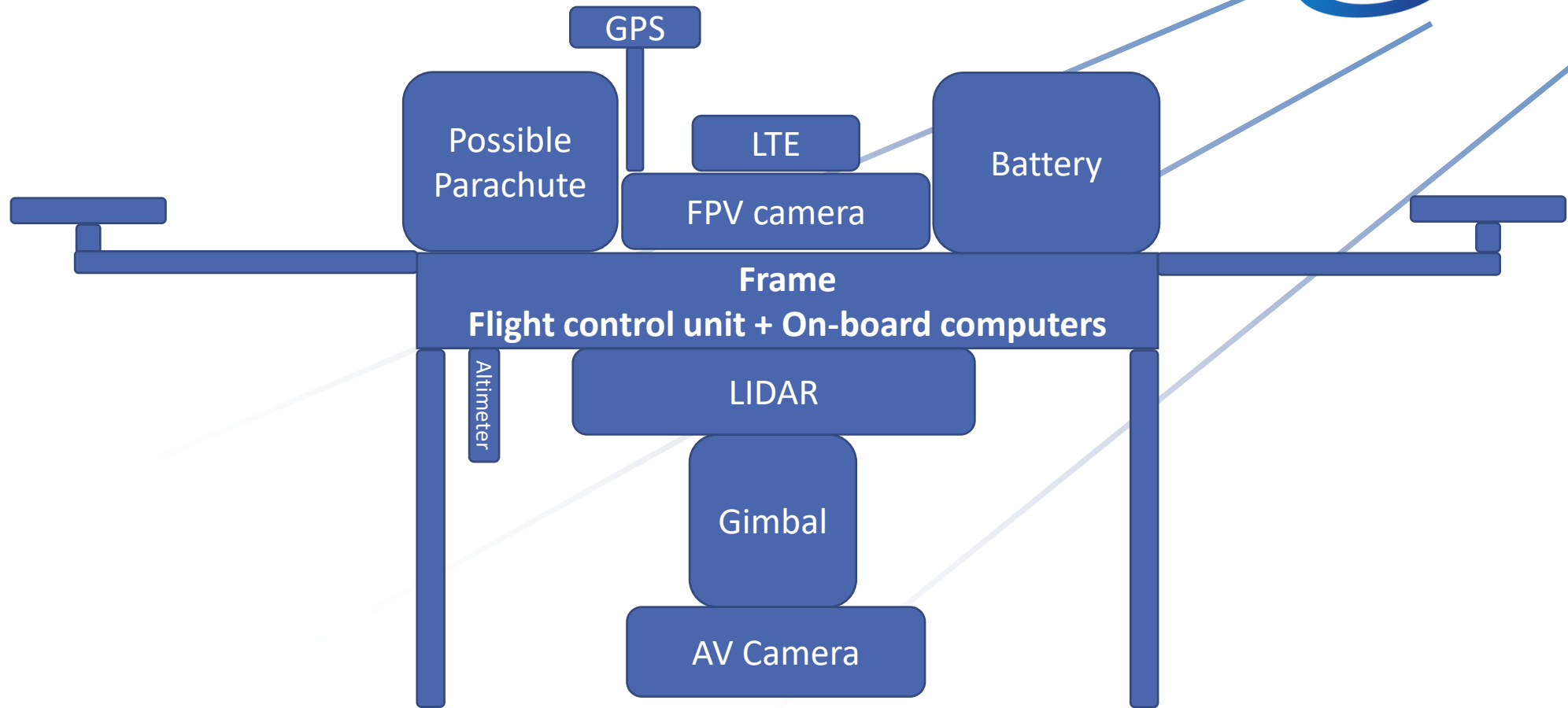
Audio-Visual Payload

Audio-Visual camera and its motorised lens, 3 axis gimbal (for stabilisation), high quality images storage

Flight Payload

Navigational camera, LIDAR, onboard computers, safety system with possibly parachute system

General Drone Architecture



Drone Communications

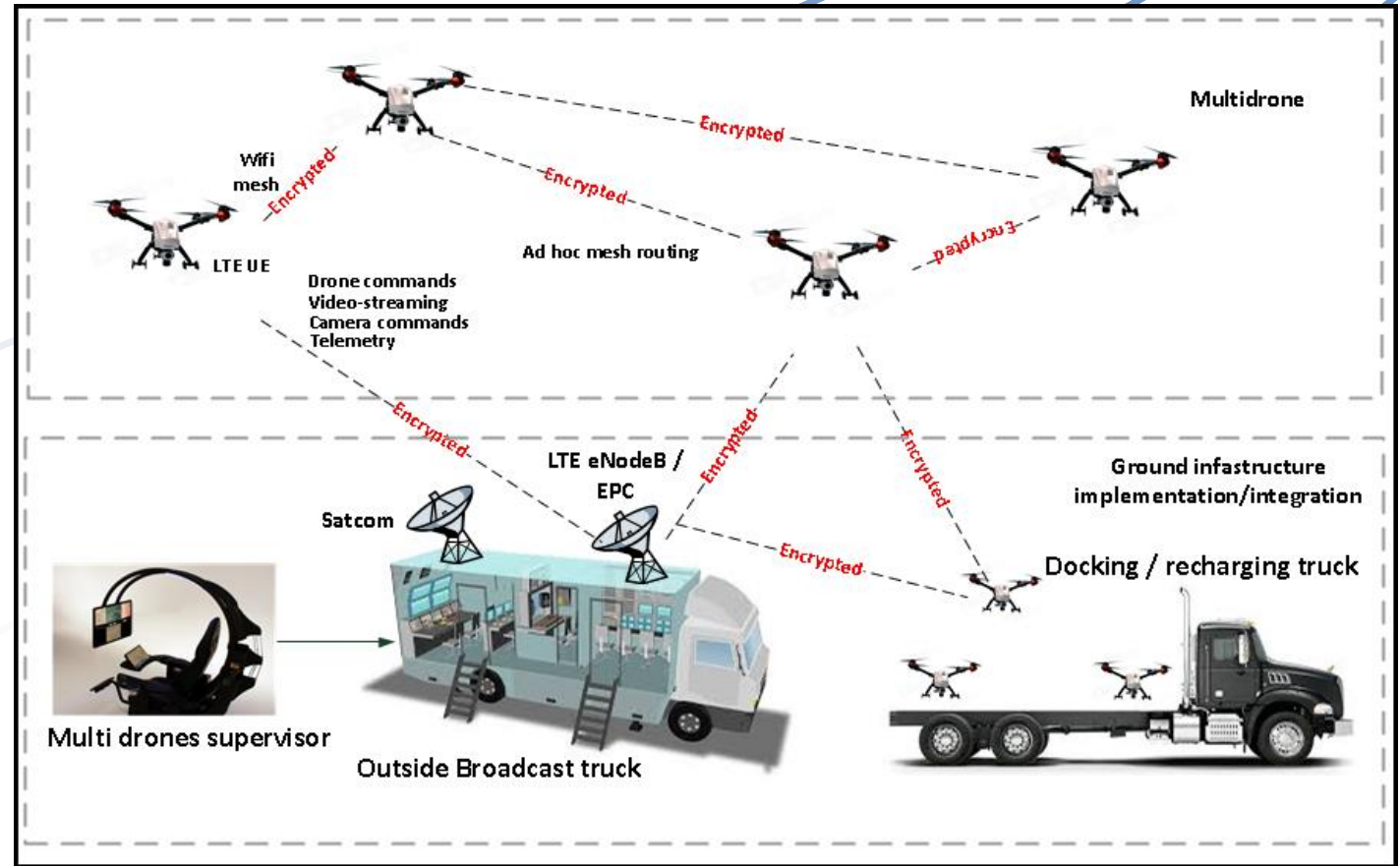
- **Communication infrastructure**
- Video streaming



Communication infrastructure



- Drone2Drone Communication.
- Drone2Ground communication.
- Live broadcasting.

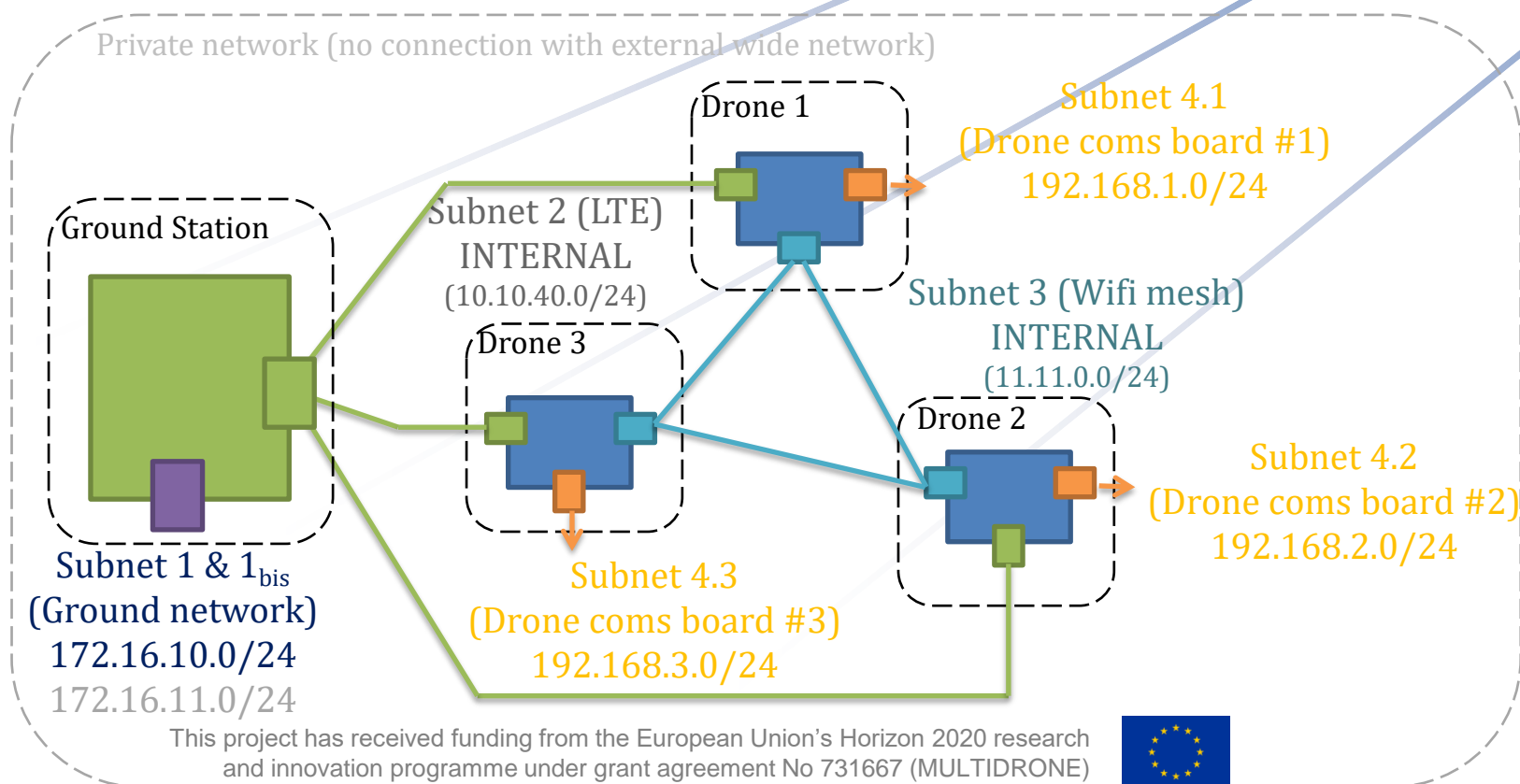


Communications Infrastructure

MultiDrone



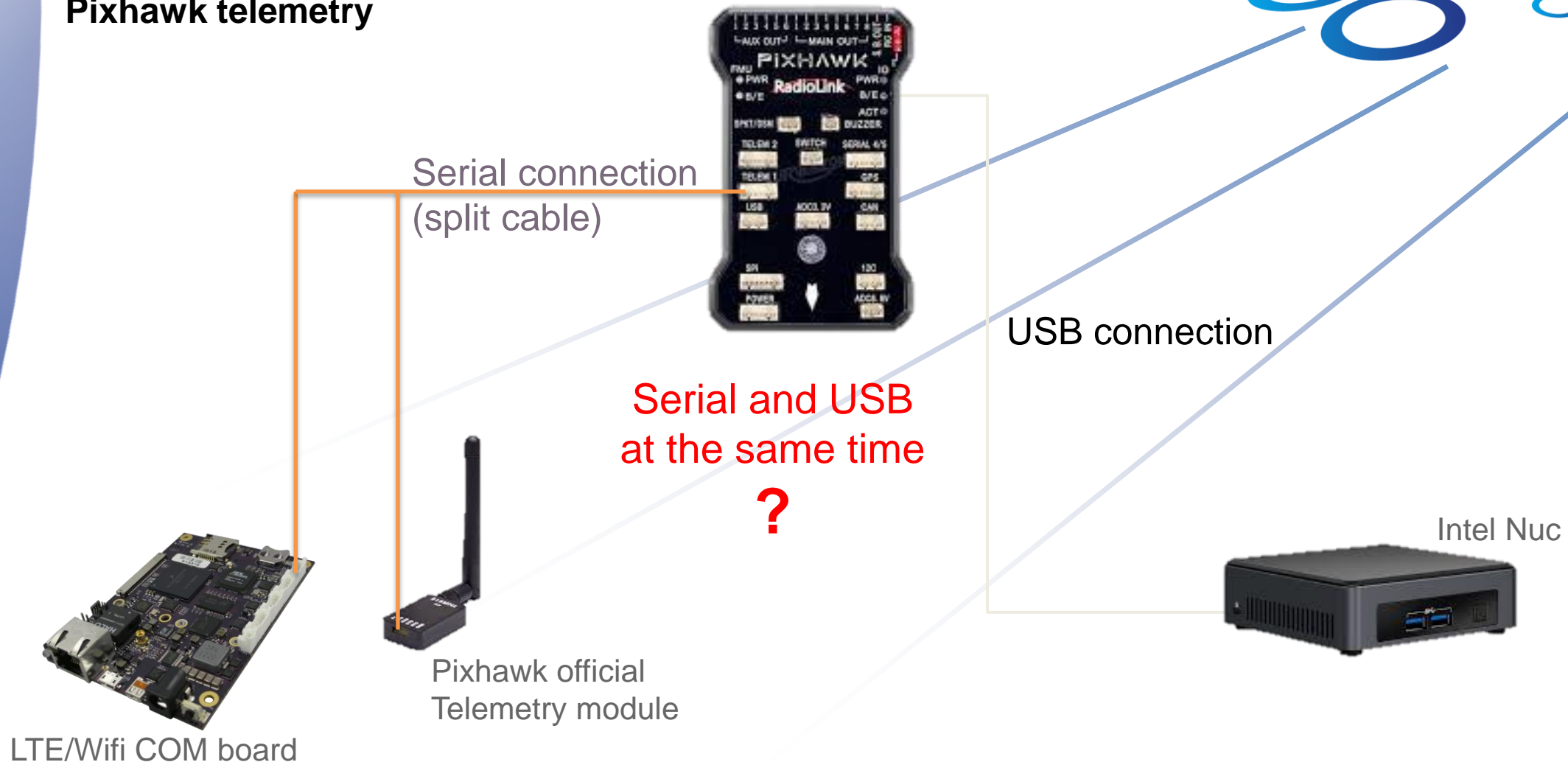
Objective: Secured and resilient transparent IP access to drones / ground station (LTE and WiFi).



Communications Infrastructure

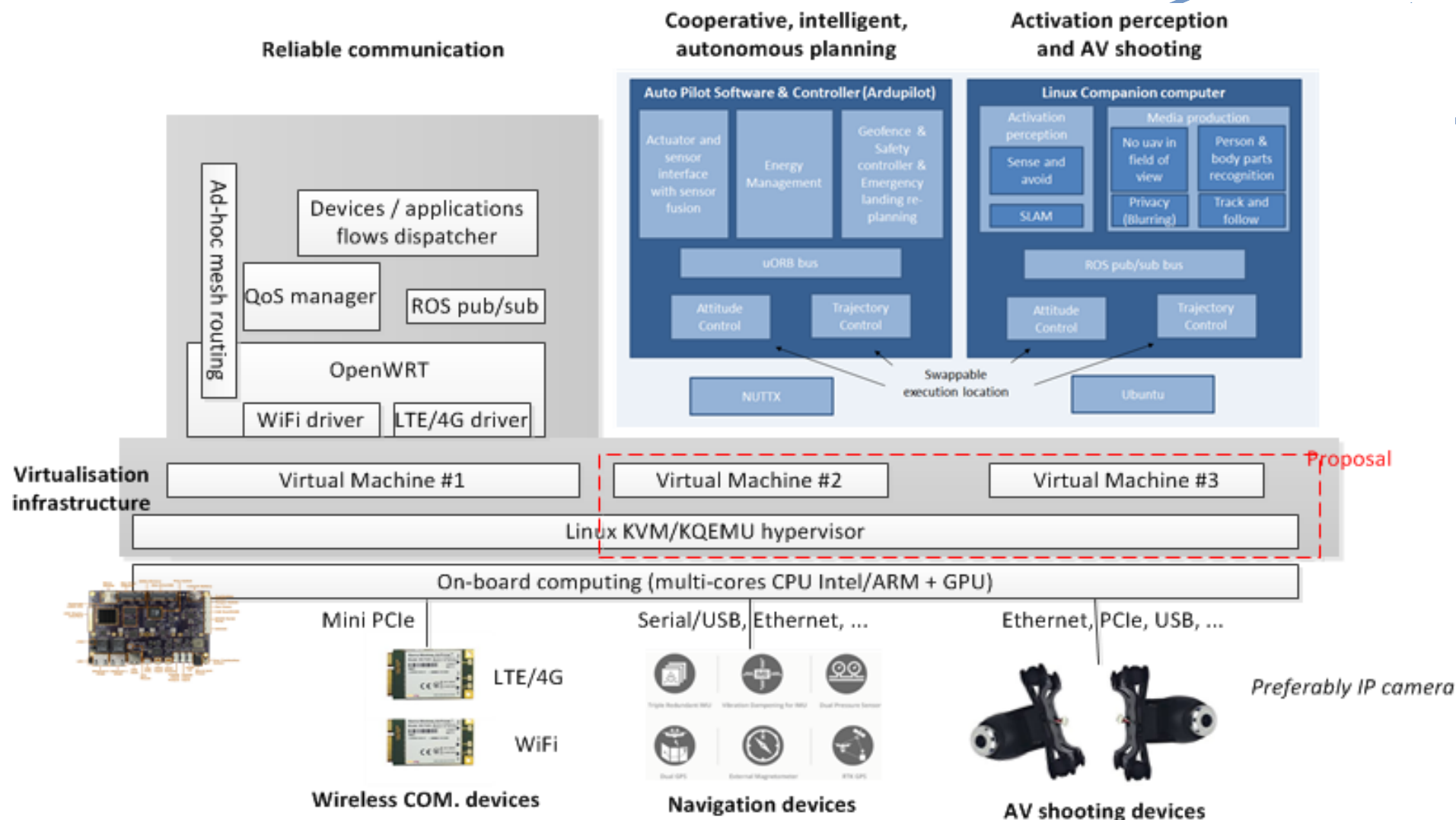
Pixhawk telemetry

MultiDrone



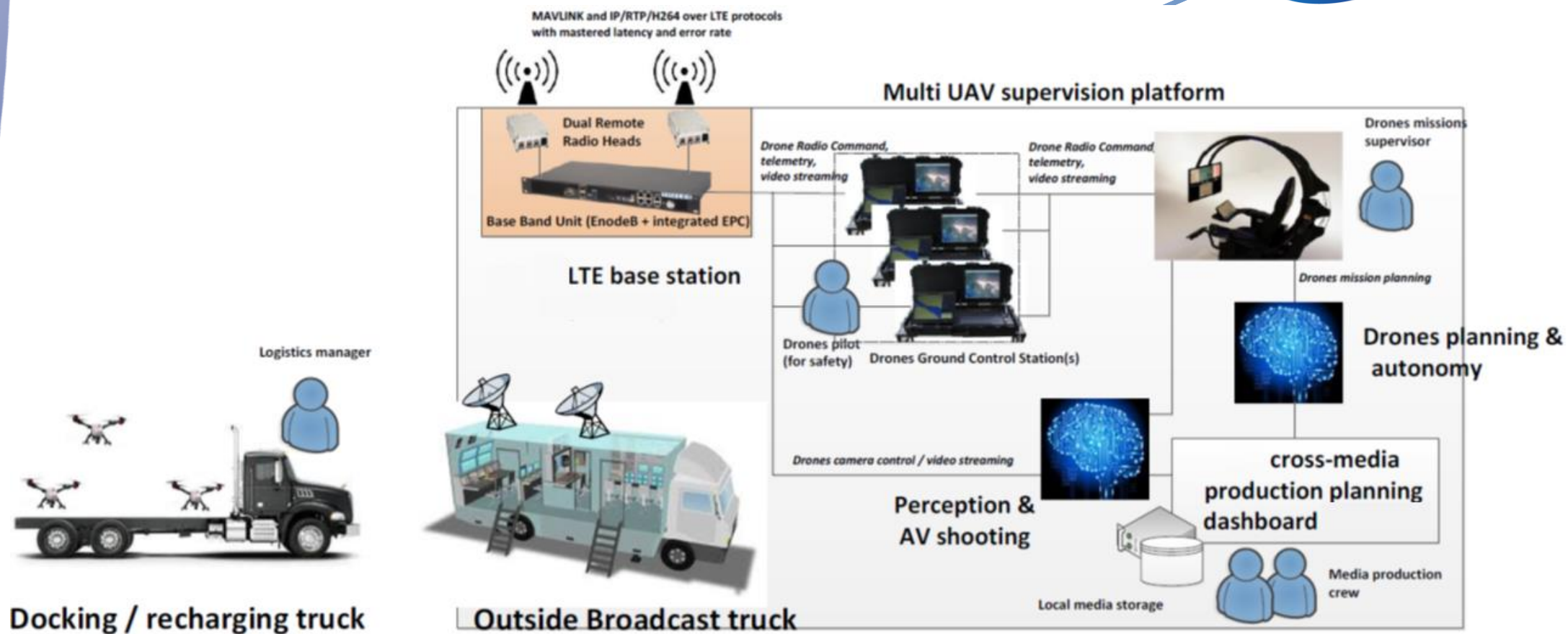
Integrated software and hardware target

MultiDrone



Multi UAV supervision platform

MultiDrone



Drone vision for cinematography: Functionalities (1)



- 1. (Multiple) drone mission planning.**
- 2. (Multiple) drone mission control.**
3. Active perception.
4. AV shooting.



Drone mission planning and control

- **Multiple drone mission planning.**
- Multiple drone mission control.
- Single drone flight control.
- Multiple drone control.
 - Drone formation control.
 - Collision avoidance.



Mission Planning Vocabulary



- MULTIDRONE **Shooting Mission**: list of **actions**
- Types of actions:
 - **Shooting Actions**: drone + camera
e.g., Lateral Tracking, Fly-Over, Orbit, ...
 - **Navigation Actions**: drone action only, does not involve shooting
e.g., Take-off, Land, Go-to-waypoint, ...
- Shooting Actions are *event-triggered*:
 - A start event is associated to each Shooting Action, which will trigger the action when it occurs.
E.g., target reaches a milestone, start of race, ...



