Multiview Object Detection and Tracking summary

K. Makroleivaditis, I. Karakostas, I. Mademlis, Prof. Ioannis Pitas
Aristotle University of Thessaloniki
pitas@csd.auth.gr
www.aiia.csd.auth.gr
Version 3.0

Artificial Intelligence & Information Analysis Lab

Multiview Object Detection and Tracking

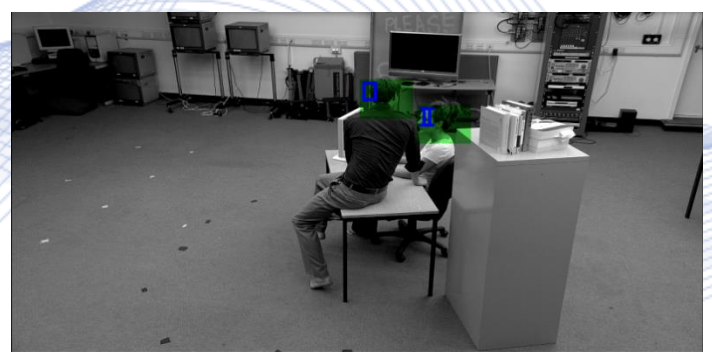
- **Multiview Human Detection**

- Multiview Object tracking

Multiview Human Detection

- Problem statement: Use information from multiple cameras to detect bodies or body parts, e.g. head.
- Applications:
  - Human detection/localization in postproduction.
  - Matting/segmentation initialization.

Camera 4



Camera 6

Multiview Human Detection

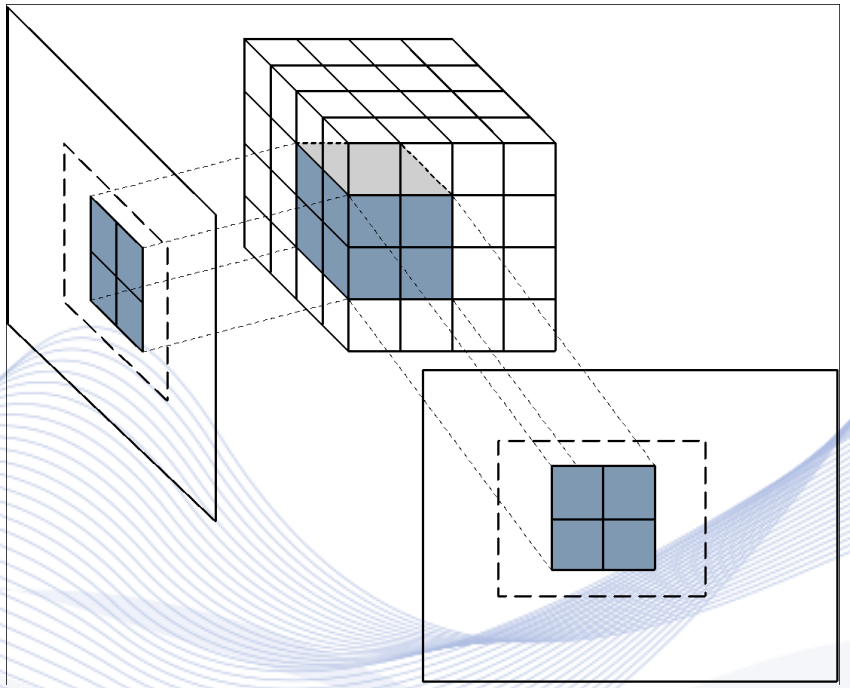
- ***Region-Of-Interests*** (ROIs) are typically bounding boxes. They are determined at a specific time t by the upper left and lower right rectangle coordinates $[x_l, y_l, x_r, y_r]$.
- Object ROI center: $\mathbf{c} = [x_c, y_c]^T$

Multiview Human Detection

- Head or body detection in two stages:
  - Use a head/face/body detector to derive ROIs in each view separately.
  - Insert these ROIs to an algorithm utilizing 3D information.
- Use of camera calibration parameters.
- Output: a rectified set of ROIs for each view that contains:
  - fewer false negatives
    - especially those due to occlusion are eliminated;
  - associations across views
    - all ROIs corresponding to the same human head/body are associated.

