



Gesture based communication with UAVs

How to communicate with UAVs when data link fails or is not available

Alexander Schelle

Institute of Flight Systems
Department of Aerospace Engineering
University of the Bundeswehr Munich
5th May 2017



© L-3 Communications



© South West Research Institute



© Institute of Flight Systems, UniBwM



Operator

Command and Control via
WiFi, LTE, LoS, SatCom, etc.



UAV



© L-3 Communications



© South West Research Institute



© Institute of Flight Systems, UniBwM



Operator



Communication via
WiFi, LTE, etc.

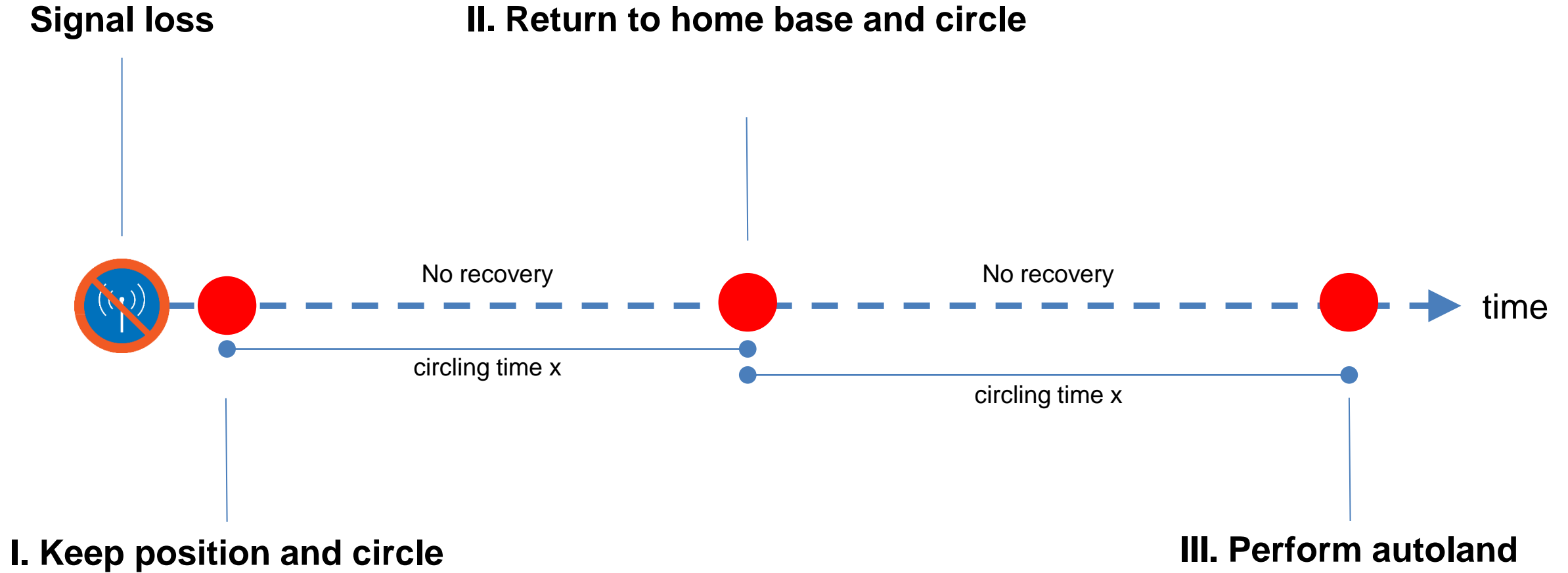


UAV





Common Behavior After Signal Loss



Photography Photogrammetry

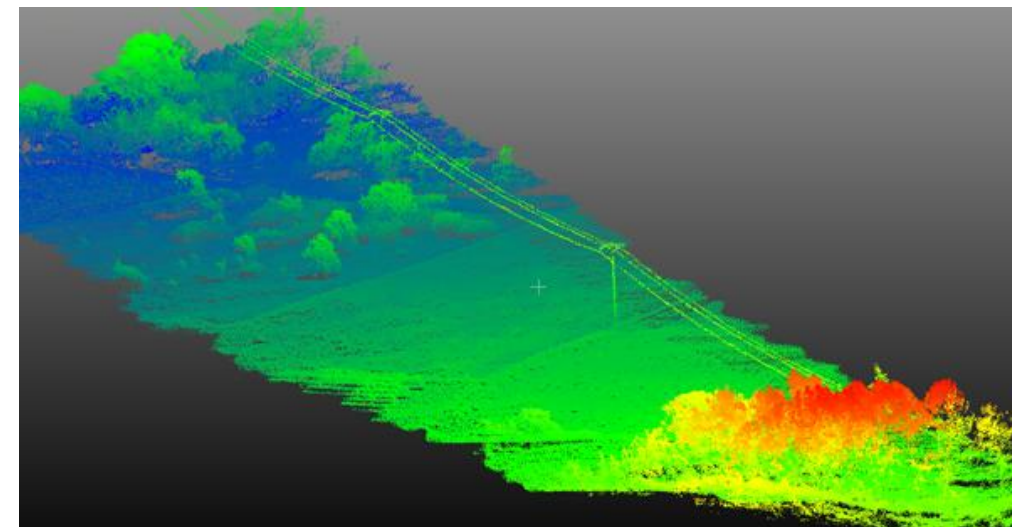
- Electro-Optical (EO)
- Light Detection and Ranging (LIDAR)



© DJI



© Phoenix Aerial Systems

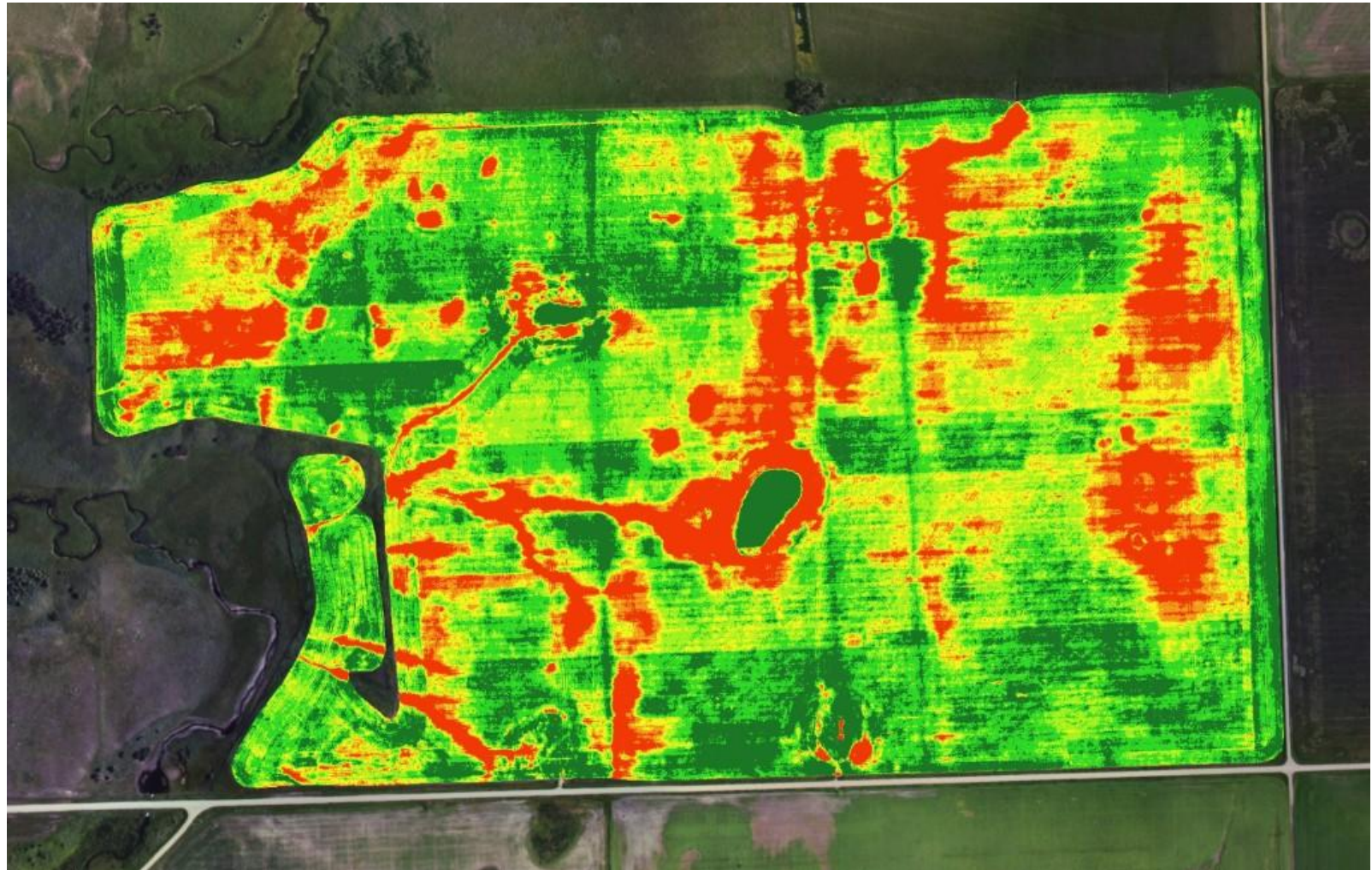


© electronicsweekly.com



Agriculture Precision Farming

- Infrared
- Multi-/Hyperspectral

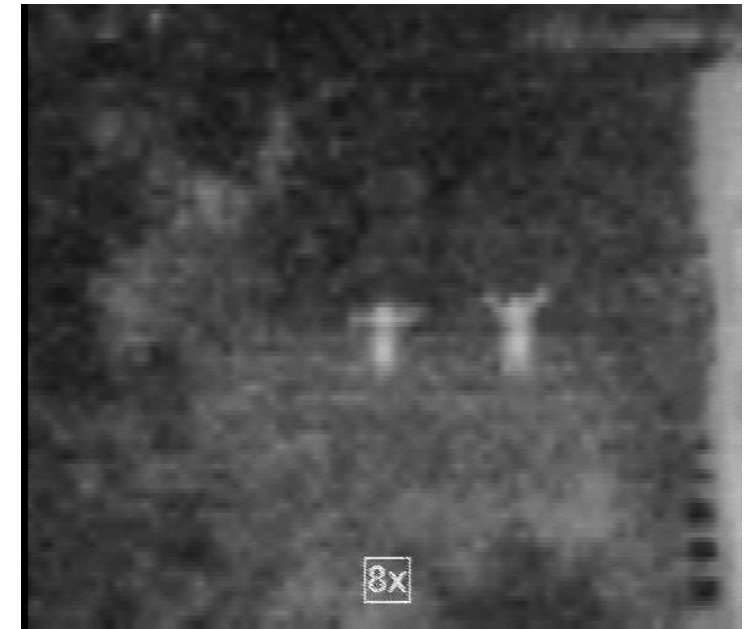


© geovantage.com



Surveillance SAR

- Thermal

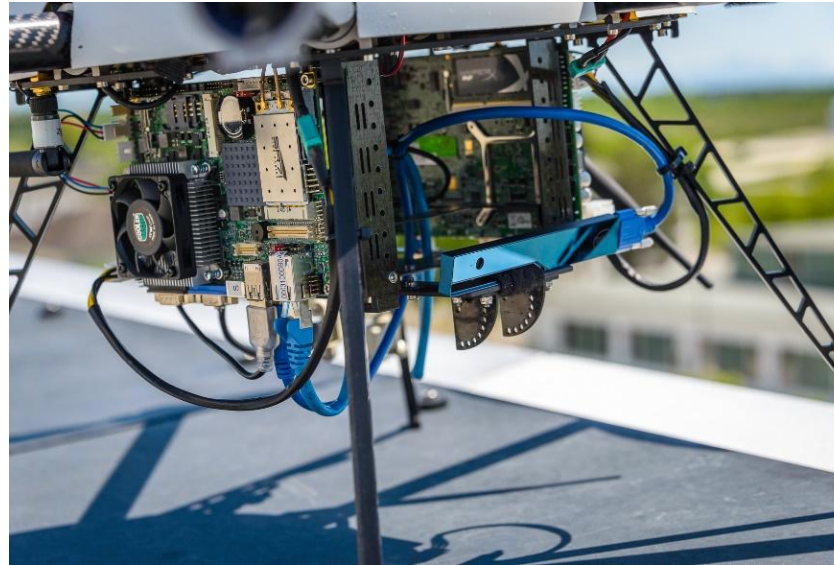


© <http://live.ece.utexas.edu>



Collision Avoidance

- Depth
- Ultrasonic



© DJI





Using available imaging sensors to establish gesture based visual communication

- **Usability**

no additional hardware on ground needed,
signal loss safe



Using available imaging sensors to establish gesture based visual communication

- **Usability**
no additional hardware on ground needed,
signal loss safe
- **Autonomous Search**
for missing people in dangerous weather conditions