Shooting Action Parameters



- **Shot type:**
 - Lateral shot, Orbital shot, etc.
- **Zoom type:**
 - Long shot, Medium shot, Close-up, etc.
- **Start position** for the drone and the camera look-at position
- **Triggering event**
- **Duration**
- **Target ID**

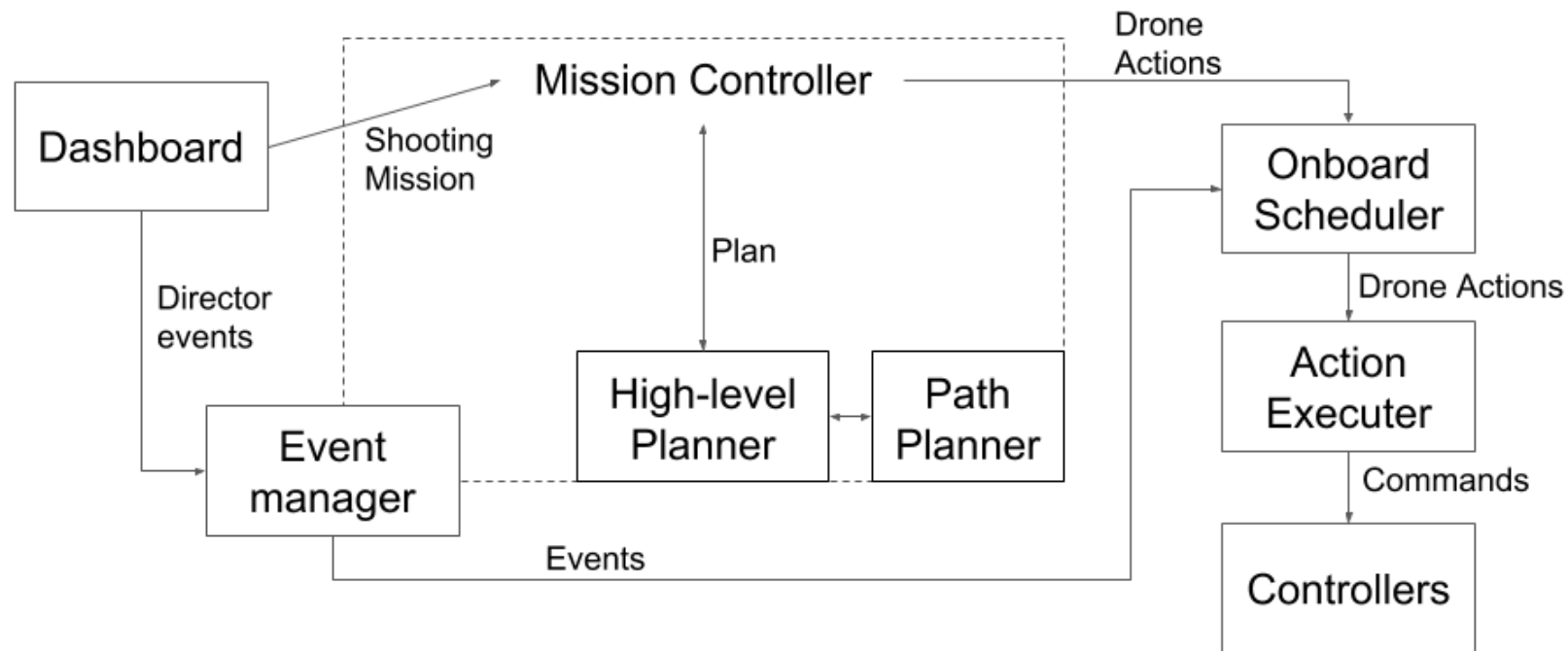


Mission Planning architecture

MultiDrone



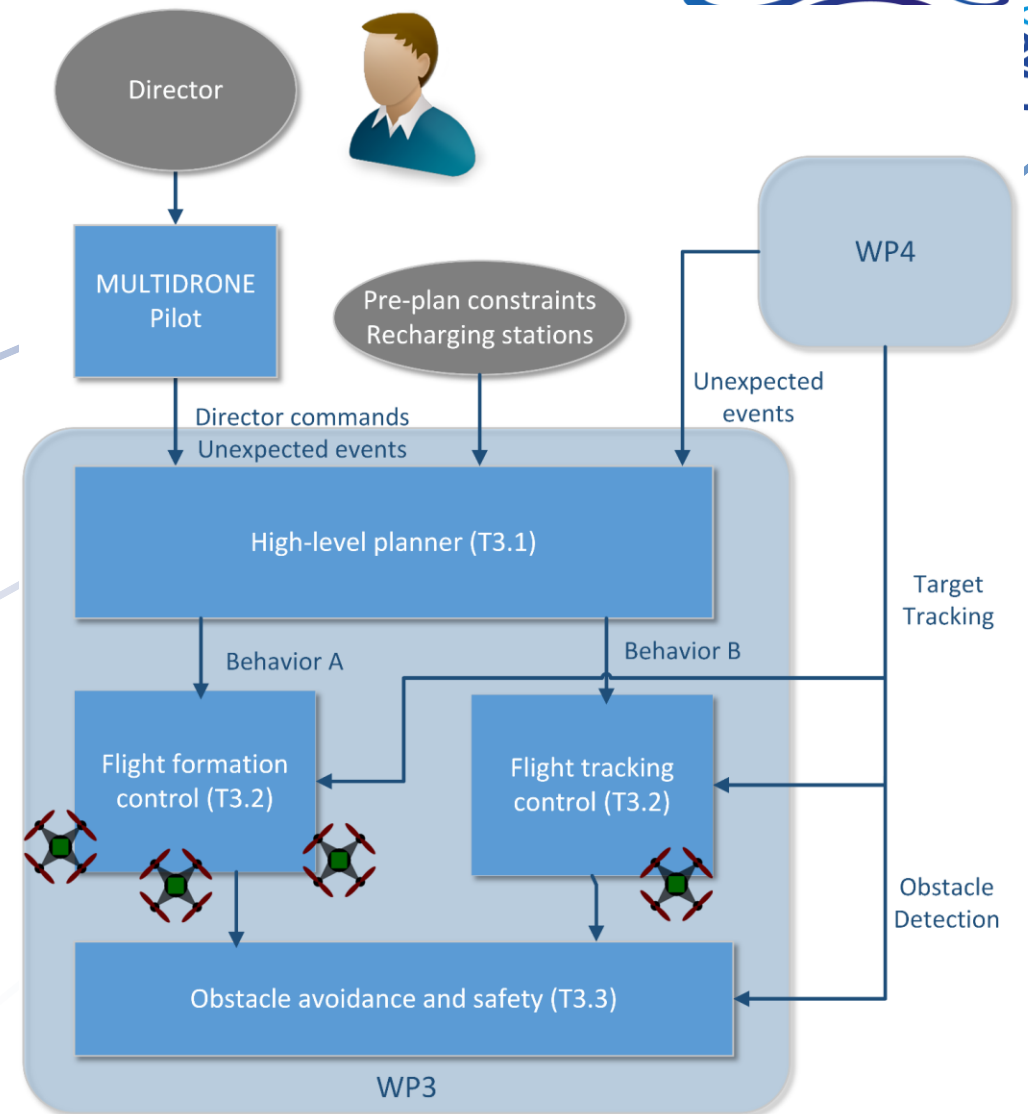
MULTIDRONE Planning



High-level pre-production/production mission planning

- **High-level planner** assigns different behaviours/tasks to the multidrone team according to director and environmental requirements.
- The multidrone planner needs to be **scalable** with multiple actors, since on-line re-planning could be needed as events happen or execution is performed.

MultiDrone



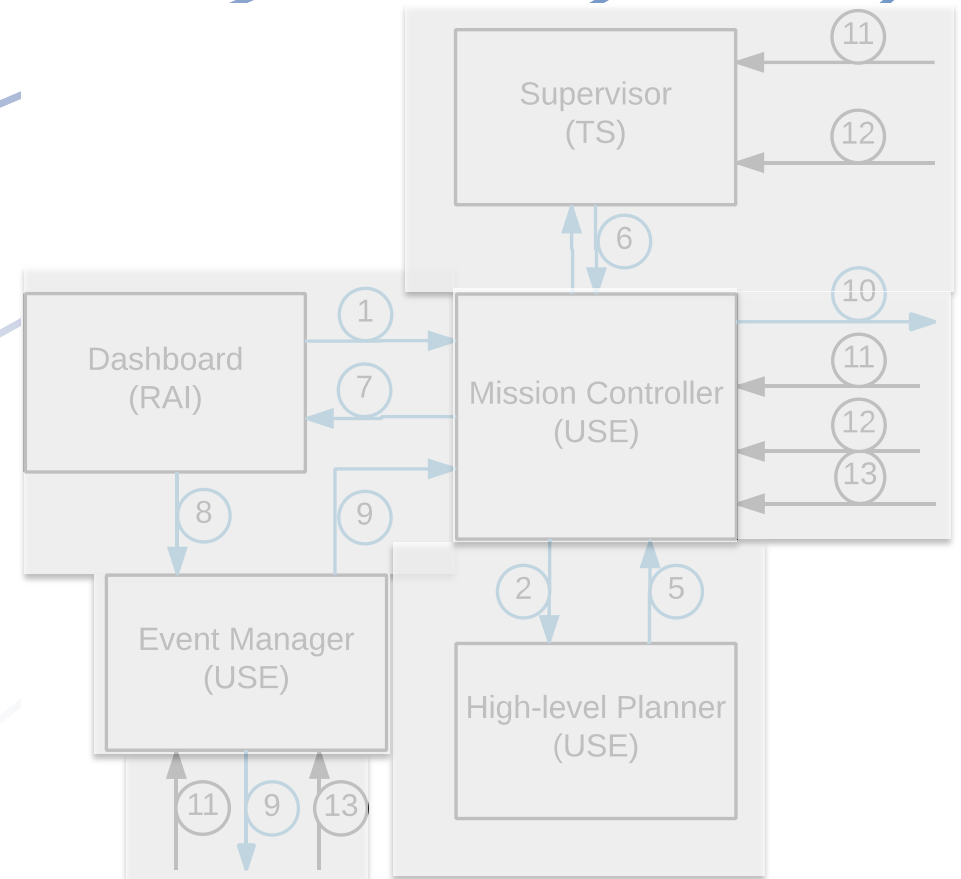
Mission Planning/Control

On ground modules

MultiDrone



- **Mission Controller:**
 - Interacts with **High-level Planner** to produce a mission plan.
 - Monitors mission execution.
 - Asks for replanning if needed.
- **Event Manager:**
 - Receives, manages, and generates events.
 - Sends events to drones to start and stop action execution.





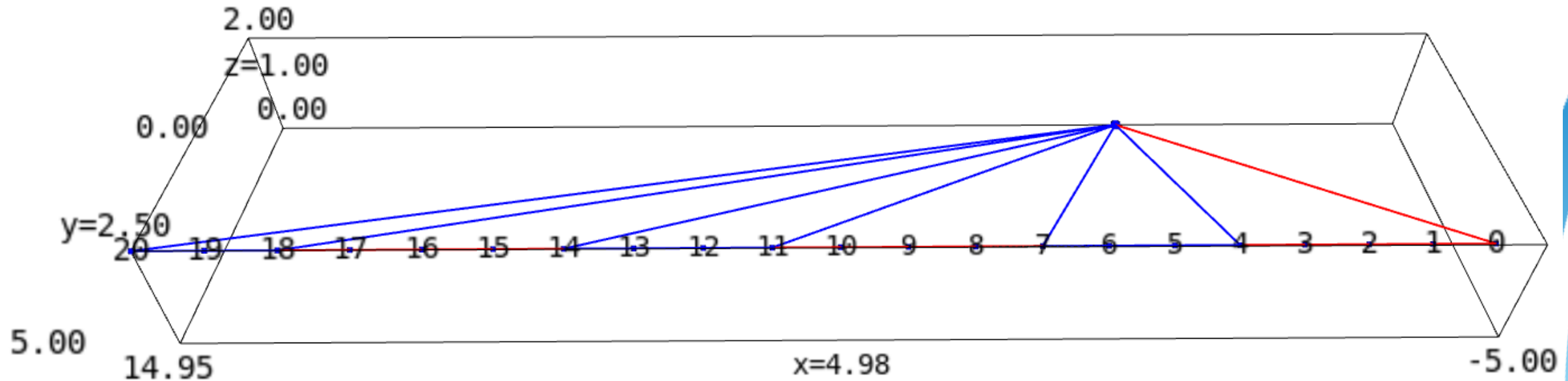
High-level planner

- Shooting Mission translated to list of Shooting Actions with triggering events.
- Tasks correspond to SAs, multi-drone SAs split into several single-drone tasks.
- Each task has a start location, start time and duration.
- Computes the plan: allocates tasks to drones fulfilling time and precedence constraints (Multi-Robot Task Allocation problem).
- MRTA problem definition:
 - N drones with known positions.
 - M single-drone tasks, each one with its time window.
 - Objective: maximize time where drones are covering (filming) tasks.

Mission Planning Example 1



- This example shows how two drones cover a long task in turns, due to the battery constraint.



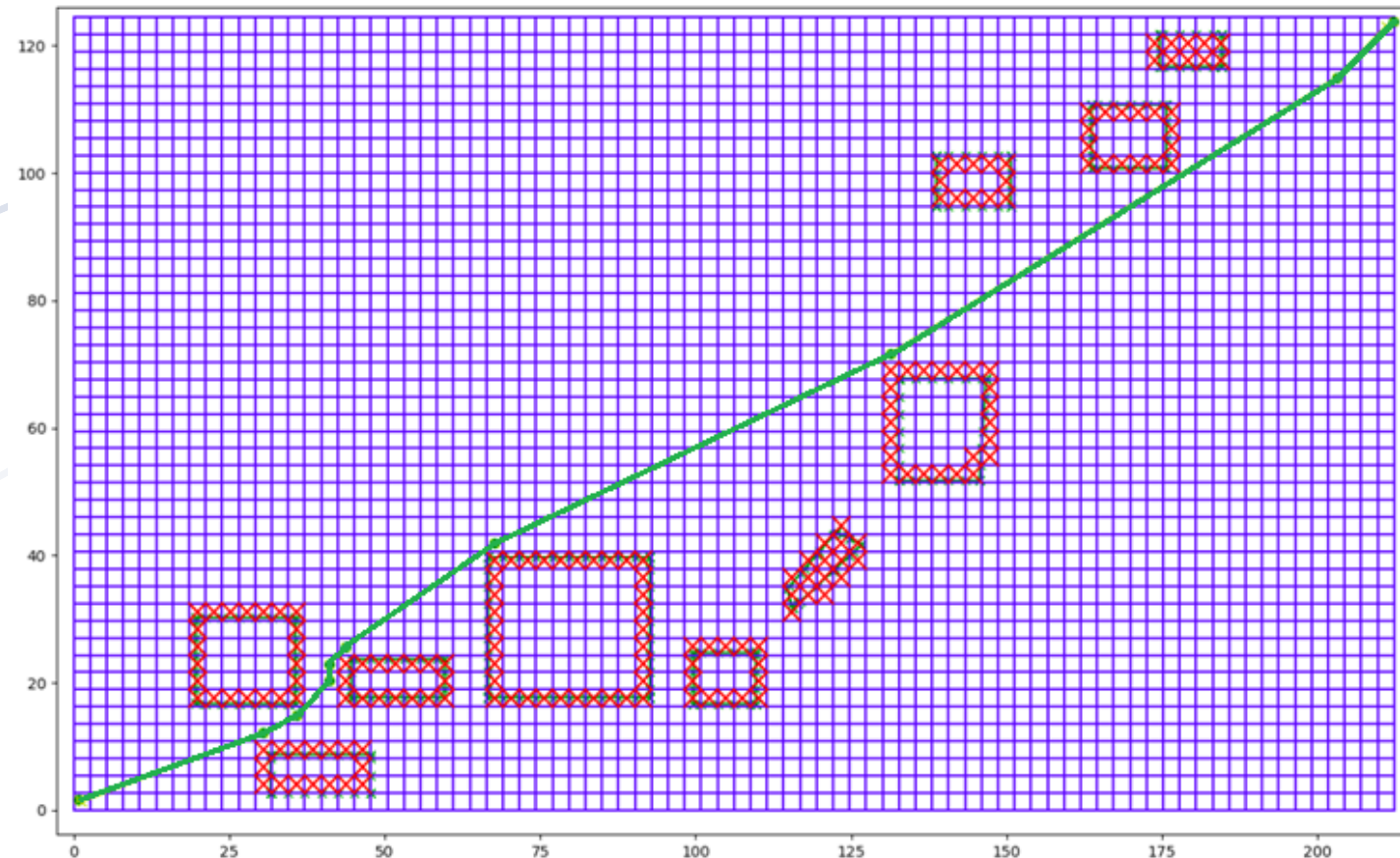
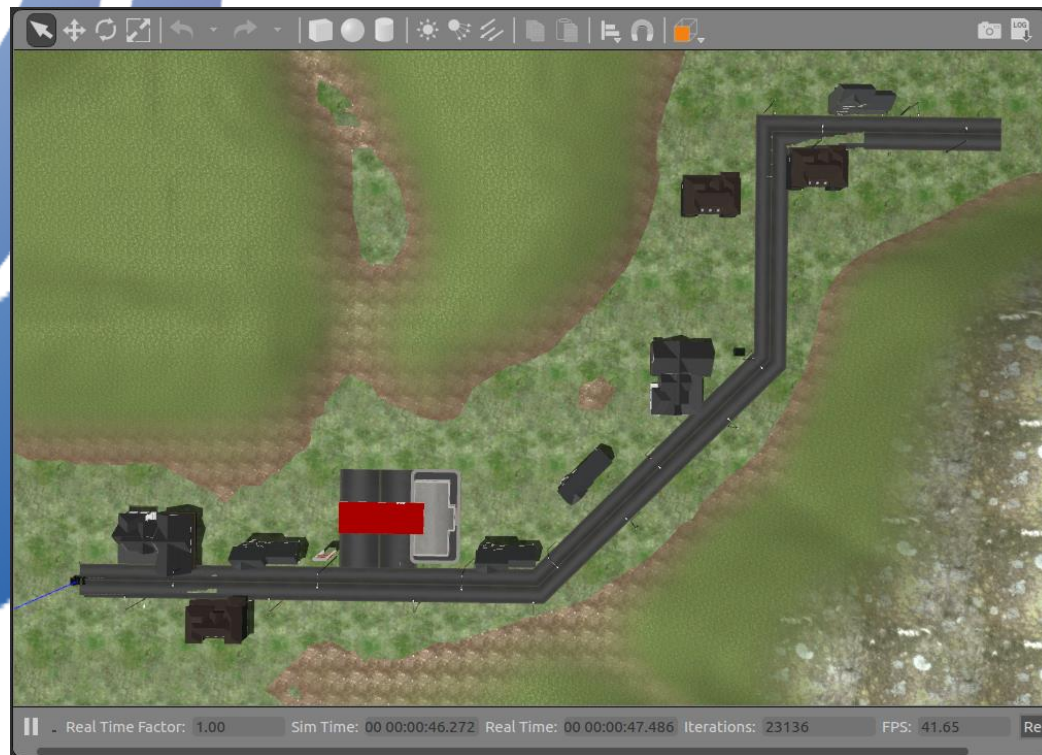


- This submodule is used by:
 - High-level Planner to estimate drone paths and flying times.
 - Onboard Scheduler to compute a path to a landing position in case of emergency.
- Navigation map implemented as a grid. Obtained from Semantic Map.
 - Semantic annotations are indicated as KML features.
 - Geodesic coordinates translated into Cartesian.
 - No-fly polygons become occupied cells in grid.
- Safe path computed using A* search algorithm. Fast for simple solution spaces.

Path Planner Example



- Path from one corner to the other. Buildings labeled as no-fly zones (obstacles represented as red crosses in the grid).
- Solved in 66 ms.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)



Drone mission planning and control

- Multiple drone mission planning.
- **Multiple drone mission control.**
- **Single drone flight control.**
- Multiple drone control.
 - Drone formation control.
 - Collision avoidance.



