Image Analysis and Computer Vision

Image Analysis and Computer Vision - staff

Teacher: Prof. Vincenzo Caglioti vincenzo.caglioti@polimi.it

Exercise classes plus some theory by:
Prof. Giacomo Boracchi giacomo.boracchi@polimi.it
Prof. Luca Magri luca.magri@polimi.it
Prof. Federica Arrigoni federica.arrigoni@polimi.it

Image Analysis and Computer Vision – weekly schedule

Tuesday	h 12.15 - 14.15	room 9.1.2
Wednesday	h 12.15 – 14.15	room 3.1.3

NO streaming
SOME recording will be available

FRAMEWORK

Observation of a static or dynamic scene in the 3D world

single frame

through

one or several 2D images

use VISUAL SENSOR to understand what is happening on the scene...

images are the data we are going to use!

sequence of images ~ 30fps

OR also ~ each seconds image to spare memory

OR also pairs of images... more than one at same time to better perceive 3RD dimension NOT perceived by 1 image only

IACV main object

2 main chapters

1 IMAGE ANALYSIS:

- extract image features (edges, corners, regions)
- match features among different images ← from different positions, match same object between the two images
- track features along an image sequence (video)

↳ the content of the image exactly — as 2D MATRIX

↓ when using sets of images, track same features

important applications, useful to understand 3D

2 COMPUTER VISION:

compute a model of the observed scene

- 2D shape reconstruction
- 3D shape reconstruction

→ observing scene, you extract shape of what you see!

" use images to reconstrum scene shapes

- Object recognition — what is present on images
- Object localization
- Action recognition

↳ content of the observed scene

‖ use the features extracted to understand what is going on on the scene..

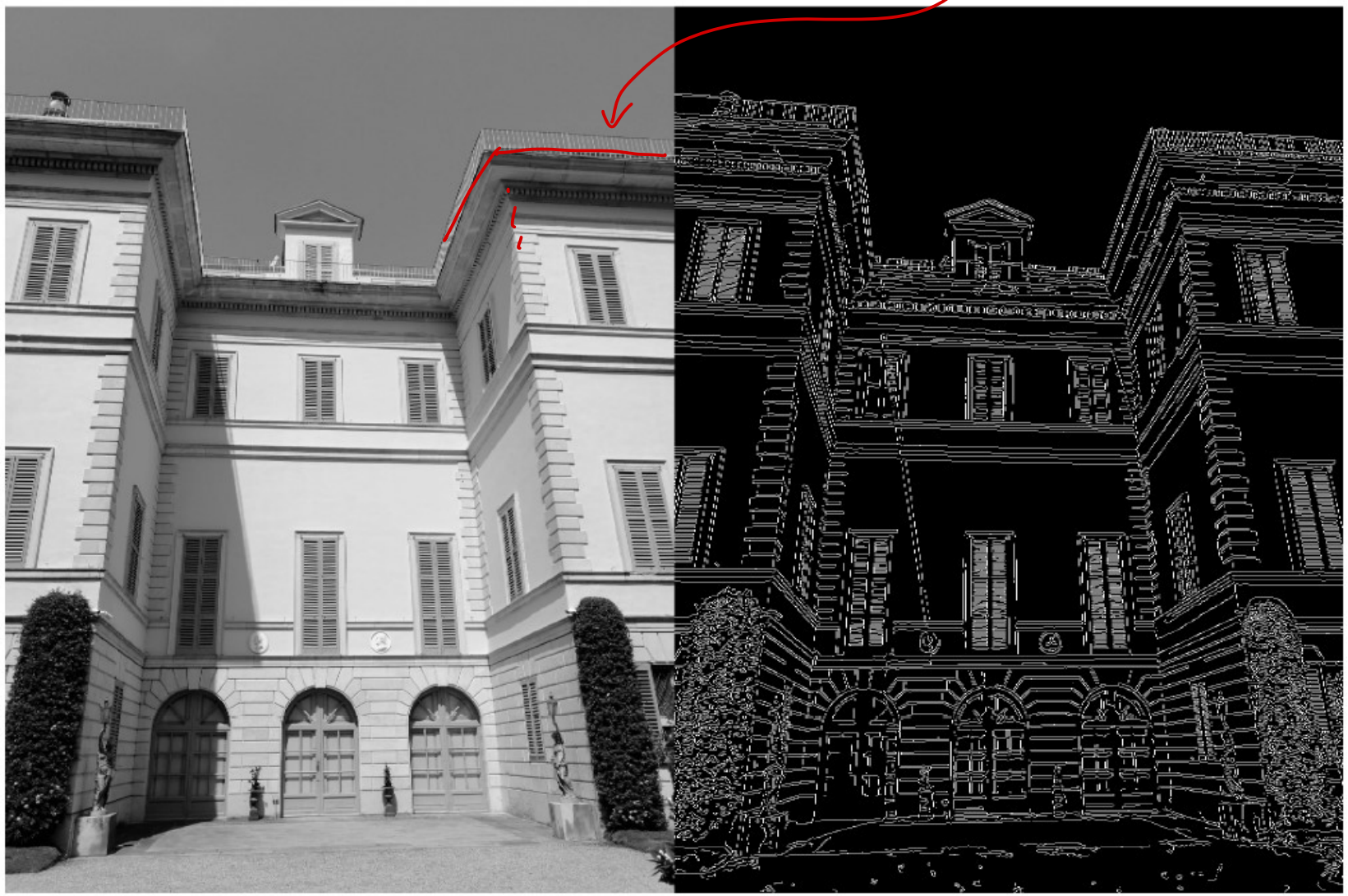
↑ (PLANAR)

" if person falls down, etch...

where are, relative position between camera and object of interest... to avoid collisions etc

extract relevant part of the image ex: board between building and sky

feature extraction: edges = where the intensity light has sudden changes between dark and light.