Multiview Human Detection

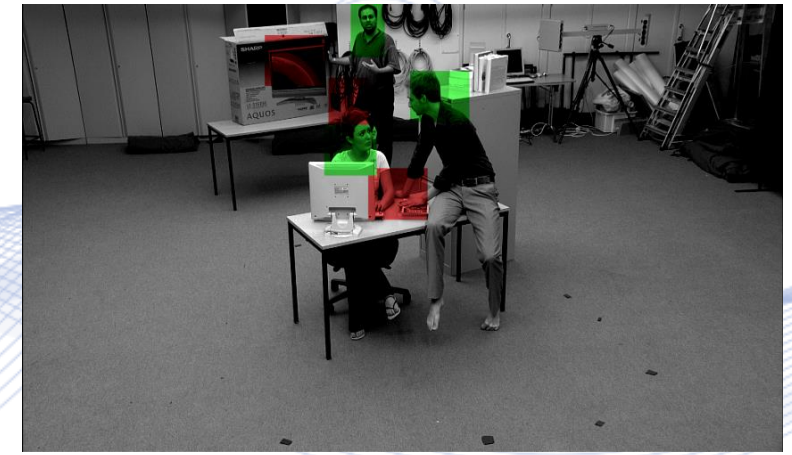
- Detected ROIs are projected back in the 3D space.
- A “probability volume” is created collecting “votes” from individual ROIs.
- High probability voxels correspond to the most probable head/body ROIs.



Multiview Human Detection

- The retained voxels are projected to all views.
- For every view we reject ROIs that have small overlap with the regions resulting from the projection.