© acmg.ca



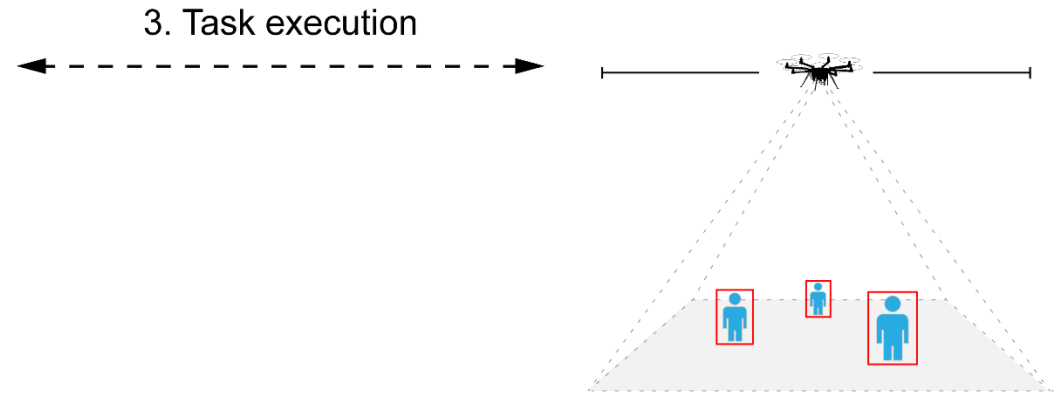
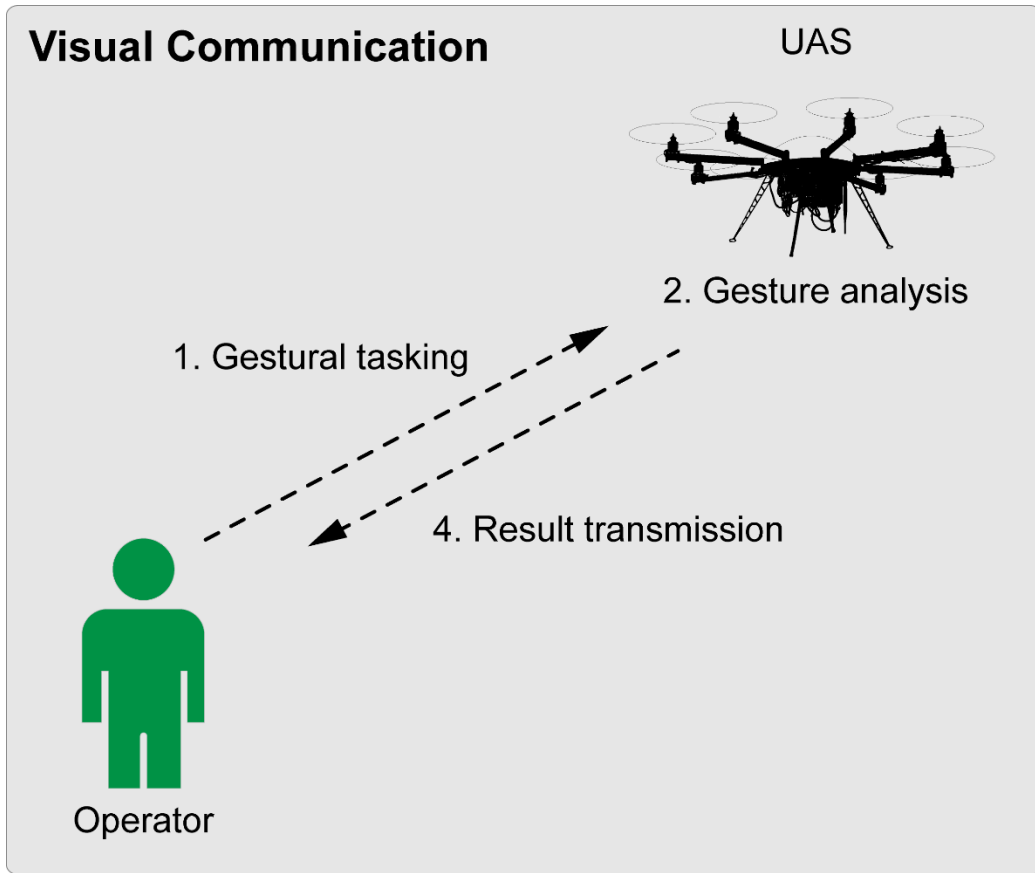


Using available imaging sensors to establish gesture based visual communication

- **Usability**
no additional hardware on ground needed,
signal loss safe
- **Autonomous Search**
for missing people in dangerous weather conditions
- **Gestural Commanding**
for authorized operators on ground in disaster
scenarios



Gestural Transmission of a Search Task





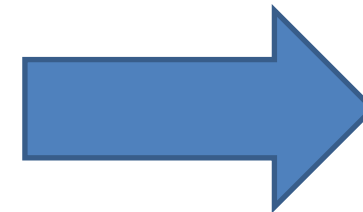
Using available imaging sensors to establish gesture based visual communication

- **Usability**
no additional hardware on ground needed,
signal loss safe
- **Autonomous Search**
for missing people in dangerous weather conditions
- **Gestural Commanding**
for authorized operators on ground in disaster
scenarios
- **Human Guidance**
Landing spot recommendations (delivery drones)



Using available imaging sensors to establish gesture based visual communication

- **Usability**
no additional hardware on ground needed,
signal loss safe
- **Autonomous Search**
for missing people in dangerous weather conditions
- **Gestural Commanding**
for authorized operators on ground in disaster
scenarios
- **Human Guidance**
Landing spot recommendations (delivery drones)



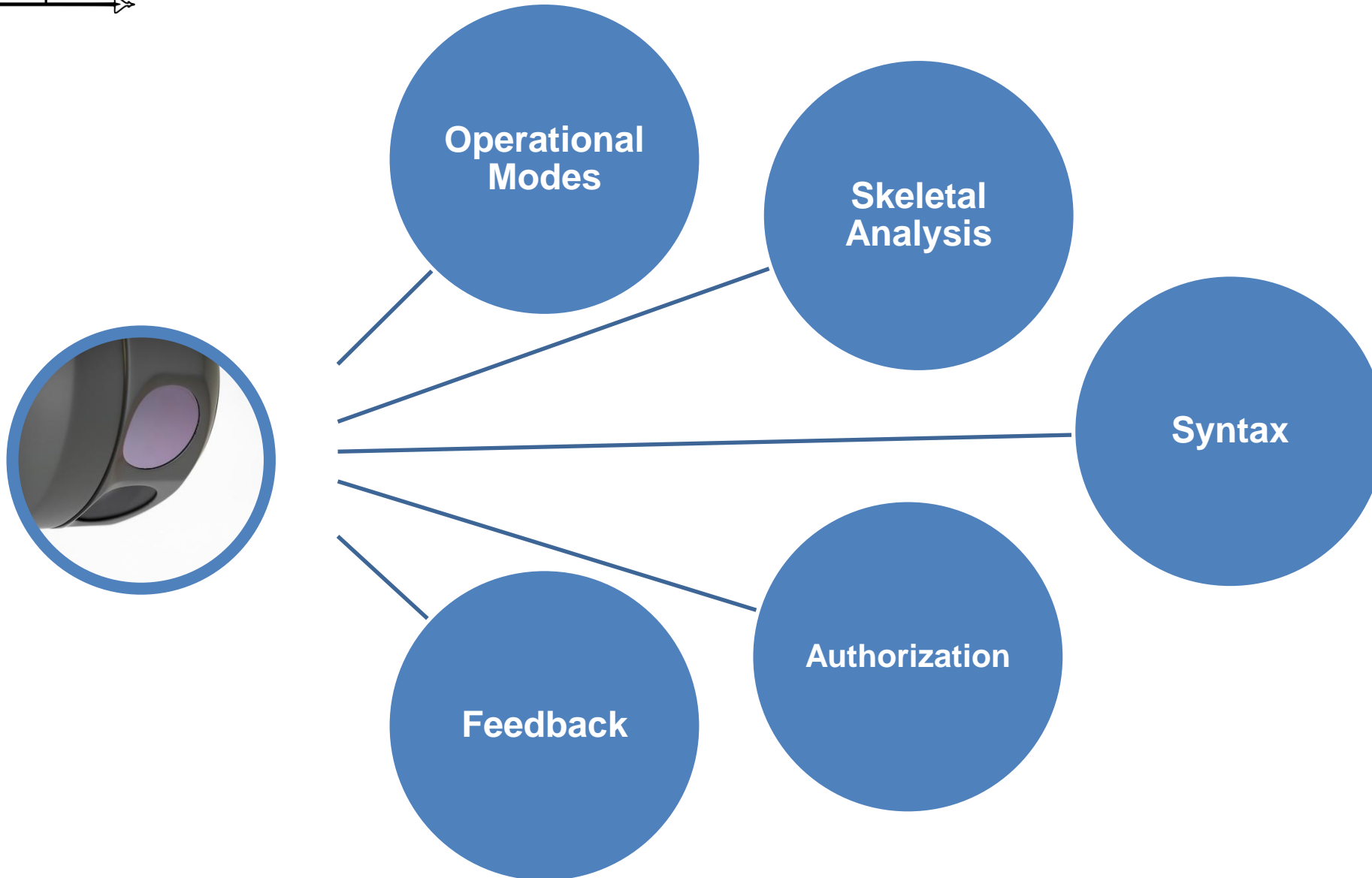
**UAV no longer
passive observer,
but active interaction
partner**





Modelling Visual Communication with UAS

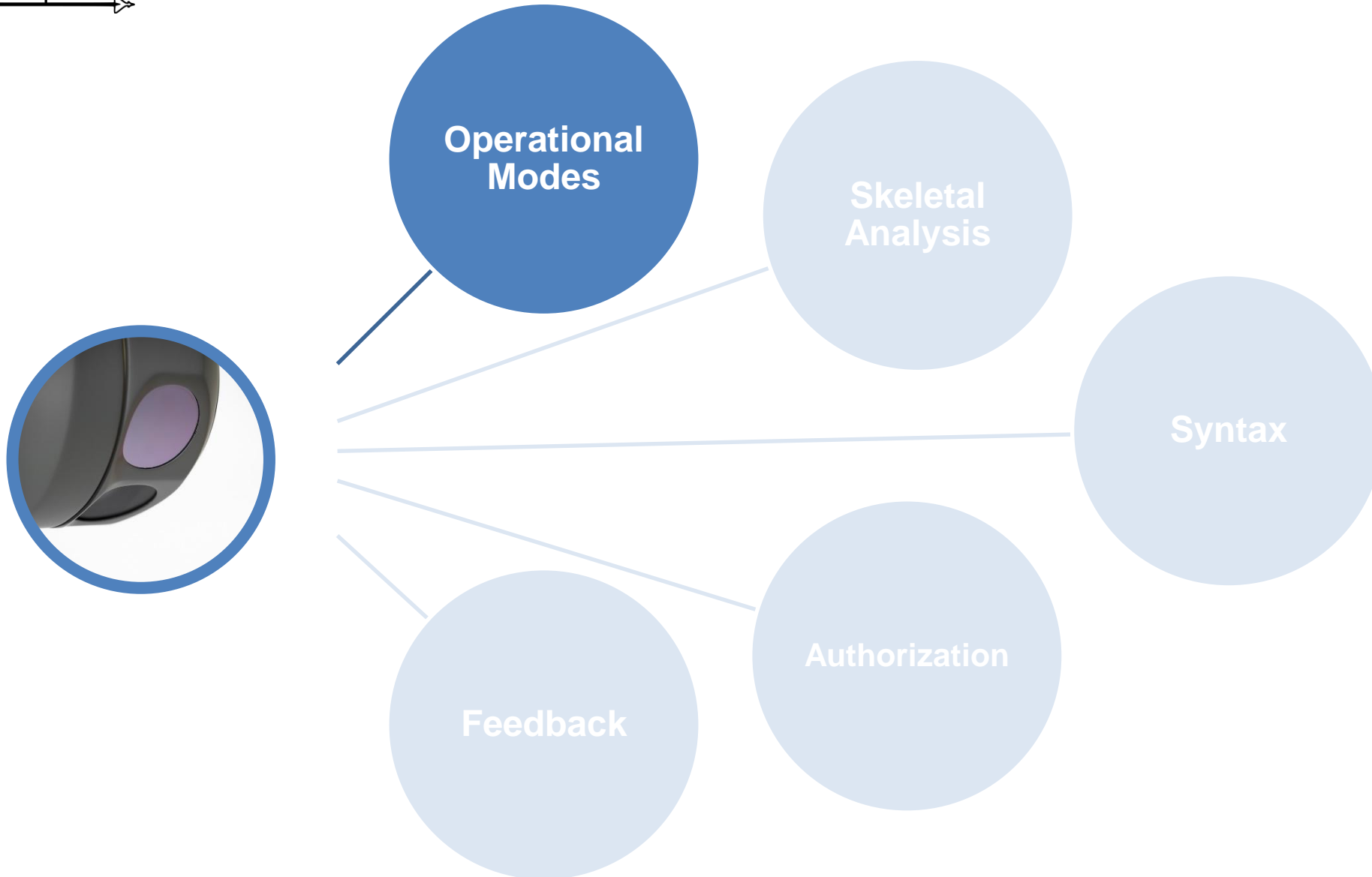
Research Topics





Modelling Visual Communication with UAS

Research Topics





- > Person detection and tracking
- > Spot potential interaction requests
(*benefit for search for missing persons*)
- > Thermal LWIR sensor
- > Low resolution operator body shape model

