

UNIVERSITY OF UDINE
DEPARTMENT of COMPUTER SCIENCE,
MATHEMATICS and PHYSICS



AViReS
research laboratory



***Beyond the Fixed Gaze: Active
Vision in the Age of UAVs***

Prof. Gian Luca Foresti

IMPROVE 2026

6th International Conference on Image Processing and Vision Engineering

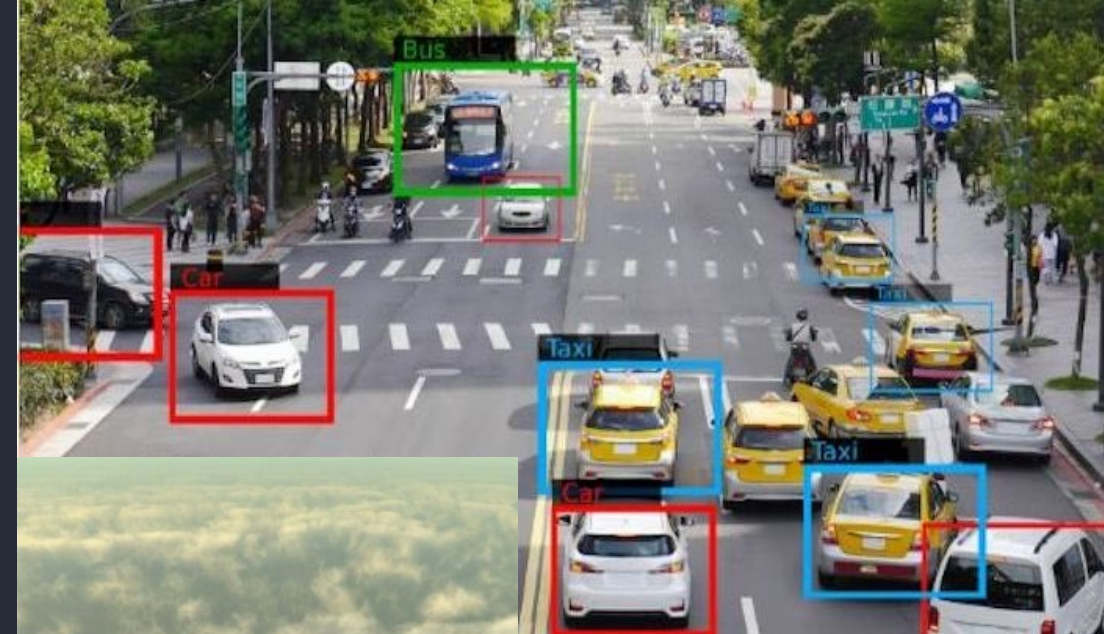
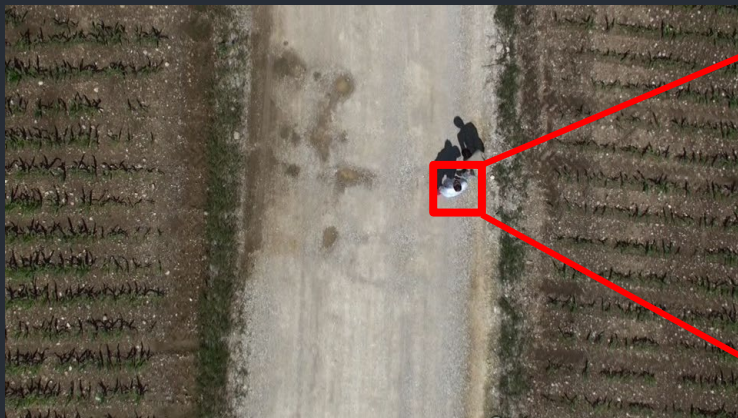
BENIDORM, SPAIN · 20 - 21 MAY, 2026



May 20, 2026

Outline

- Introduction – From Active Surveillance Systems to UAV
- Physical and Logical Architecture of UAV Systems for Situational Awareness
- Real Applications



...from Active Surveillance to UAV

ACTIVE VIDEO SURVEILLANCE



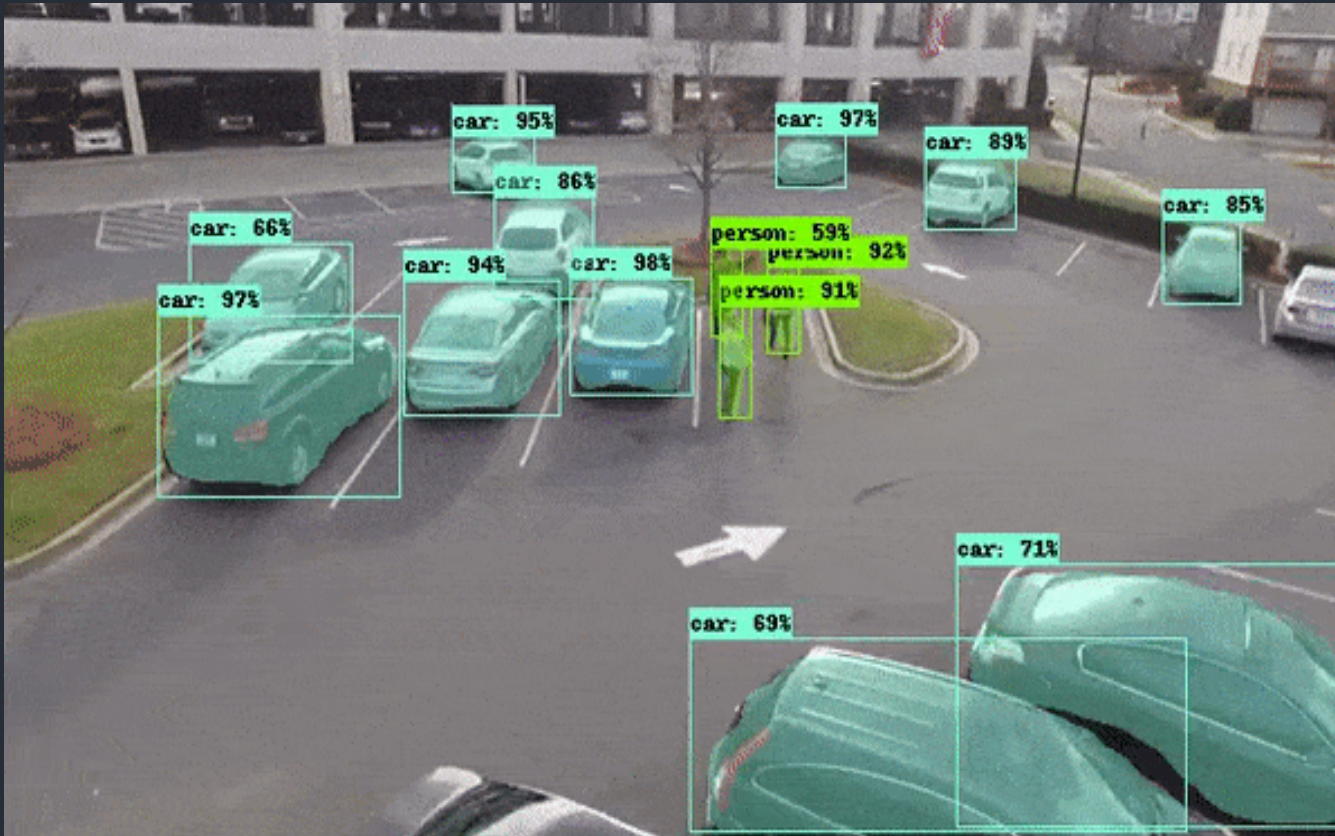
- The term **Active Video Surveillance** refers to the implementation of Automatic Systems able to:
 - Monitor and Understand **behaviors**, **activities** and **events**, carried out by humans and/or vehicles and/or natural and artificial agents.



- ACTIONS
 - **Alarms** (Sound Alarms, Operations Centers, Emergency Calls, etc.)
 - **Defense** (Countermeasures, Defense Systems/Closures, etc.)

ACTIVE VIDEO SURVEILLANCE (2)

Object/Target Recognition



- The ability of an algorithm to determine, in a certain spatial position, a "target" and also to be able to classify it within a predefined number of classes



- Knowing how to determine the **Where**...
- Knowing how to determine the **What**...
- The **When** is implicit!

ACTIVE VIDEO SURVEILLANCE (3)

Person Re-identification (Detection, Classification, Identification)

AViReS
research laboratory

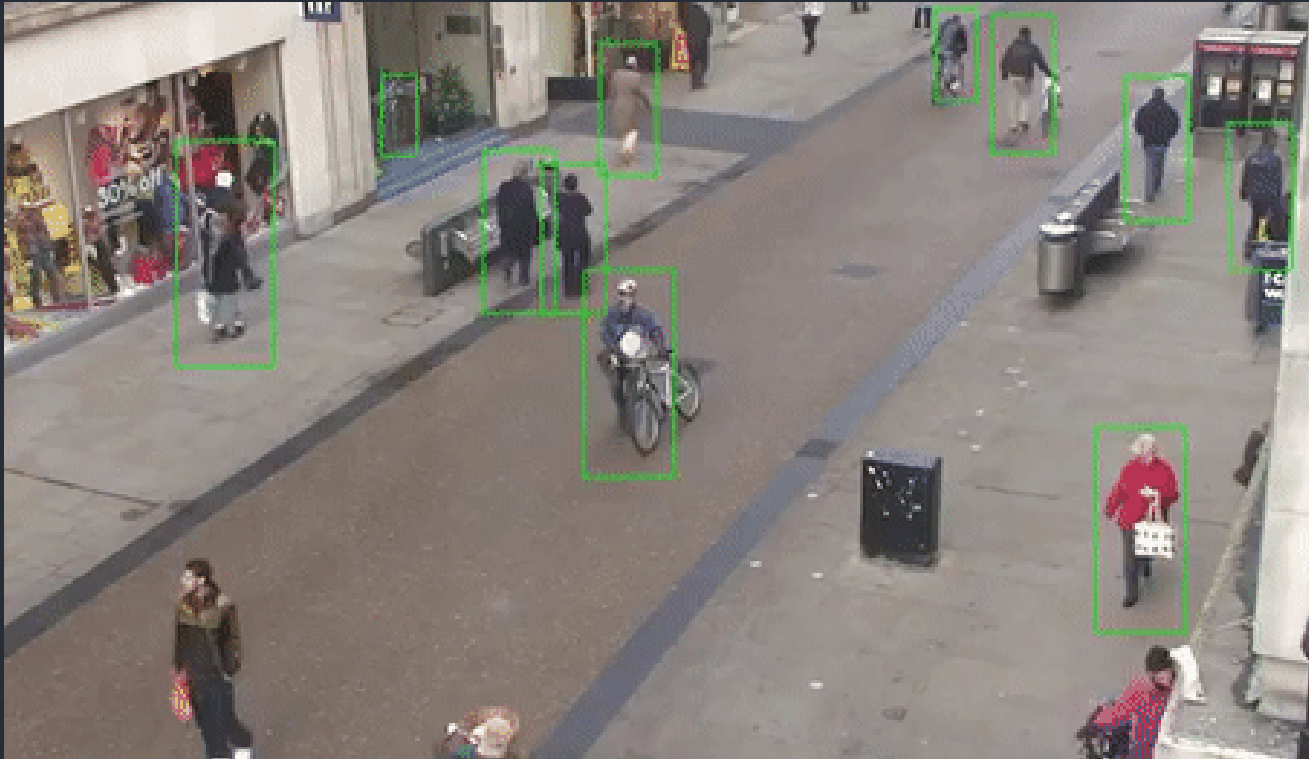


- The ability of an algorithm to determine, in **different places and at different time**, the **visual identity of a person** (e.g., we follow “Mario Rossi” around the city).