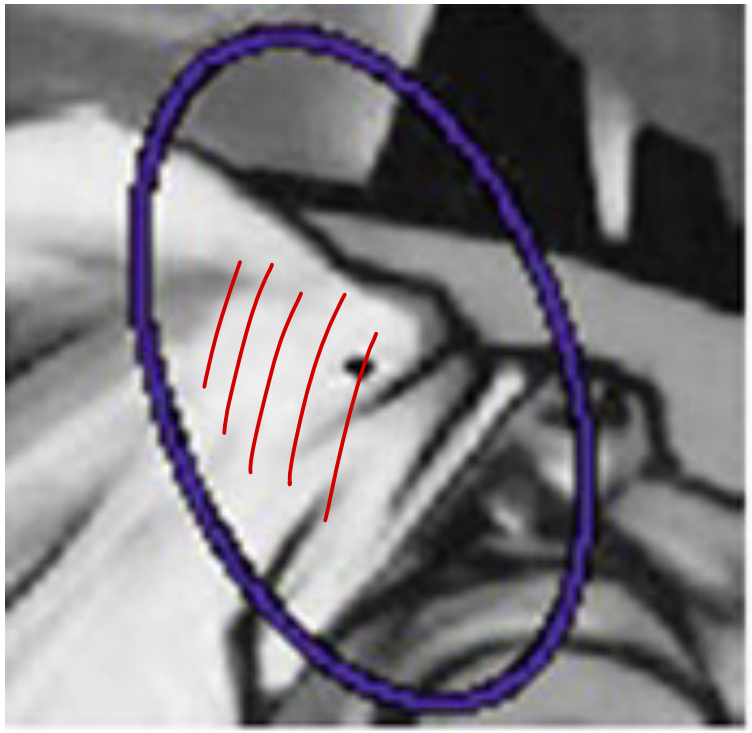


borders with different values of intensity
EDGES

NOT easy! with typical sw available in imy processing, eat of false positive are extracted...

feature extraction: corners

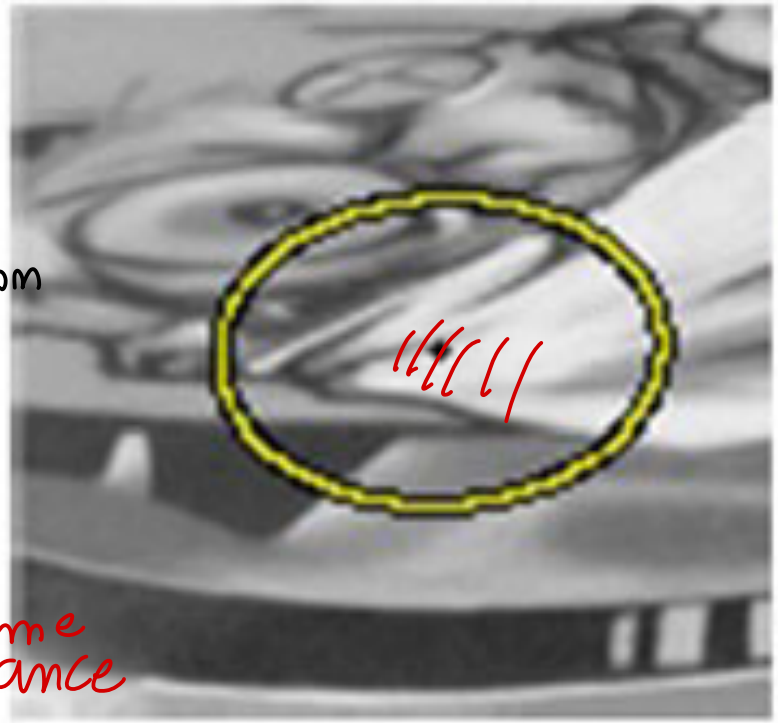
feature extraction: regions



different orientation

SAME REGION!