Detection Mode





First Experiment in Detection Mode

Utilized Sensor System

- > **2 Axis gyro-stabilized gimbal incl. GEO-lock**
DST OTUS-L170
- > **IR camera FLIR TAU2 640**
640x480 @ 25 Hz, HFOV 69° (9mm lens)
- > **GigE** interface for video
- > **RS485** interface for control



> Test setup in detection mode

- Cloudy day
- Height ~30 m
- Distance ~105 m
- Offline processing

f=9mm



2x digital zoom

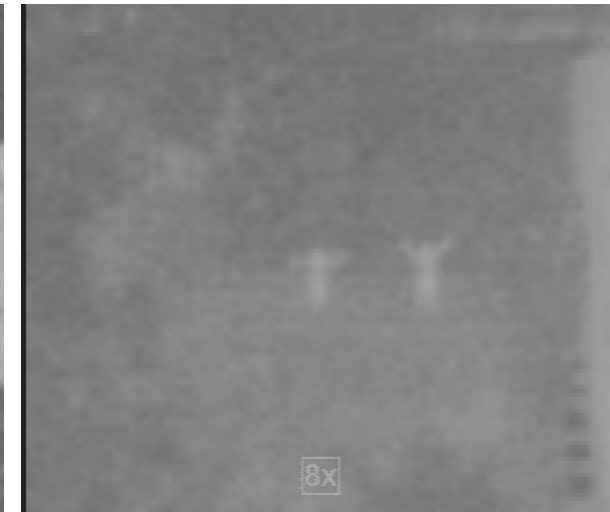


105m

30m



4x digital zoom



8x digital zoom



Person and Waving Gesture Detection with LWIR sensor





- > Operator detection and tracking
- > Spot potential interaction requests
(*benefit for search for missing persons*)
- > Thermal LWIR sensor
- > Low resolution operator body shape model

Detection Mode



Interaction Mode



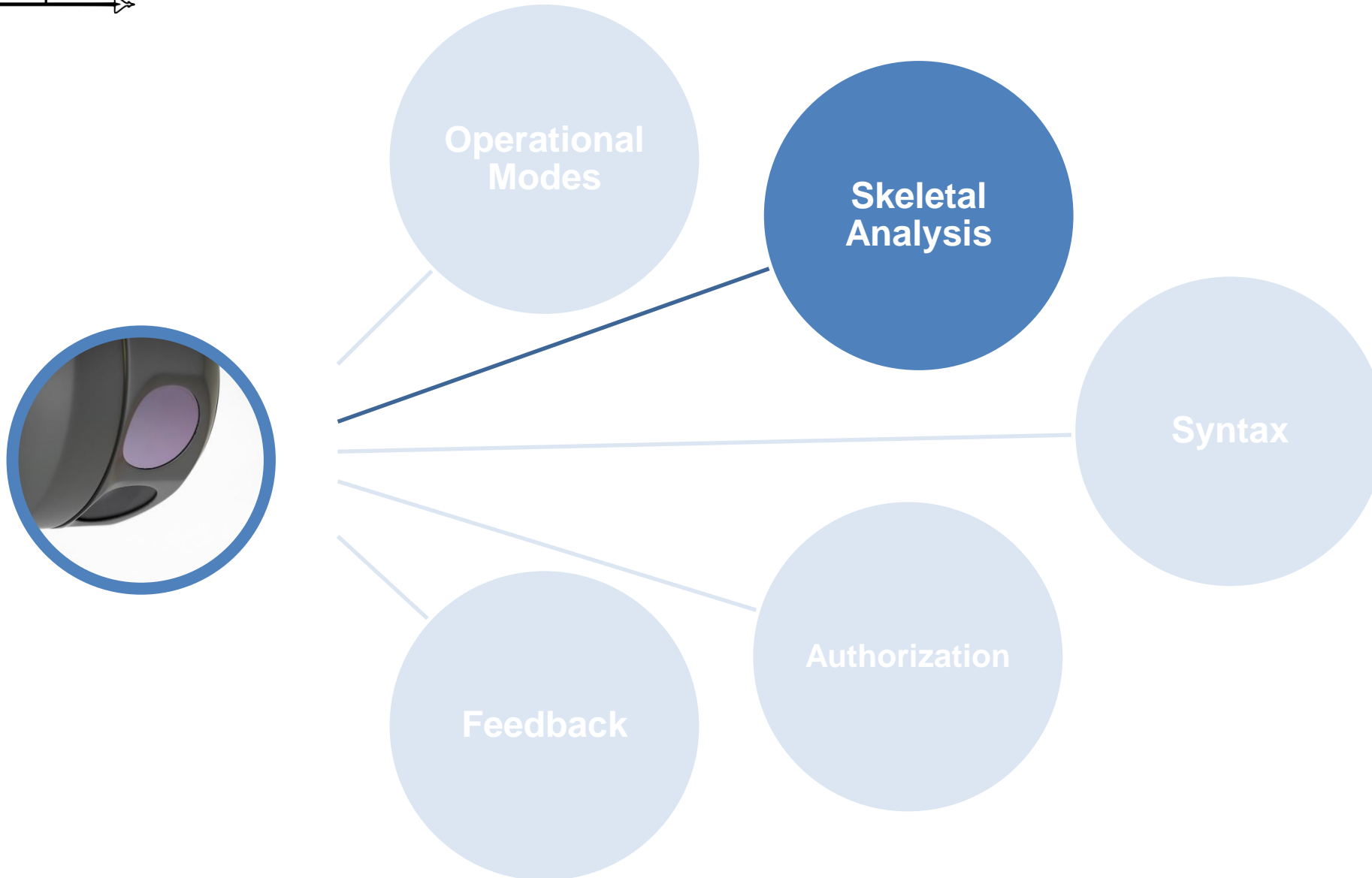
- > Gesture recognition
- > Translation into gestural command components
- > Depth sensor, EO with high zoom
- > Medium or high resolution operator body shape model