ACTIVE VIDEO SURVEILLANCE (4)

Simple Events



- **Simple Events** describe a series of "primitive" actions performed by subjects (e.g., people or vehicles).
- Subjects perform actions by interacting with the environment, but not by interacting with other subjects.
- Actions performed by a subject, if temporally and spatially correlated, trigger a Simple Event.



Unmanned Aerial Vehicles (UAVs)

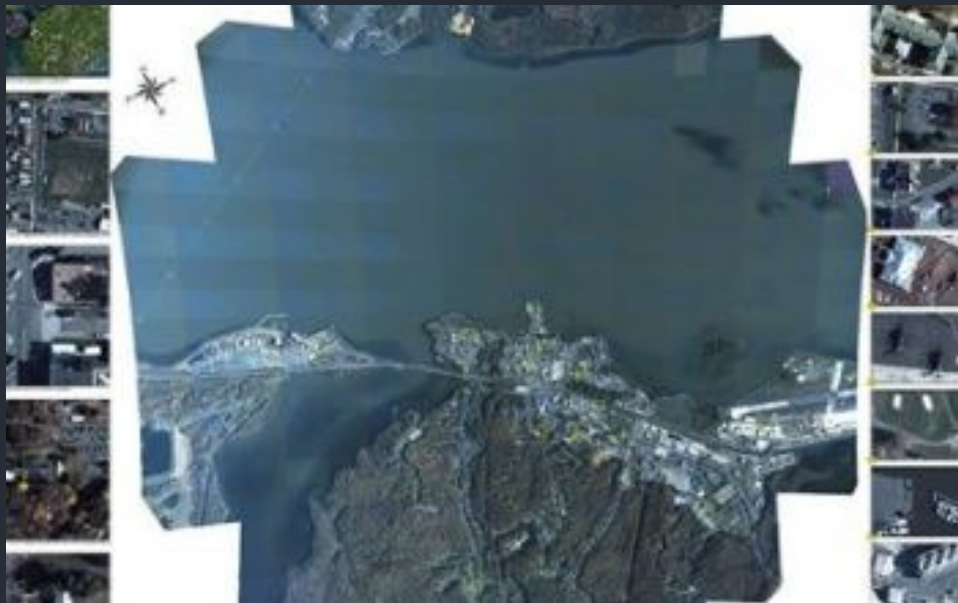
Unmanned Aerial Vehicles (UAVs)

UAV-based missions

AViReS
research laboratory

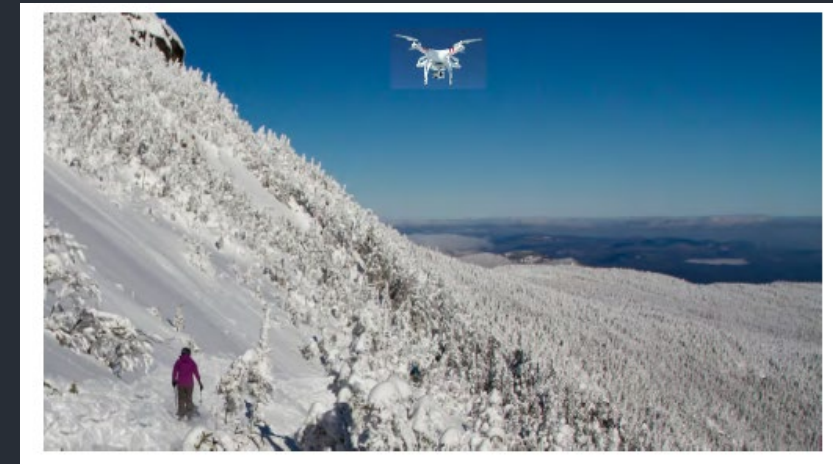
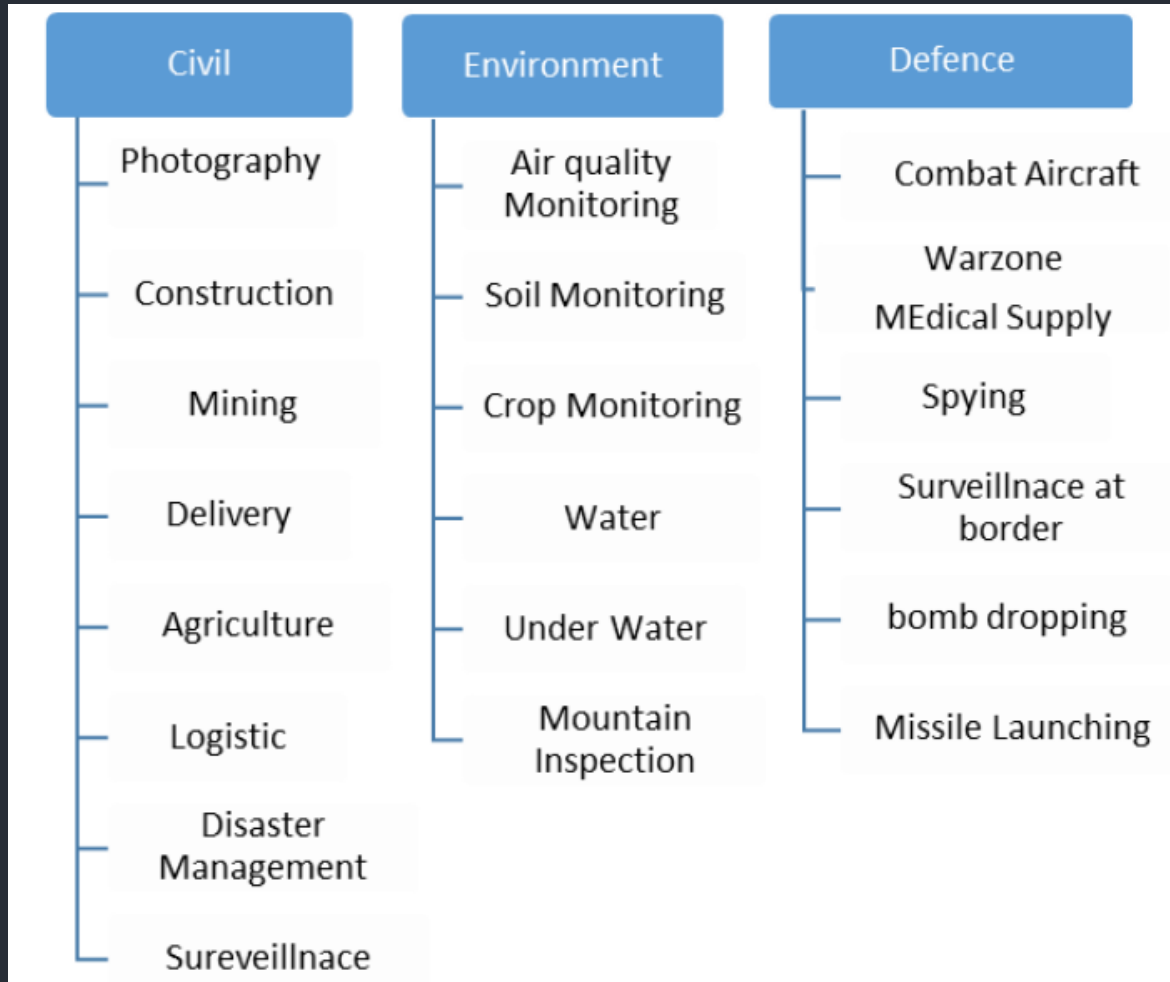


- Capturing data and images in high-risk areas (Fukushima)
- Creation of maps through images (*mappings*) for search, safety, rescue operations, etc.