even if NOT same appearance

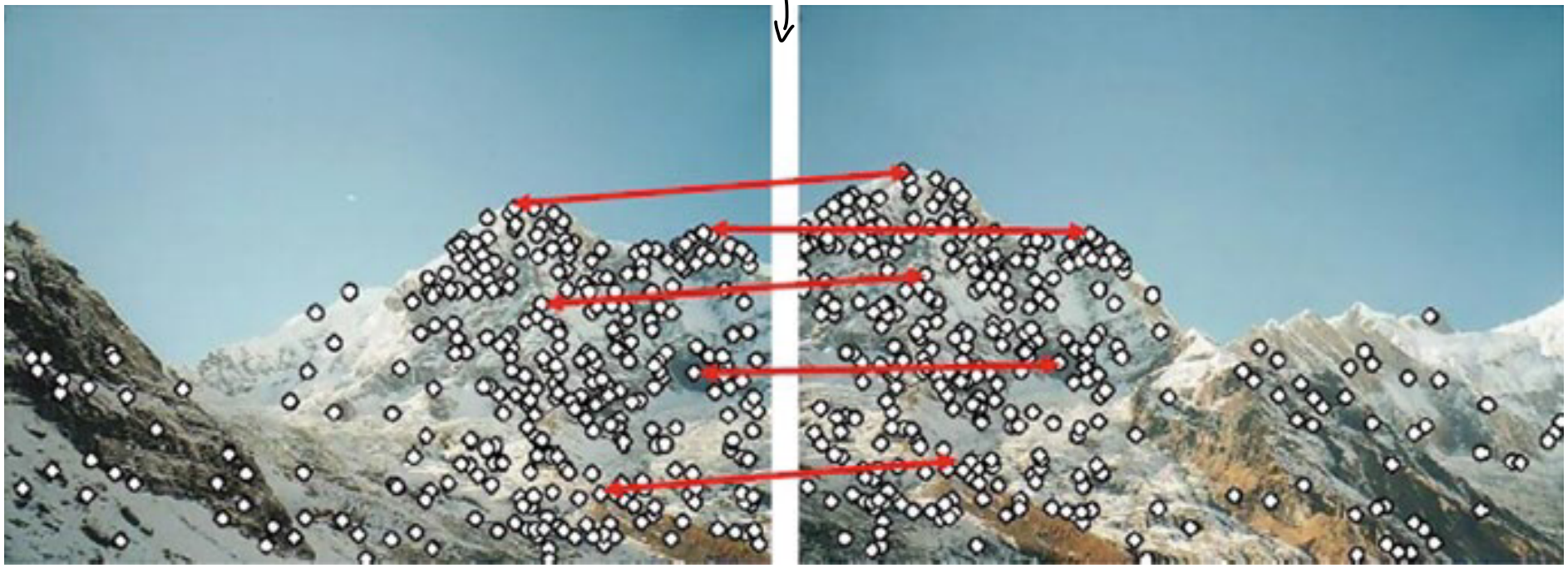


same region observed by different positions

feature matching

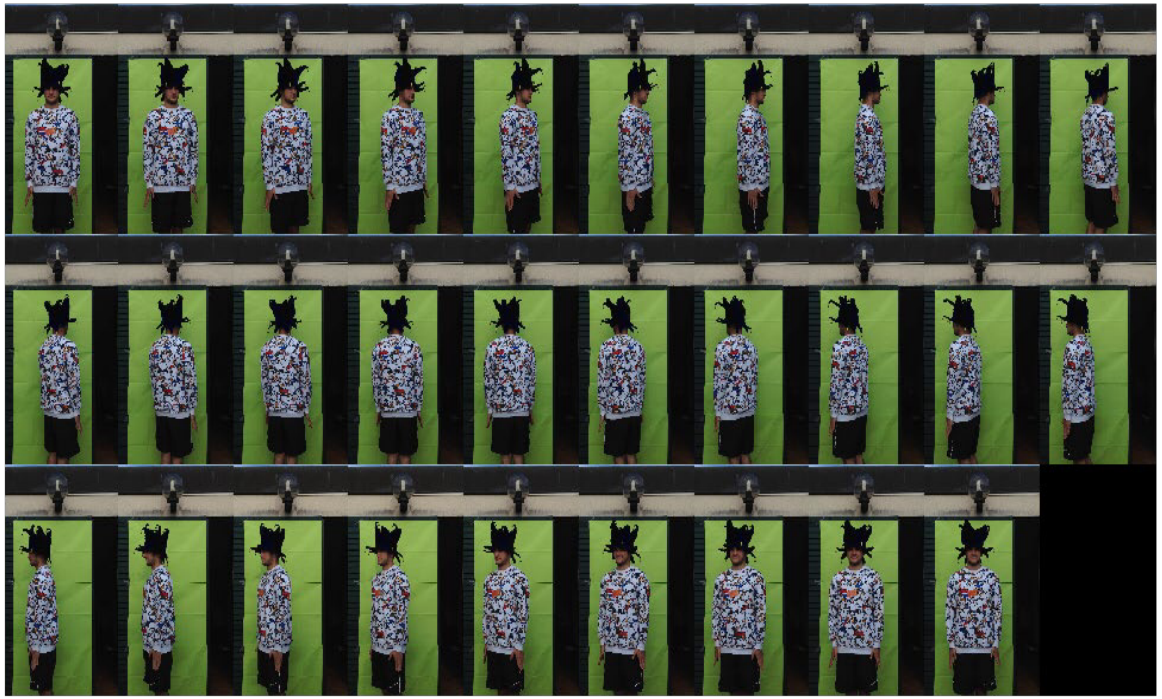
from 2 images from different view point... Match features that are corresponding...