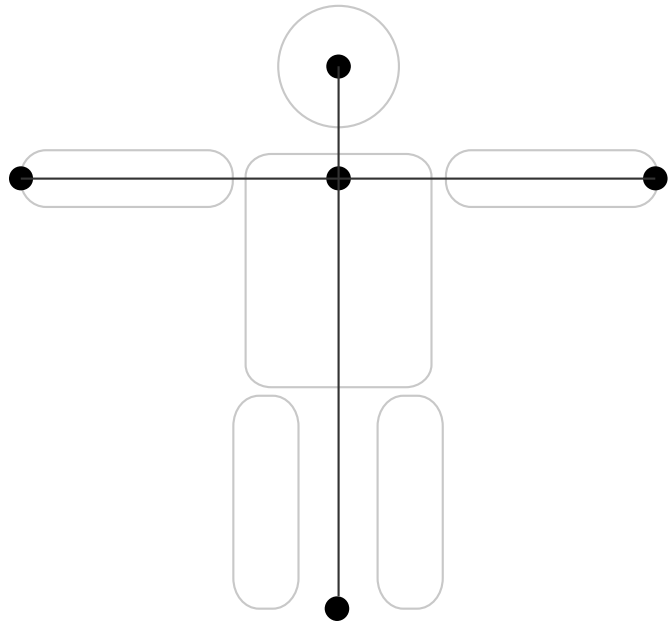


Modelling Visual Communication with UAS

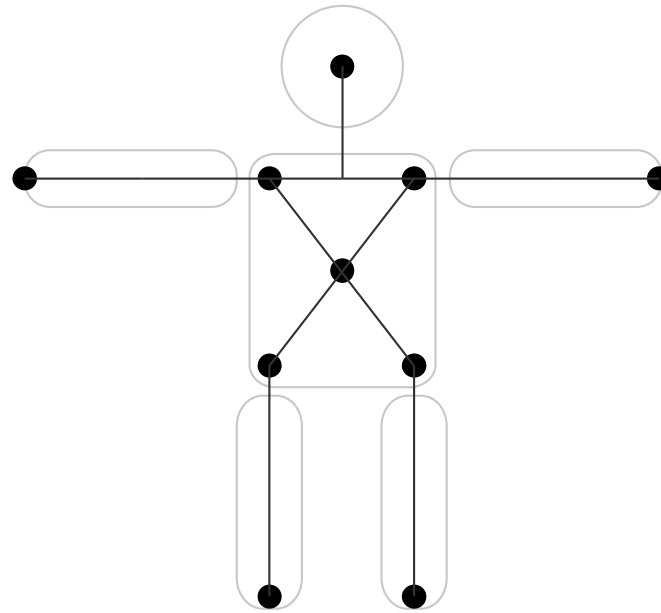
Research Topics



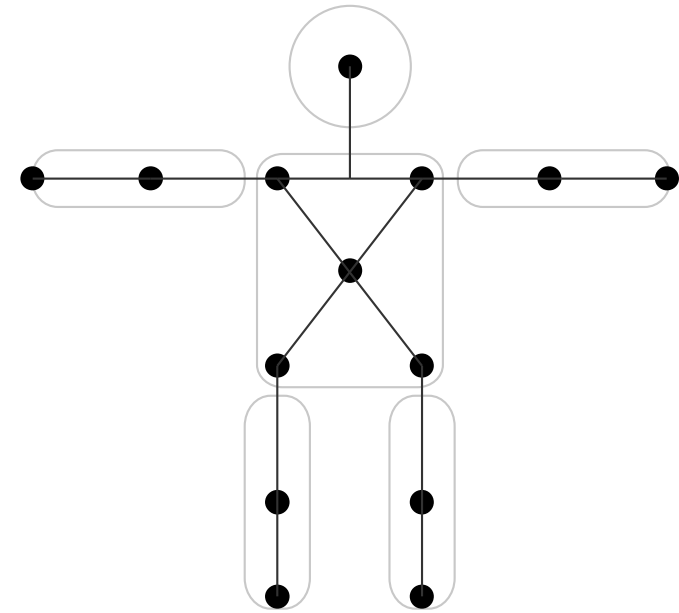
Ground sample distance



Low Resolution



Medium Resolution



High Resolution

LWIR-Sensor

EO-Sensor

Depth-Sensor,
High Zoom EO-Sensor



Number of recognizable gestures



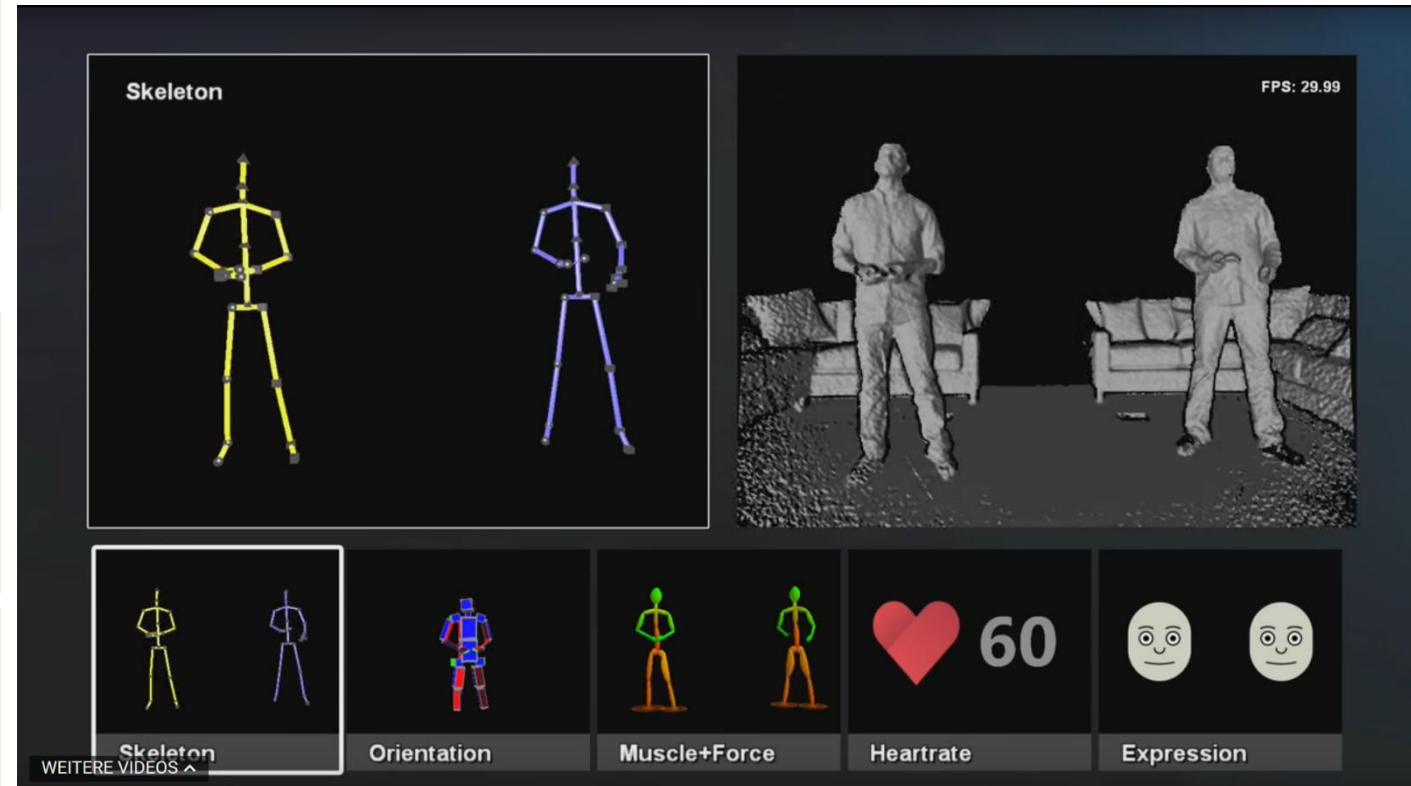
Common Tools for Skeletal Analysis

- **Kinect v2**

Depth sensor + big learned dataset
+ hardware processing

+ Fast
+ Accurate

- Hardware dependent
- Limited sensor range
- Indoor use only



© youtube, wired



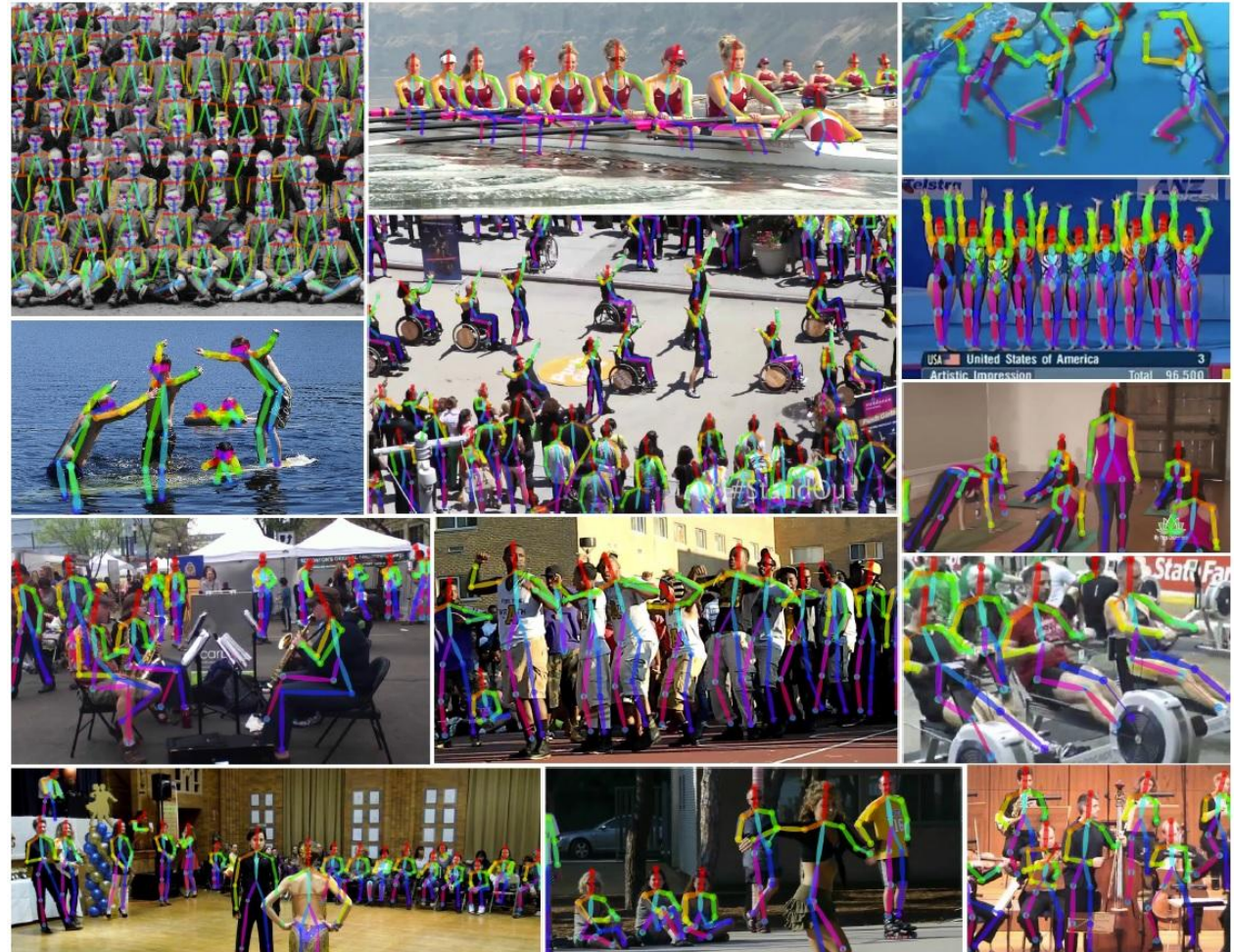
New Advances in Pose Estimation using Deep Neural Networks

- **Zhe et al., 2016**

“Realtime Multi-Person 2D Pose Estimation using part Affinity Fields”

- + Bottom up approach
- + Accurate
- + No depth sensor needed

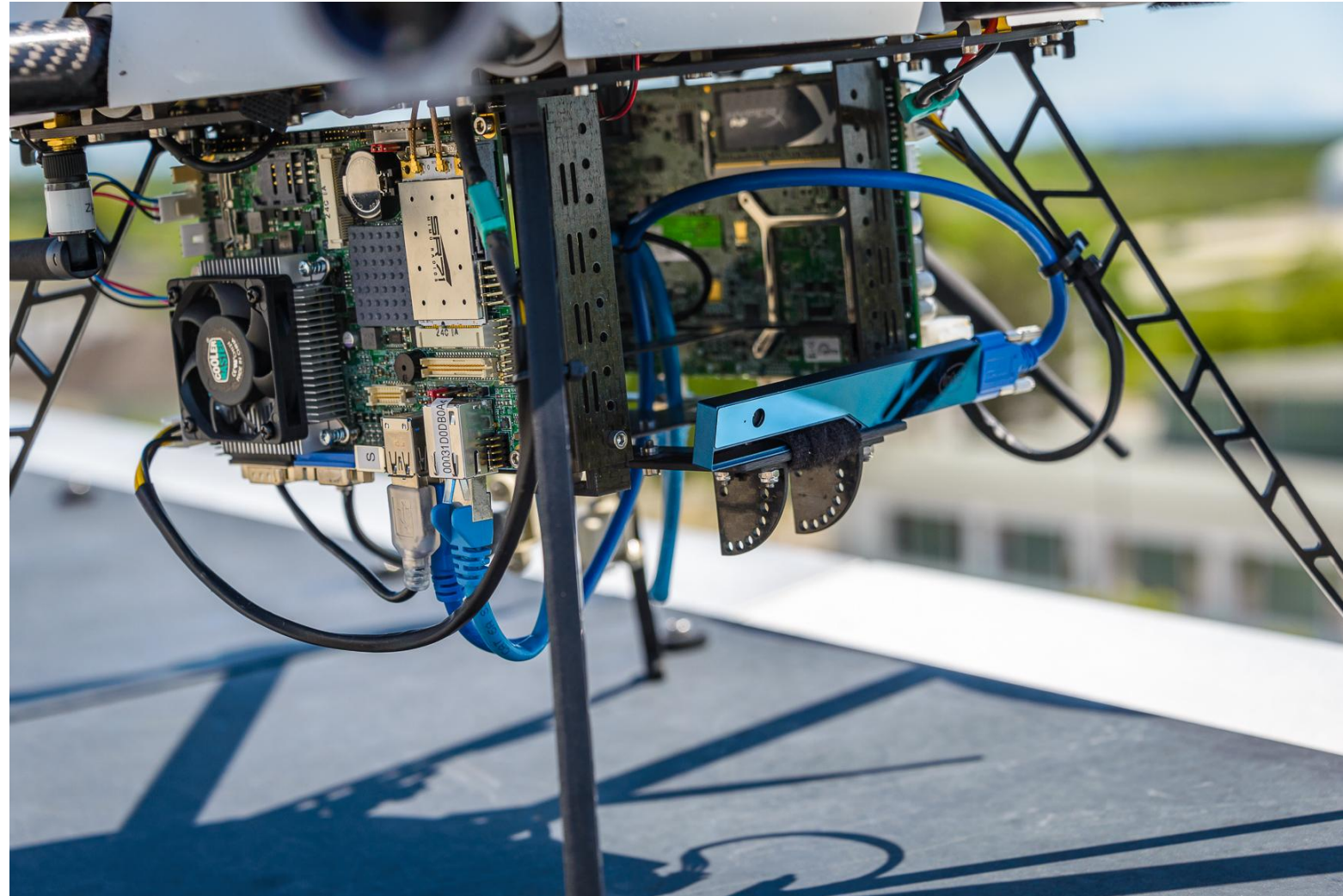
- Computational costs
- Real-time only with GPU acceleration



© Zhe et al., 2016



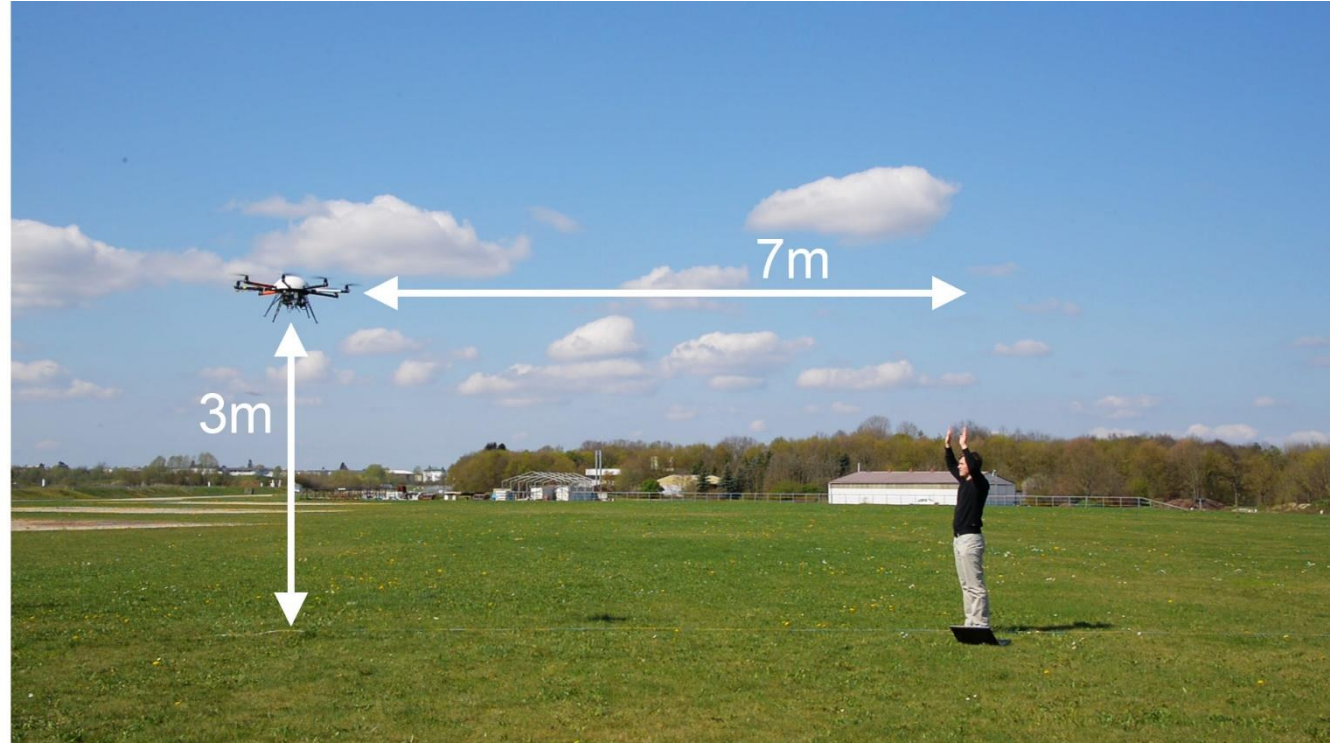
- > **Stereoscopic depth camera**
Intel RealSense R200
- > **Color stream**
1920x1080 @ 60 Hz, HFOV 70°
- > **IR stream**
628x468 @ 60 Hz, HFOV 59°
- > **Depth stream**
628x468 @ 60/90 Hz, HFOV 59°
- > **Depth range**
0.6...10m (outdoor)
- > **USB 3.0 interface**





> Octocopter sensor platform

- Industrial 3.5" Mainboard
- Intel i7 3860 multicore CPU
- 8GB RAM
- 512 GB mSATA SSD



> Performed gestures

- Waving with both arms
- Attention!/I'm in command
- Move left, move right
- Come closer, back up