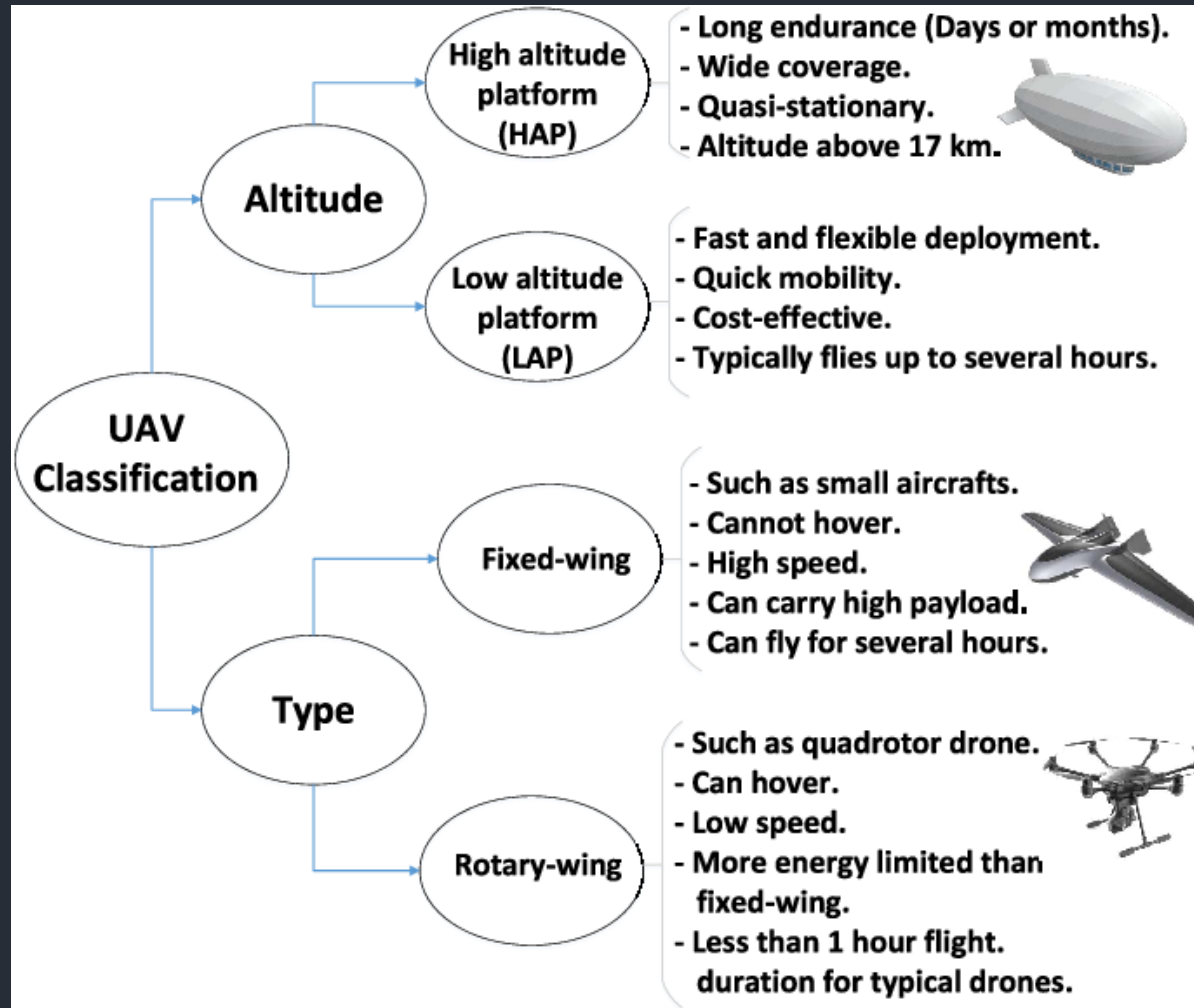
Applications

Application Domains (Scientific Research, Civilian and Military missions)



Use of single UAV systems in SAR operations, mountain avalanche events

UAV Classification



Micro/Mini UAVs

Main Characteristics

- Simple structure
- High maneuverability
- 6 Degrees of Freedom (DoF) (4 controlled inputs)
- Components:
 - Frame
 - Propellers
 - Flight controller (+ *Inertial Measurement Unit (IMU) sensors*)
 - Motor control
 - RF Link
 - GPS
 - Advanced sensor (*Proximity/Optical/Acoustic/...*)



Computer Vision: UAV Advantages

- Very high spatial resolution (up to 1cm/px)
- 3D modeling
- Not affected by cloud coverage
- High temporal resolution
- Technology
- Costs
- Innovations



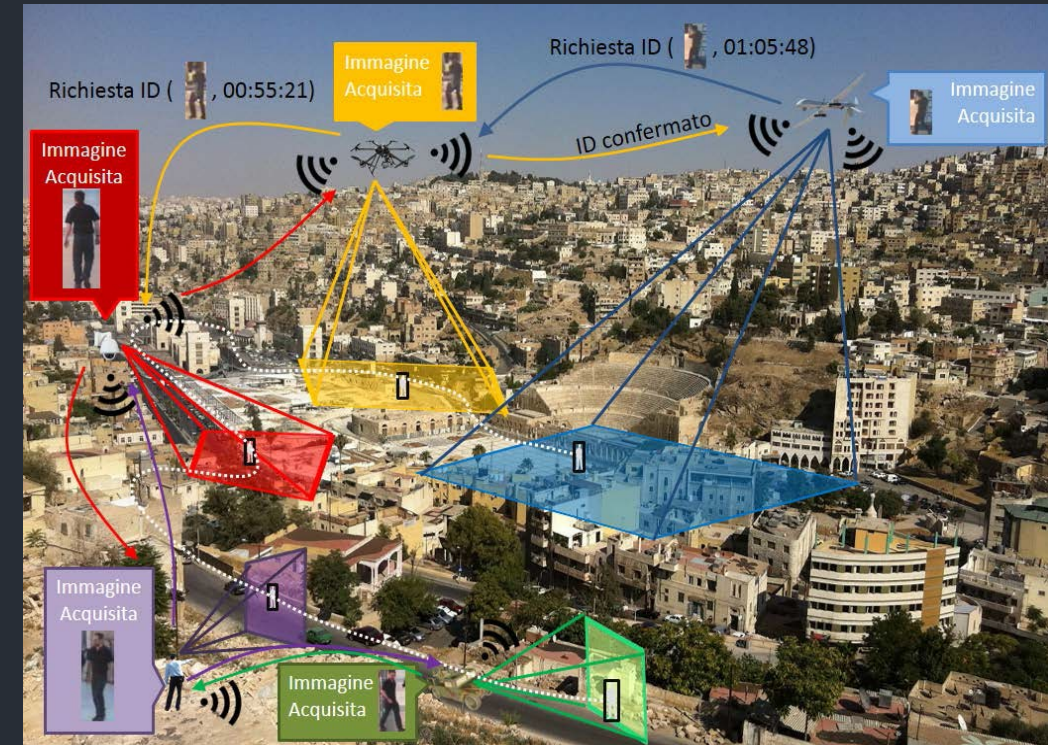
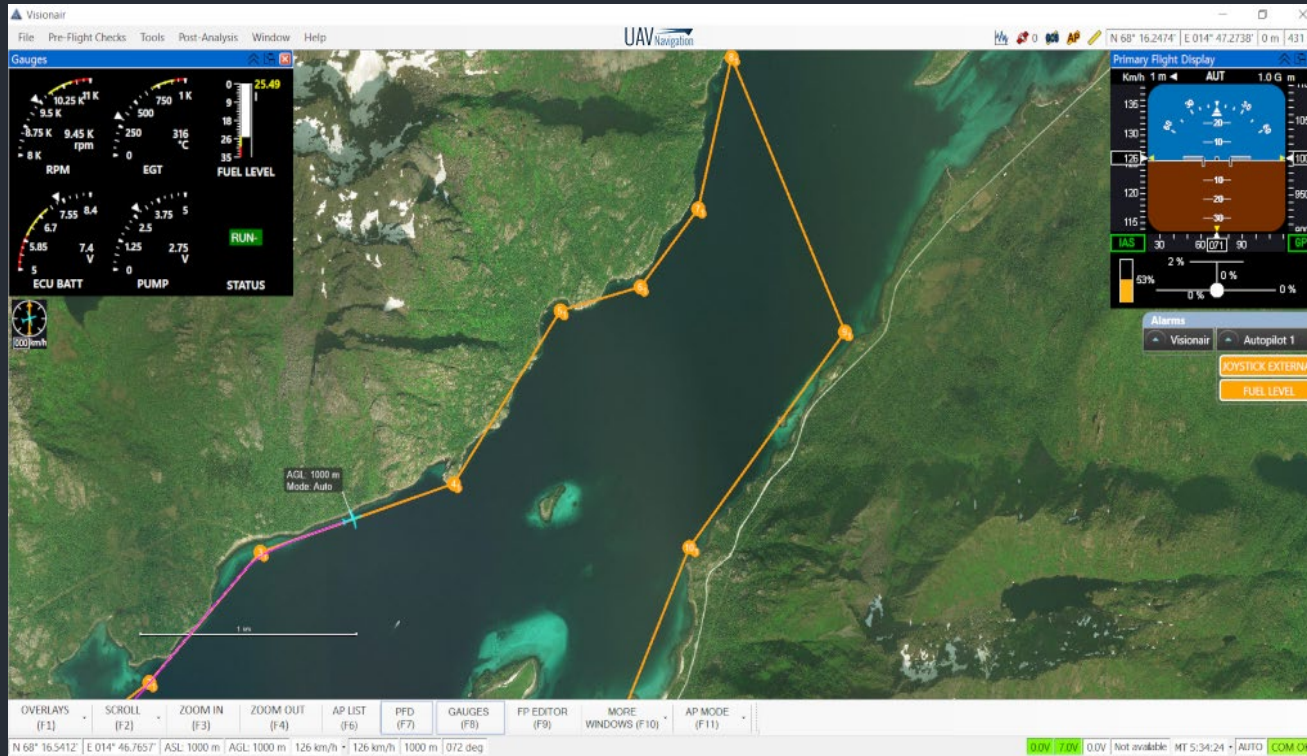
Operational Scenarios

Civil/Military

AViReS
research laboratory



- Situational awareness through sensor networks able of capturing and processing data from the area of operations



Objectives

Technology/Research trends

AViReS
research laboratory



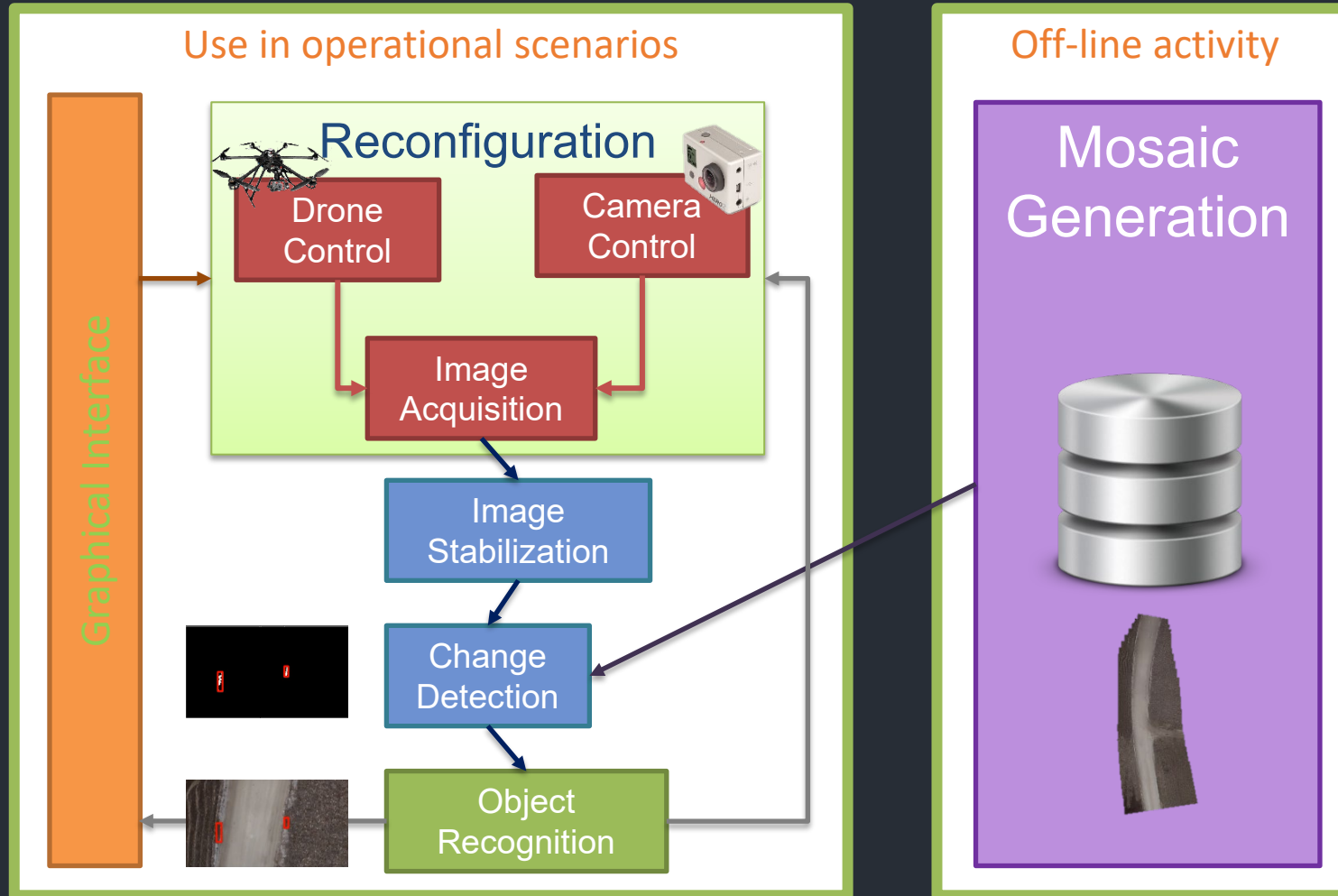
- *Real-time object detection, recognition, and identification*
- *Automatic reconfiguration of UAV parameters and sensors*