Mission Execution/control

On-drone modules

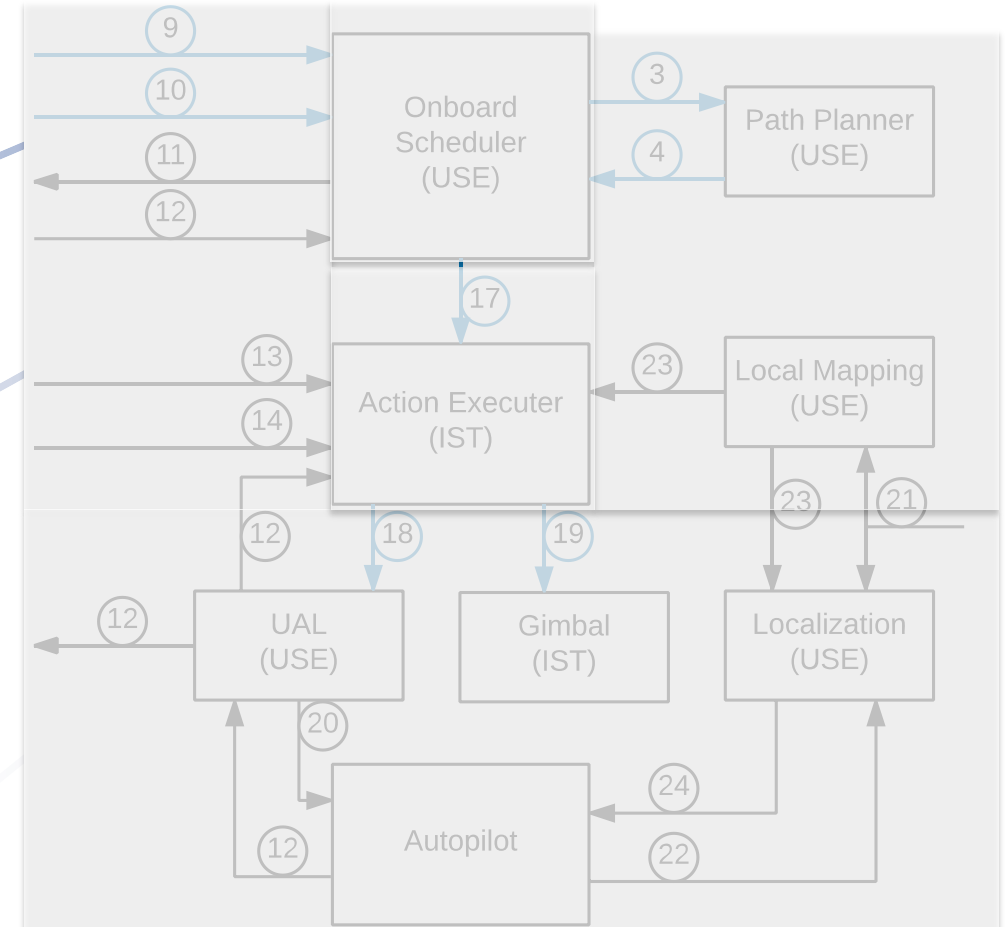


- **Onboard Scheduler:**

- Receives list of actions
- Receives events to trigger action execution
- Activates the Action Executer
- Sends drone status to ground

- **Action Executer:**

- Translates Shooting Actions into desired drone+camera configurations
- Interacts with other modules to produce commands for autopilot, camera and gimbal





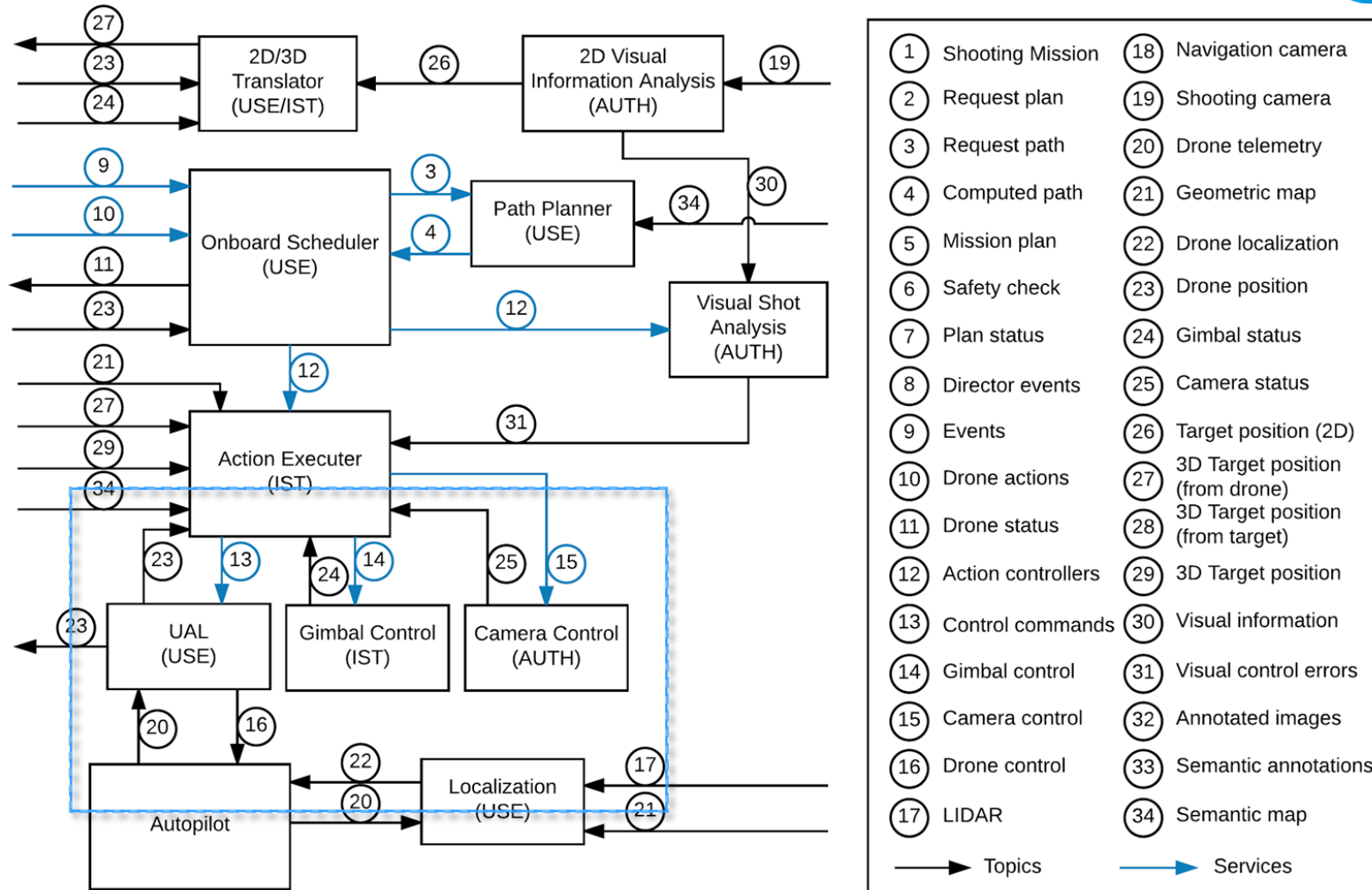
Action Execution

- Onboard Scheduler activates execution of individual drone actions:
- *Navigation Actions*: Take-off, Land, Go to Waypoint, etc.
- *Shooting Actions (SA)*: Lateral Tracking, Chase, Still, Orbit, etc.
- Shooting actions involve drone control + gimbal control for target tracking.

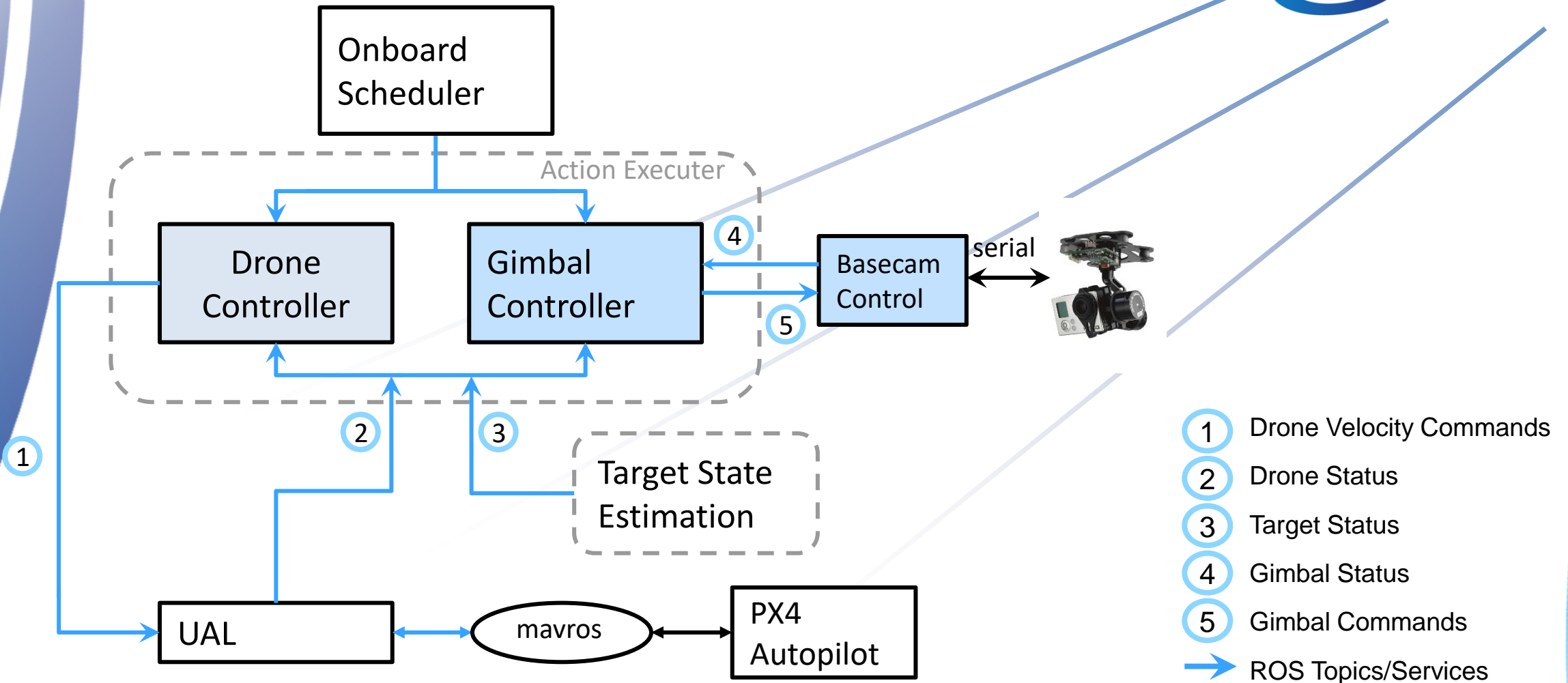


On-drone functional architecture

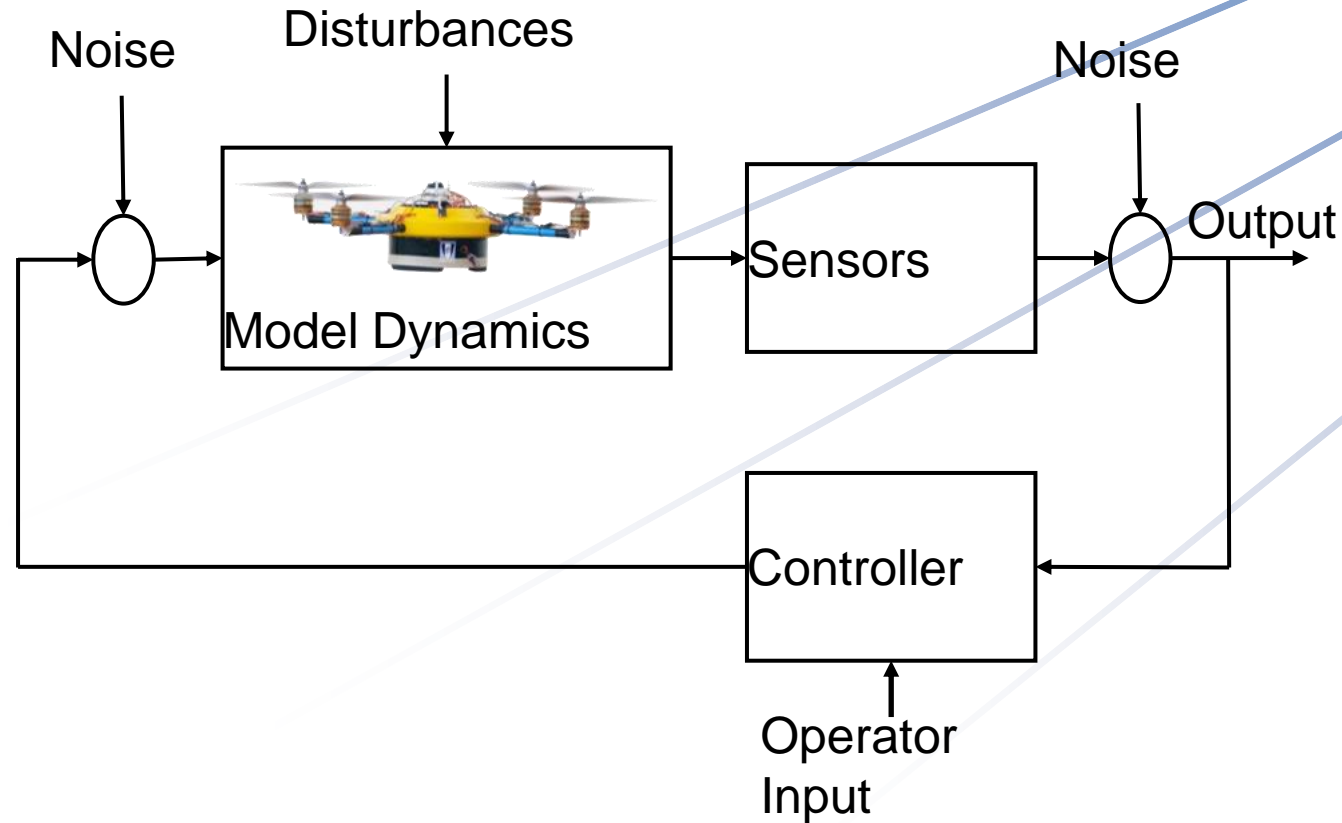
MultiDrone



Action Execution

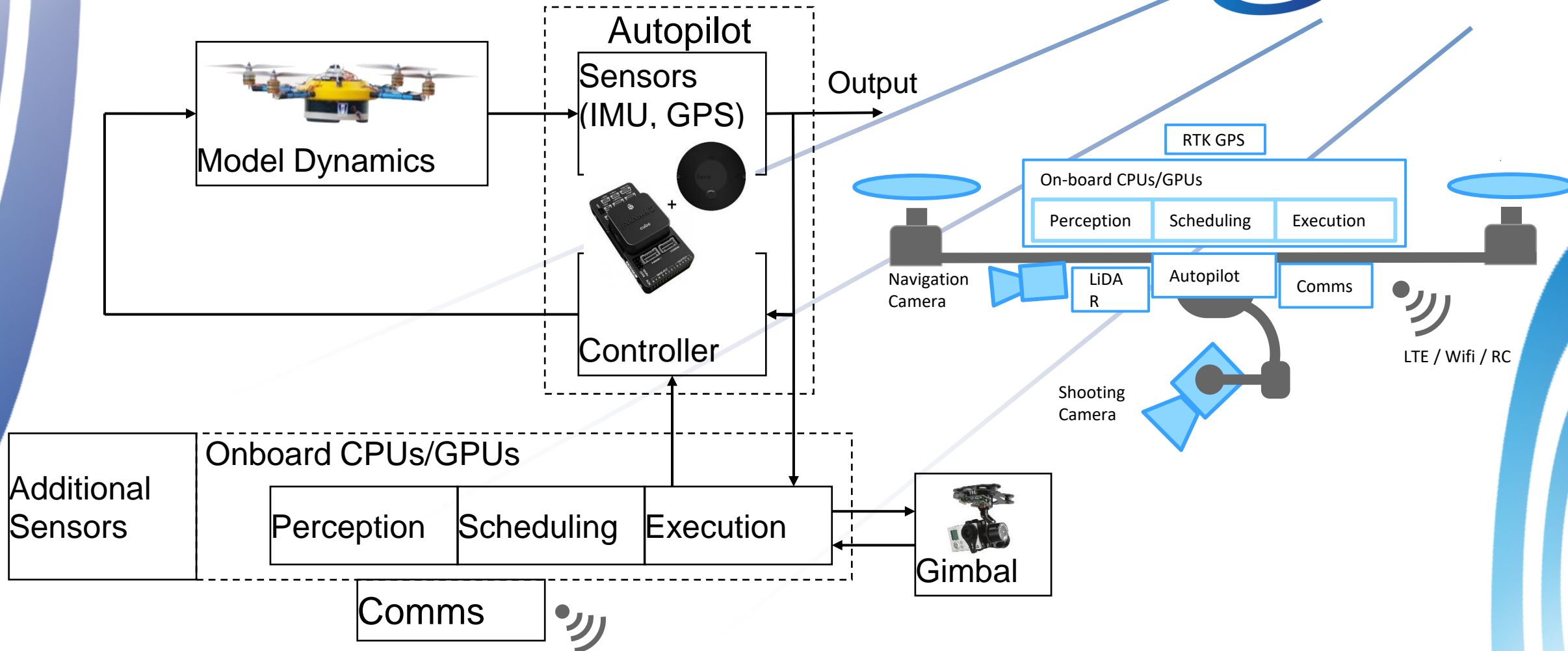


Automatic drone operation: Drone as Control System



Multidrone Onboard Architecture

MultiDrone



Gimbal control

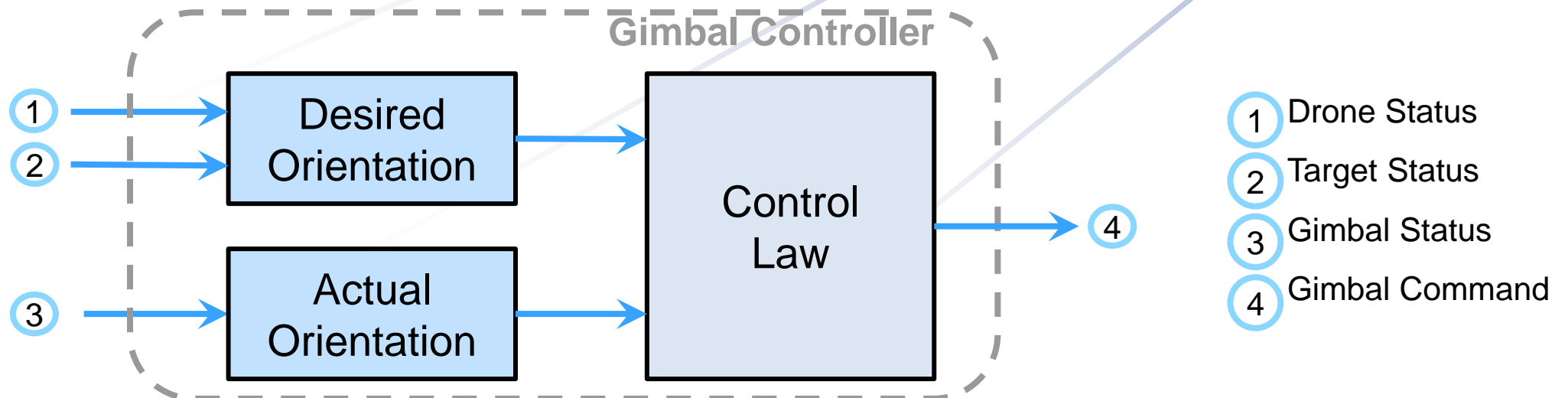


Control objective:

Point towards the target.

Approach:

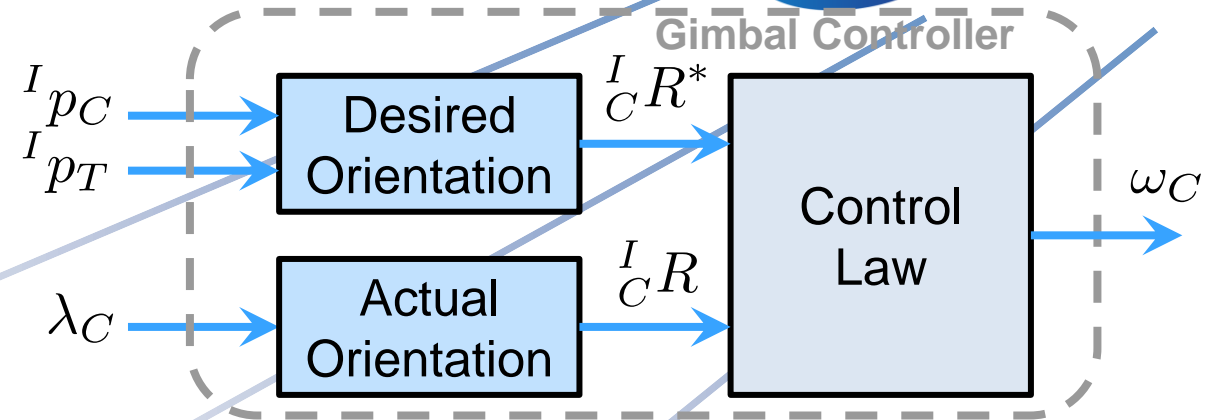
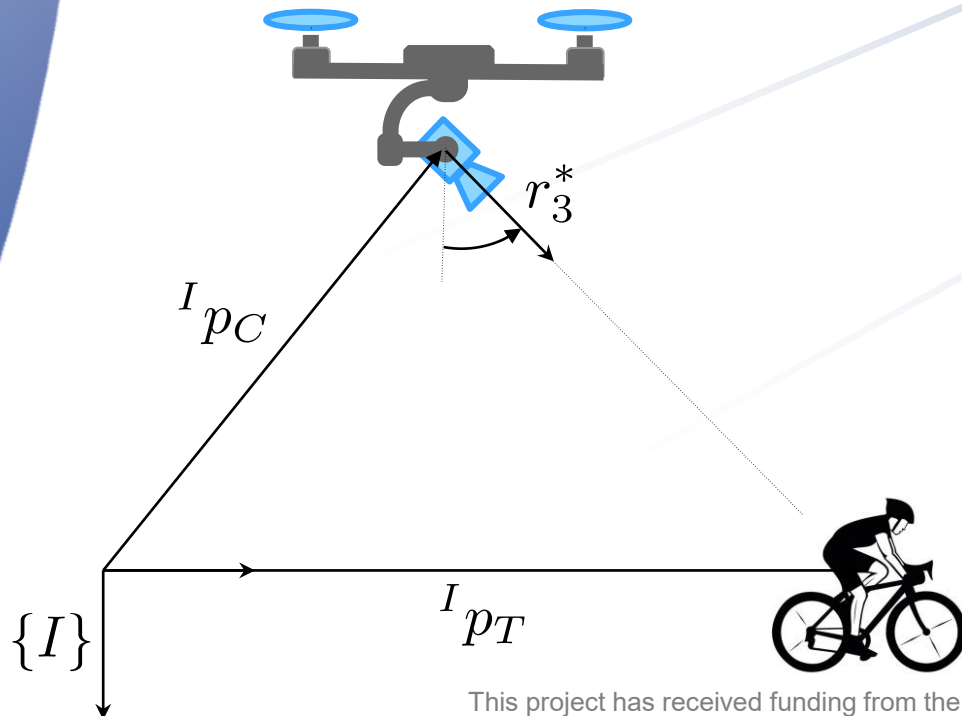
Treat gimbal control independently from drone control.



Gimbal control

Desired orientation

$${}^I_C R^* = \begin{bmatrix} r_1^* & r_2^* & r_3^* \end{bmatrix} \in \mathbb{SO}(3)$$



Desired optical axis
in inertial coordinate
s

$$r_3^* = \frac{{}^I p_T - {}^I p_C}{\| {}^I p_T - {}^I p_C \|}$$

Camera provides

$${}_I^C R \frac{{}^I p_T - {}^I p_C}{\| {}^I p_T - {}^I p_C \|}$$



Gimbal control. Experimental tests.