point in 3D scene matched

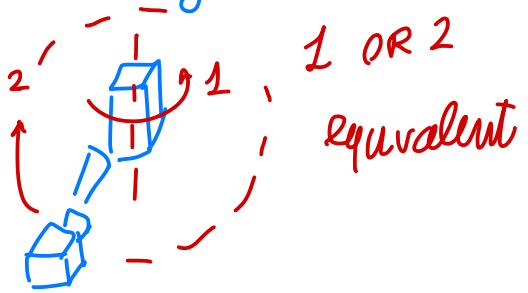


3D shape reconstruction

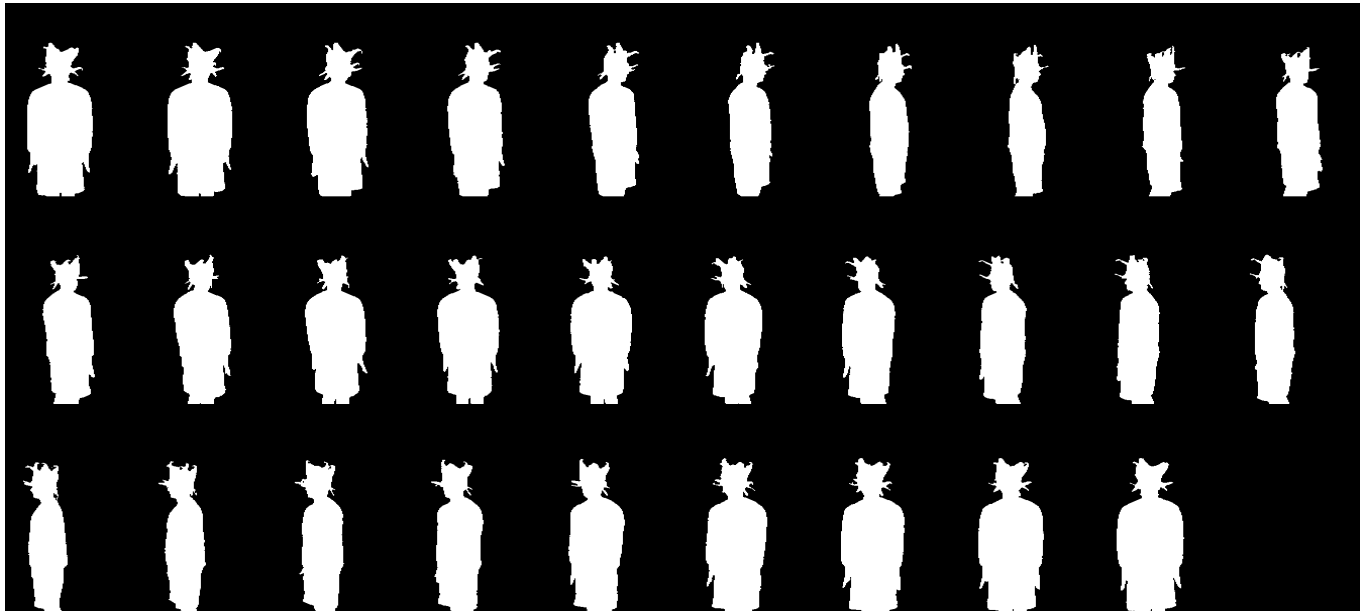
to reconstruct 3D shape...
NOT just 1 image, BUT several images are used!



rotating imagery by camera moving OR person moving...



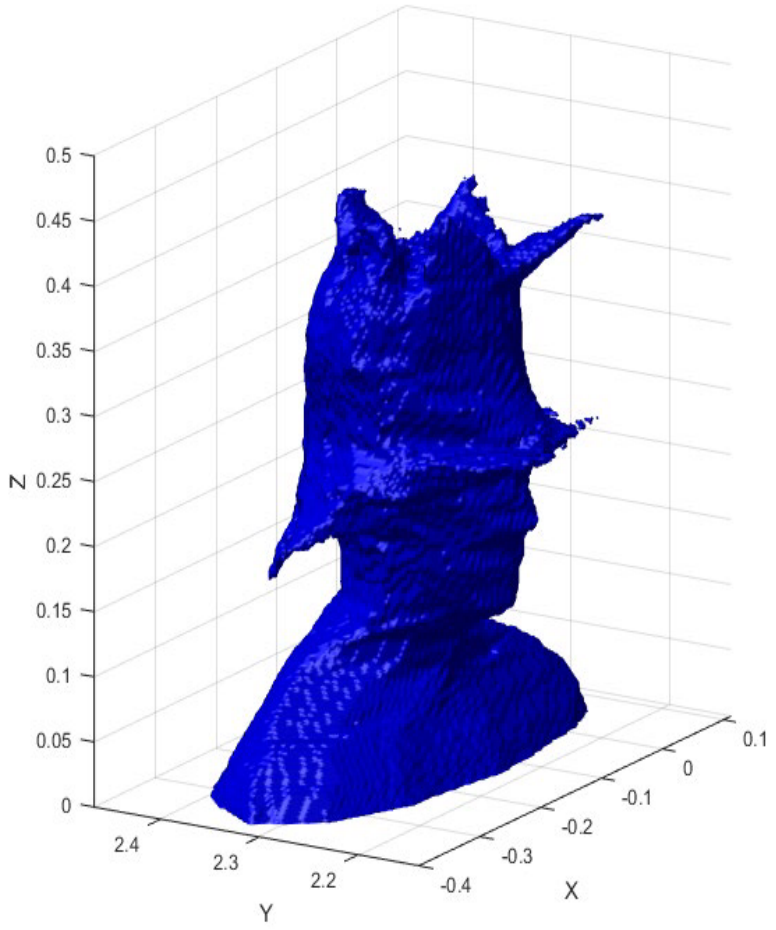
extract silhouette / between object and background