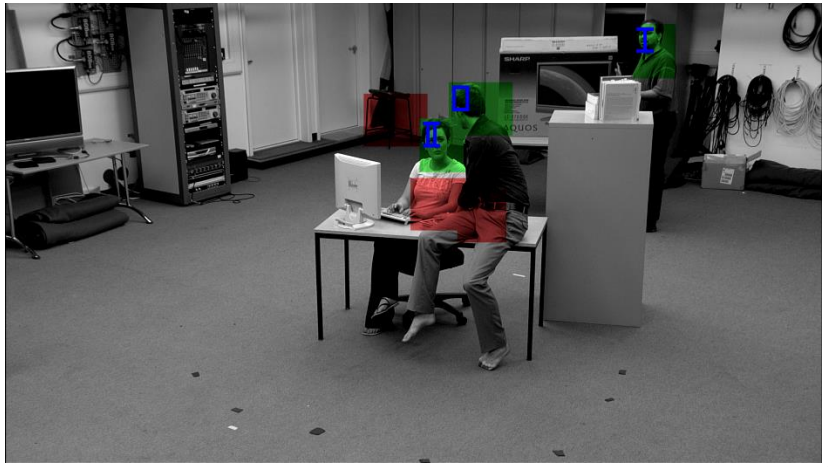
Camera 2

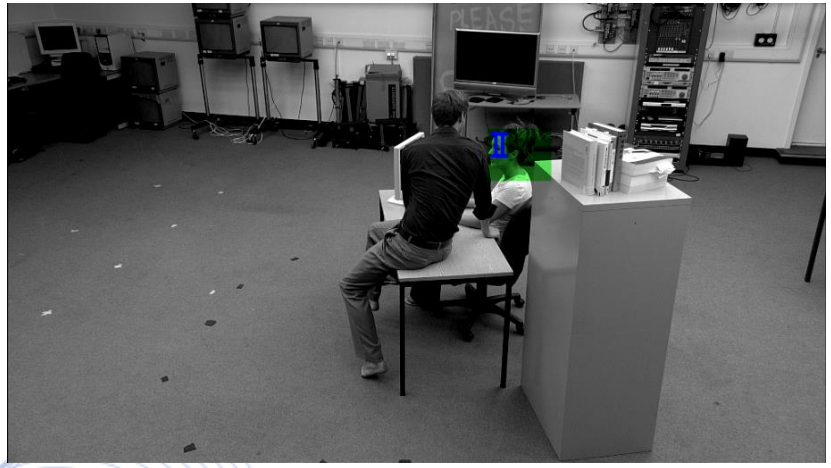


Camera 4

Multiview Human Detection

Camera 4

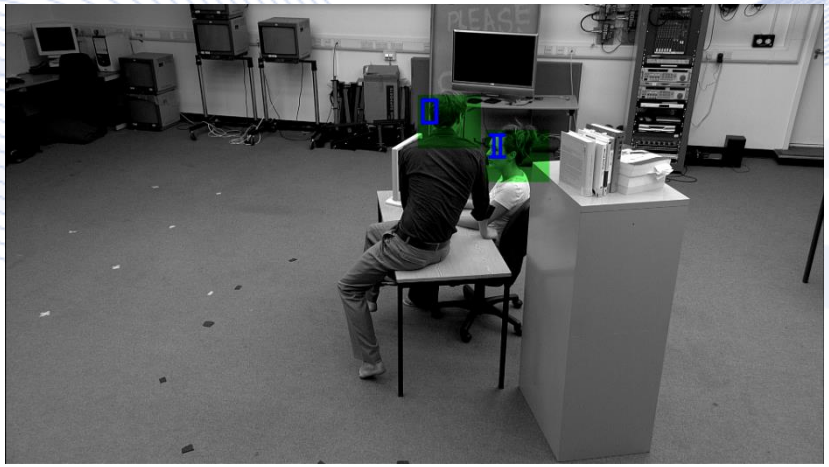




Camera 6

After ROI Rectification

Camera 4



Camera 6

Multiview Object Detection and Tracking

- Multiview Human Detection

- **Multiview Object tracking**

Multiview Object tracking

- Motion blurs, partial occlusions, background clutter and image motion are the main issues that an object tracker has to overcome, in order to precisely track a moving object.

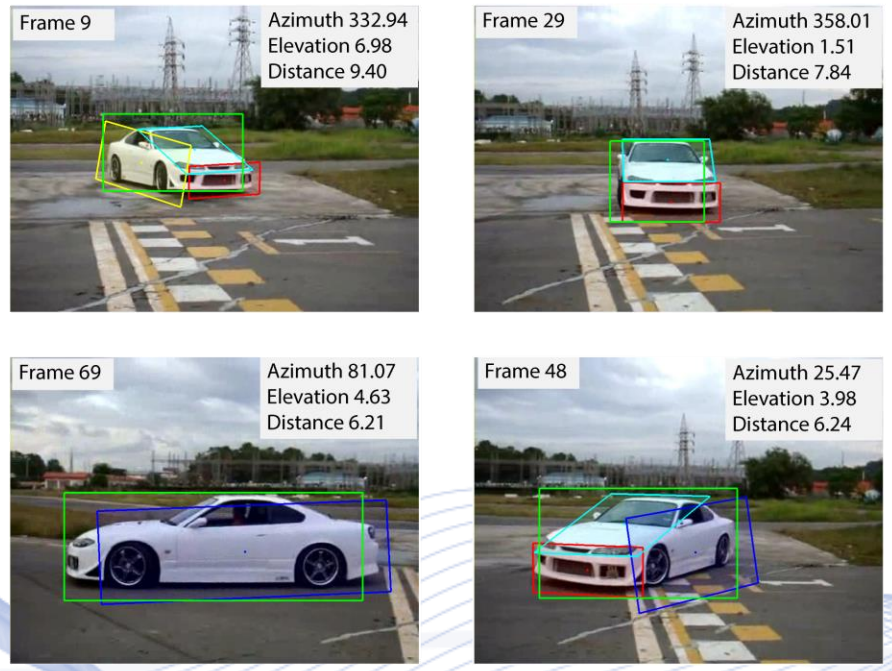
- Tracking a detected object in frame $t+1$.
  - Predicted object position $[x, y]^T(t+1)$
  - Compute ROI parameter vector $\widehat{\mathbf{y_1}}(t+1) = [x, y, w, h]^T$ within a *search region* on frame $t+1$
  - Retain object ID $\mathrm{J}(t+1) = \mathrm{J}(t)$