BMMCC Camera control



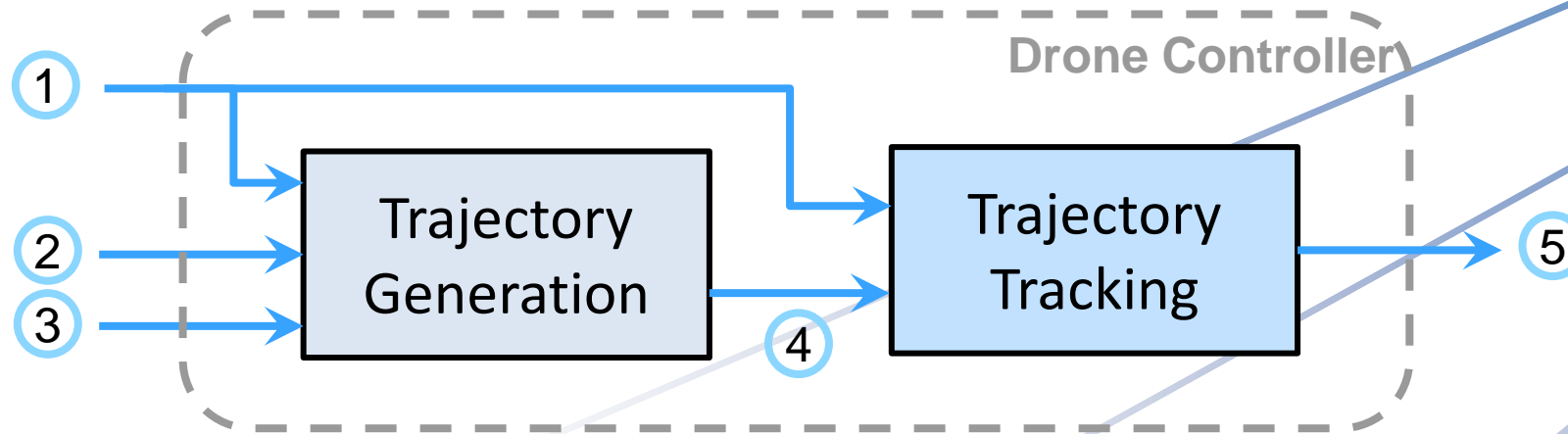
- No feedback from the camera
- Two ways of mapping the commands to the controls:

Select a particular setting	Increment / decrement
Iris	Zoom
Focus	Start/Stop Recording
Audio	Auto Focus
Frame Rate	ISO
Codec	Shutter Angle
	White Balance



Drone Controller

MultiDrone



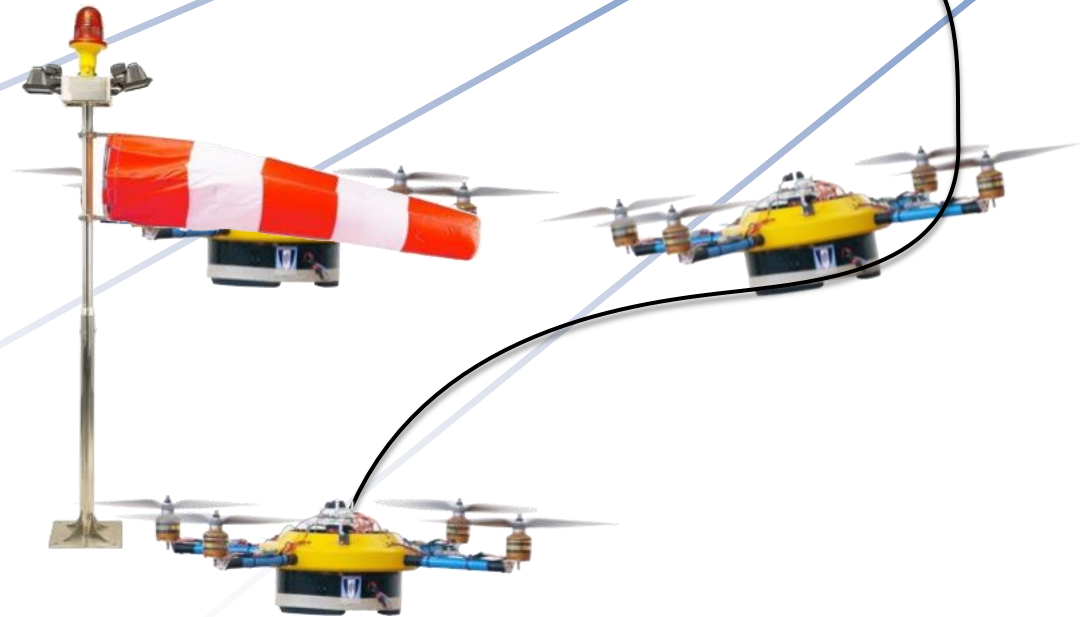
- 1 Drone Status
- 2 Target Status
- 3 Shooting Action parameters
- 4 Reference
- 5 Drone Velocity Command



Control Objectives – Trajectory Tracking



Track a trajectory.
Realistic model.
Robustness to disturbances.
Bounded actuation.
Large basin of attraction.



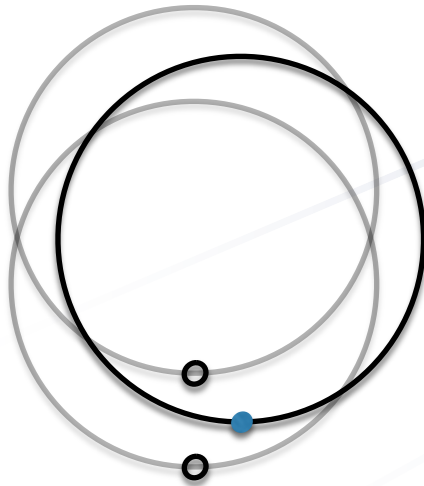
Drone mission planning and control

- Multiple drone mission planning.
- Multiple drone mission control.
- Single drone flight control.
- **Multiple drone control.**
Drone formation control.
Collision avoidance.

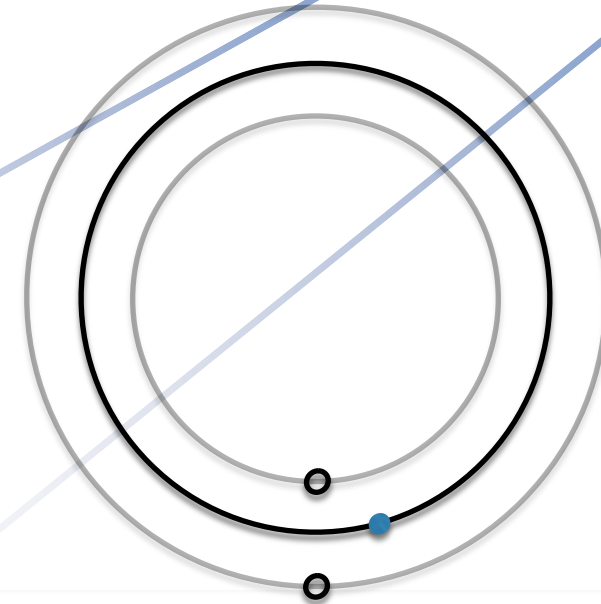


Leader-following for formation control

Main idea:
Trailer-like behavior for the followers.



In inertial frame:
Translated identical paths



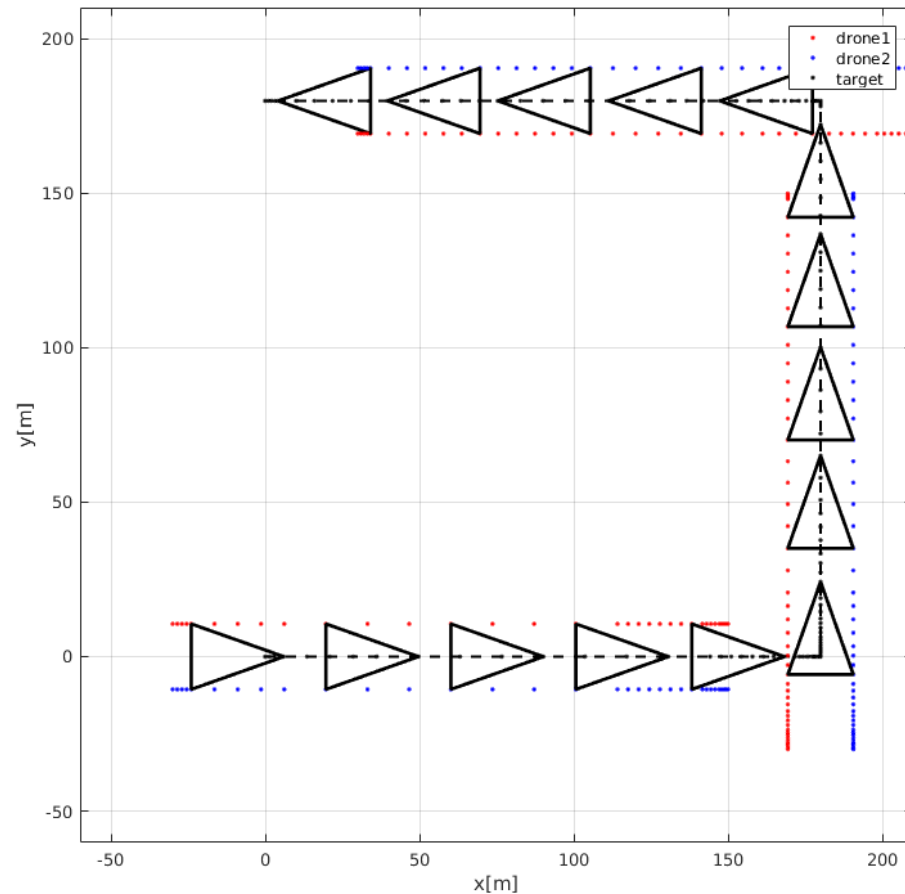
In trailer frame:
Different paths, no superposition

Trailer approach properties

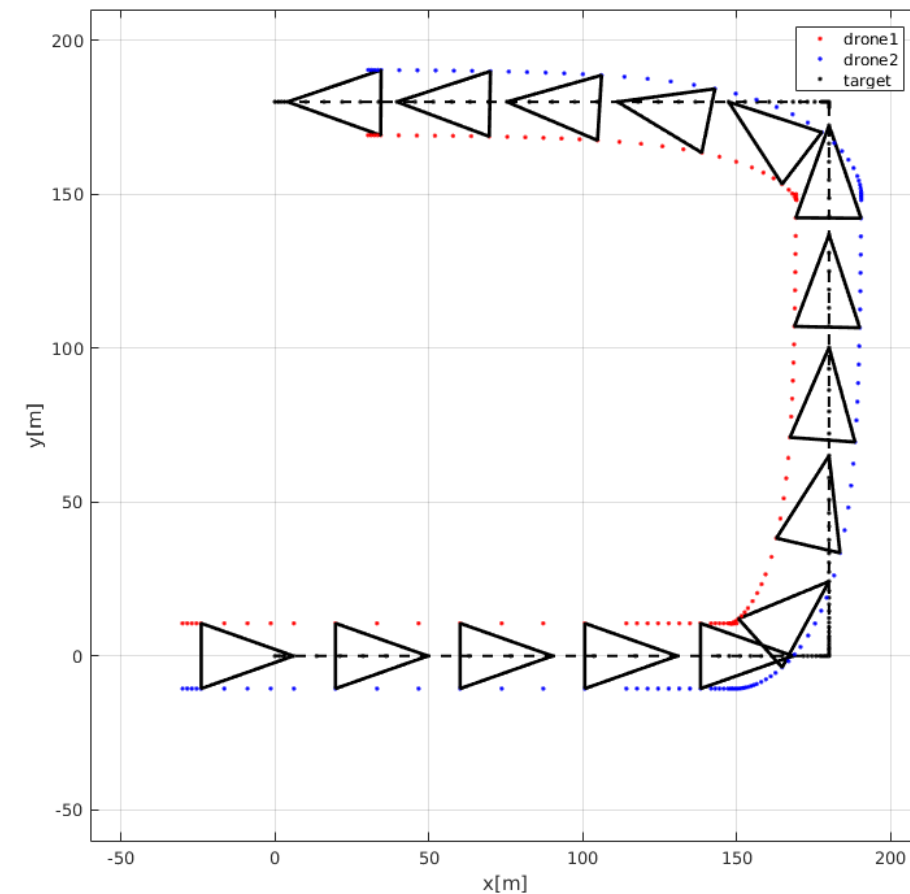
MultiDrone



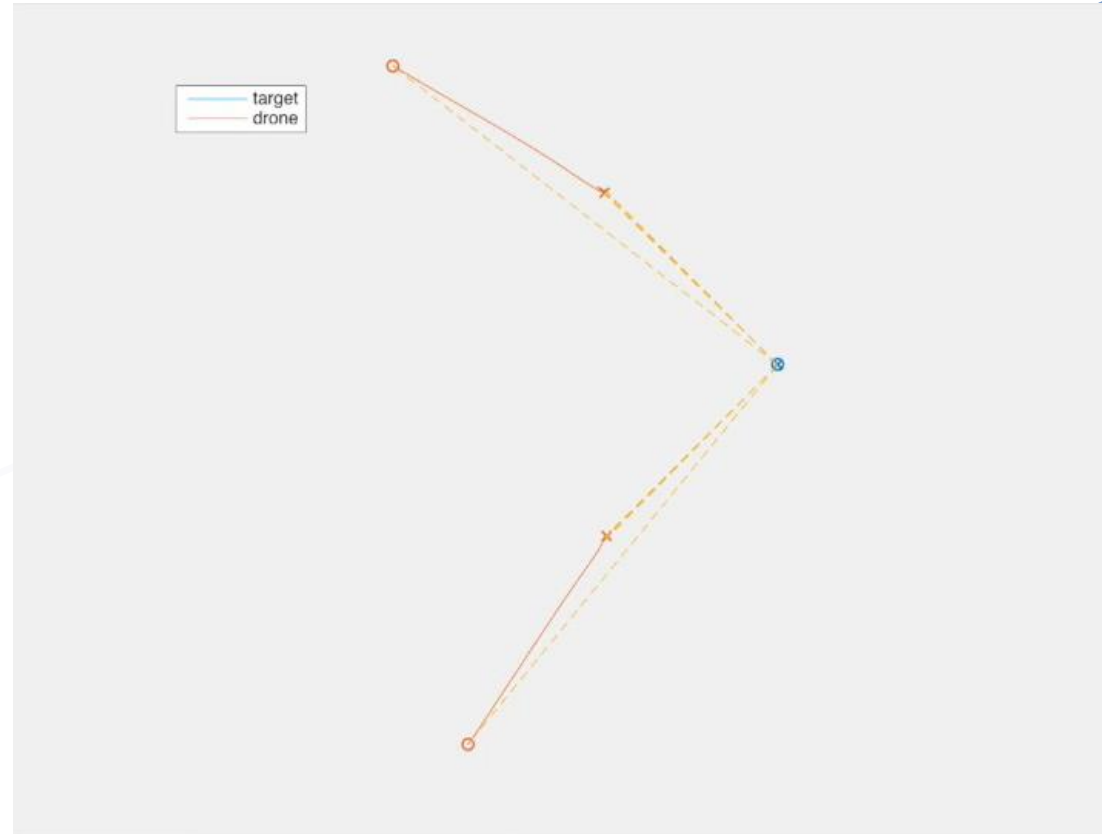
Without Trailer



With Trailer

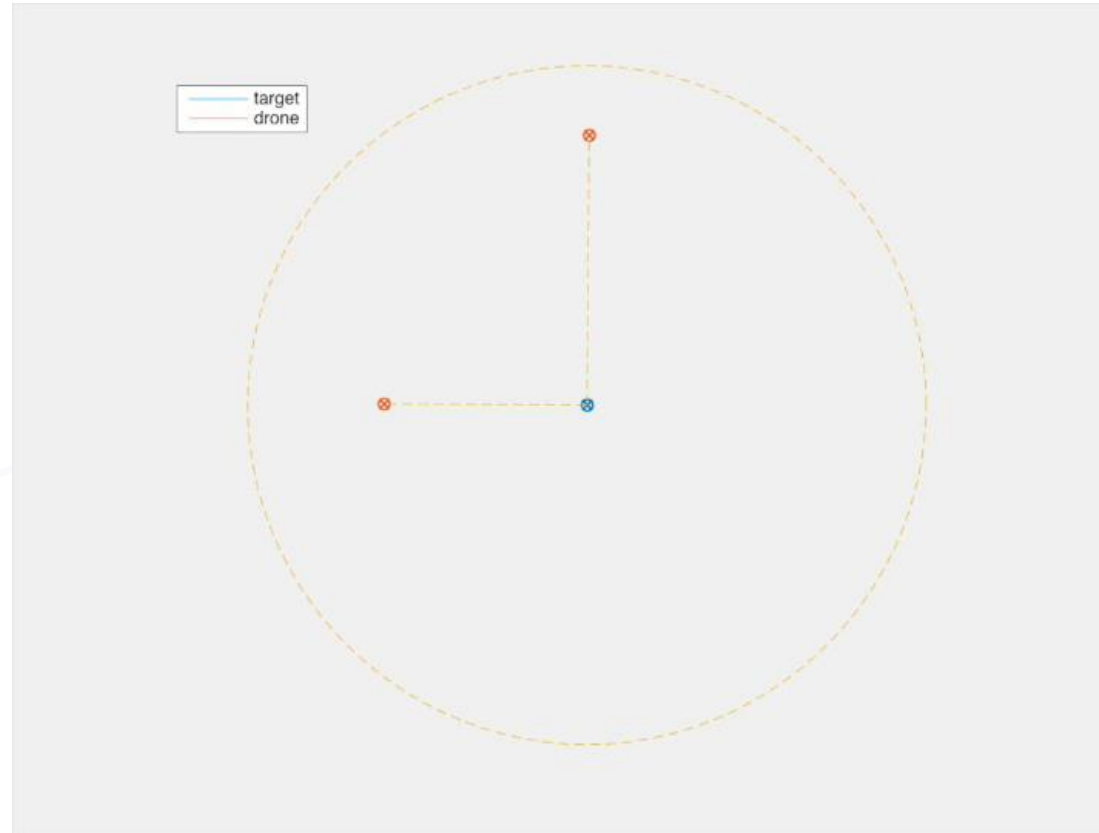


SA1 - Constant relative positions



SA2 - Orbit trajectory

MultiDrone



Drone mission planning and control

- Multiple drone mission planning.
- Multiple drone mission control.
- Single drone flight control.
- Multiple drone control.
 - Drone formation control.
 - Collision avoidance.**



Multi-drone Conflict Resolution Problem Definition



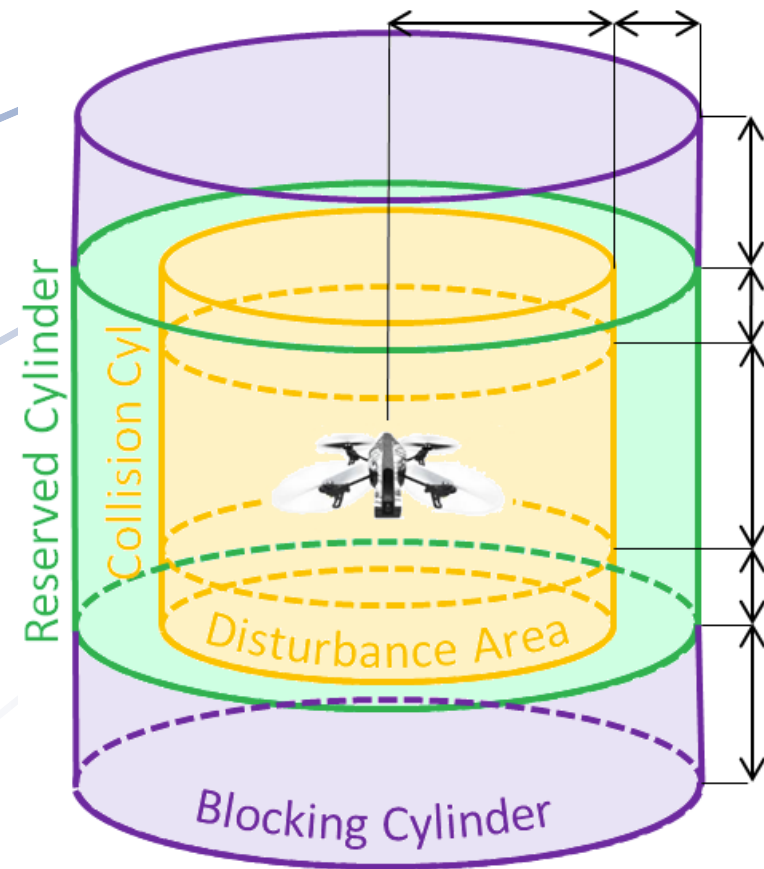
- Navigate a team of drones in a shared 3D space without collision.
- Starting configuration to a goal configuration.
- Drones must detect and resolve conflicts in a decentralized manner.



Decentralized 3D collision avoidance

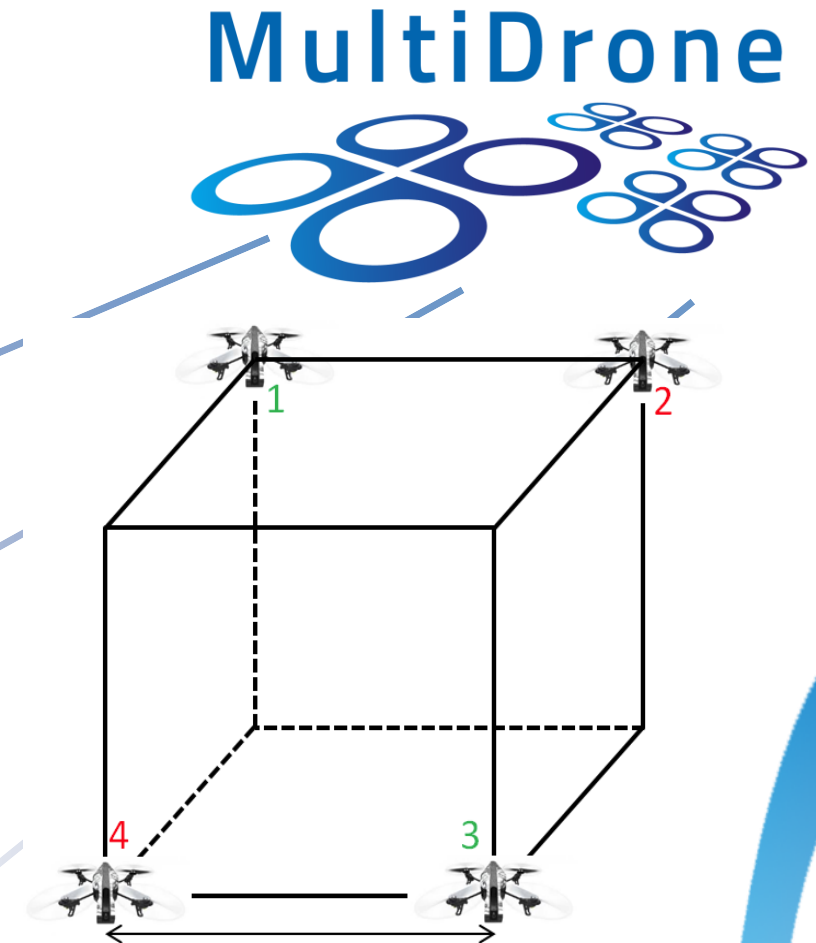
- Collision hull defined as a cylinder (yellow).
- Horizontal conflict when reserved cylinder (green) overlaps with others.
- Vertical conflict when blocking cylinder overlaps with others.
- Cylinders allow drones to brake on time and maneuver to avoid collision.

MultiDrone



SITL Simulations

- Simulations with a SITL scheme. Noisy GPS measurements.
- Drones in cube exchanging positions.
- Drones in conflict surround each other creating a virtual roundabout.
- Clearance level: minimum distance of each drone to its nearest neighbor.