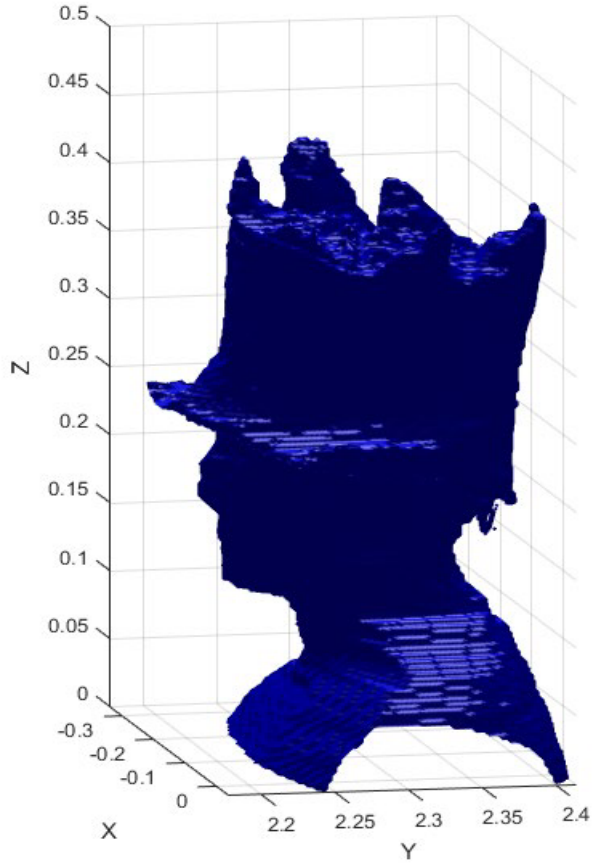


⇓

to extract 3D shape of object

3D shape reconstruction!





you don't need exact knowledge of camera position! Person moving around object!

extract features/points with frequency of aquisitions

from 1 image

next

small displacement, similar object in image ⇒

Microsoft

easy to do matching between features ⇒ triangulation of features

shape reconstruction

Object recognition

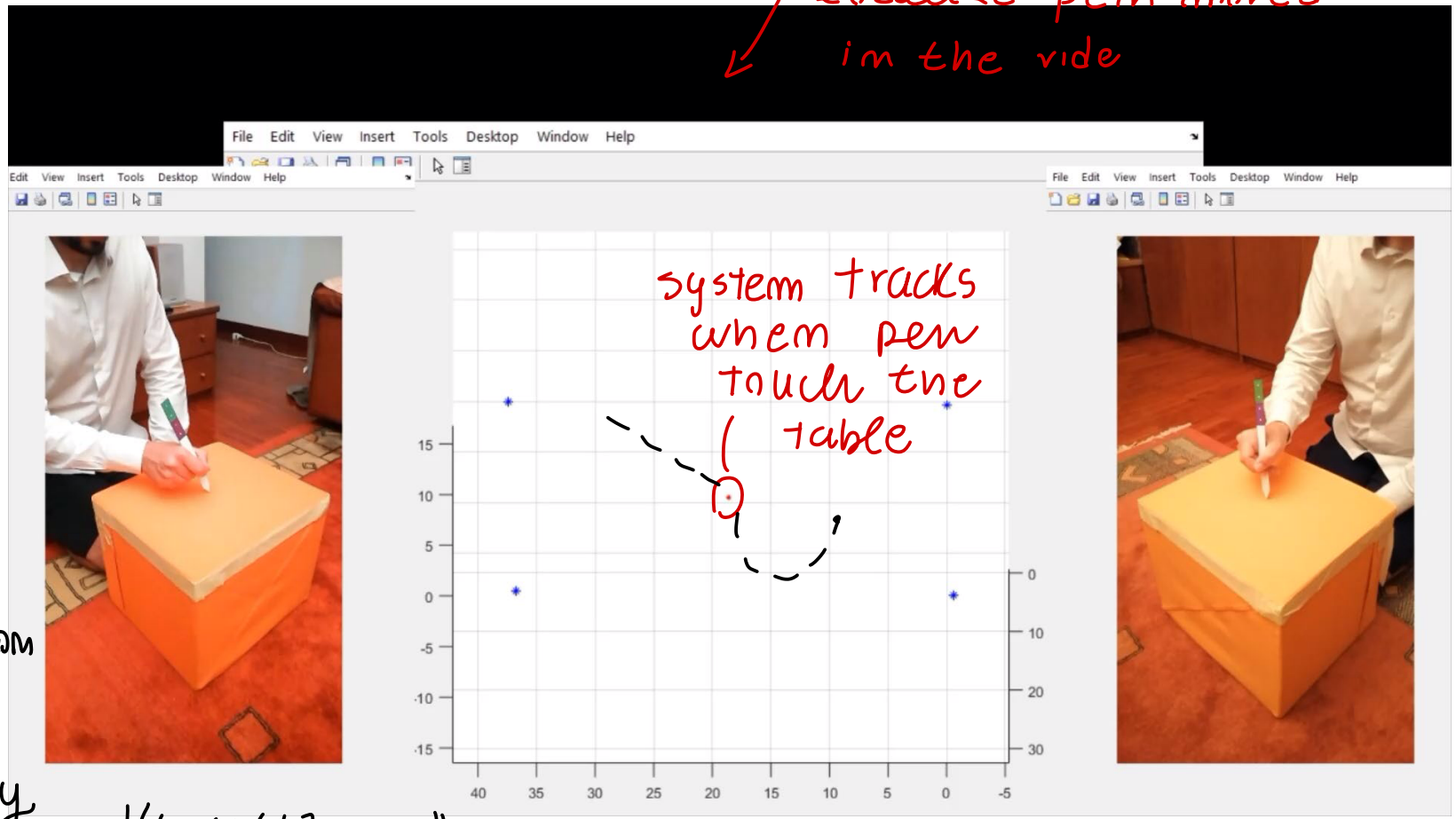
We expect some specific object on the image ⇒ from a model of each expected object