UAV ARCHITECTURE

Logical Architecture

SW



Unmanned Aerial Vehicles (UAVs)

Mosaicking

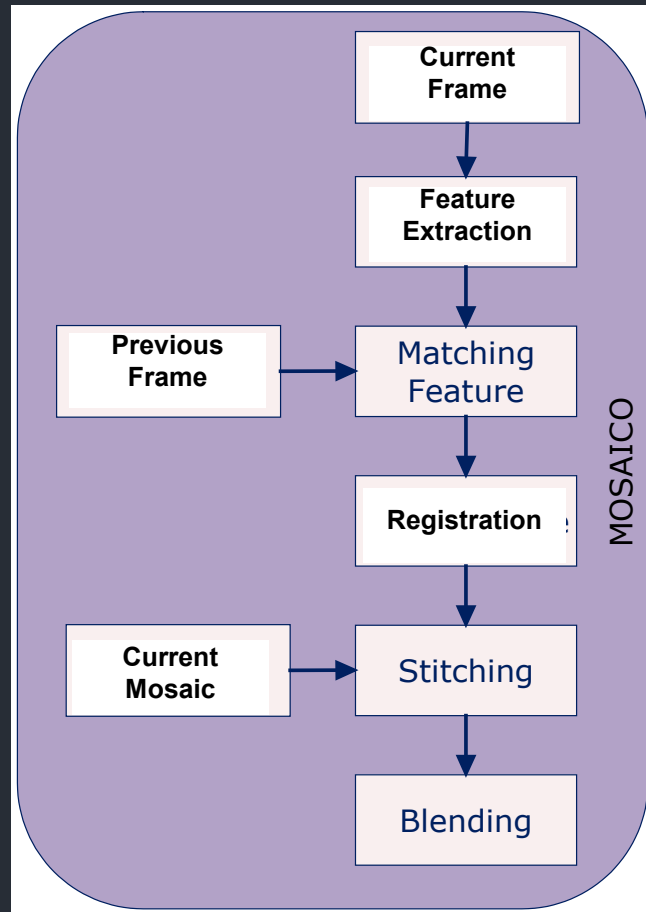
AViReS
research laboratory



MOSAICKING

Mosaic Generation

Off-Line Activity



Mosaic



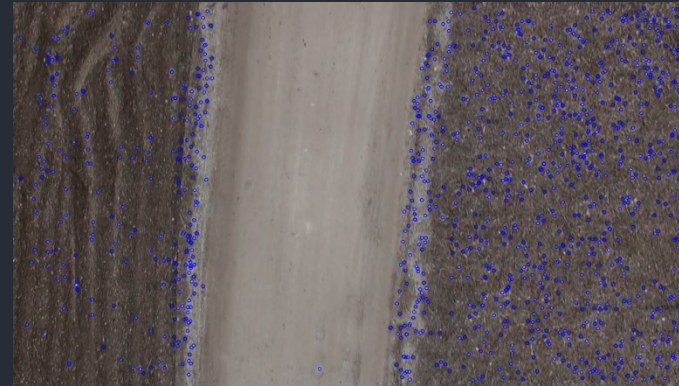
Mosaic Generation (2)

Off-Line Activity - Feature Extraction

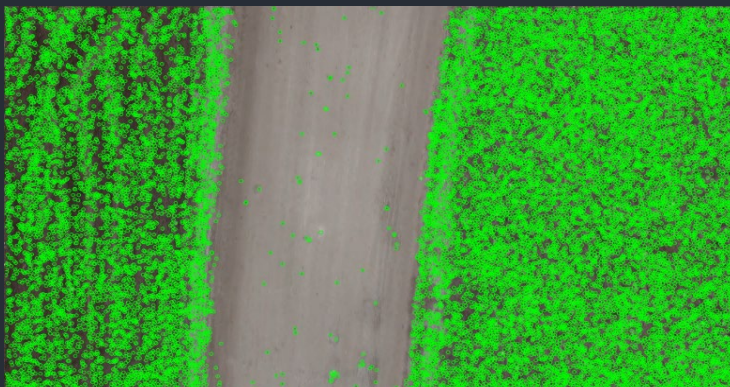
SIFT $G(x, y, k\sigma) - G(x, y, \sigma) \approx (k - 1)\sigma^2 \nabla^2 G$



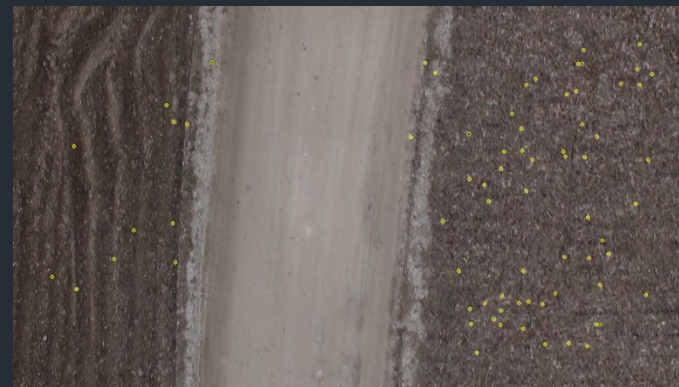
SURF $H(x, \sigma) = \begin{bmatrix} L_{xx}(x, \sigma) & L_{xy}(x, \sigma) \\ L_{xy}(x, \sigma) & L_{yy}(x, \sigma) \end{bmatrix}$



FAST $C_k = \frac{g_{xx}g_y^2 + g_{yy}g_x^2 - 2g_{xy}g_xg_y}{g_x^2g_y^2}$

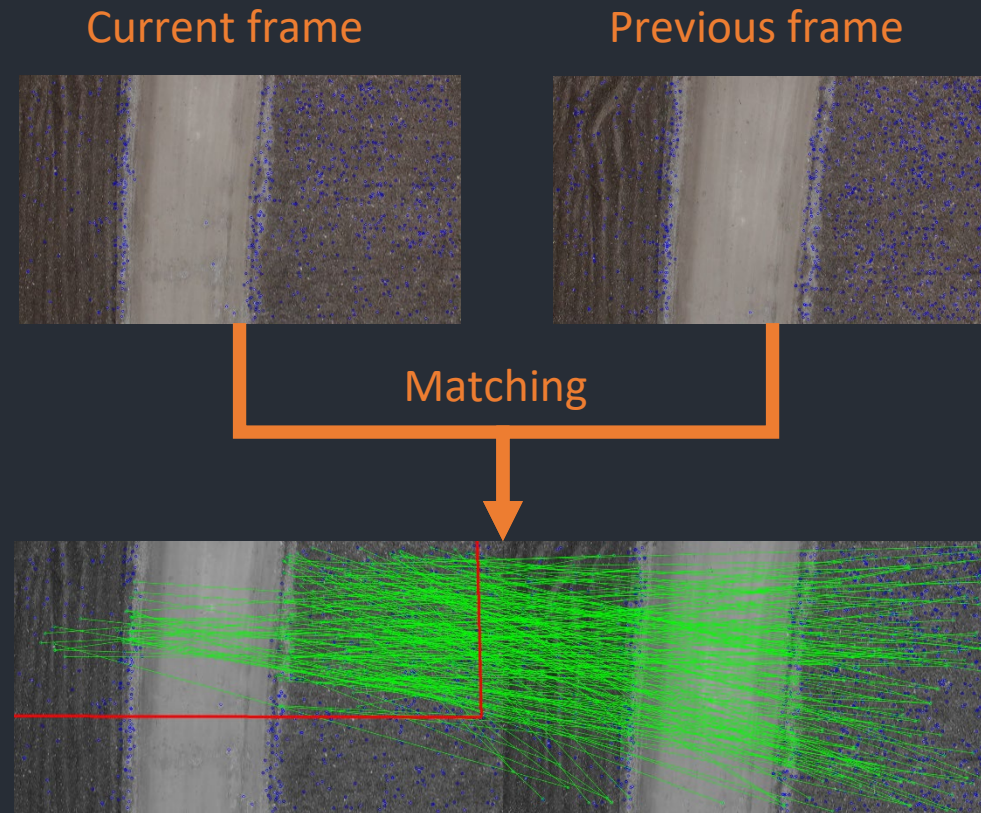


STAR $I_\alpha(x, y) = \sum_{y'=0}^y \sum_{x'=0}^{x+\alpha(y-y')} N(x', y')$



Mosaic Generation (3)

Off-Line Activity - Feature Matching



- Iterative process
- It is assumed, given the distance of the UAV from the ground, that all points lie on a plane
- The transformation between the current and previous frame can be assumed as a linear homography