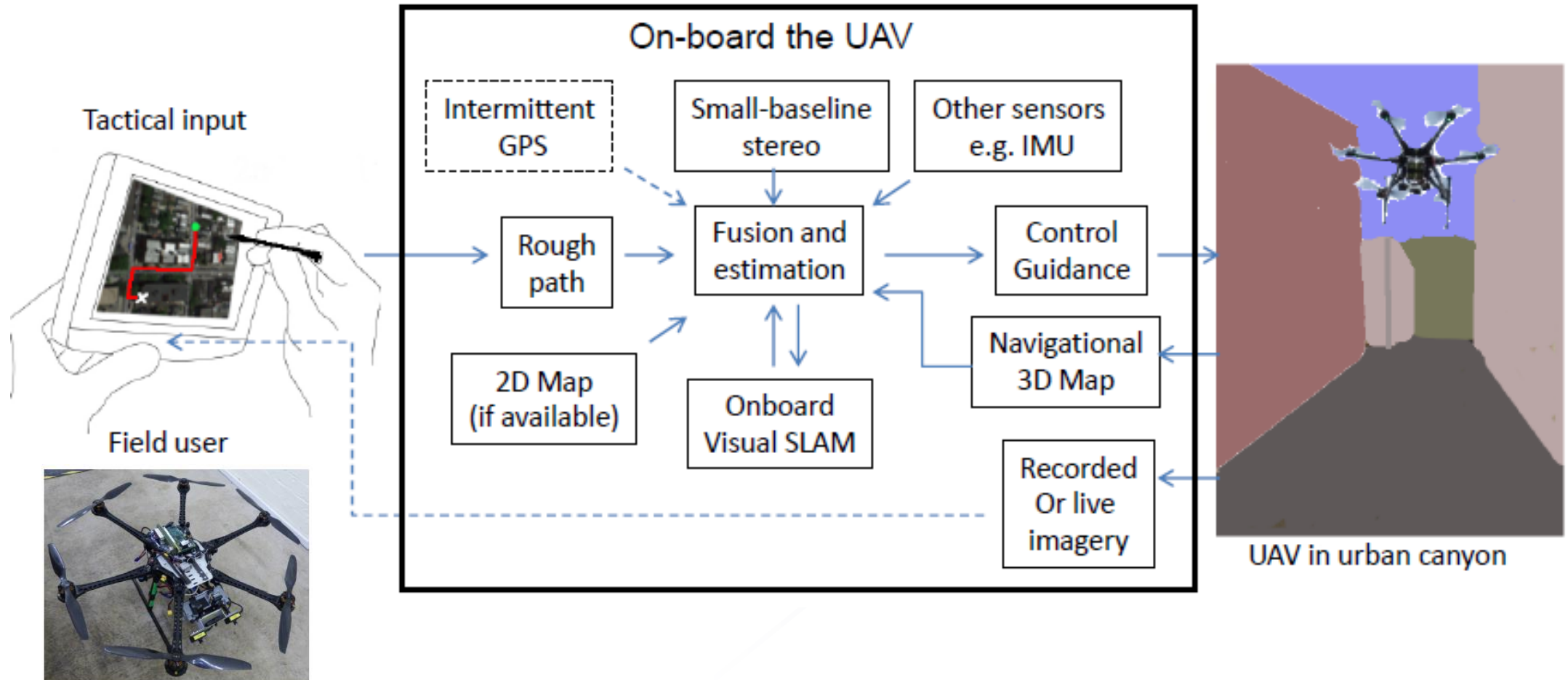
Drone vision for cinematography: Functionalities (2)



- 1. Perception and localization. SLAM:**
 1. Semantic 3D world mapping
 2. Drone localization.
- 2. Visual and perception data analysis for AV Shooting:**
 1. 2D target (athlete, boat, cycle) detection and tracking
 2. 3D target localization and following
 3. Drone cinematography
 4. Target pose estimation.



UAV Simultaneous Localization and Mapping



3D Drone Localization and Mapping



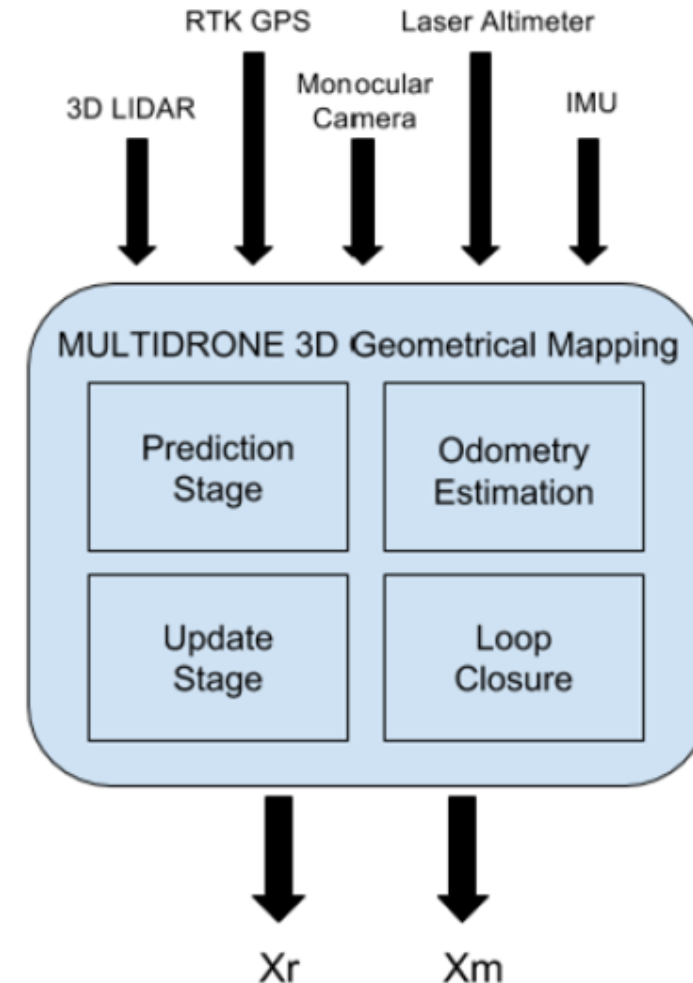
- Drone localization:
 - **Sources, sensors.**
 - Mapping.
 - Localization.
 - SLAM.
- Data fusion in drone localization.
- Semantic mapping.



Geometrical mapping



- Sensors:
 - Velodyne HDL-32E
 - Monocular camera
 - IMU
 - laser altimeter
 - RTK D-GPS
- Processing:
 - Intel NUC NUC6i7KYK2 i7-6770HQ
 - Jetson TX2



Geometrical mapping



- 3D LIDAR.
 - SLAM-like algorithm based on Prediction-Update Recursions
 - Extract from the LIDAR measurements: corner and surface points
 - Prediction: Estimate LIDAR-based odometry from different scans using the ICP algorithm
 - Update: Matching of the LIDAR scan with the estimated map
 - Good estimate of robot 6 DoF pose and geometrical map
- Visual camera
 - Extraction of features using detectors such as SURF, SIFT or ORB
 - Estimation of visual odometry
- Robot odometry:
 - Combination of:
 - LIDAR-based odometry
 - Visual odometry
 - IMU

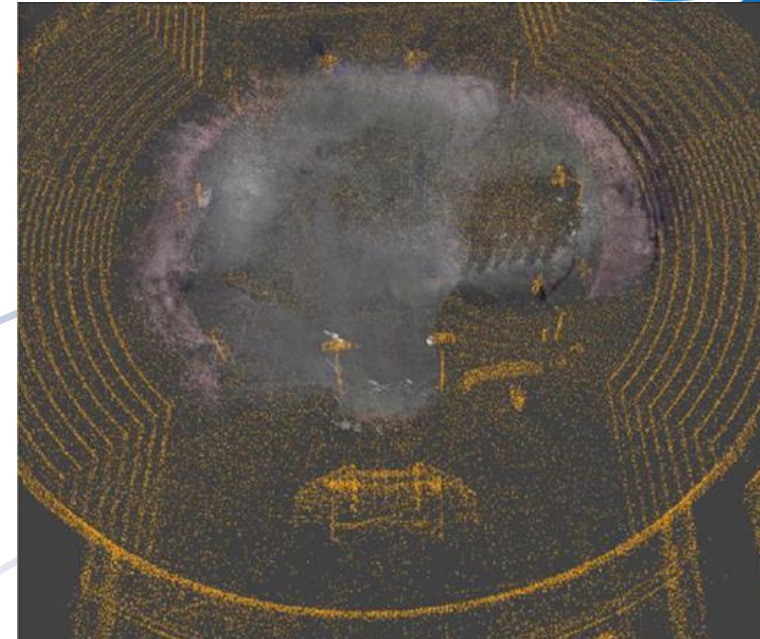


Geometrical mapping

MultiDrone



Experiments

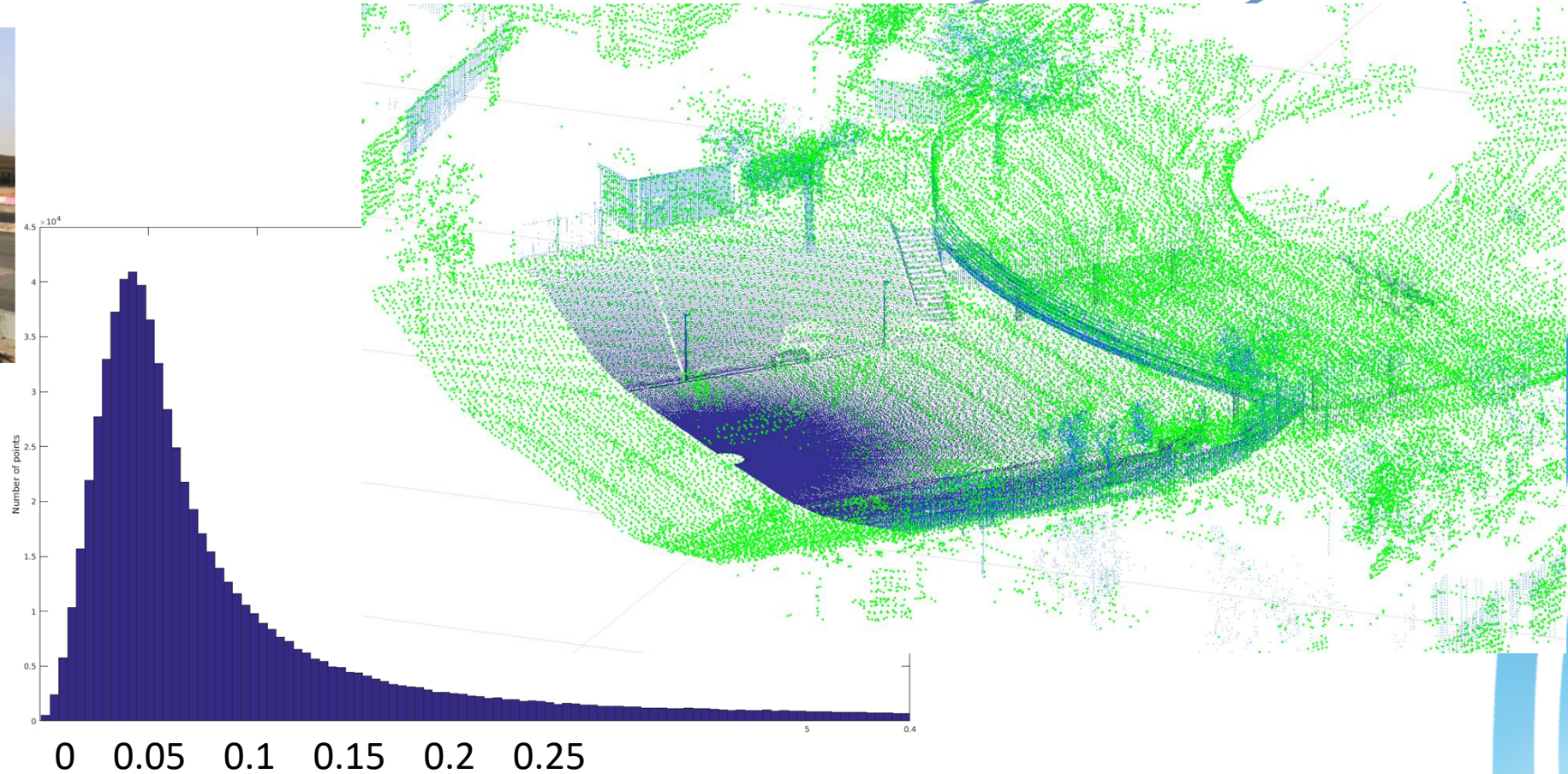


Repeatability

Dataset	Mean Error (m)	Median Error (m)	Min Error (m)
1	0,1377	0,1073	0,00098
2	0,1053	0,0769	0,00045
3	0,0847	0,0578	0,00083
4	0,1074	0,0792	0,00078
5	0,1722	0,1560	0,00130

Geometrical mapping

Validation with a TOTAL STATION



3D Drone Localization and Mapping

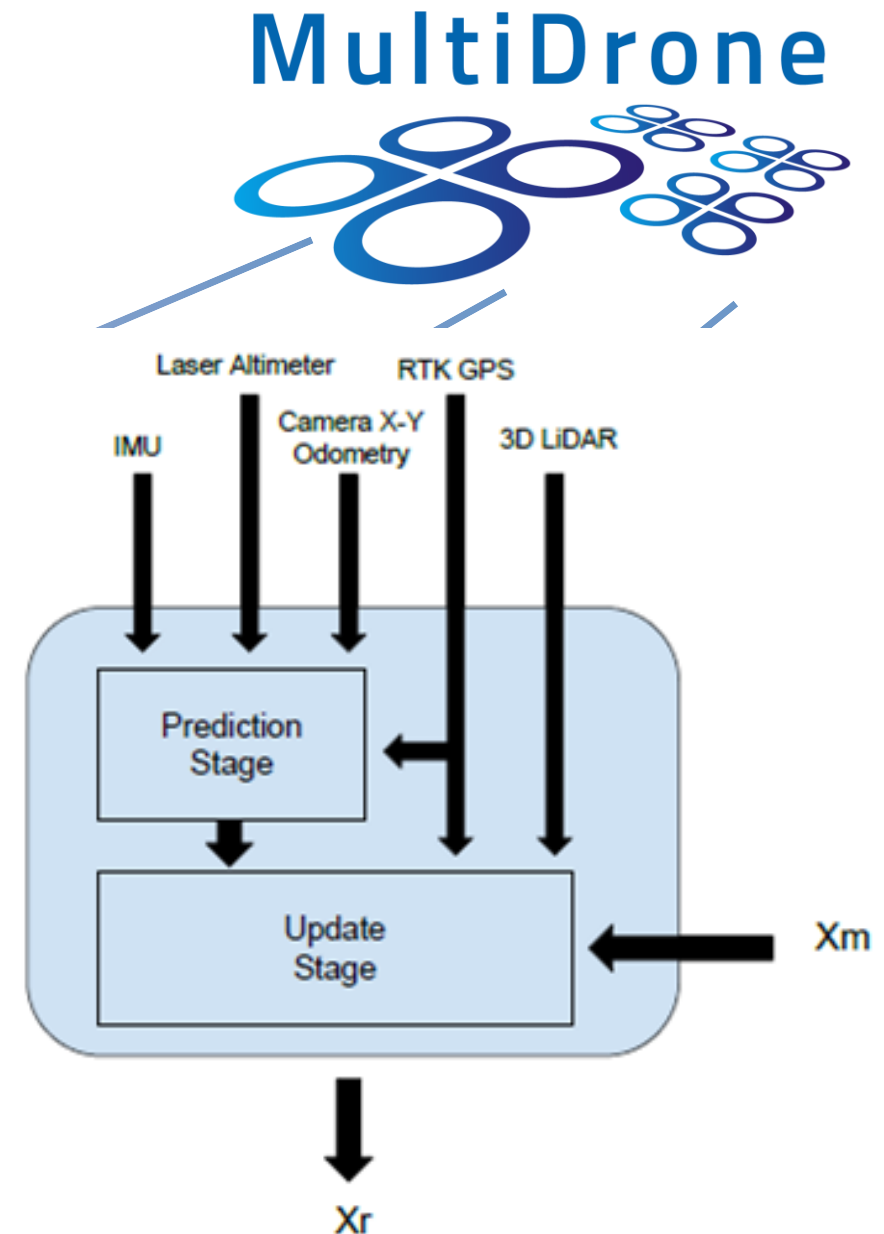


- Drone localization
 - Sources, sensors
 - Mapping
 - **Localization**
 - SLAM
- Data fusion in drone localization
- Semantic mapping



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



Object detection



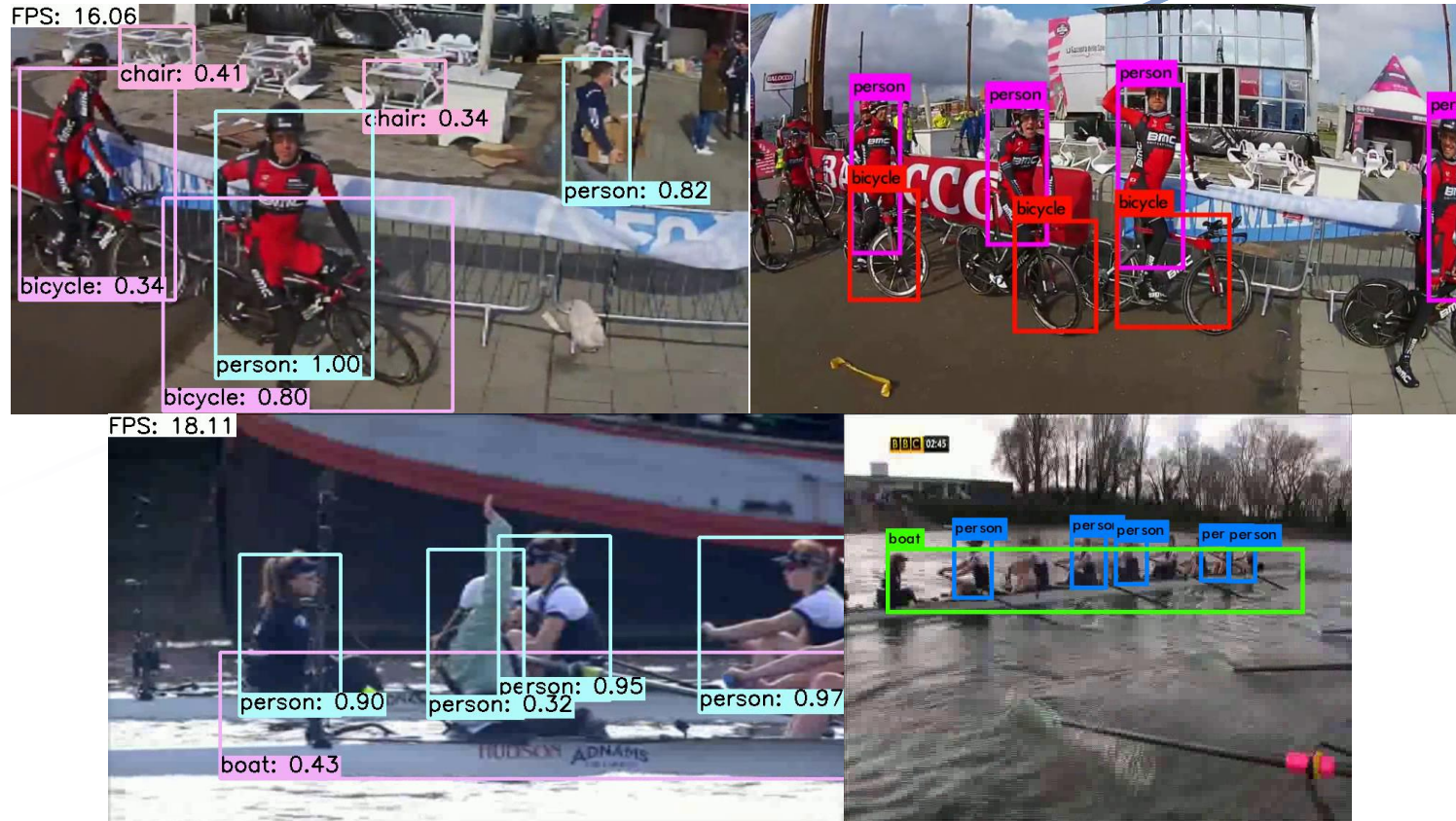
- **State-of-the-art** object detectors are based on **very Deep** and **multiple-channel CNNs**.
- **Multiple** layers of **many** convolutional filters are applied to the input image, forming a *very deep* architecture of successive convolutions and optionally some fully connected components.
- Trained on large-scale datasets, such as
 - VOC2007 with 10k images depicting **~24k objects** belonging to **20 classes**
 - VOC2012 with ~11k images depicting **~27k objects** belonging to the same **20 classes** as VOC2007
 - COCO with 328k images, about **2.5 million objects** belonging to **91 classes**.



Object detection



- Current neural detectors are very capable of accurately detecting objects



SSD

YOLO



Object detection acceleration



- Examples of acceleration techniques:
 - Input size reduction.
 - Specific object detection instead of multi-object detection.
 - Parameter reduction.
 - Post-training optimizations with TensorRT, including FP16 computations.



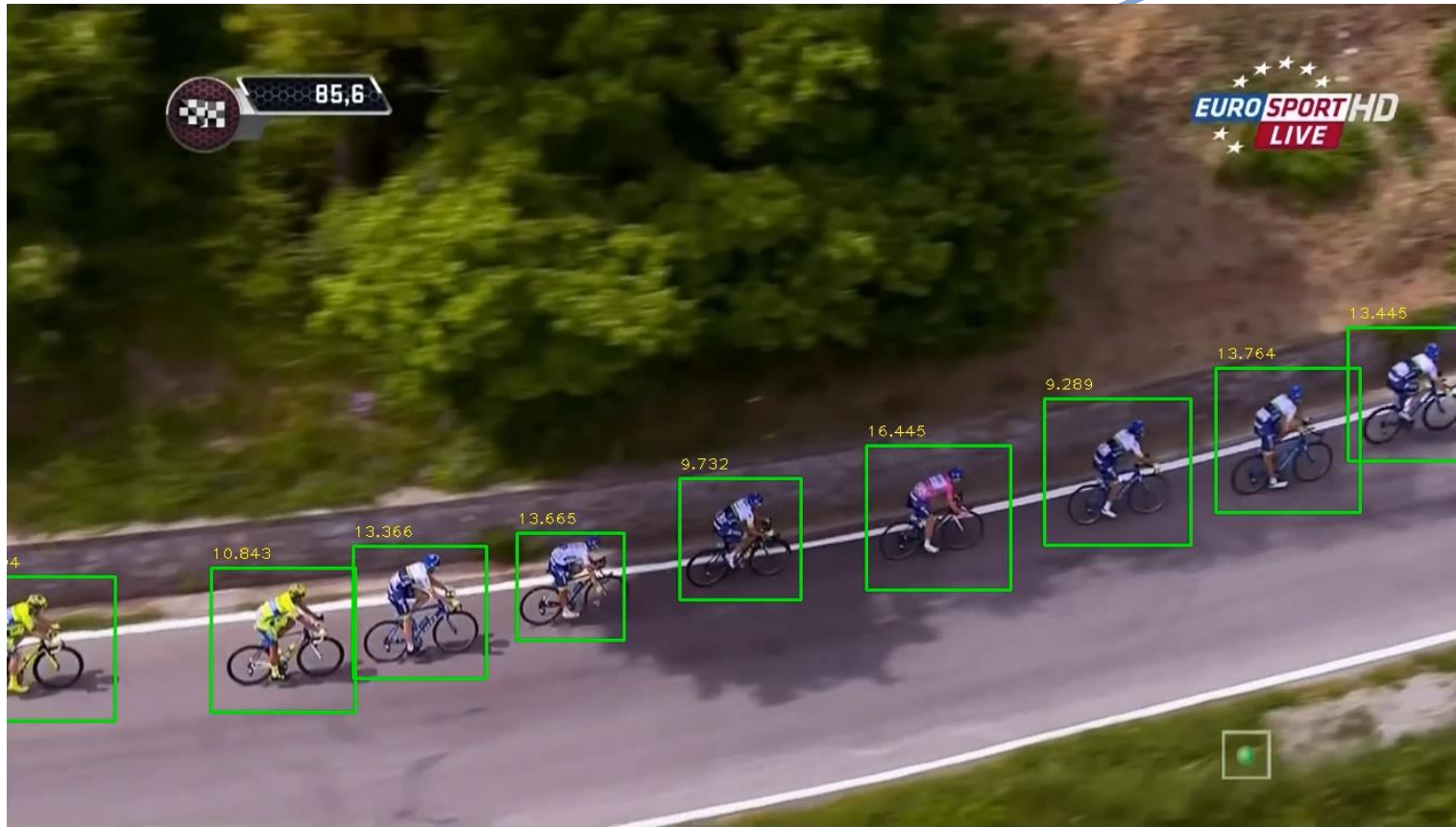
Object detection acceleration

- YOLO: good precision in general, but too heavyweight
 - small objects are more challenging to detect.
- Evaluation on VOC (Mean average precision, time):

Input Size	FPS	mAP	Forward time (ms) No TensorRT	Forward time (ms) TensorRT	Forward time (ms) FP16
608	2.9	71.26	241.5	128.8	69.3
544	3.2	73.64	214.4	121.2	64.3
480	5.4	74.50	155.4	62.3	35.7
416	6.4	73.38	155.3	56.5	32.5
352	7.8	71.33	111.0	45.0	24.3
320	8.5	70.02	103.0	40.4	22.8

Object detection

- Detection with **Light weight** deep CNNs



Object detection

- Detection with **Light weight fully** CNNs
- The network is trained with training samples of size **32x32**
- It is **fully convolutional** and accepts images of **arbitrary** size
- The network outputs a classification heatmap containing probability scores for each **32x32** region of the input





UAV Object Tracking specs

- 2D visual tracking will be employed for target following.
- Satisfactory performance in UAV sports footage is required.
- Target tracking should be performed in real-time, i.e., $> 25 \text{ fps}$.
- On drone implementation might be required as well, thus low computational complexity is preferred.
- Parallel or parallelizable methods (e.g., with CUDA implementations) should be preferred as well.
- Assuming 2D target tracking methods operate faster than combining target detection and recognition methods, long-term UAV tracking is also preferred.