image → identify model object in image

Object localization

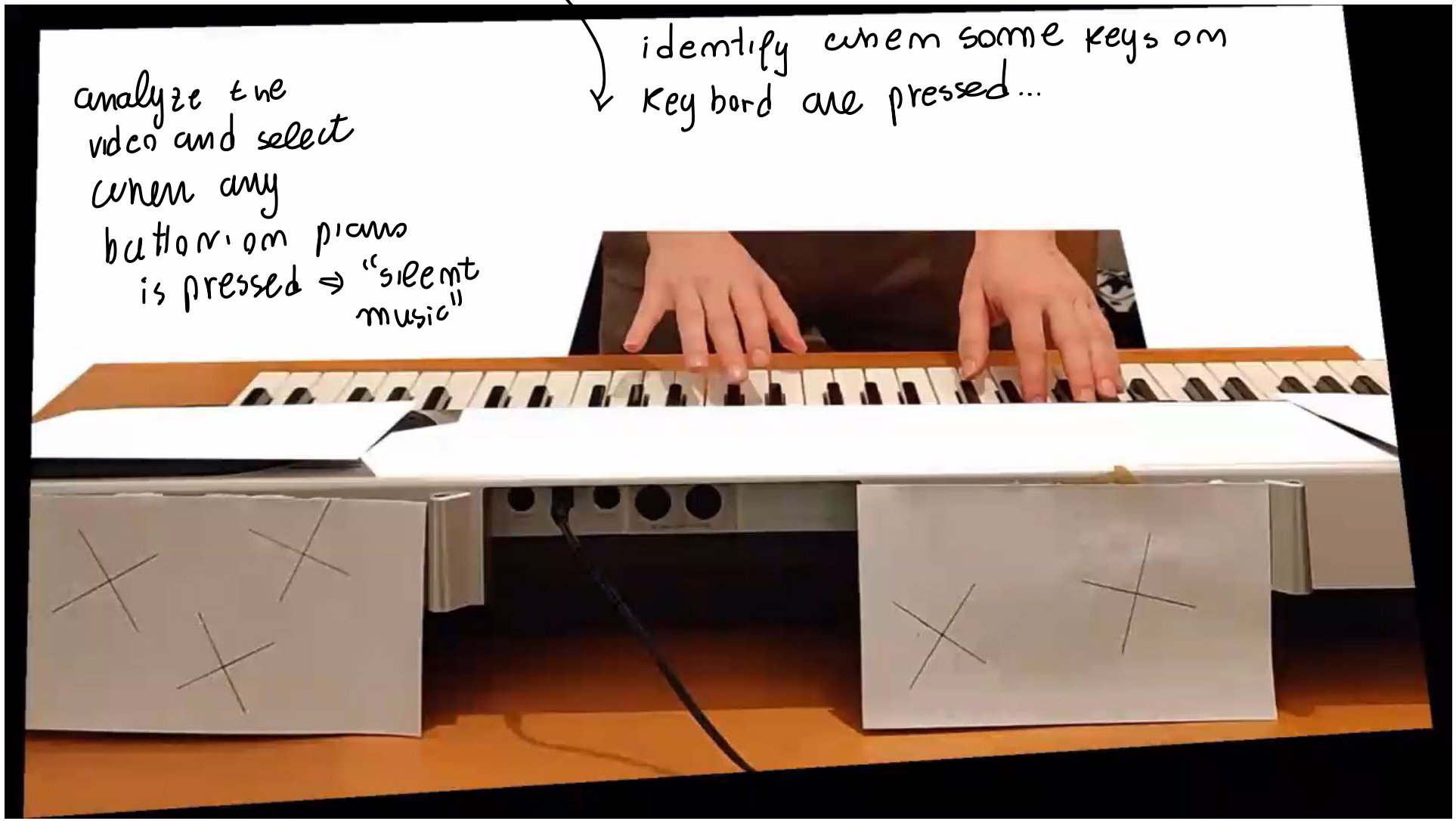
Localize pen moved in the vide

Distinguish when pen tip is on the table and when NOT, produce graphic representation of what is writing



"LOCALIZATION"

Action recognition



Human pose detection

segments of human body and estimate ORIENTATION

Keypoints detection

in acquired signal, img is NOT a perfect view... light energy ⇒ electrical : NOISE (thermal noise)

current collected is NOT exact representation of the scene...