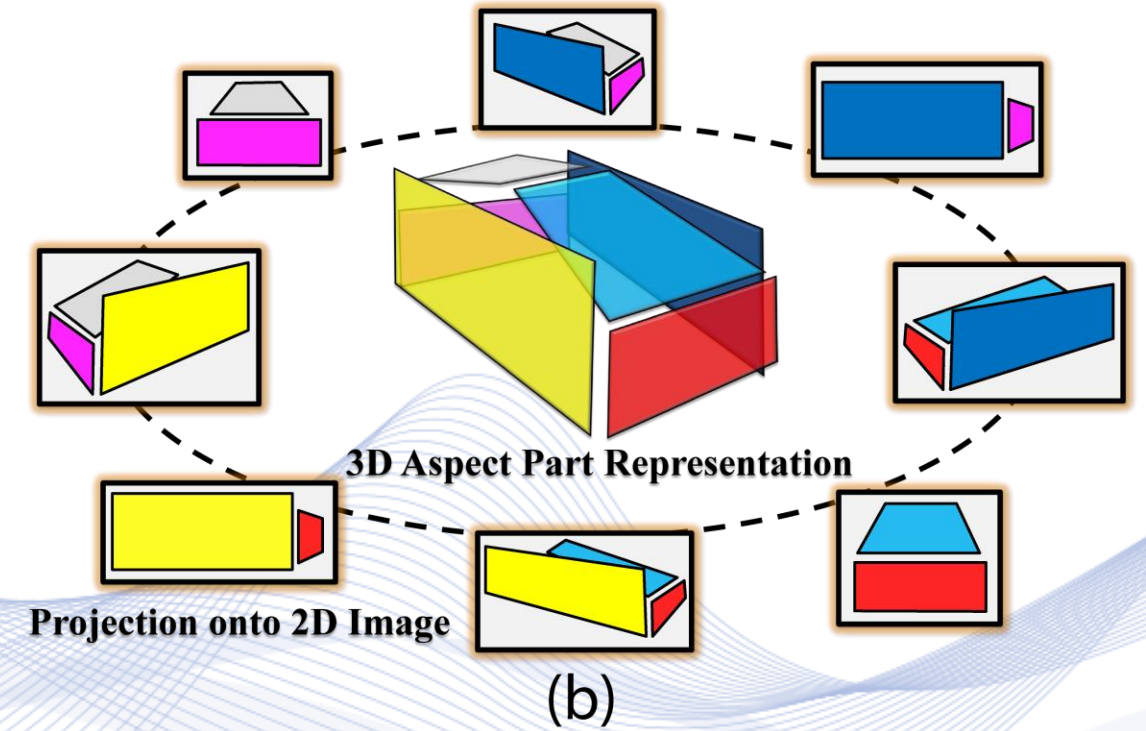
- Tracking failures can occur due to occlusions to the background and in those cases ***object-redetection*** techniques are implemented.

Multiview Object tracking

- Different and multiple viewpoints pose a challenge to traditional object tracking methods.
- Variations on topological appearance are dealt with by viewpoint transformations; representing objects with 3D aspect parts and modeling the connection in part-based particle filtering frameworks.
- The tracking framework can handle the appearance alternation and accurately predict the visibility and shape of a part.

Multiview Object tracking





(a): Example output for the tracking framework; (b): 3D aspect part representation of the car, and its projections from different viewpoints

Multiview Object tracking

Multiview 3-UAV ORBIT

(a) Video frame from UAV 0.

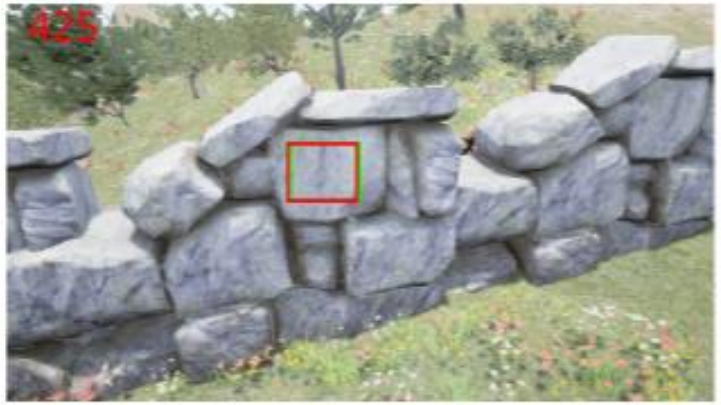
(b) Video frame from UAV 1.

(c) Video frame from UAV 2.

Multiview Object tracking

Multiview 3-UAV TRACK

(a) Video frame from UAV 0, following the desired target with an LTS CMT.



(b) Video frame from UAV 1, following the desired target with an LTS CMT.

(c) Video frame from UAV 2, following the desired target with a VTS CMT.

Multiview Object tracking

Multiview Object tracking

Baseline 3-UAV TRACK (without multiview fusion)

(a) Frame from UAV 0 following the desired target with an LTS CMT.

(b) Frame from UAV 1 following the desired target with an LTS CMT.

(c) Frame from UAV 2 following the desired target with an VTS CMT.

Multiview Object tracking

- After acquiring the MULTIDRONE experimental media production data, multiview UAV-captured footage depicting a rowing boat race were successfully synchronized by AUTH using the exhaustive search approach.

Bibliography

[PIT2017] I. Pitas, “Digital video processing and analysis” , China Machine Press, 2017 (in Chinese).
[PIT2013] I. Pitas, “Digital Video and Television” , Createspace/Amazon, 2013.
[PIT2021] I. Pitas, “Computer vision”, Createspace/Amazon, in press.
[NIK2000] N. Nikolaidis and I. Pitas, “3D Image Processing Algorithms”, J. Wiley, 2000.
[PIT2000] I. Pitas, “Digital Image Processing Algorithms and Applications”, J. Wiley, 2000.

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