UAV Object Tracking benchmarking



- 14 top performing 2D trackers [VOT 2016] were implemented in MATLAB using the UAV123 dataset interface.
- Performance was evaluated in 26 UAV videos obtained from UAV123 and YouTube, including long term videos as well.
- 3-fold evaluation:
 - Precision plot (the ratio of successful frames, where the tracker output is within the given threshold (x -axis of the plot, in pixels) from the ground-truth, measured by the center distance between bounding boxes)
 - Mean time before success rates falls below $y\%$, $y = 10, \dots, 100$
 - Operation speed.
- Evaluation platform: Ubuntu 16.04, 8GB ram, i7

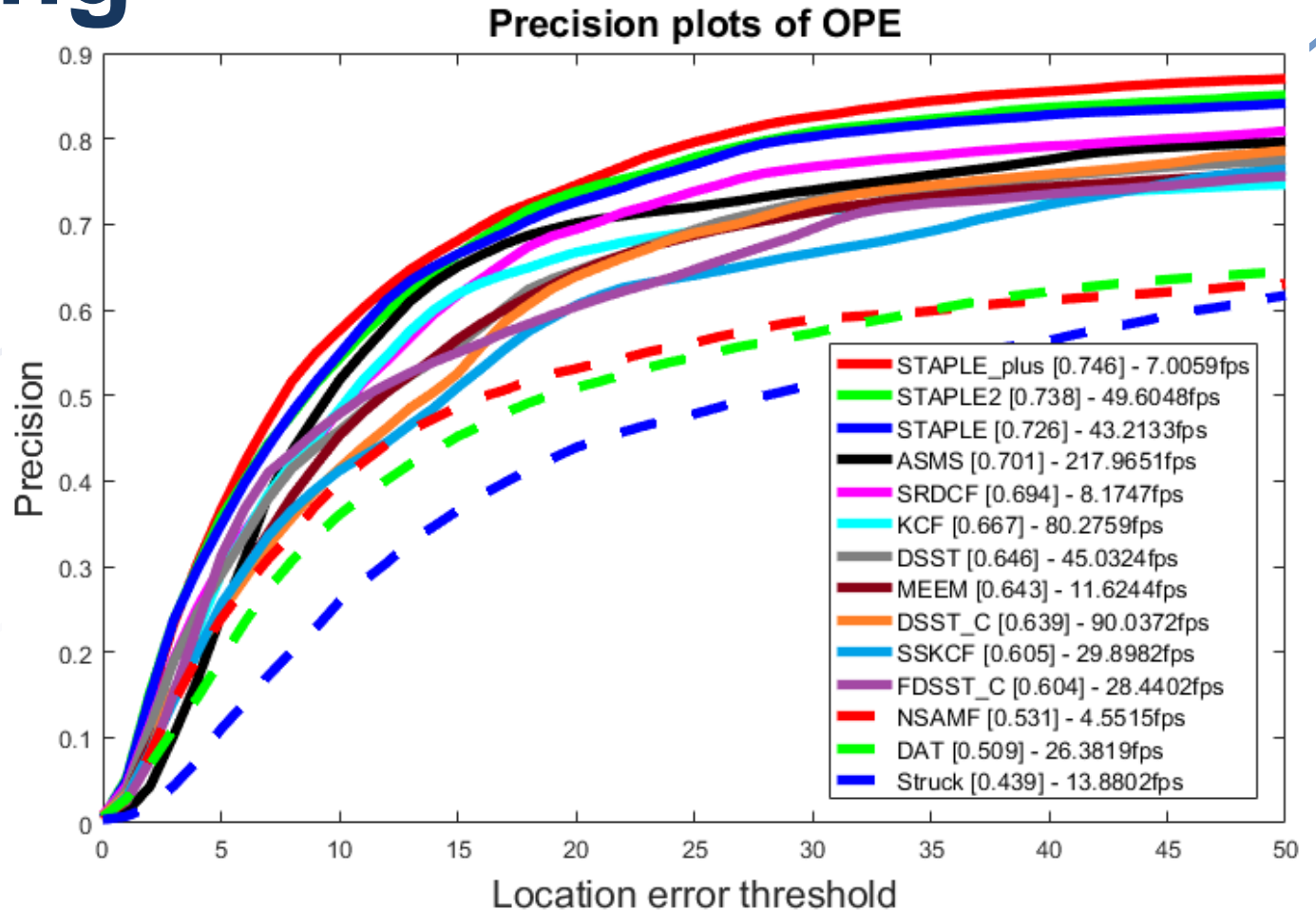


UAV Object Tracking benchmarking

MultiDrone



- ASMS provides a good compromise between accuracy and fps rate



Correlation filter based tracking

- Long- term 2D visual tracking faces 3 challenges:
 - Scale changes (SoA trackers handle it pretty well).
 - Target Occlusions (The biggest problem).
 - Target Rotations (KCF handles it well).
- Our goals:
 - Long-term tracking precision (low distance from the actual target center).
 - Real-time on-drone performance.
- Our contributions:
 - Background filter (+~1% precision*) – Helps with rotations.
 - Occlusion detection based on PSR metric (+2.5% precision*) – Helps with Occlusions.
 - Learning occlusions from tracker responses using an advanced classifier (SVM) (+~4% precision*).
 - Extending the framework to other trackers.

*When compared with standard KCF tracker

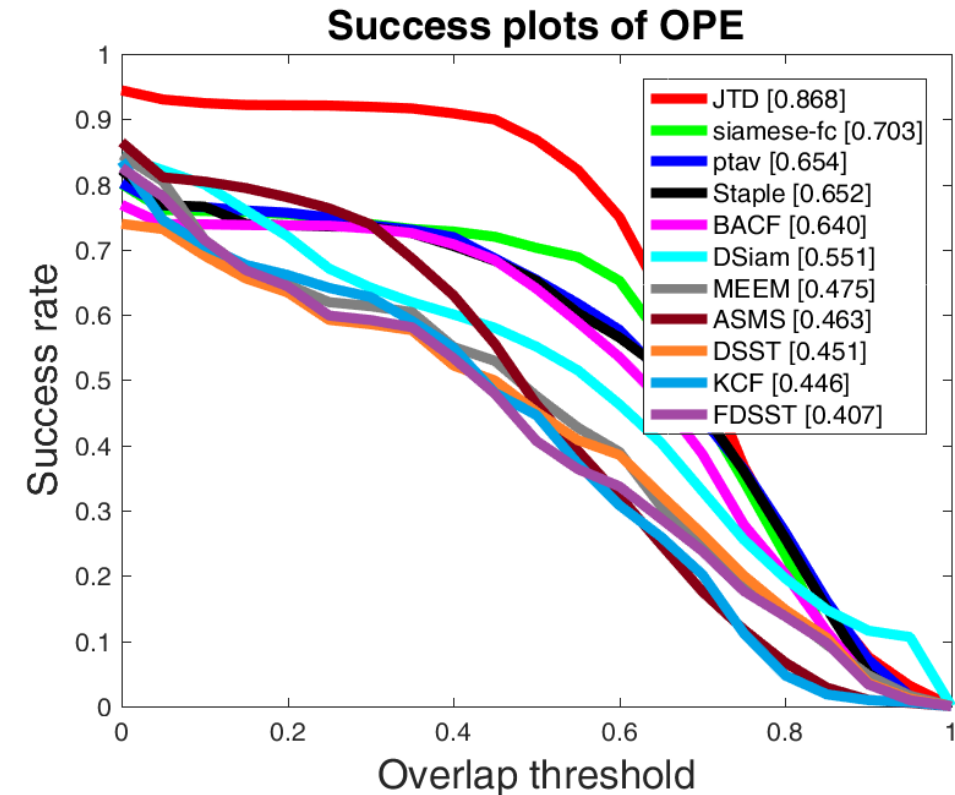
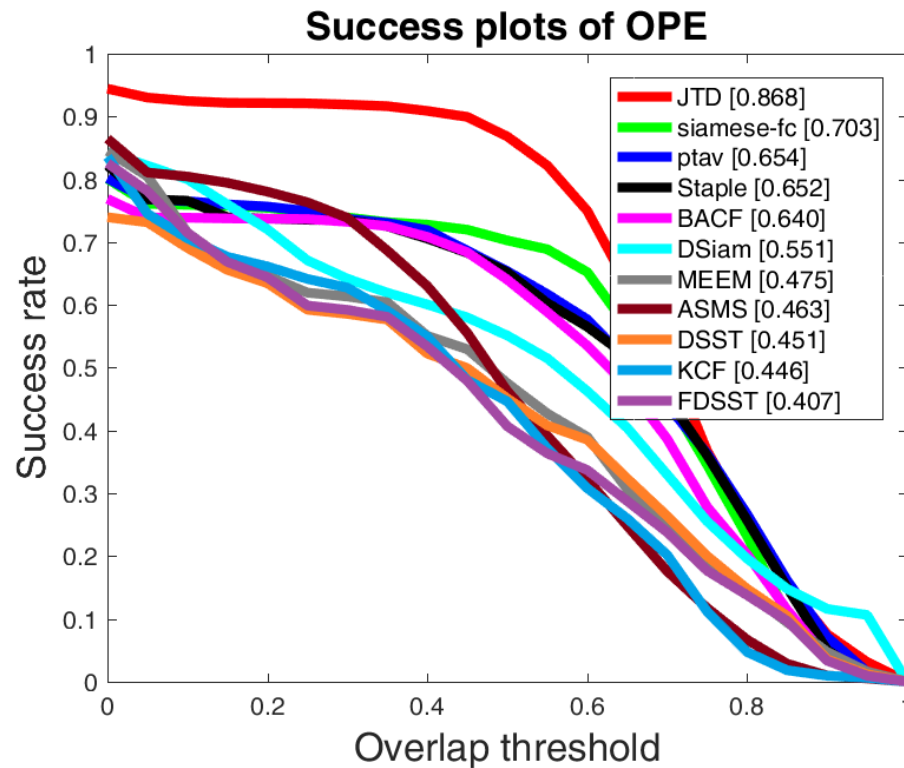
Joint Detection & Tracking

- **Tracker:** Given the initialized position of a target, the tracker T is responsible for estimating the bounding box of the target in the subsequent frames.
- **Detector/Verifier:** Given a bounding box defining the target in a specific frame produced by the tracker, the detector D is responsible for verifying this result, and then provide the appropriate feedback to the system. If the verification fails this module is responsible for detecting the target in a local search area and provide the correct bounding box to the master node M
- **Master:** M is responsible for the coordination of the two aforementioned modules. The node provides the necessary services to control the verification, the detection and the tracking tasks and controls the communication between the different parts of the system.



Joint Detection & Tracking (JTD)

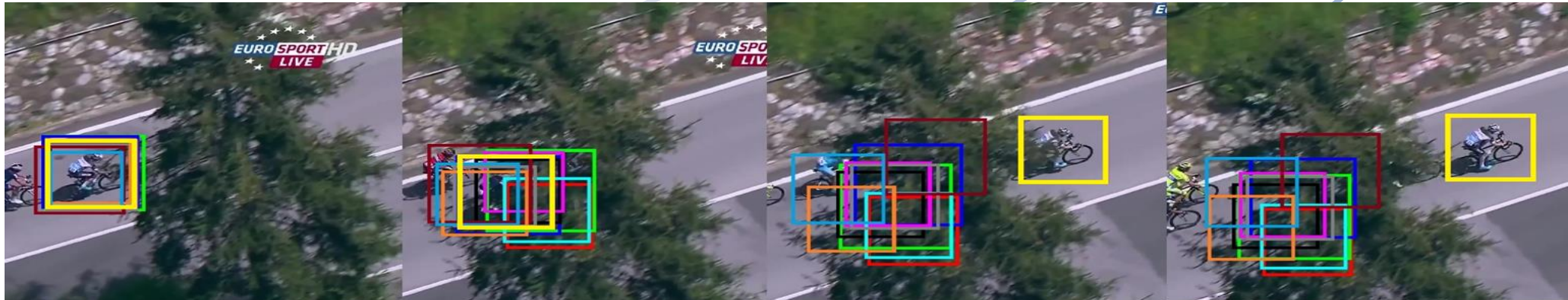
Evaluation against other state-of-the-art trackers on 13 aerial videos



Joint Detection & Tracking



- Target reinitialization by the detector in hard tracking cases when tracking algorithms fail



Joint Detection & Tracking



- Target reinitialization by the detector in hard tracking cases when tracking algorithms fail



Multi-Target Tracking



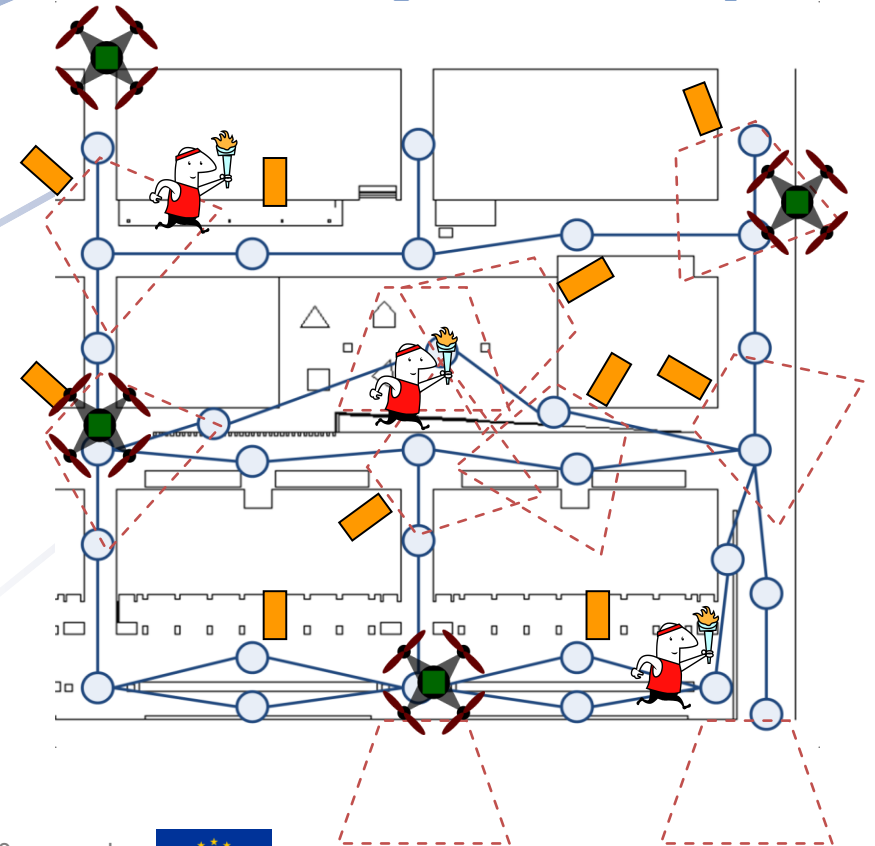
- The implementation is extended to support the tracking of multiple targets while maintaining real-time performance



Optimal multi-sensor multi-drone 3D target localization, tracking & following



- Problem: maximize a merit metric resulting from multi-drone object tracking
- Assumptions:
 - M drones
 - All cameras can be oriented
 - Drones motion is assumed known
- Objective: maximize a merit metric resulting from multi-drone





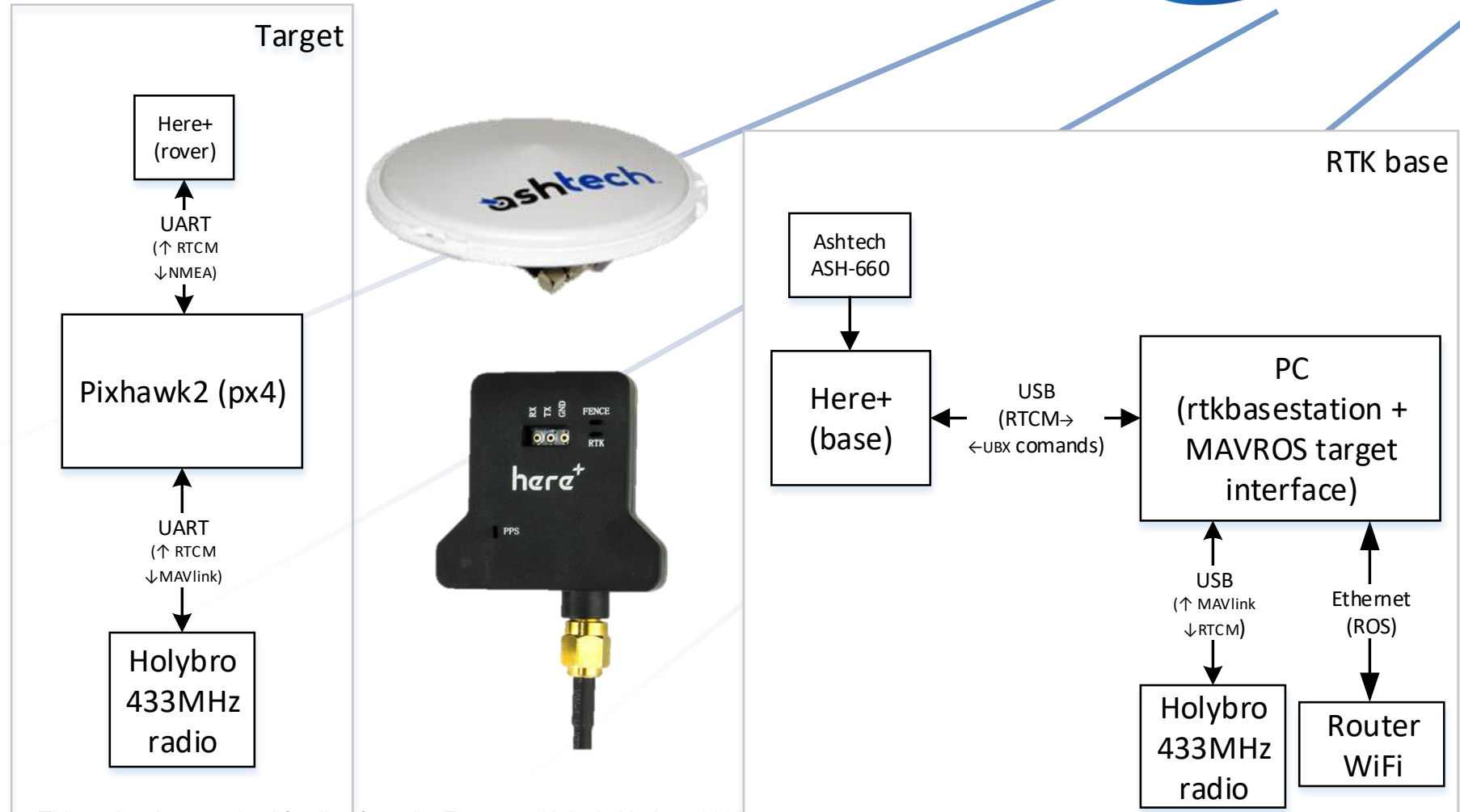
3D target localization

- 3D target localization using 3D maps
- **GPS target localization**



Target RTK GPS

MultiDrone

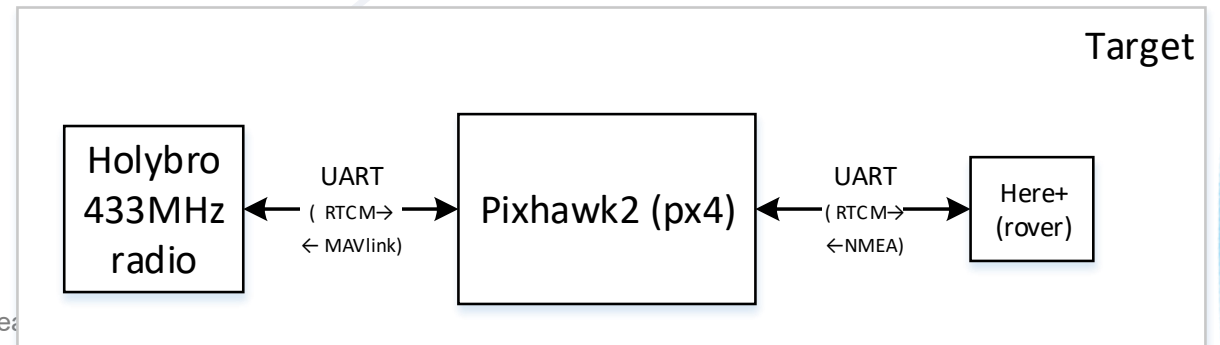
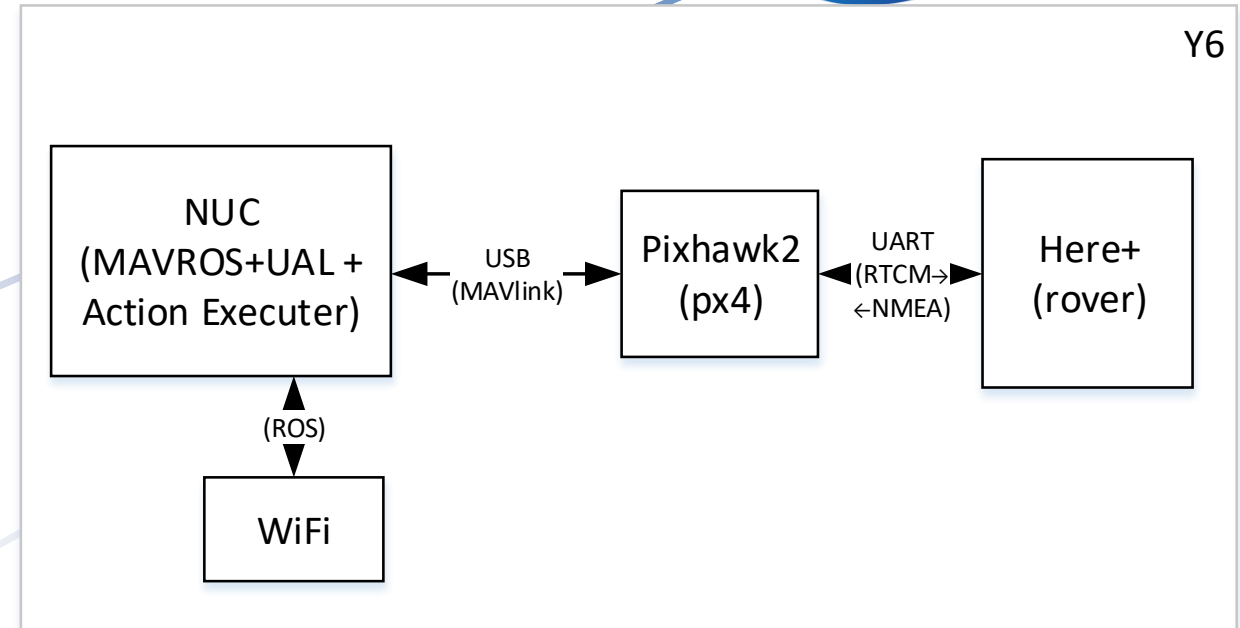
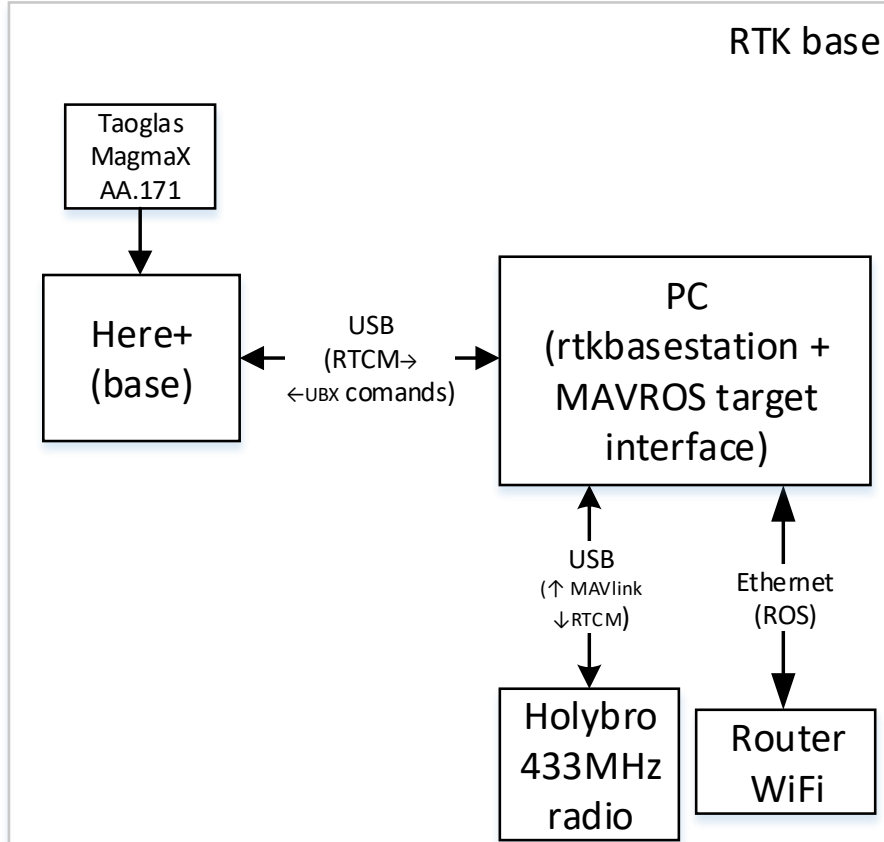