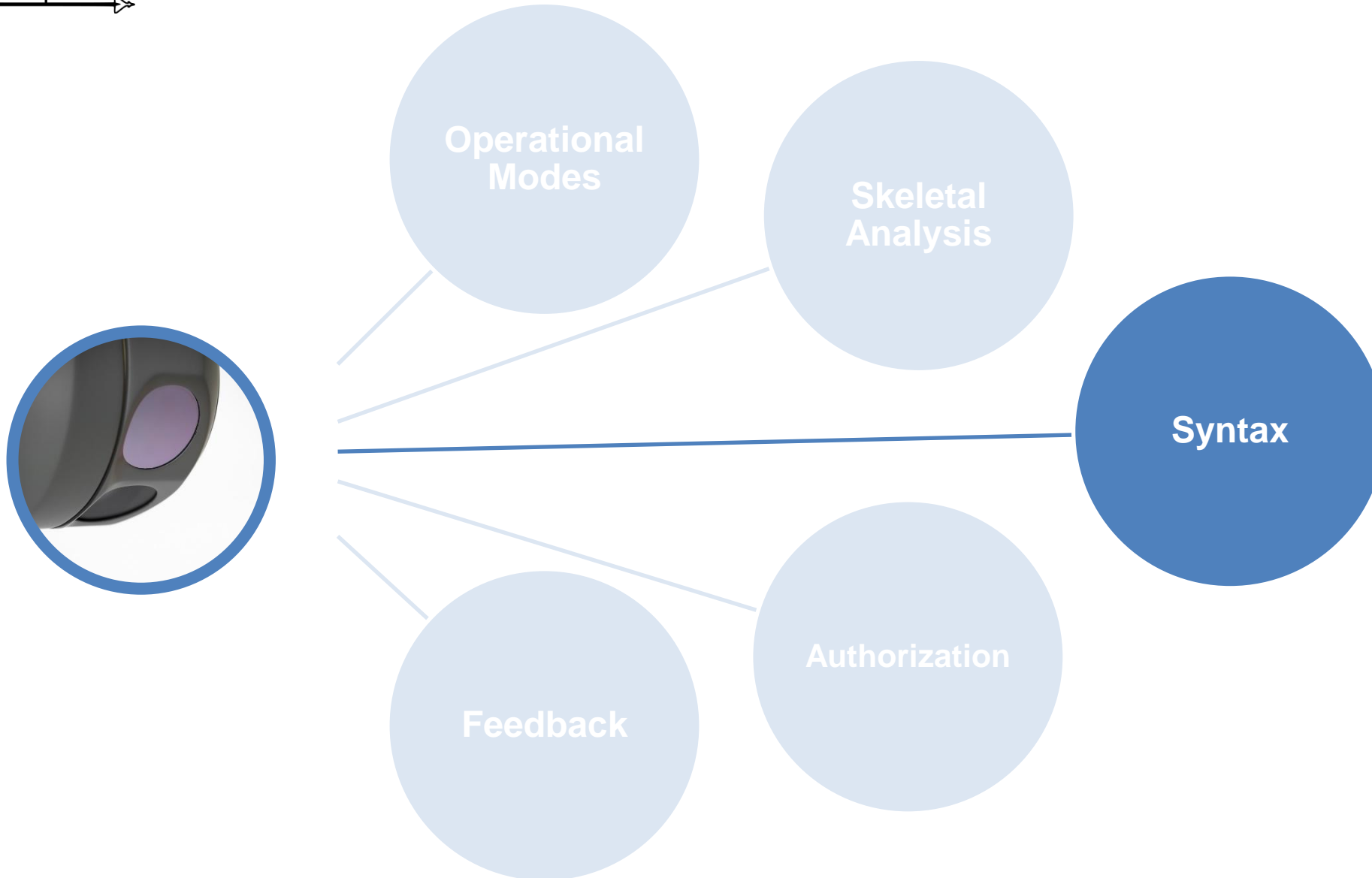
Experiment in Interaction Mode

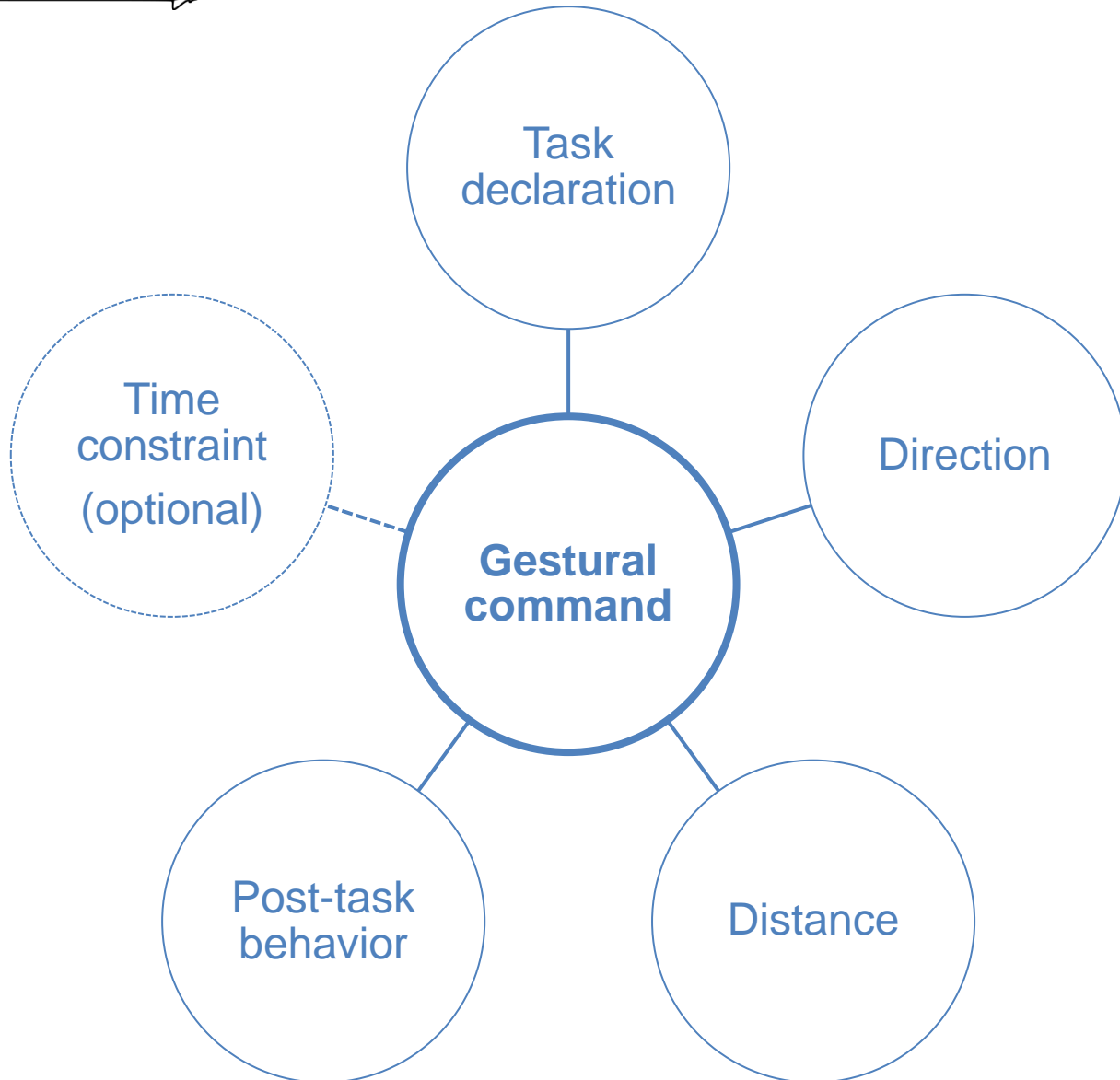
Results for small distance

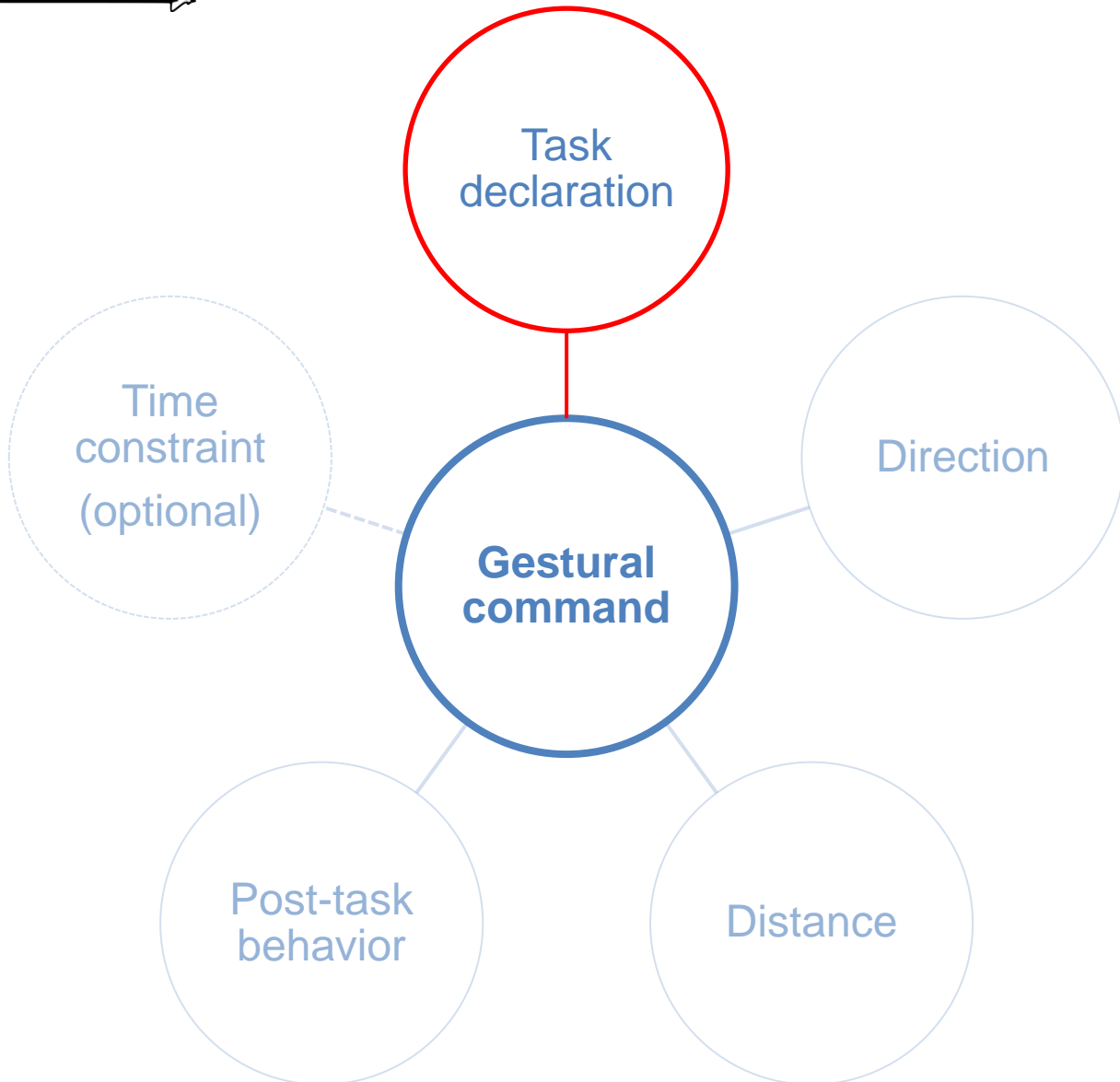
Distance to Operator: 7m

Stage I	Stage II	Stage III
<p>Attention!</p>	<p>Dynamic gesture</p>	<p>DNN approach based on "Zhe et al. <i>Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields</i>, 2016"</p>
<p>Gesture spaces approach</p>	<p>Gesture type recognition</p>	