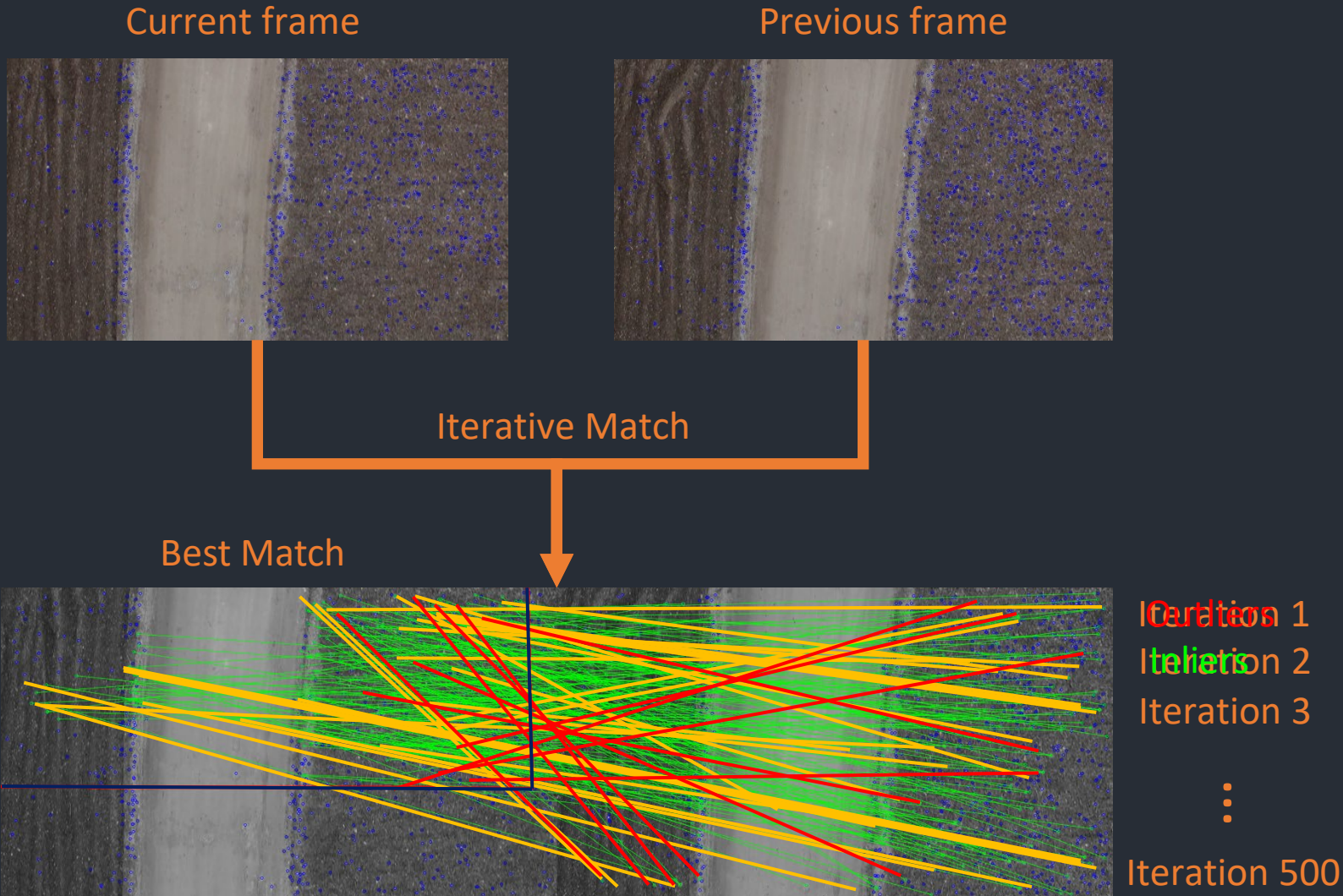
Mosaic Generation (4)

Off-Line Activity – Algorithm for Inliers Identification

- Match computation
- Repeat ($n=0$)
 - Select r matches ($r = 4$)
 - Using the selected features for computing the omography matrix H
 - Compute inliers (features correctly matched by H)
 - $n=n+1$
- until $n = 500$
- Select the omography matrix H that maximize the number of inliers

Mosaic Generation (5)

Off-Line Activity - Inliers Identification (Running Example)



Mosaic Generation (6)

Off-Line Activity – Iterative Match - Example



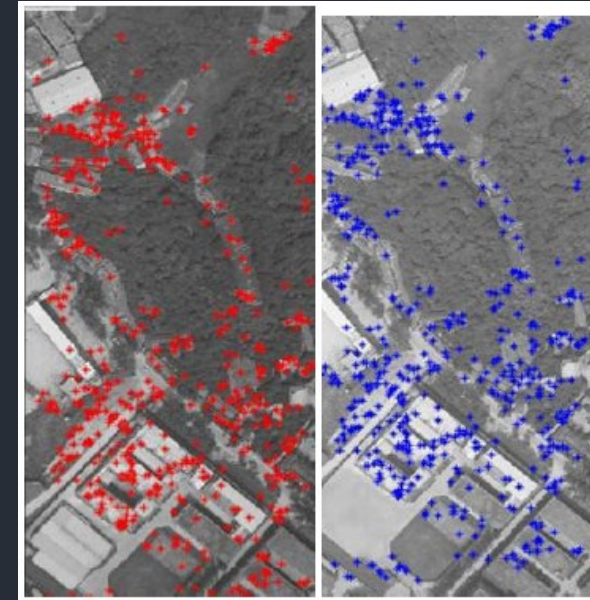
Iterative Match



Mosaic Generation (7)

Running example - Matching

- Feature extraction (current frame)
- Matching feature (previous frame)
- Registration
- Stitching
- Blending



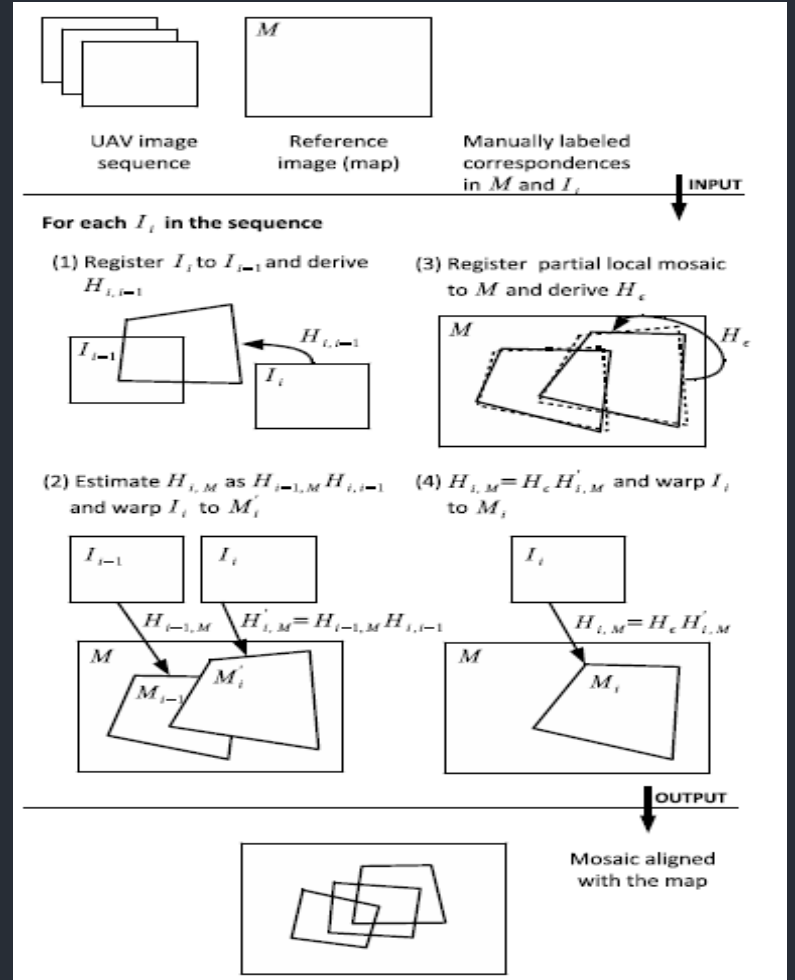
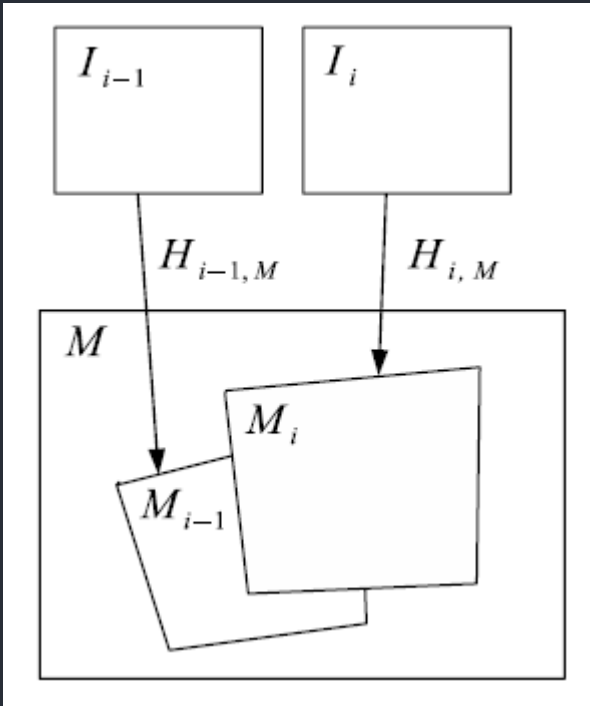
Mosaic Generation (8)

Running example - Registration

- Feature extraction (current frame)
- Matching feature (previous frame)
- **Registration**
- Stitching
- Blending

$$H_{i,j}I_i = I_j$$

$$H_{i,M}I_i = M_i$$



Mosaic Generation (9)

Running example - Stitching

- Feature extraction (current frame)
- Matching feature (previous frame)
- Registration
- **Stitching** →
- Blending



Mosaic Generation (10)

Running example - Blending

- Feature extraction (current frame)
- Matching feature (previous frame)
- Registration
- Stitching
- Blending



Mosaic Generation in Real-Time

Using Image Features + GPS data

AViReS
research laboratory



Image feature



GPS + Image feature

Mosaic Generation in Real-Time (2)

Running example in real-time

AViReS
research laboratory



Unmanned Aerial Vehicles (UAVs)