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)



Target tracking with a y6 drone

- Ongoing experiments





UAV Shot Type Taxonomy

- There is no prior, comprehensive work on identifying aerial shot types suitable for UAV-based cinematography
- In this context, shot type refers to a combination of camera/UAV motion (with regard to the target) and target composition/framing
- 26 motion types have been identified, each one compatible with a number of framing types
- Not all of these motion types include shooting a specific target/subject (e.g., establishing shots)





UAV Shot Type Taxonomy

- Example 1
 - Motion Type: Fly-By
 - Framing Type: Long Shot



- Example 2
 - Motion Type: Chase
 - Framing Type: Long Shot



UAV Shot Type Identification



- Example:
CHASE



UAV Shot Type Simulation

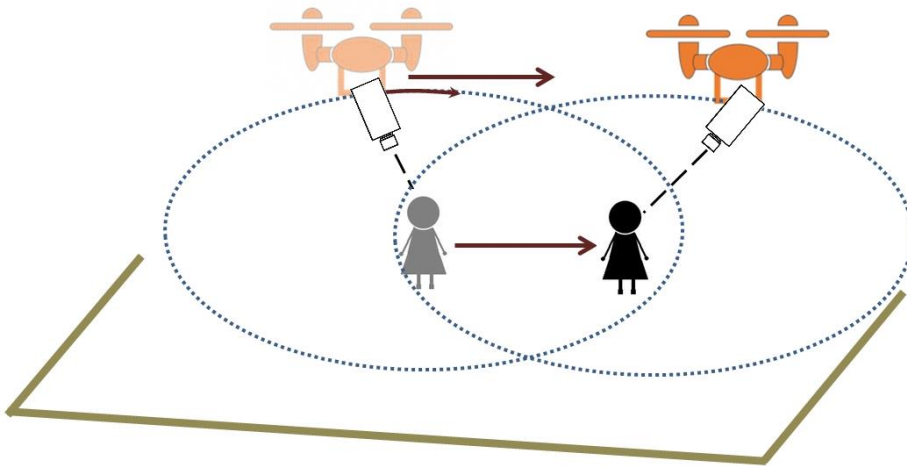


- Example:
CHASE



UAV Shot Type Identification

- Example: ORBIT



$\mathbf{P}_t = [p_{t1}, p_{t2}, p_{t3}]^T$: Target 3D position at time instance t

$\mathbf{X}_t = [x_{t1}, x_{t2}, x_{t3}]^T$: Look-From Point at time instance t

\mathbf{L}_t : Look-At Point at time instance t

$$t \in [0, \frac{T\theta}{\omega}]$$

$$\theta_0 = \arctan\left(\frac{x_{02}}{x_{01}}\right)$$

$$x_{t3} = x_{03}, \forall t$$

$$\lambda = \sqrt{\lambda_{3D}^2 - x_{t3}^2}$$

$$\mathbf{X}_t = [\lambda \cos(t\frac{\omega}{T} + \theta_0), \lambda \sin(t\frac{\omega}{T} + \theta_0), x_{t3}]^T$$

$$\mathbf{L}_t = \mathbf{P}_t$$

Shot type feasibility for intelligent UAV AV shooting



Determining the desired focal length to achieve specific shot types (constant distance between UAV and target)

Motion type	$\min f_{max}$	f_s when $c_s = 50 \%$ (medium shot)	f_s when $c_s = 80 \%$ (closeup)
LTS	77.8 mm	103.4 mm, not feasible	150.7 mm, not feasible
CHASE	162.9 mm	103.4 mm, feasible	150.7 mm, feasible
ORBIT	128.8 mm	103.4 mm, feasible	150.7 mm, not feasible

A shot type is feasible if the $f_{max} > f_s$



Target Pose Estimation

- **Computer Vision Approach**

- Relies on detecting a set of *predefined points* (e.g., facial landmarks) and then using a method for solving the respective *Perspective-n-Point (PnP) problem*, i.e., estimation of the camera position with respect to the object.

- **Limitations:**

- The 3-D coordinates for the landmark points must be known, i.e., a 3-D model of the object is needed
- The landmarks points must be precisely tracked, i.e., the texture of the object must allow for setting enough discriminative landmarks





Target Pose Estimation

- **Machine Learning Approach**

- A neural network receives the object and directly *regresses* its pose
- Only a set of pose-annotated object pictures are needed
 - There is no need to manually develop 3-D models
 - The models are more robust to variations of the object for which we want to estimate its pose
 - The pose estimation can run entirely on GPU and (possibly) incorporated into a unified detection+pose estimation neural network
- Very few pre-trained models are available
 - Models must be trained for the objects of interest (faces, bicycles, boats, etc.)





Target Pose Estimation

- **Machine Learning Approach**
 - We integrated a pre-trained yaw estimation model of facial pose (DeepGaze library) into the SSD-300 object detector (trained to detect human faces)
 - Varying illumination conditions seem to affect the estimation.



Drone vision for cinematography: Functionalities (3)



- 1. Visual and perception data analysis for safety and security:**
 - 1. Obstacle detection.**
 - 2. Event detection.**
 - 3. Privacy protection.**
 - 4. Emergency landing site detection.**
 - 5. Crowd detection.**
 - 6. Semantic 3D map annotation.**



Privacy Protection, Ethical and regulatory issues



- Legal, ethical, safety and security issues arise upon scheduling professional UAV flight/capture sessions.
- Challenging constraints, restriction on the shooting mission are imposed, deriving from:
 - flight regulations
 - safety and security considerations
 - data privacy rules





Flight regulations

- Different flight regulations apply depending on employed UAV types and the application.
- Restrictions include:
 - Maximum UAV weight
 - Permitted flight radius
 - special prerequisite conditions (e.g., licensed pilot requirements and insurance policies).
- An important issue is that flight restrictions vary over different countries, while professional pilot licenses and insurance policies may not be internationally valid.
- Adjusting/replacing components, may impact the category classification (weight is calculated with payload)

Typical UAV flight regulations in EU



- UAVs < 2kg are allowed within a 50m flight radius without professional pilot license.
- Pilot license and drone insurance are required for all professional applications.
- UAVs > 2kg of weight may be required to carry emergency parachutes (France).
- UAVs exceeding 15kg of weight might require special license or even be prohibited (Germany).



Typical UAV flight regulations in EU



- Maximum **flight altitude** is typically **restricted to 120m or 150m** (400ft or 500ft) within several European countries.
- **Line of sight must be maintained** by the licensed pilot of the UAV **at all times**, either physically, or using visual aids (e.g., VR-goggles).
- **Horizontal distance** between the drone and the pilot is typically **limited to** specific meters (e.g., **500m**).
- Outdoor UAV **flight is restricted/prohibited above congested areas**, crowds of people **and airports**, leading to permissible flight zones delineated by law.
- Inherently complying with such a complex and varying web of regulations (geo-fencing) is a challenge for all autonomous UAV applications (e.g., DJI app automatically downloads and determines permitted flight zones).



Safety, Security and ethics

- Misuse avoidance
 - no specific legislation prescribes protective measures against misuse and vulnerability exploitation.
 - Drone hacking, GPS signal jamming, weak security in communications can also allow obtaining the video captured by the drone, or its intended flight path.
 - Redundant active perception methods (drone localization), secure and signed autopilot firmware updates, as well as autopilot input commands filtering, can be employed to this end.
- Data security
 - Footage data collected by UAVs raise privacy concerns.

UAV data security requirements



- The types of data that must be protected are:
 - data stored within drones;
 - On-drone data encryption, allowing access to authenticated people only.
 - data stored in ground infrastructure.
 - data transmitted over the air;
 - Wifi and radio data transmitted in commercial drones are not encrypted.
 - Data protection can be achieved with ciphering and authentication mechanisms, e.g. IPSec over LTE for transmitted data.
 - Data that are to be publicly distributed (e.g., AV datasets)





Privacy protection

- Protection of personal data must be ensured in the acquired video and/or images.
- The EU's General Data Protection Regulation 2016/679), repealing the 1995 Data Protection Directive.
- *“Member States shall protect the fundamental rights and freedoms of natural persons and in particular their right to privacy, with respect to the processing and distribution of personal data.”*



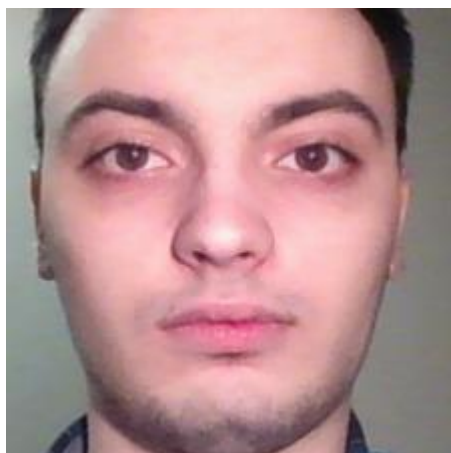
Privacy Protection

- An issue of ethics and security
- Post-production stage
- Approaches
 - Face de-detection (Face detector obfuscation)
 - Naïve approach
 - SVD-DID
 - Face de-identification (face recognizer obfuscation)
 - Gaussian blur
 - Hypersphere projection

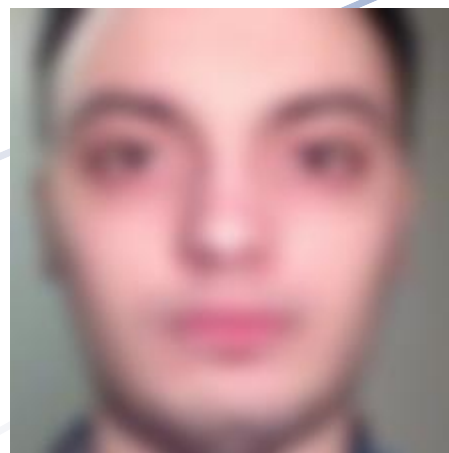


Privacy Protection: acceptable facial image quality?

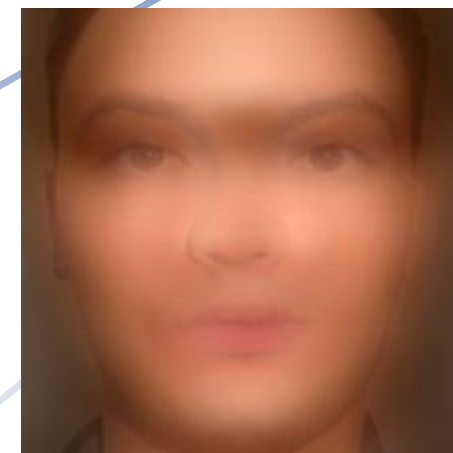
MultiDrone



Original Image



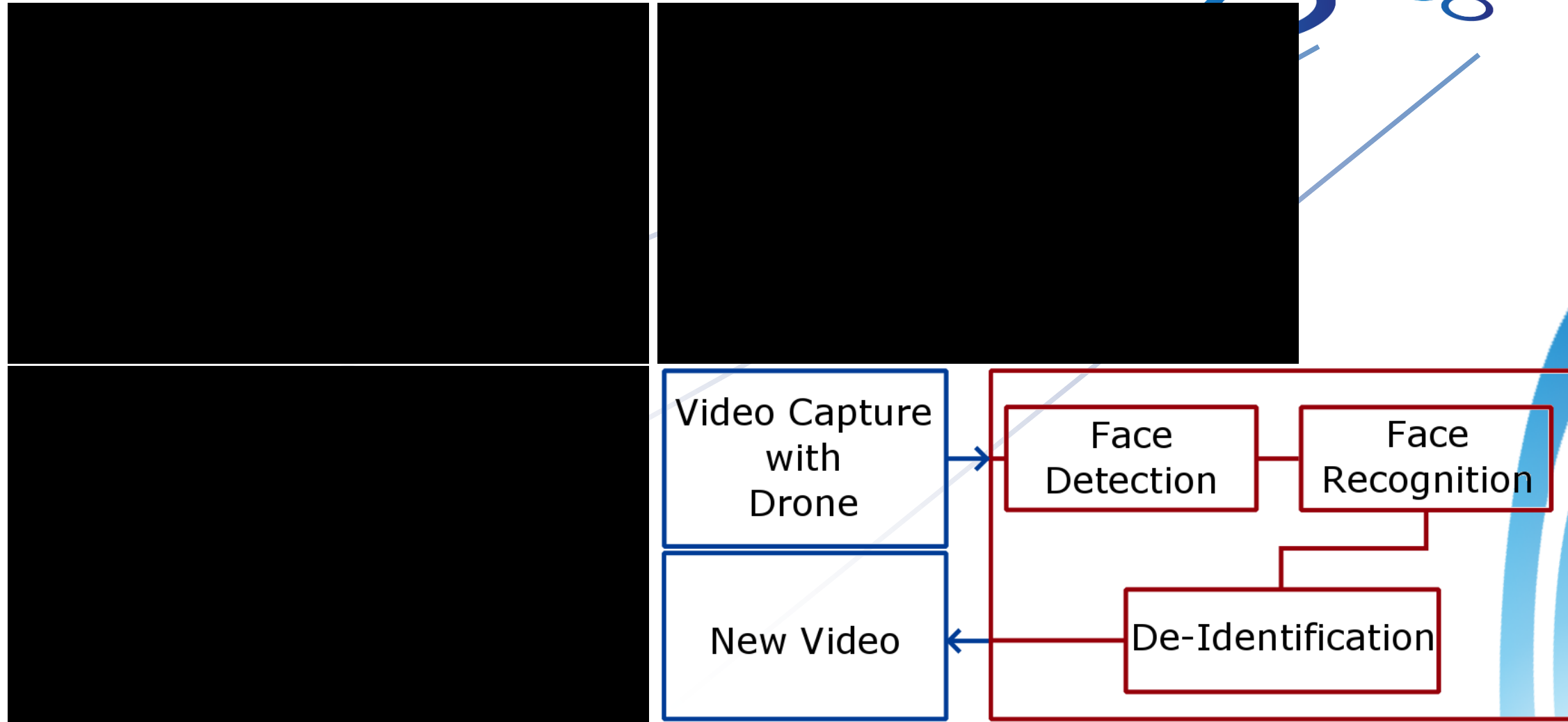
Gaussian blur with
std. deviation of 5



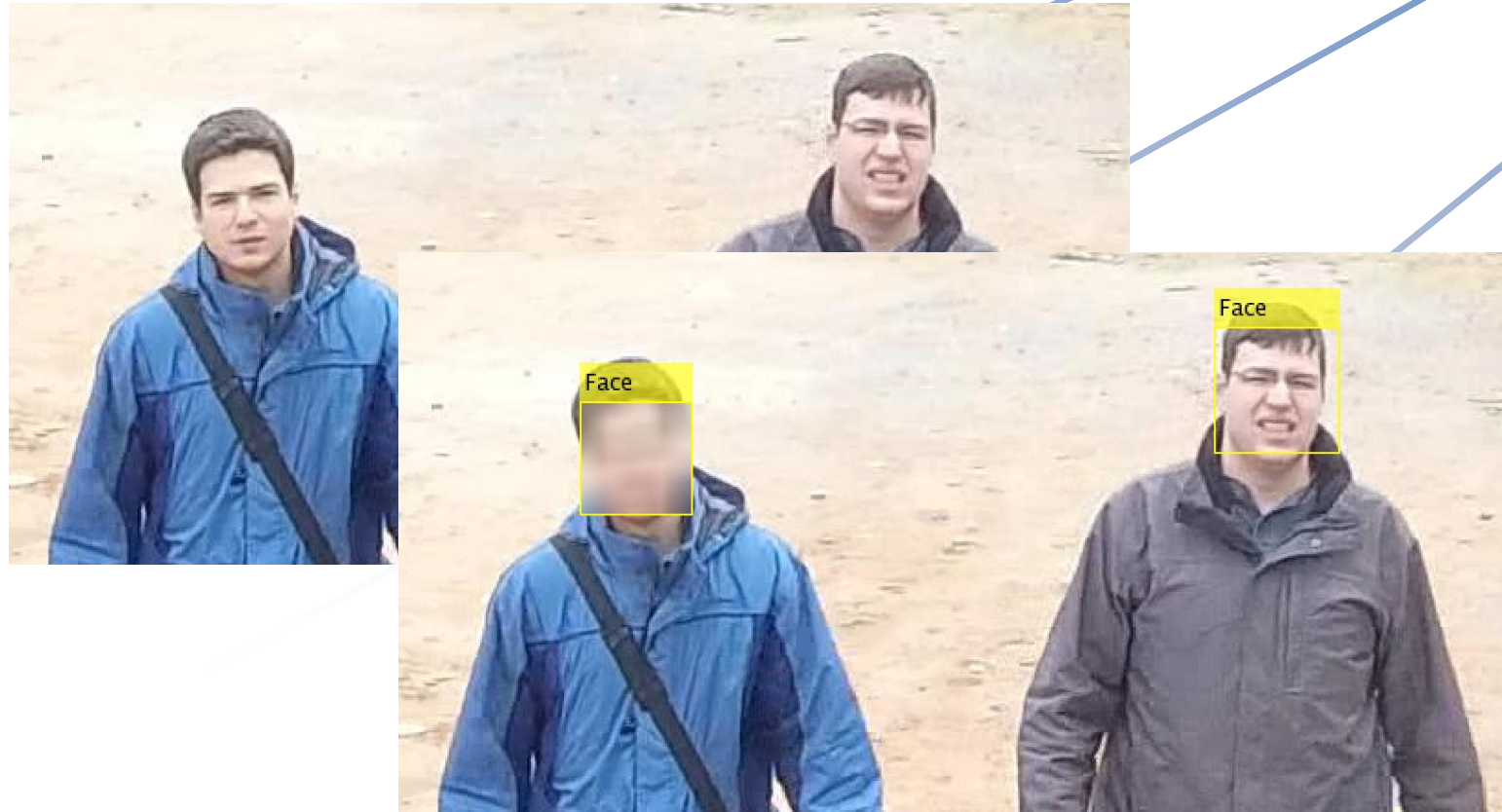
Hypersphere
projection with
radius of 8



Application on drone videos



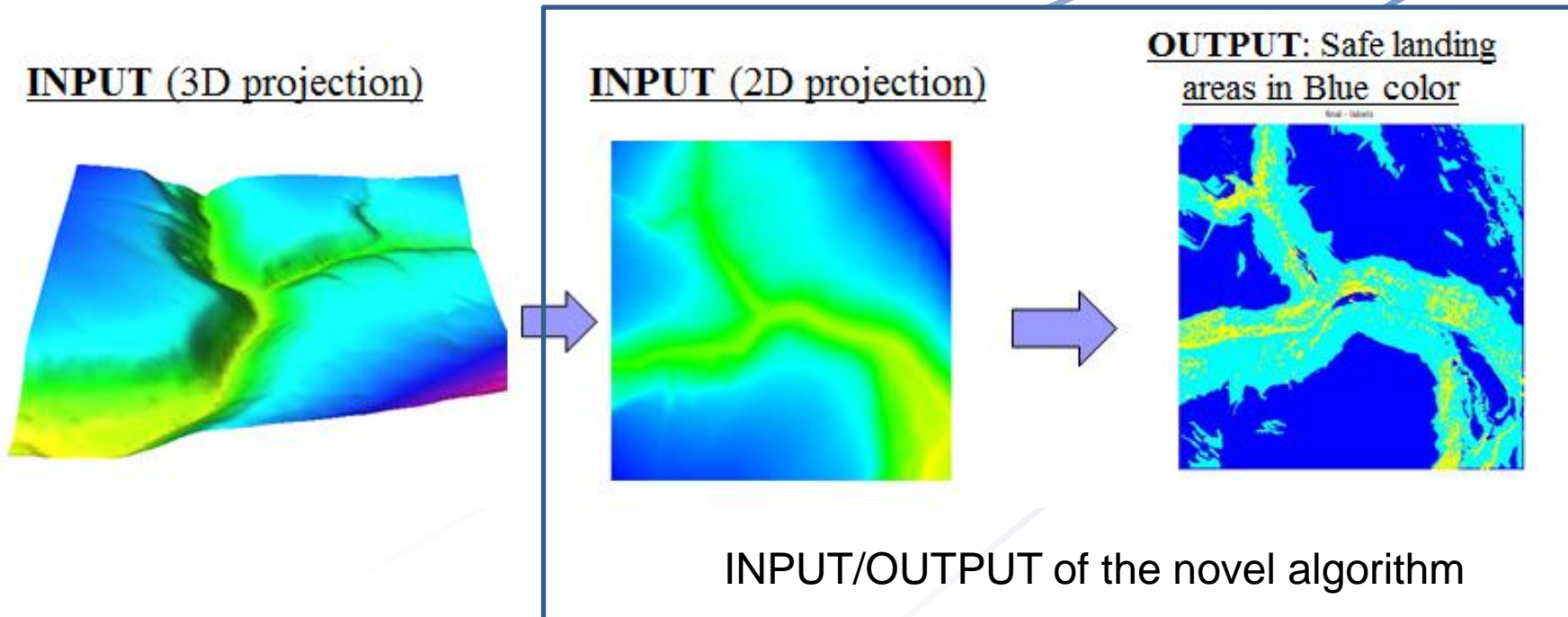
Face recognition/de-identification/privacy protection