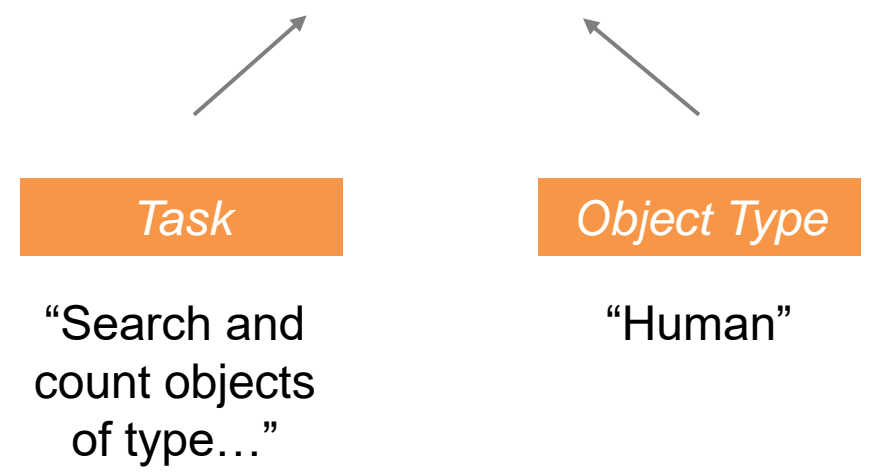




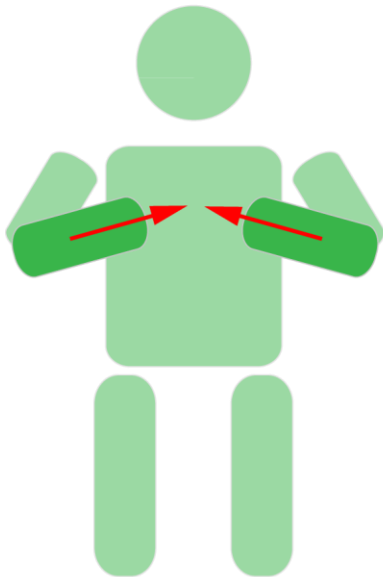




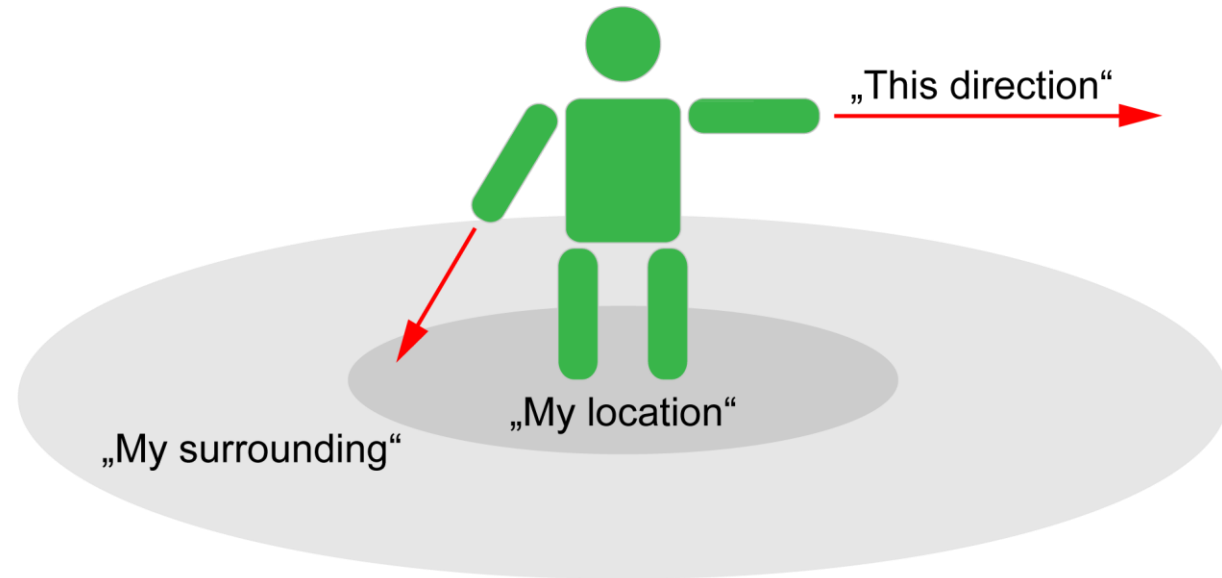
TASK DECLARATION

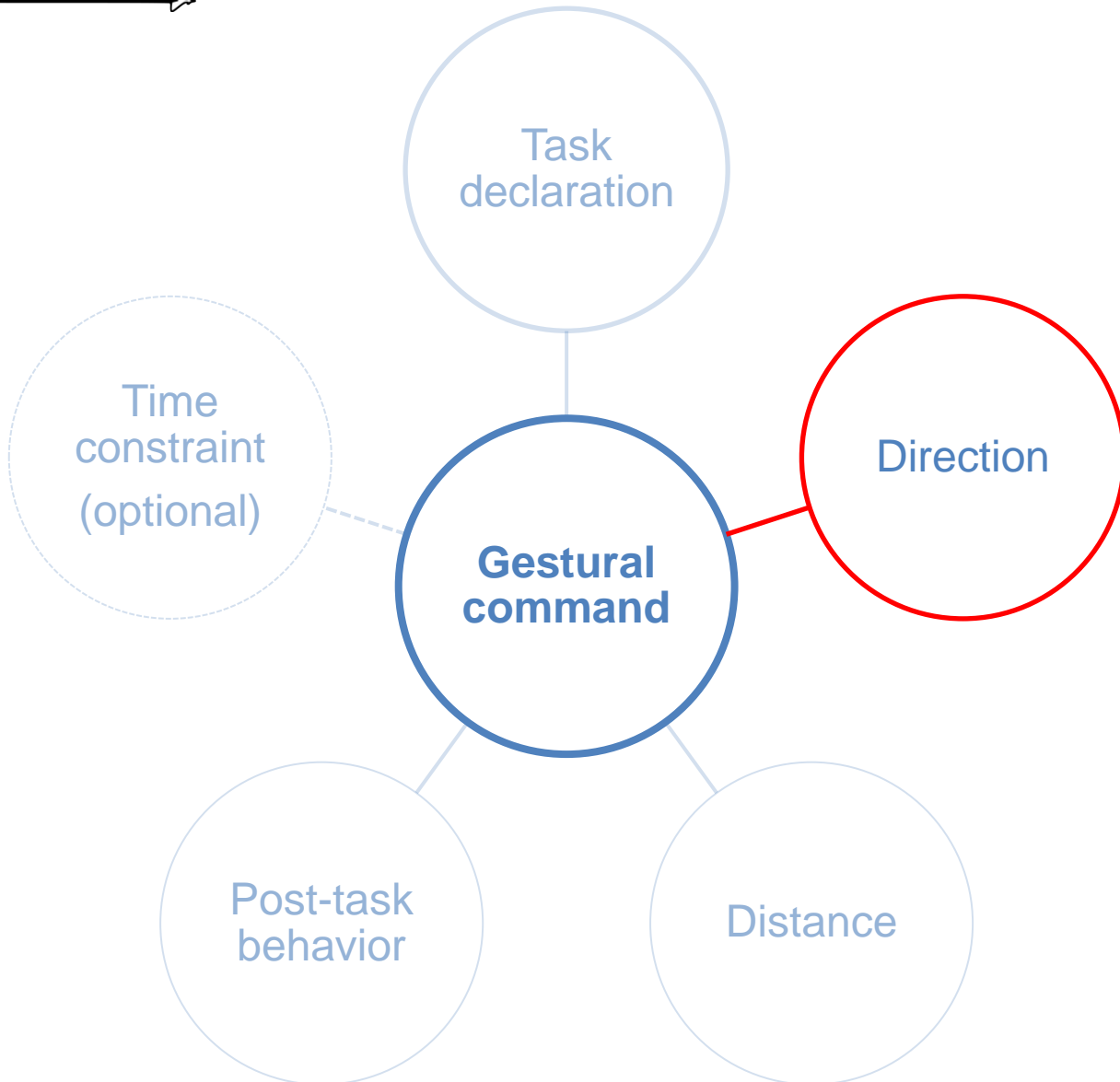


„Human“

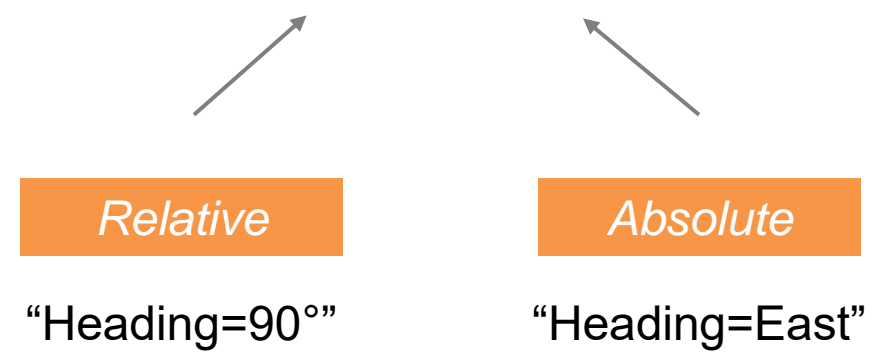