DIFFICULTIES

1. • Noise, degradation, blur, and ILL-POSEDness

   → feature extraction and feature matching difficult

   (light/dark division is NOT so easy! → only when large gap

   ↑ difference in pixel can be noise or transition...)

   (slight gap can be confused with noise..

   noise affect quality of image analysis....)

2. • Projection -> information loss (dimensionality reduction from 3D to 2D)

   → 3D reconstruction is difficult

   ↓ geometric nature problem 3D → 2D loss 1 dimension...

   ↑ NOT enough information!

   plane on sky: it is a play OR a real plane? you lose a dimension! NOT from a single image possible to discriminate

→

we need to study:

extract features even in the presence of noise

↑

- IMAGE SIGNAL ANALYSIS
- PROJECTION GEOMETRY

↳ geometry / dimensionality loss, how to compensate information loss

im CHAPTERS.

TOPICS

"calibration" := operation done to measure, estimate characteristic camera parameter

- Introduction :

1) • Optical Sensors (camera) ← simplified model of optical sensor — How camera works

2) • 2D Projective Geometry (image plane, planar scenes) } → study geometry

3) • 3D Projective Geometry (3D scenes)

4) • Camera Geometry (3D → 2D projection) and Calibration ↘ combination of 2D×3D as camera works in a combination

feature extract.

5) • Digital image Filters and Image Morphology ↑ allows you to neglect noise.. remove noise artifact! → reduce noise effect

6) • Edge and Corner Detection

7) • Image Feature Detection and Description and Matching ↑ match features

8) • Model Fitting to Noisy Data := how to create high level models from noisy data ⇒ How to extract high level models