Change Detection

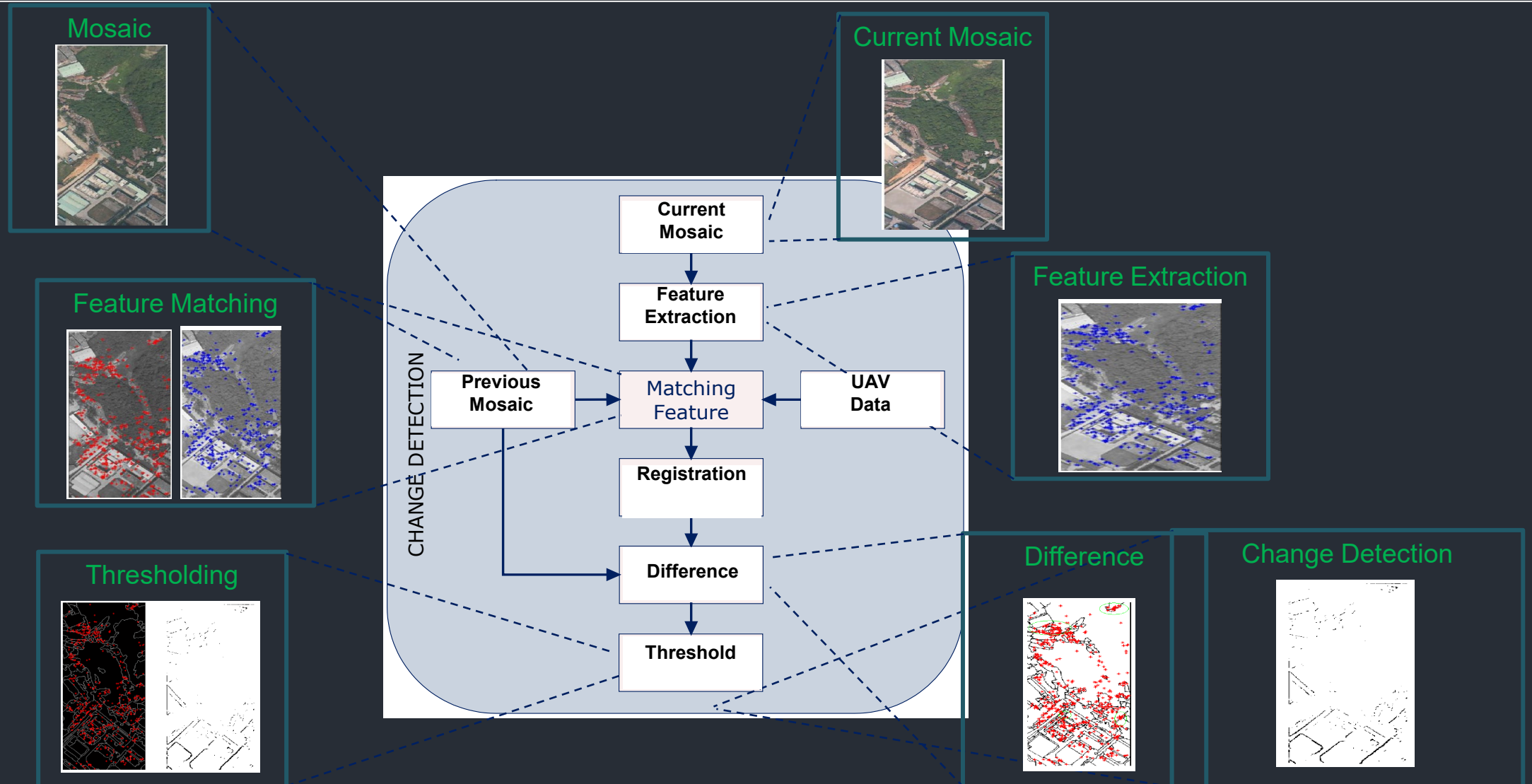
AViReS
research laboratory



CHANGE DETECTION

Change Detection

Logical Architecture

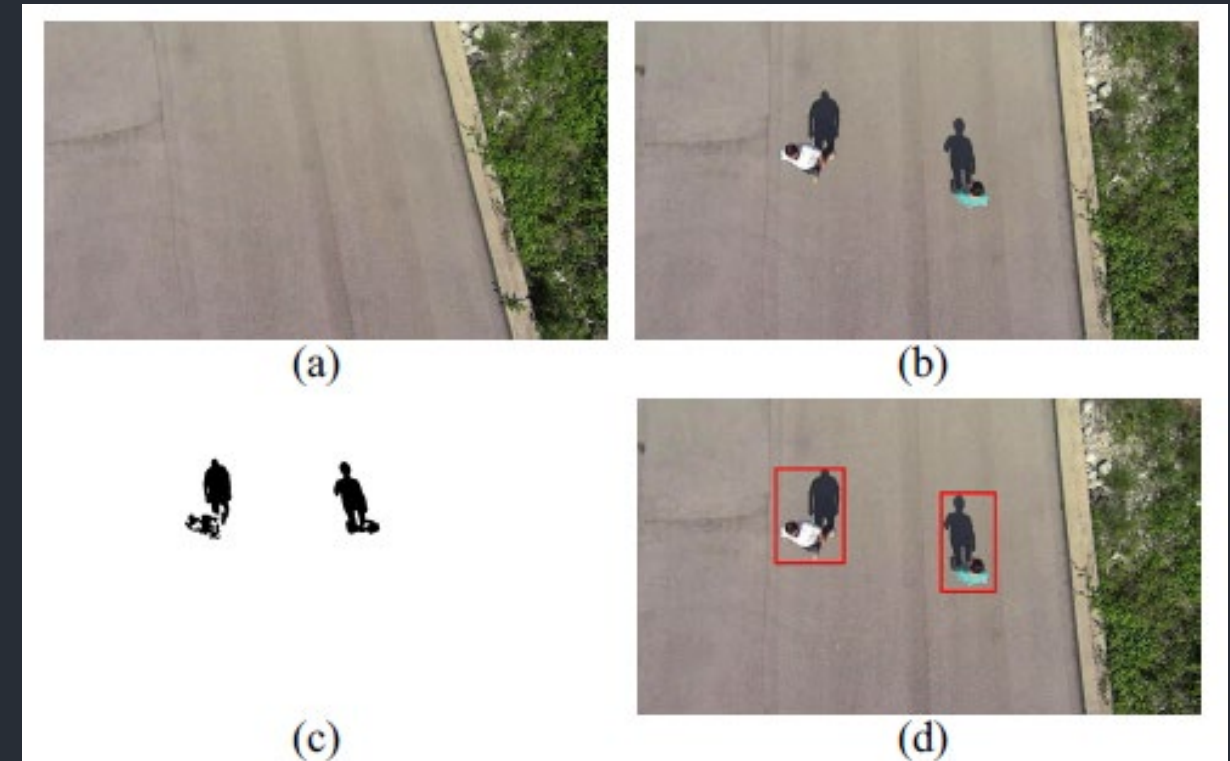
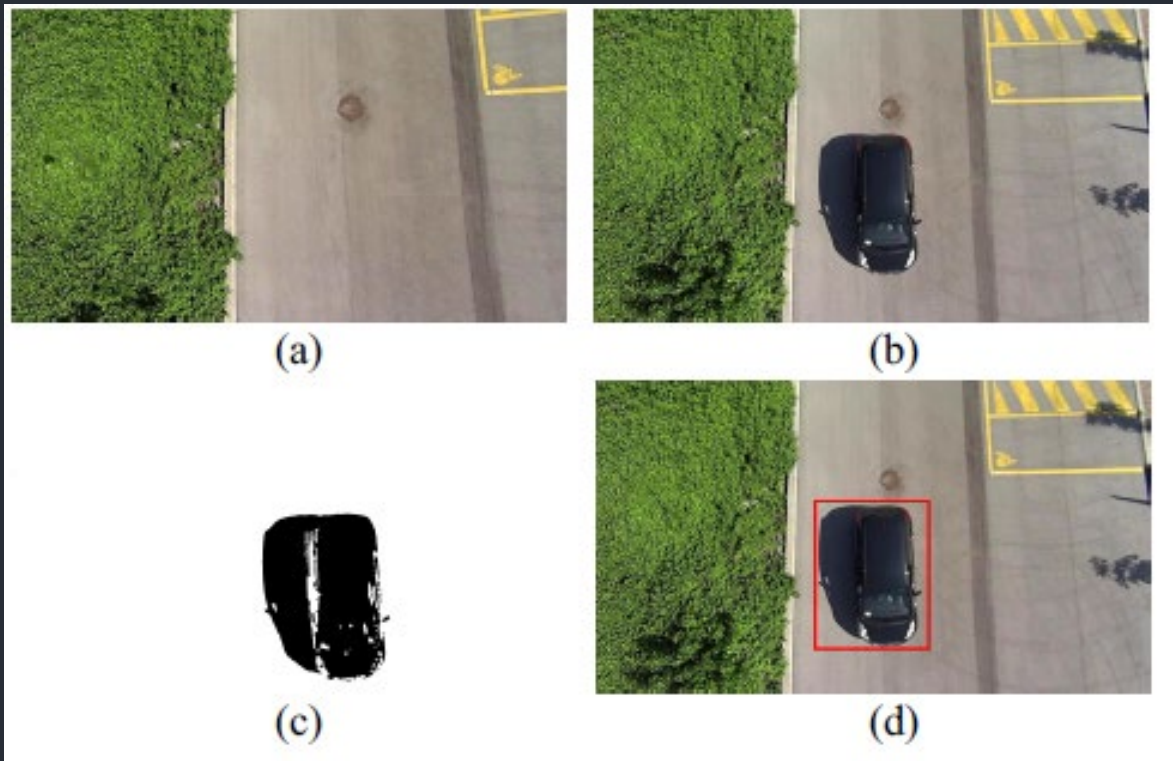