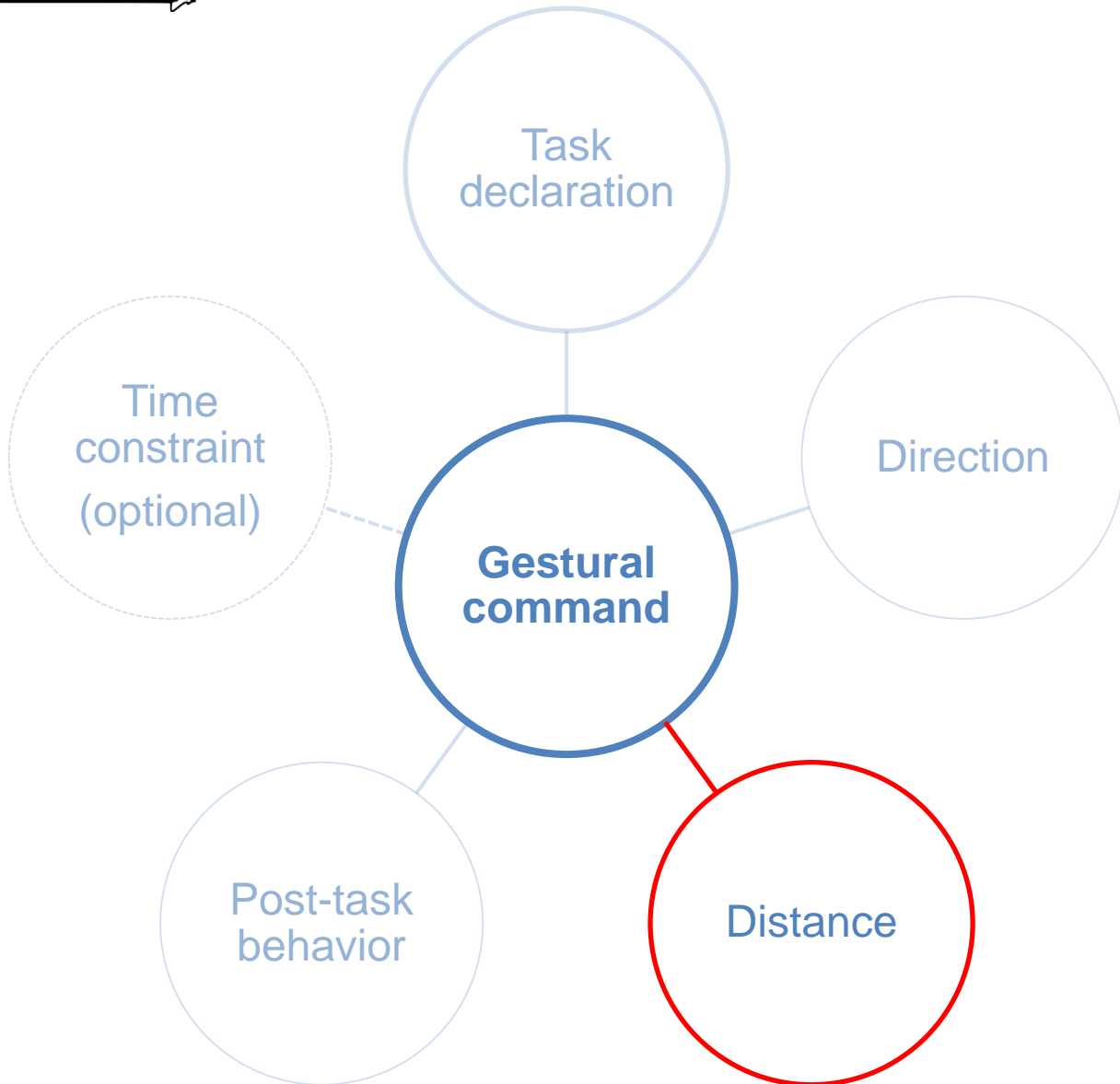
„This direction“





DIRECTION



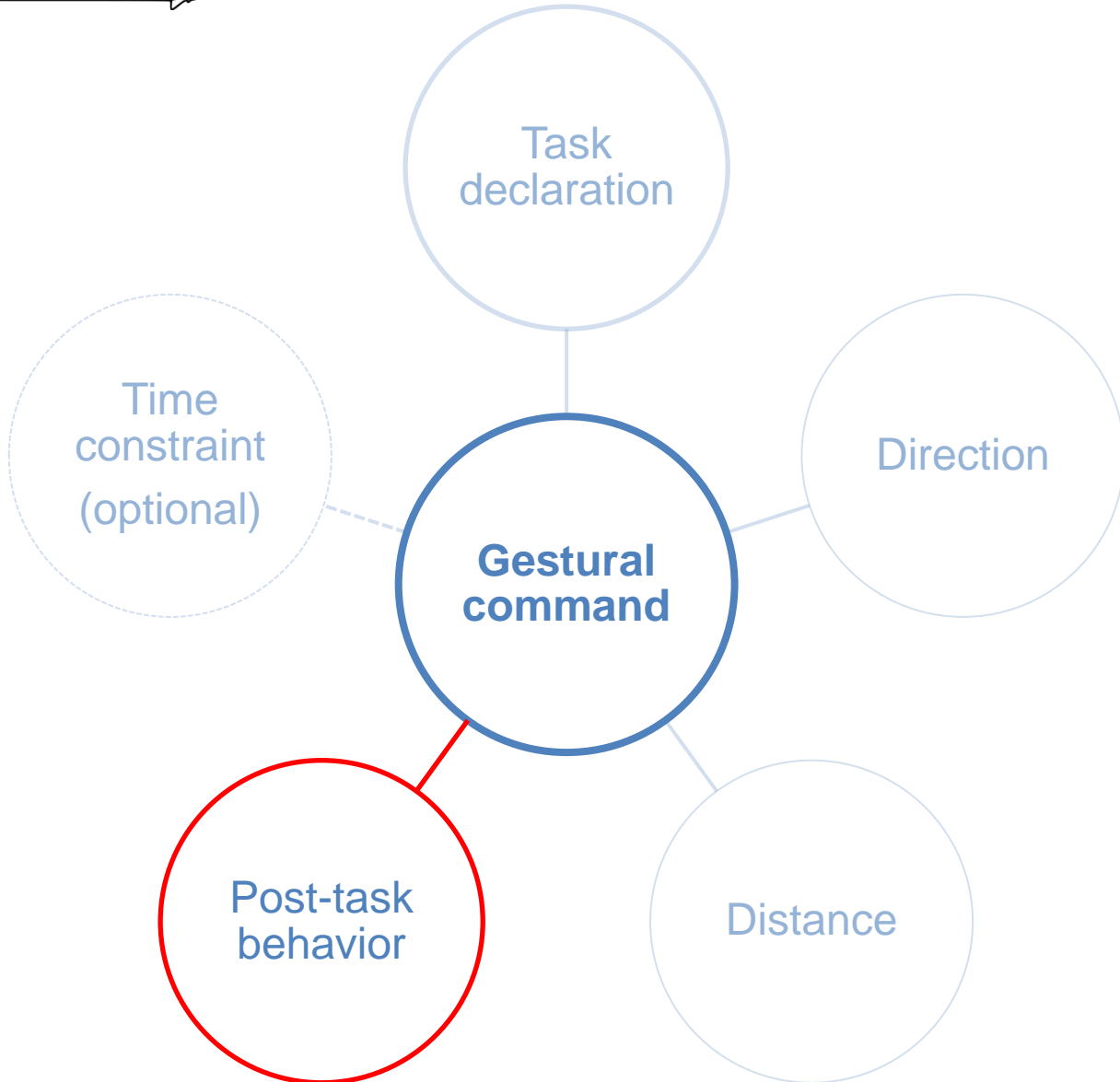


DISTANCE

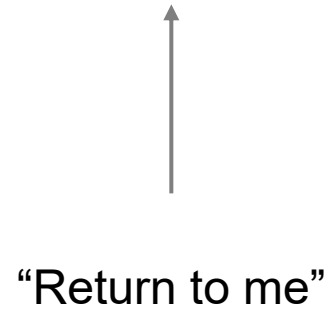


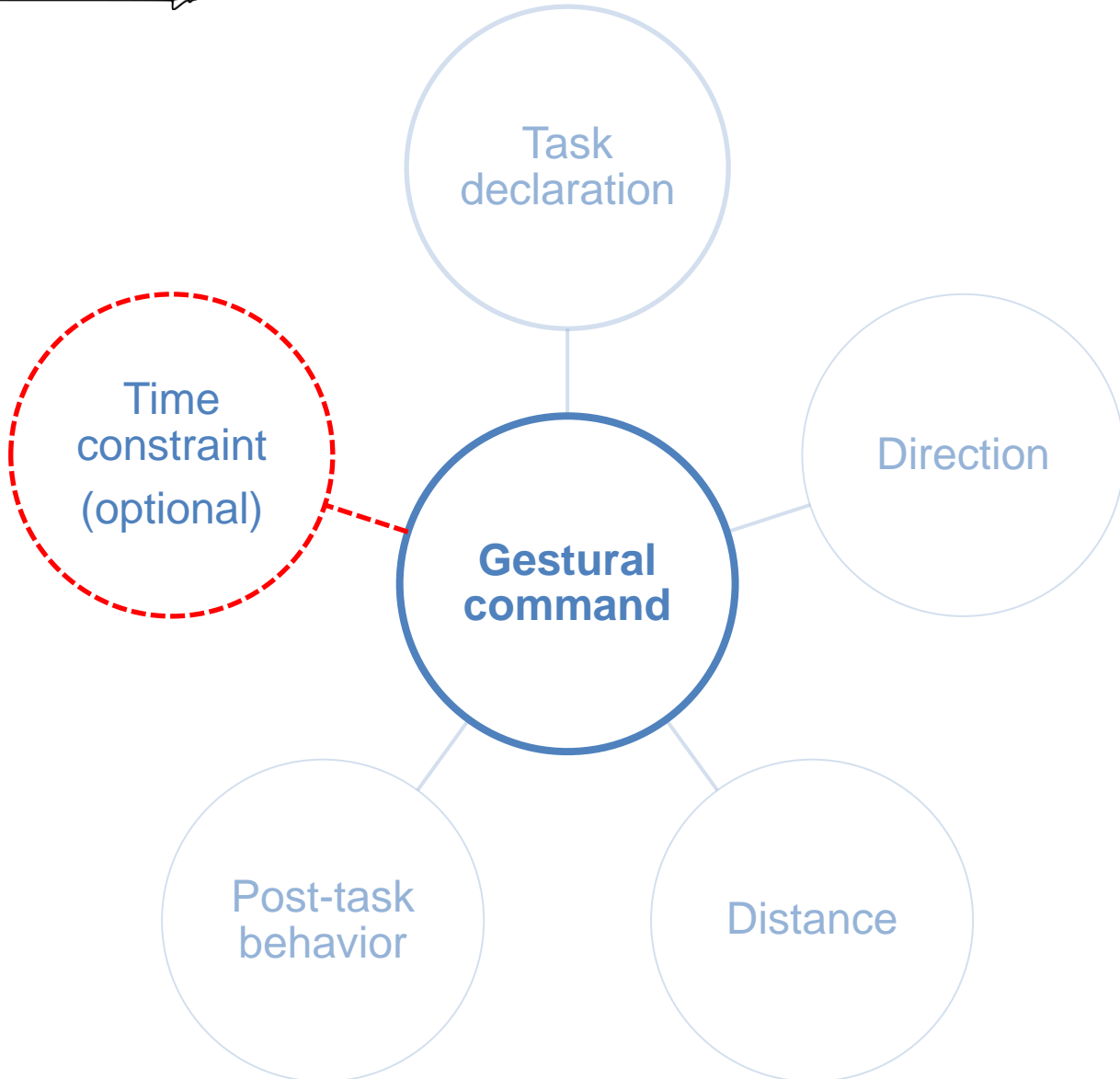
“Distance=500m”