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)



Potential Landing Site Detection

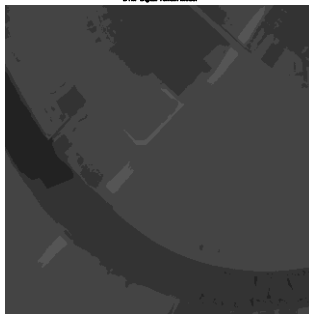


Potential Landing Site Detection

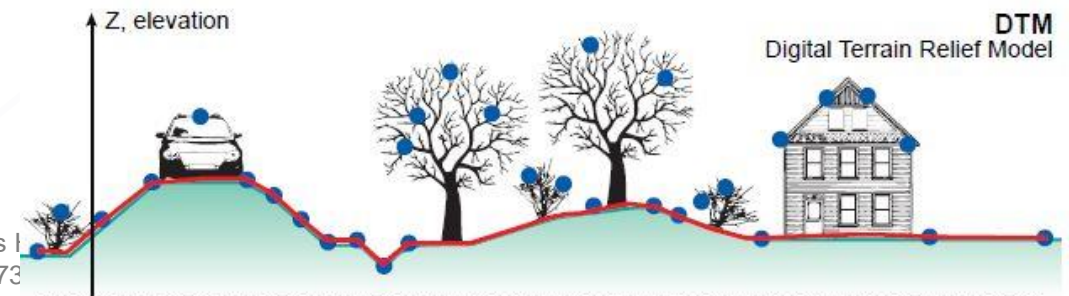
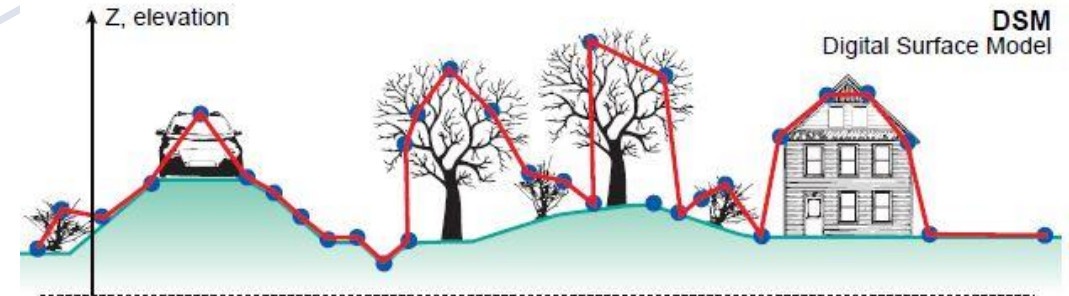
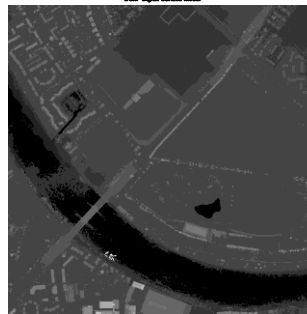


- Algorithm for identification of potential landing areas using digital elevation models (DEM) consists of five discrete steps:
 - input consists of two digital elevation models namely
 - the digital surface model (DSM) and
 - the digital terrain model (DTM) of a region
 - in raster format, i.e., as a regular grid of elevation values of a depicted terrain.

DTM



DSM



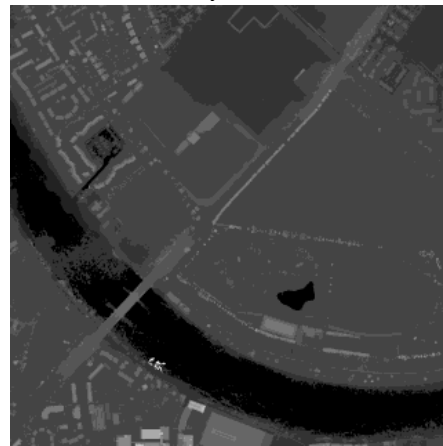
This project has received funding from the European Union's
and innovation programme under grant agreement No 73

Potential Landing Site Detection

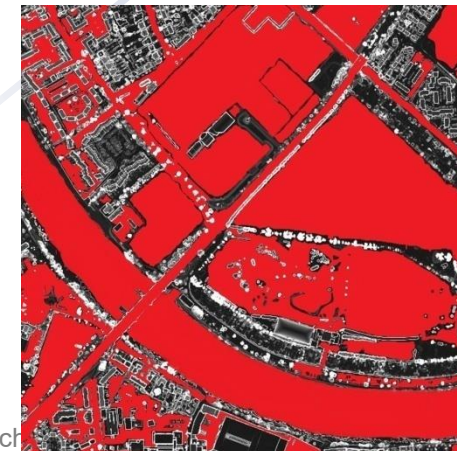


- Number of areas depicted in the DEM data from the publicly available dataset provided by UKs Environment Agency
 - DEM covering urban, suburban, rural and bush areas
 - Spatial resolutions (pixel size per dimension) ranging from 0.25m to 2m
- Areas no 1 and 2, resolution 0.25m refer to an urban environment with many obstacles such as buildings and trees
- Area no 3, resolution 2m is a rural environment with steep downhill descent parts
- Ground truth (potential landing sites, areas not suitable for landing) was manually constructed through visual inspection of the DEMs and satellite images (the latter were obtained by Google Maps).

DSM



Ground Truth



This pr
ar

Union's Horizon 2020 research
grant No 731667 (MULTIDRONE)



Crowd Detection



- Can effectively distinguish between crowded and non-crowded scenes.
- Provide crowd heatmaps that can be used to semantically enhance the flight maps by defining no-fly zones.
- Provide lightweight NN models, as imposed by the computational restrictions of the application.



Crowd Detection



- Propose to adapt a pre-trained CNN on our task, transforming it to a fast fully convolutional network.
- Second, we propose a two-loss-training model, which aims to enhance the separability of the crowd and non-crowd classes



Crowd Detection

MultiDrone



- A *Fully Convolutional Neural Network* can be trained for Crowd Detection
- The result is a heatmap



Semantic Map Annotation types (navigation/logistics)

MultiDrone

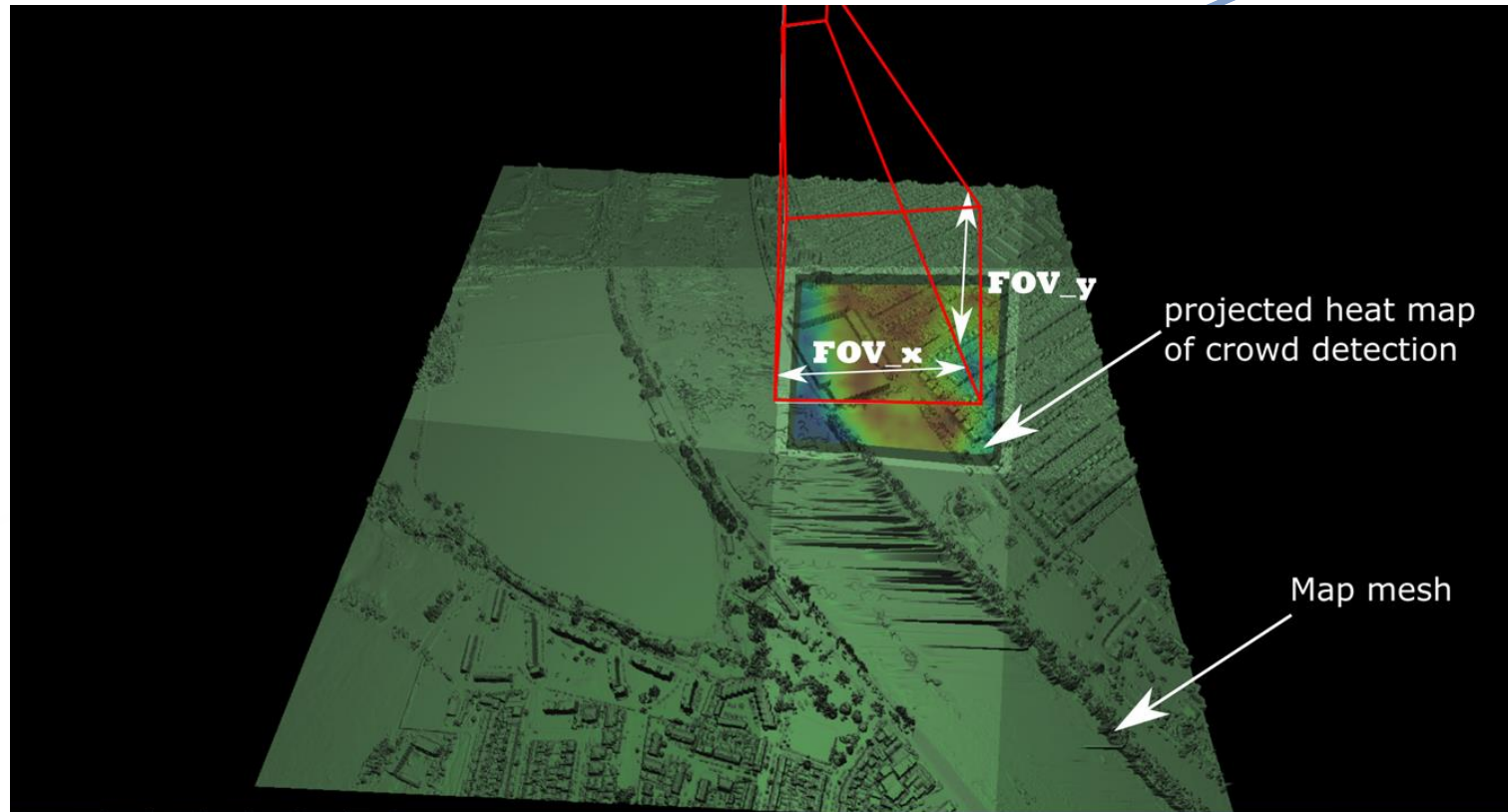


Type	Static/dynamic	Who	How	Geometric entity type
Regular takeoff and landing sites	Static	Supervisor	Manually	Point
No flight zones	Static	Supervisor	Manually or imported from a file, if available	Polygon (2D coordinates, longitude-latitude)
Potential emergency landing sites	Static	Supervisor	Manually	Polygon
Crowd gathering areas	Dynamic, during production	Visual Semantic annotator, Semantic map manager	Automatically	Polygon (2D coordinates, longitude-latitude)
Points of interest	Static		Manually	Point

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE)



Semantic 3D Mesh Map Annotation

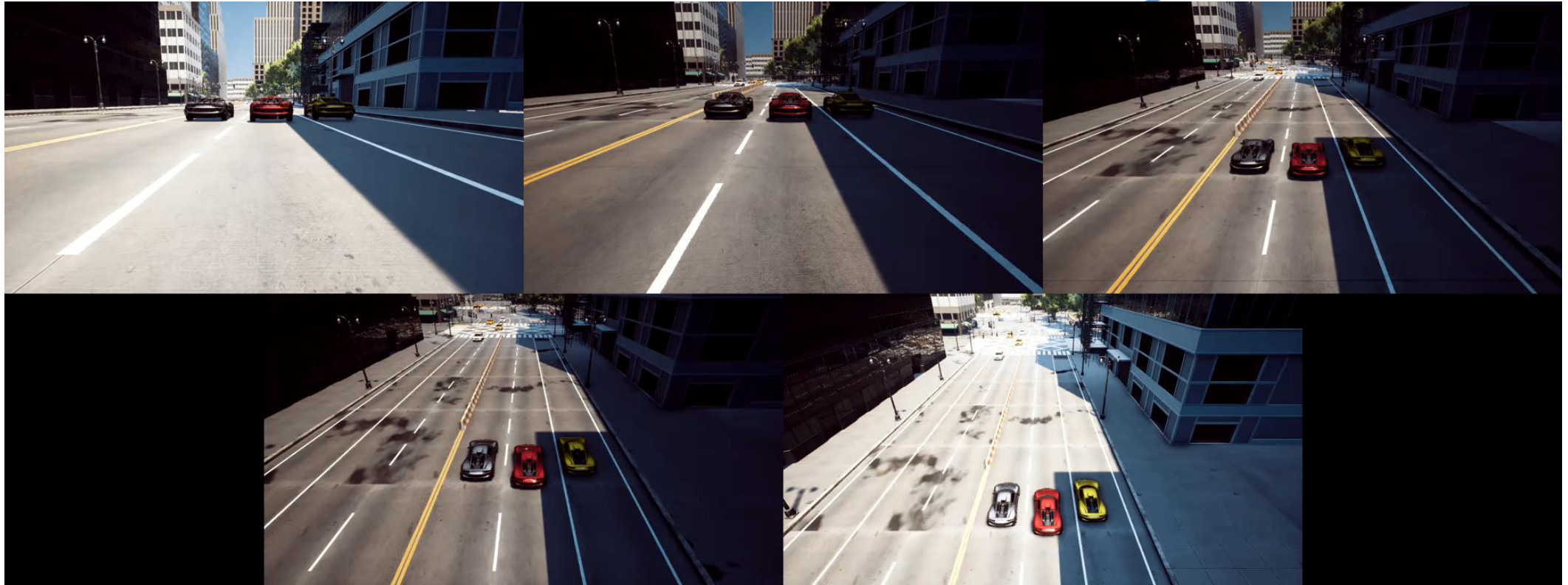




Drone mission simulations

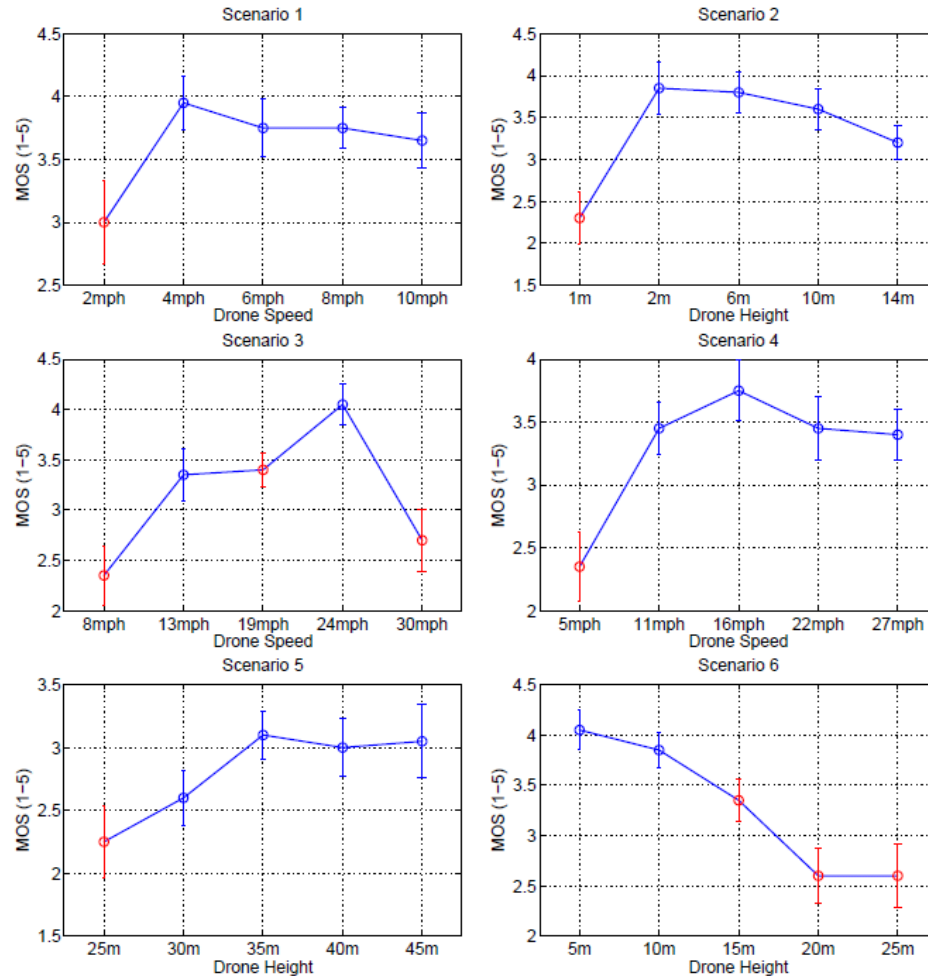
- **Subjective Evaluation on Viewing Experience of Drone Videos**
- Simulations for training data generation
- Simulations in Gazebo

Test Sequence Example II: S2



VIDEO: Scenario 2 with drone height of 1, 2, 6, 10 and 14m.

Analysis and Discussion



- The preference of drone parameters is content dependent and very subjective.
- Informal interviews after subjective test show significantly different scoring criteria.
- The optimal parameter range can be identified for each scenario, but needs to be related to relative motion or size.
- Short video duration can create a wow factor shot, but a longer duration may upset the viewer.
- This is only a pilot experiment and a much more comprehensive study is underway.

The Winner of Scenario 2



The Winner of Cycling 2



Demo Videos 2 from DW@Muncheberg Footage



GPU and multicore CPU architectures. Algorithm mapping

- **NVIDIA embedded processing boards**

NVIDIA Jetson TX2

NVIDIA Jetson Xavier

- GPU and multicore CPU architectures

Multicore CPUs

GPUs

- Algorithm mapping:

Convolutions



Introduction to fast CNN convolution algorithms

MultiDrone



Typical 2D convolutional layer l of a CNN:

$$x(i, j, c_{l+1}, l+1, k) = f(b(l, k) + \sum_{c=1}^{C_l} \sum_{i'=0}^{H_{l,k}} \sum_{j'=0}^{W_{l,k}} h(i', j', l, k) x(i-i', j-j', c, l, k))$$

input feature map \mathbf{X}_l : $N_l \times M_l \times C_l$ -dimensional 3D tensor

$\mathbf{W}_{l,k}$: $N_{l,k} \times W_{l,k} \times C_l$ -dimensional 3D tensor

$b(l, k)$: bias term

f : nonlinear activation function



Fast 1D convolution algorithms with minimal computational complexity

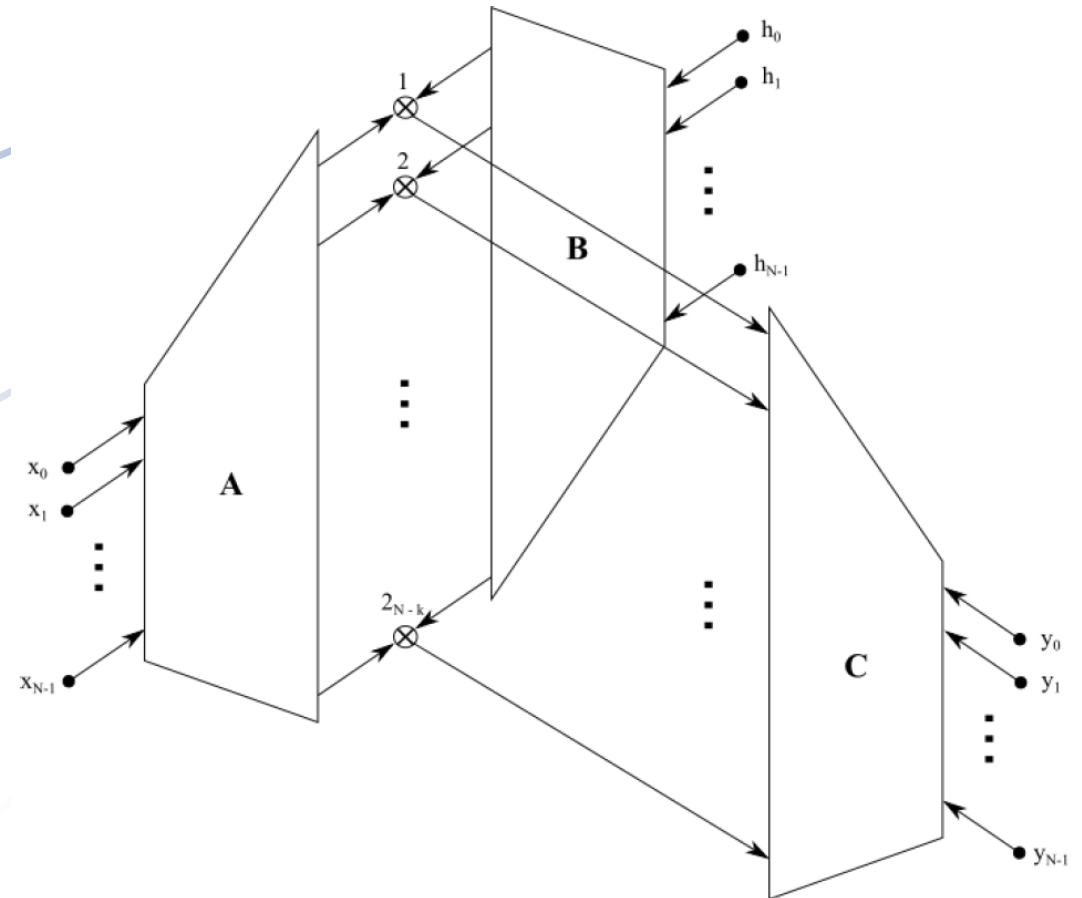
MultiDrone



Winograd convolution algorithms

$$Y = C(Ax \otimes Bh)$$

Require only $2N - v$ multiplications
in their middle vector product, thus
having minimal multiplicative complexity



Multidrone Consortium

1. Aristotle University of Thessaloniki, Greece (Coordinator)
2. Thales Communications & Security SAS
 - a. Thales Services
3. University Of Bristol
4. The University of Seville
5. Deutsche Welle (DW)
6. RAI Radiotelevisione Italiana (RAI)
7. Alerion
8. Instituto Superior Técnico, Portugal



Cooperation with other labs, R&D groups and projects



- **Have a look at www.multidrone.eu**
- Icarus.aiia: student forum on drone technologies.
- Creation of a SIG on drone technologies.
- Competitions on drone cinematography.
- Organization of an open call workshop in 2019.
- Special sessions and special issues.
- **Open to any new idea and cooperation options!**
 - **Send message to pitas@aiia.csd.auth.gr**



Q & A



Thank you very much for your attention!

**Contact: Prof. I. Pitas
pitas@aiia.csd.auth.gr
www.multidrone.eu**