Change Detection + GPS



Change Detection

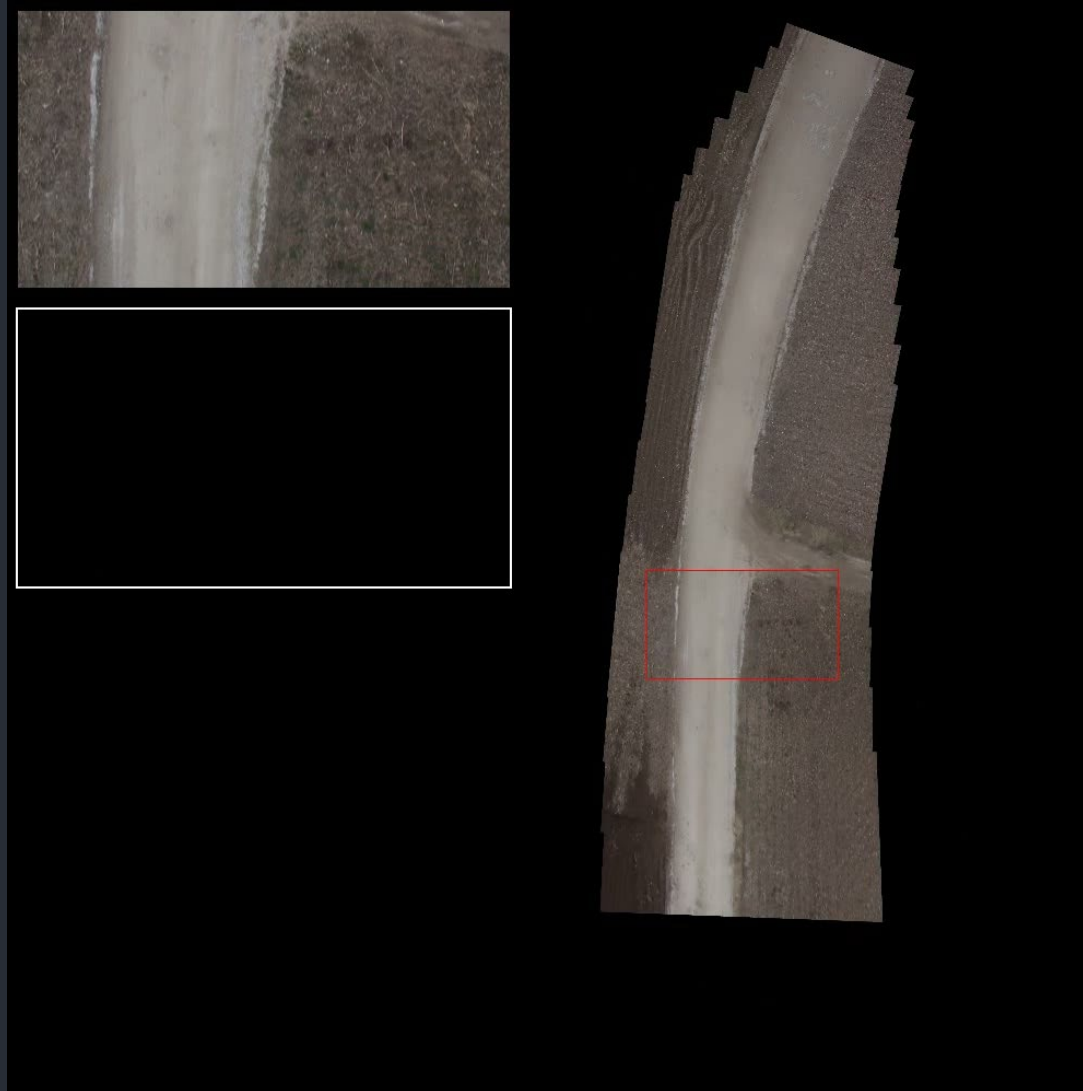
Change Detection and Object Detection



Change Detection (5)

Running example in real-time

AViReS
research laboratory



Unmanned Aerial Vehicles (UAVs)

Reconfiguration

AViReS
research laboratory



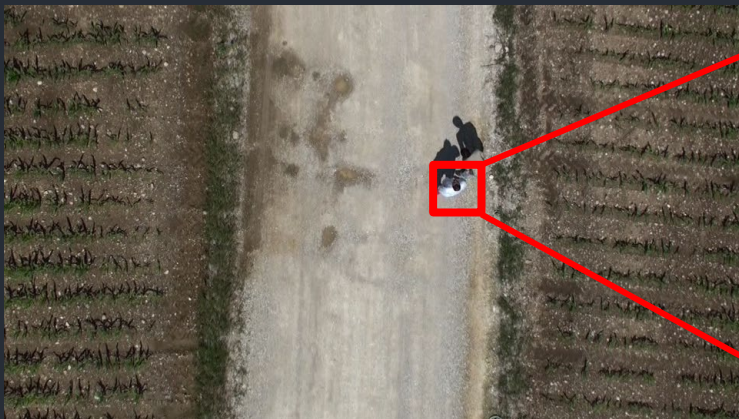
UAV RECONFIGURATION

Image Resolution

Image and Scene Dimensions



- The size of a digital image is given by the number of pixels (*height x width*)
- The size of the observed scene is its real measure in meters



- The ratio (image size / observed scene size) is called **spatial resolution**, and is measured in **pixel/meter**

Image Resolution (2)

Different spatial resolutions



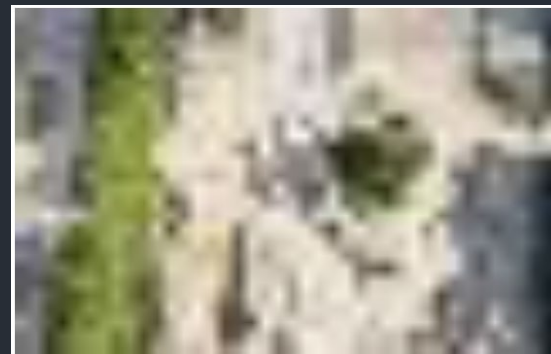
264x176



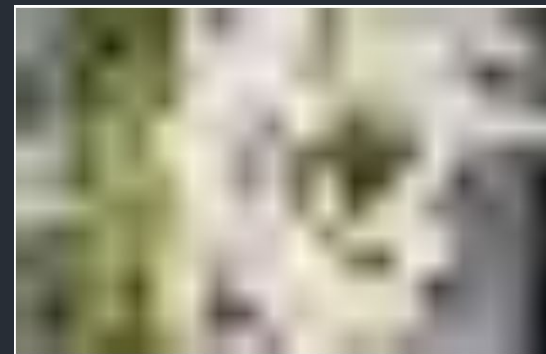
132x88



76x44



38x22



19x11

Image Resolution (3)

Different spatial resolutions – Object Type



What is this object?

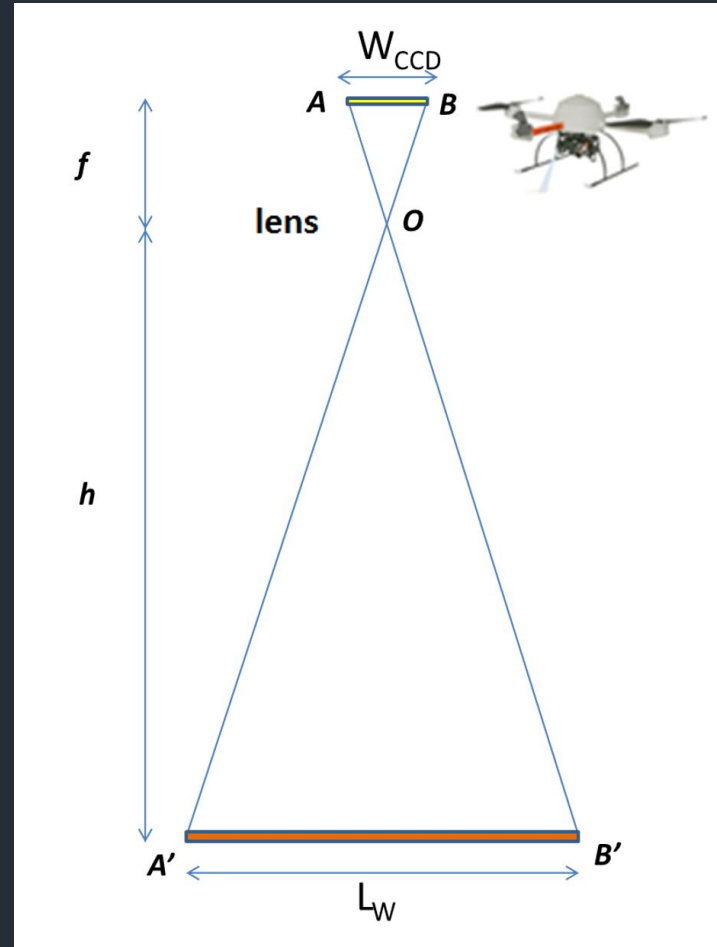
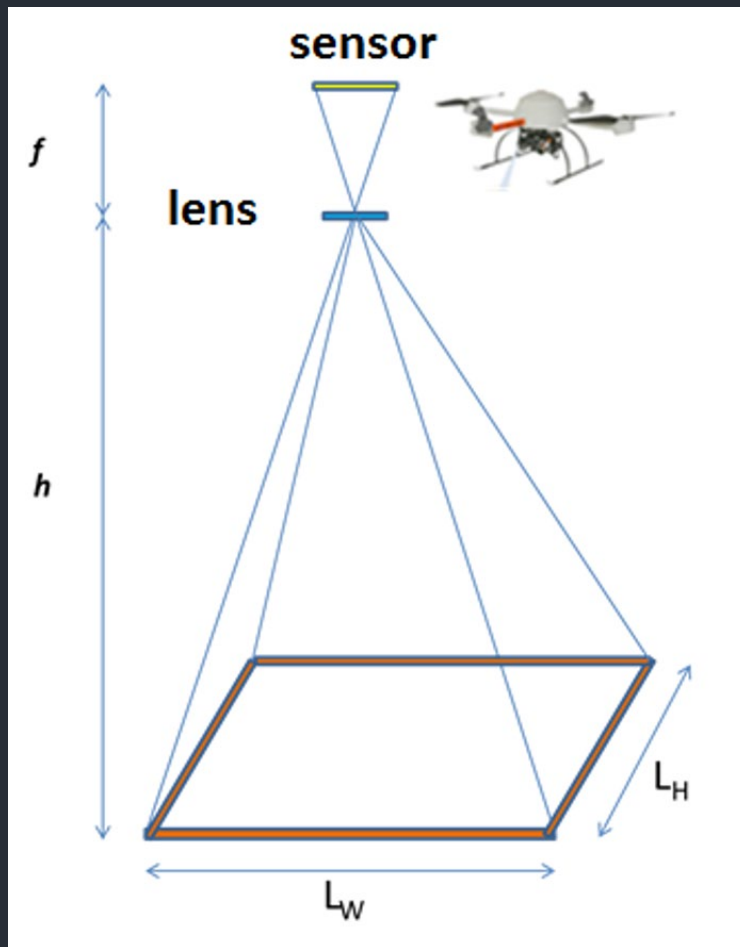


At a higher spatial resolution, you can recognize a metal briefcase



Image Resolution (4)

UAV flight parameters



Intrinsic Parameters

f : focal length

W, H : image size [pixels]

W_{CCD}, H_{CCD} : CCD size [mm]

Extrinsic Parameters

h : flight height

L_W, L_H : size of the observed scene [m]