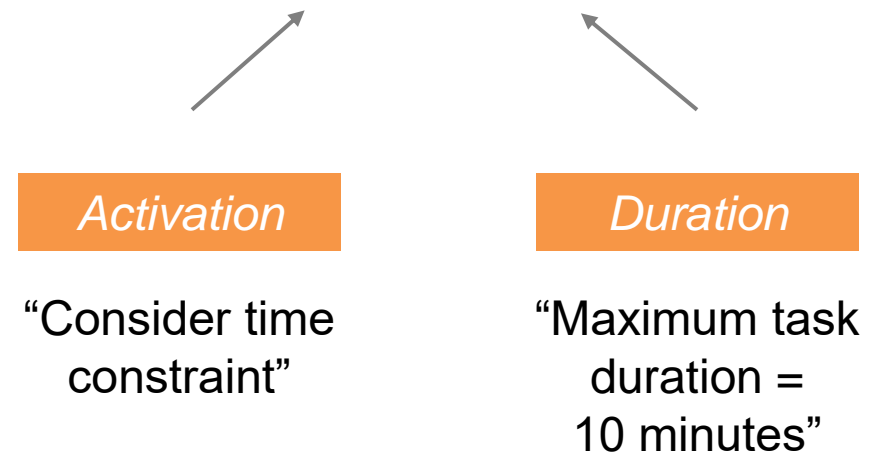


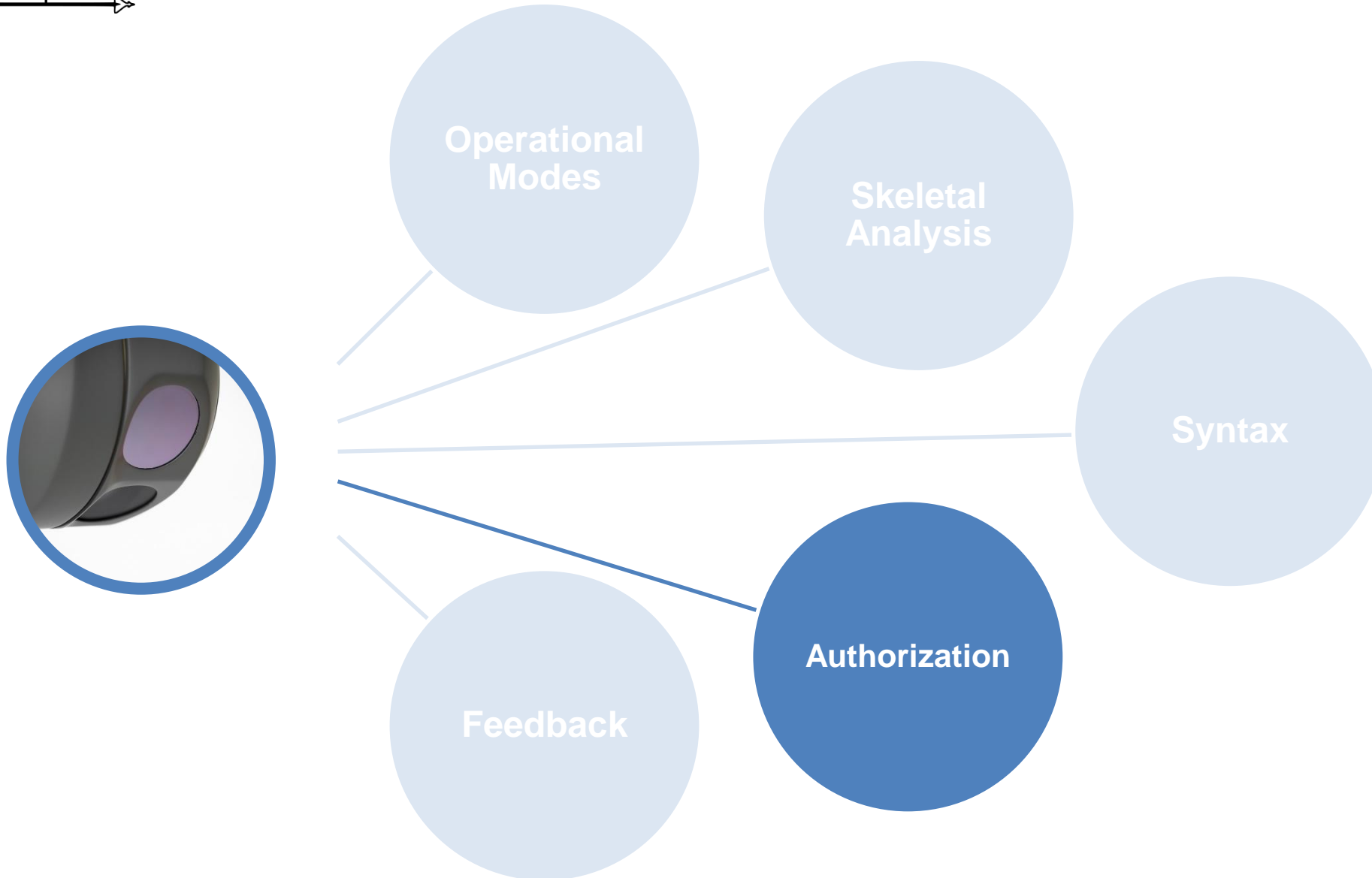
POST-TASK BEHAVIOR



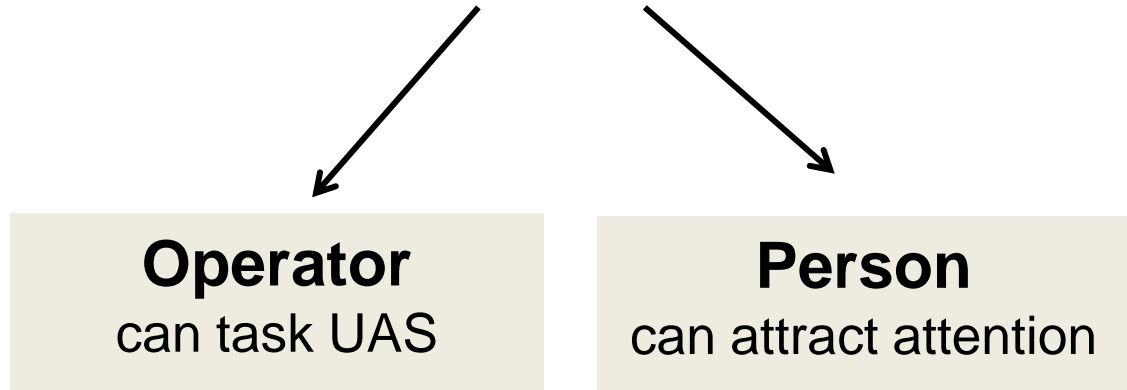


TIME CONSTRAINT

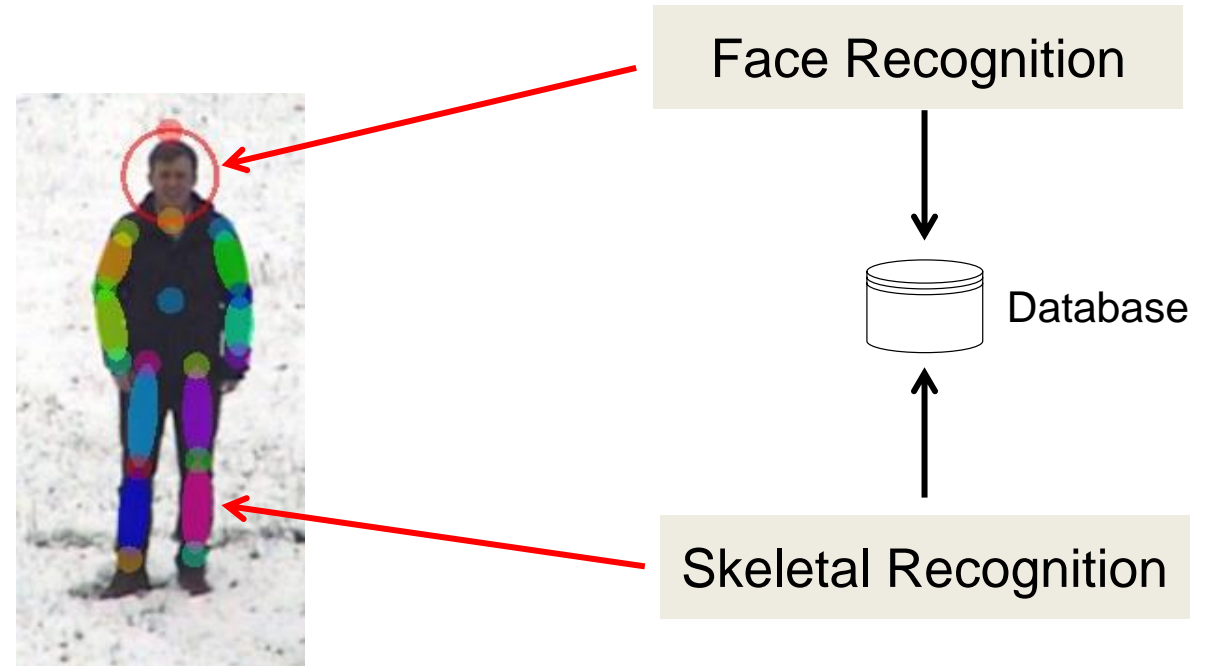


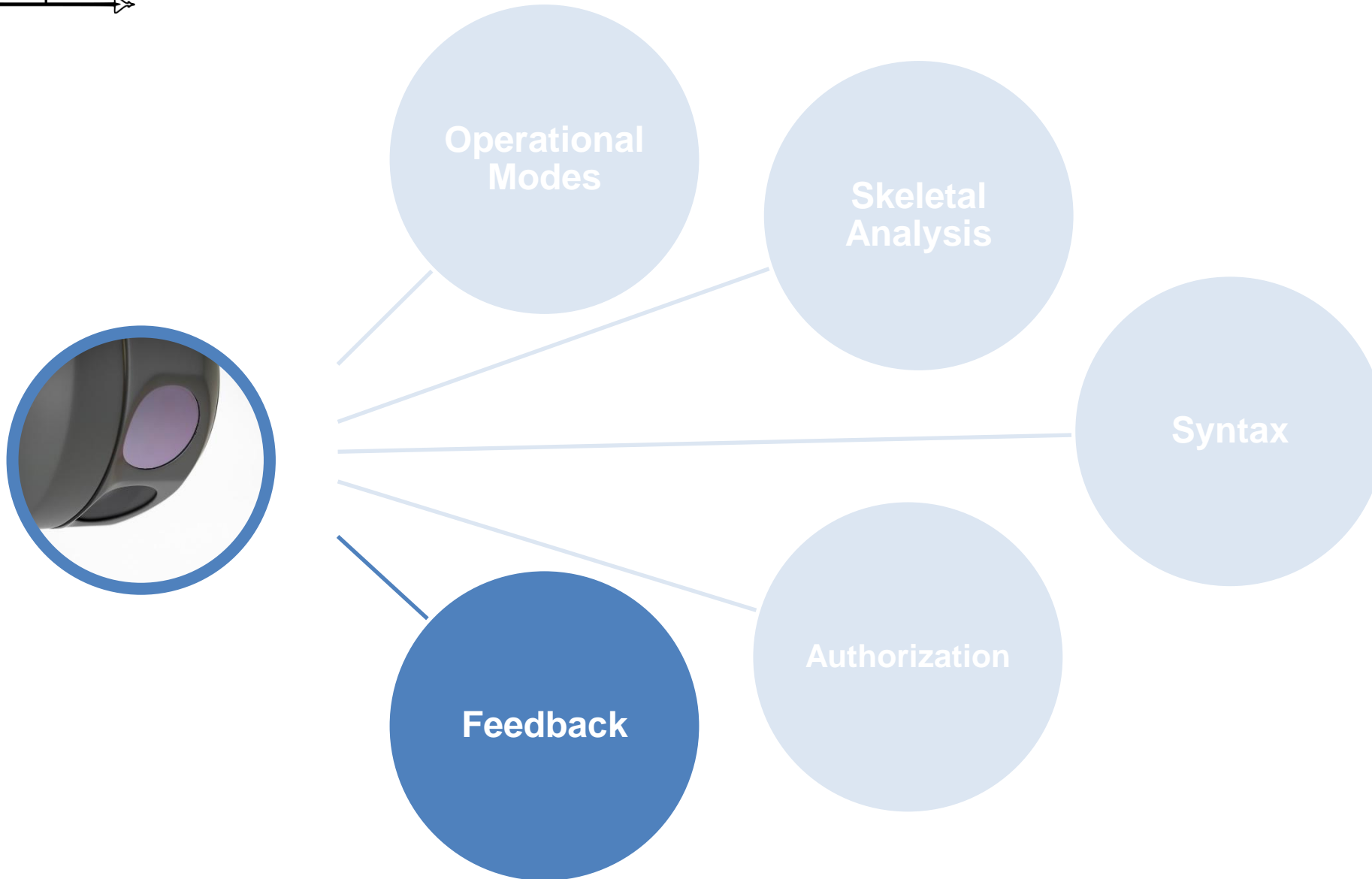


Authorization levels




Authorization features






Flight Maneuvers

MESSAGE RECEIVED AND UNDERSTOOD
Aircraft will indicate that ground signals have been seen and understood by—




Day or moonlight: Rocking from side to side.




Night: Making green flashes with signal lamp.

MESSAGE RECEIVED BUT NOT UNDERSTOOD
Aircraft will indicate that ground signals have been seen but not understood by—



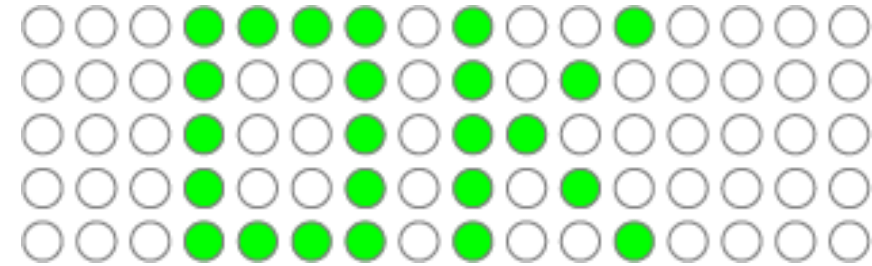
Day or night: Making a complete right hand circle.



Night: Making red flashes with signal lamp.

© aircav.com

Text



More low bandwidth feedbacks:

- DIRECTION?
- DISTANCE?
- TASK?
- TYPE?
- REPEAT





Experiment in Interaction Mode

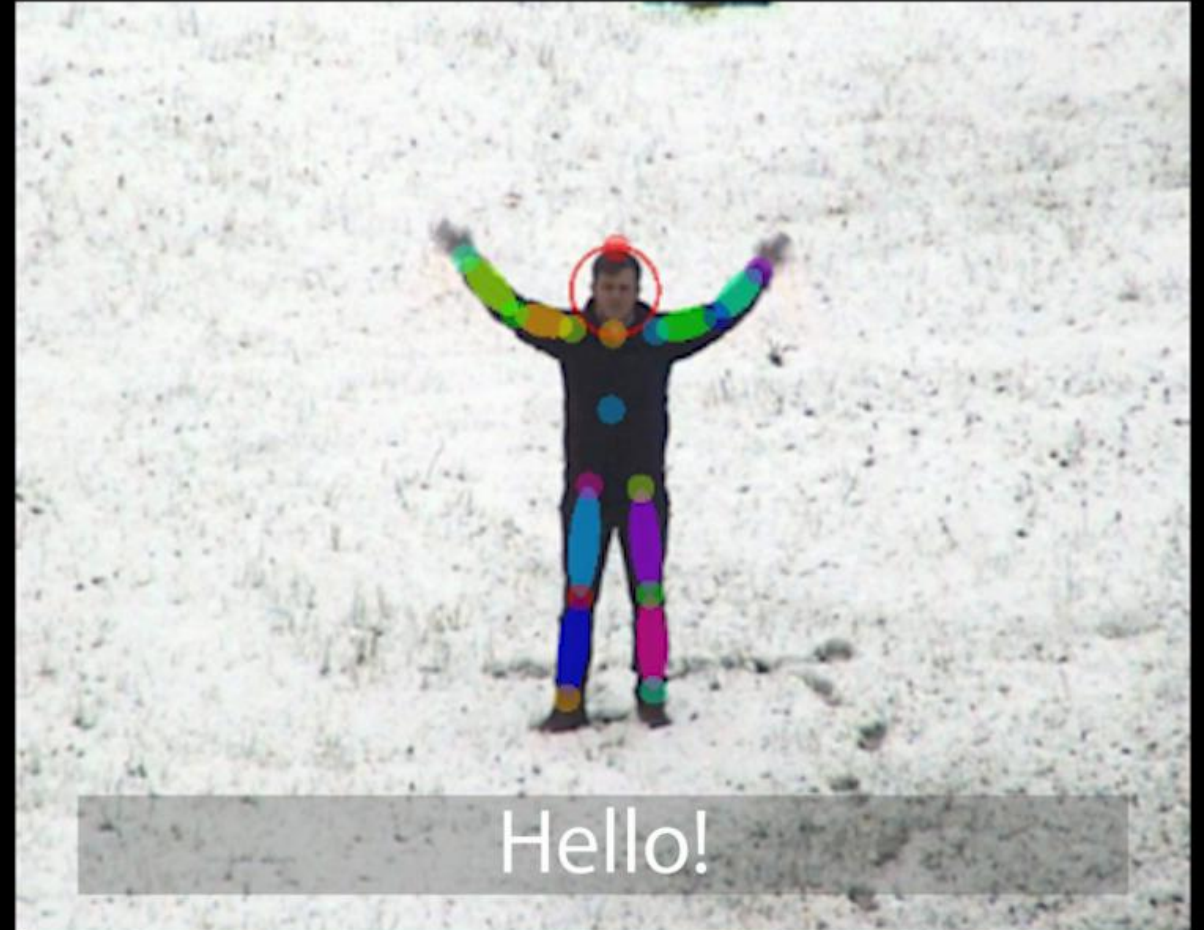
Result

Distance to Operator:
105m

LWIR
17° HFOV, 4x digital zoom



EO
3° HFOV, 16x optical zoom





Visual Communication with UAVs

- Enables bidirectional interaction
- Allows a more natural interface
- Improves usability
- Opens new fields of application







- [1] A. Kendon, „Gesticulation and speech: Two aspects of the process of utterance“, *The relationship of verbal and nonverbal communication*, vol. 25, pp. 207-227, 1980
- [2] E. Fricke, „Grammatik Multimodal: Wie Wörter und Gesten Zusammenwirken“, Boston: Walter De Gruyter Incorporated, 2012
- [3] N. Dalal and B. Triggs, „Histograms of Oriented Gradients for Human Detection“, in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, 2005, pp. 886-893
- [4] D. McNeill, „Hand and Mind: What gestures reveal about thought“, University of Chicago Press, 1992