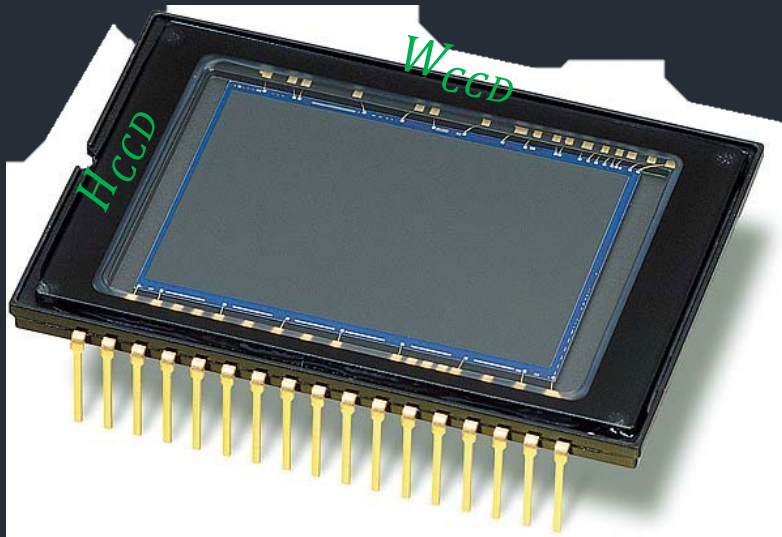
$$L_W = \frac{h W_{CCD}}{f}$$

$$L_H = \frac{h H_{CCD}}{f}$$

Image Resolution (5)

Pixel size

- The size k of the pixel is defined as the ratio between the real size of the CCD [mm] and the number of pixels of the image ($W \times H$)



$$k_w = \frac{W_{CCD}}{W}$$

$$k_H = \frac{H_{CCD}}{H}$$

Normally (cameras with square pixels), you have that:

$$k_w = k_H = k$$

Image Resolution (6)

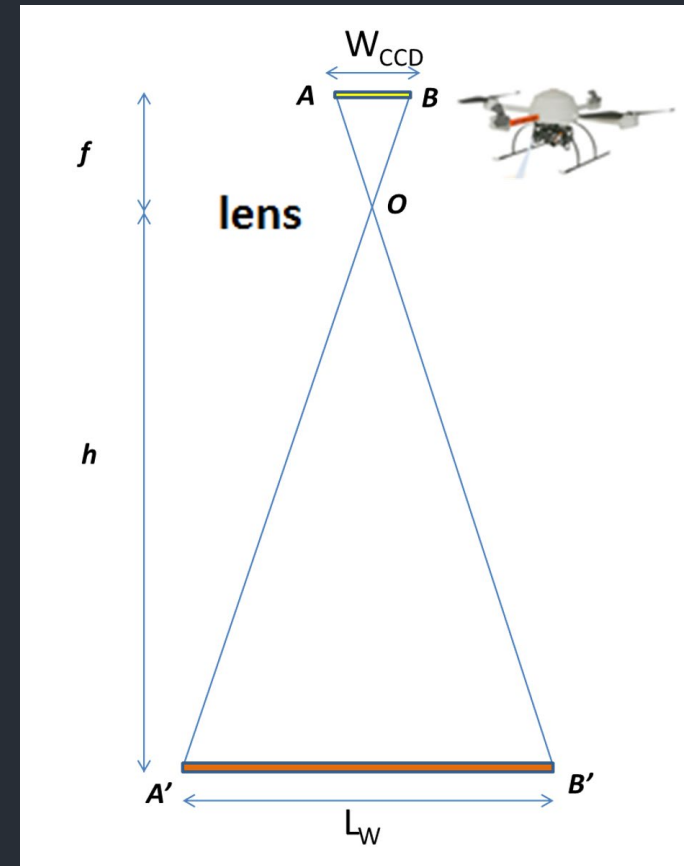
Spatial Resolution

$$\frac{W}{L_W} \text{ [pixel/mm]} \quad \frac{H}{L_H} \text{ [pixel/mm]}$$

■ It follows that: $L_W = \frac{h W_{CCD}}{f}$ $L_H = \frac{h H_{CCD}}{f}$

$$\frac{W}{L_W} = \frac{f W}{h W_{CCD}} = \frac{f W}{h k_W W} = \frac{f}{h k_W}$$

$$\frac{H}{L_H} = \frac{f H}{h H_{CCD}} = \frac{f H}{h k_H H} = \frac{f}{h k_H}$$



where f , k_H and k_W are constant depending on the sensor

The main image processing tasks are:

- **Detection:** Detect that something is present in the scene
- **Recognition:** classify the detected object (e.g. to distinguish if it is a person or a vehicle)
- **Identification:** Identify the specific instance of the object (e.g. recognize the specific type of a vehicle or discriminate between enemy and friends vehicles)

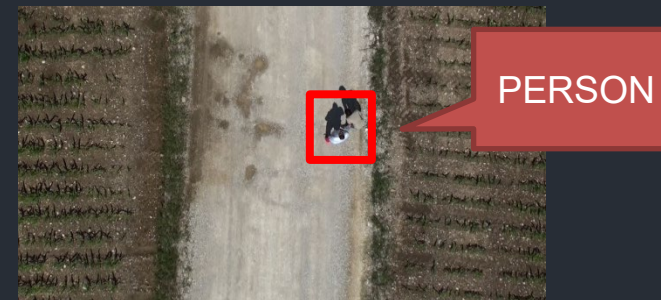
Image Processing Tasks (2)

Task vs Spatial Resolution

Detection: Detect that something is present in the scene



Recognition: Classify the detected object (e.g. to distinguish if it is a person or a vehicle)



Identification: Identify the specific instance of the object (e.g., recognize the vehicle's type or classify a friend/enemy vehicle)

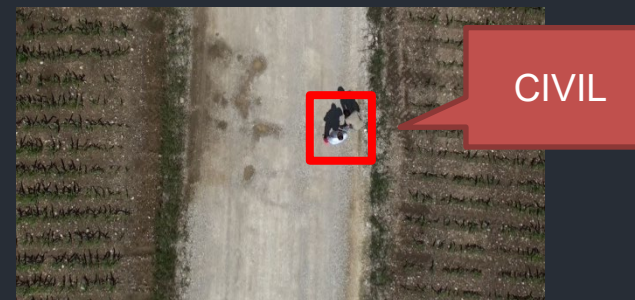


Image Processing Tasks (3)

Task vs Spatial Resolution (Johnson's criteria)

- The problem was initially studied by J.B. Johnson
- Johnson has defined some criteria that are accepted for computing the minimum resolution required to perform a given task.
- The critical dimension D of an object is defined as:

$$D = \sqrt{W \cdot H}$$

- Example - The critical size of a person with height $H=1.80\text{m}$ and wide $W=0.50\text{m}$ is

$$D = \sqrt{1,80 \cdot 0,50} = 0,9 \text{ m}$$

Image Processing Tasks (4)

Task vs Spatial Resolution – Critical Dimensions

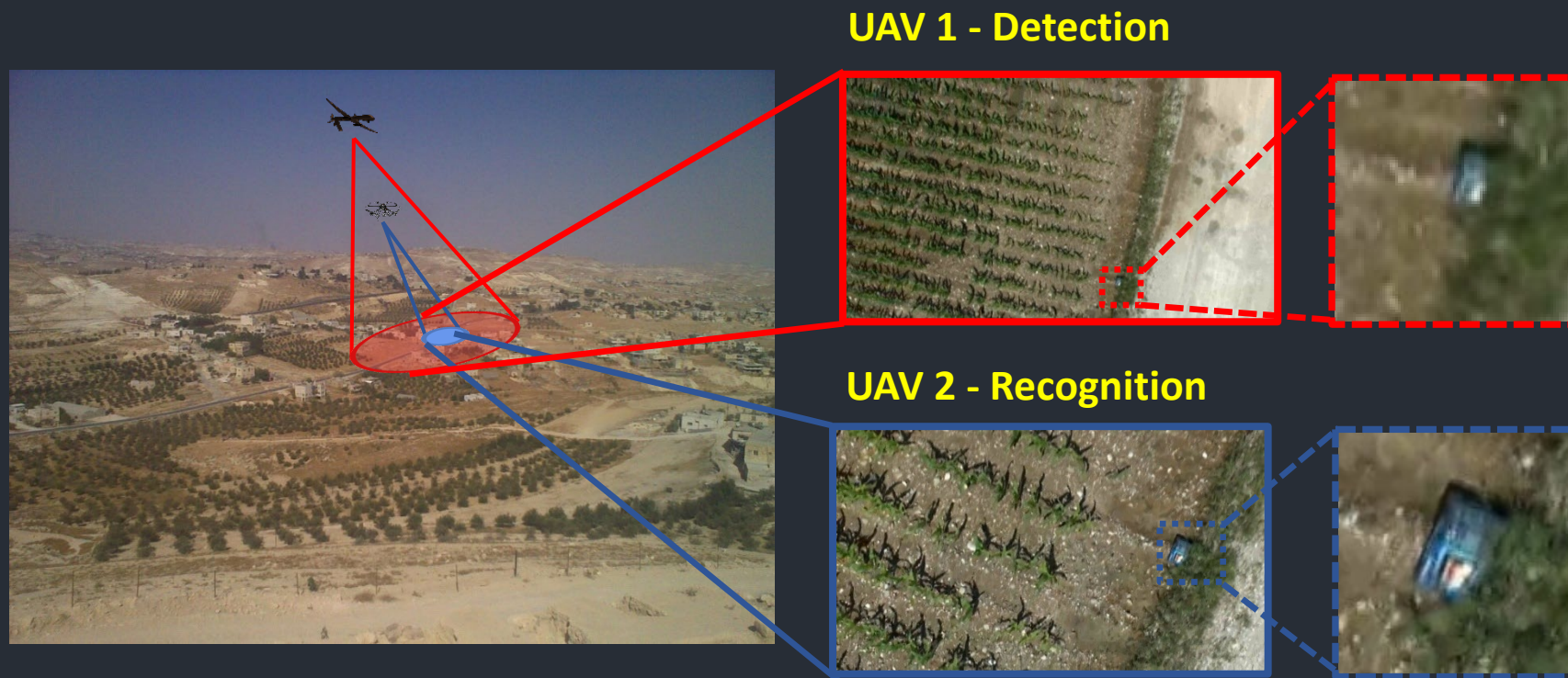
- Based on the critical dimension, Johnson defines the minimum resolution to perform the three main tasks of image processing:
 - **Detection** : $1.5 / D$ [pixel/m]
 - **Recognition** : $7 / D$ [pixel/m]
 - **Identification** : $14 / D$ [pixel/m]
- At this resolution, the probability of successfully performing the required task is about 50%.

FUTURE TREND and CONCLUSIONS

Future Trends - Assigning different tasks

Adaptive Sensing

- Ability to assign different tasks to different UAVs:



Future Trends – Process Parallelization

- Off-Line for mosaic construction
- Real-Time for scene analysis



Future Trends – UAV Swarms

- UAV Swarms cooperating to monitor a given areas



Thanks for your attention