9) • Multi-view Geometry and 3D Shape Reconstruction

↓

needed for shape reconstruction. due to dimensionality reduction

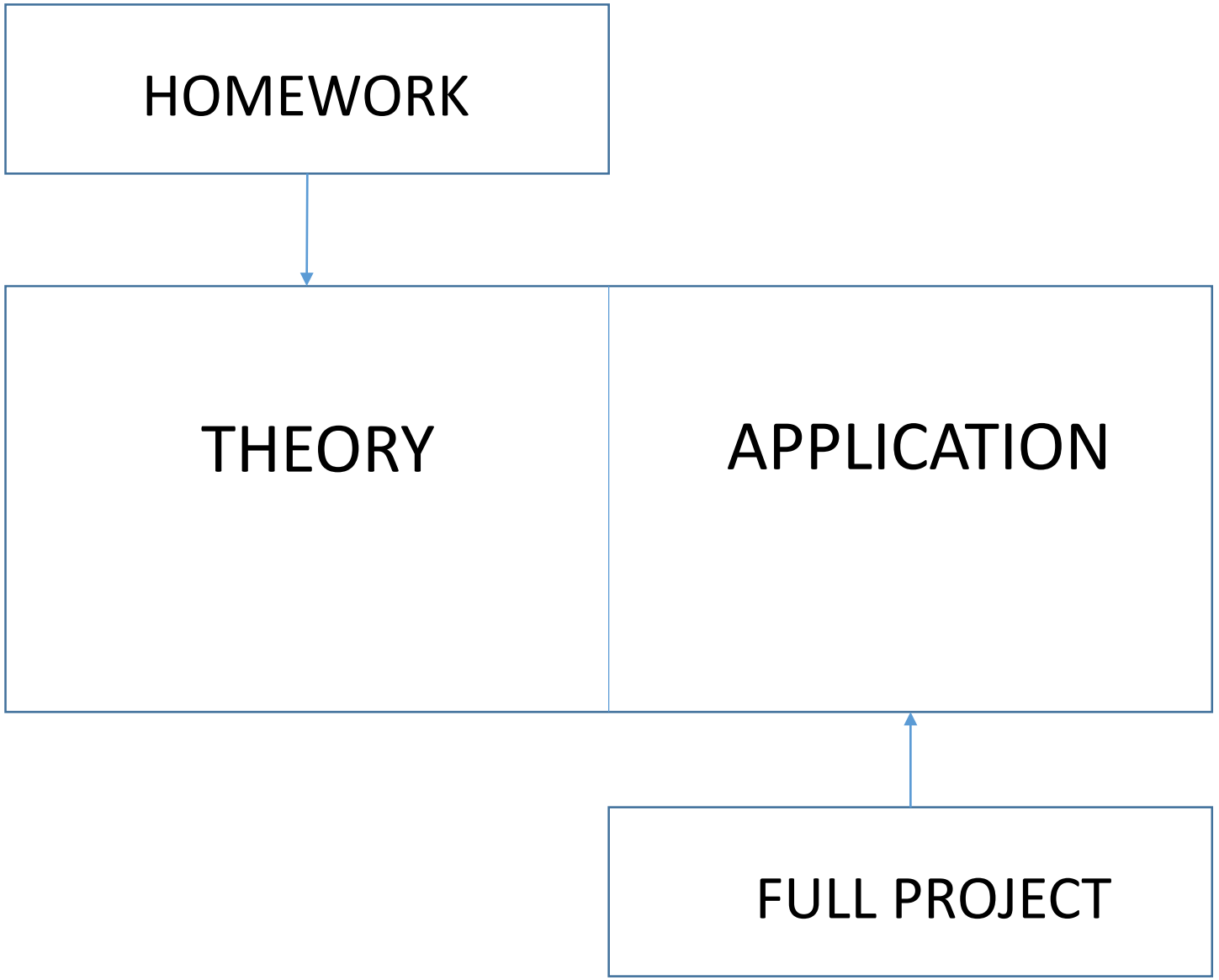
textbooks

- Hartley & Zisserman «Multi-view Geometry in Computer Vision»
- Forsyth and Ponce «Computer Vision: A Modern Approach»
- Course Slides on the WeBeep Platform

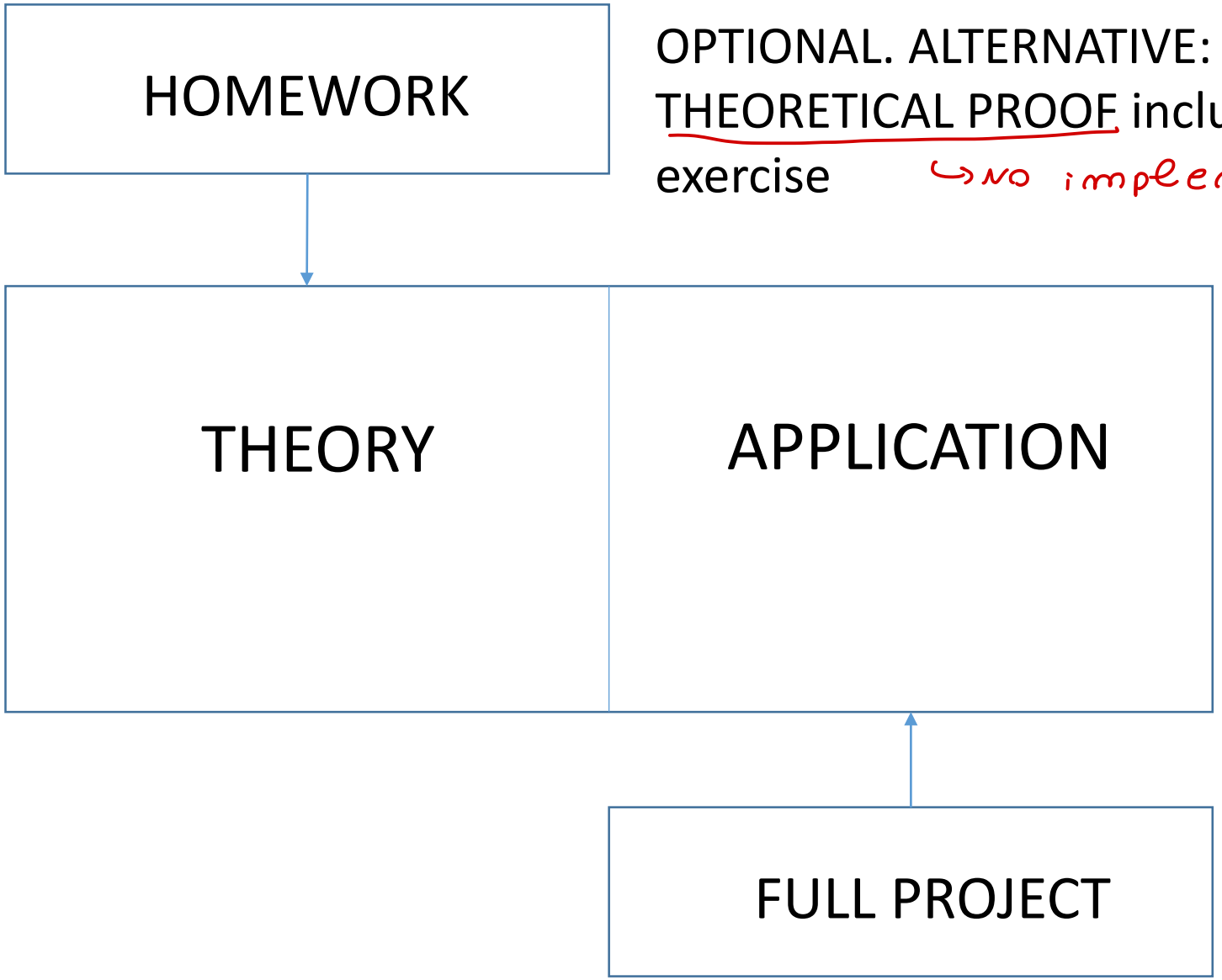
STANDARD EXAM: 2 parts → up to 30 cum laude

THEORY	APPLICATION

STANDARD EXAM: 2 parts → up to 30 cum laude



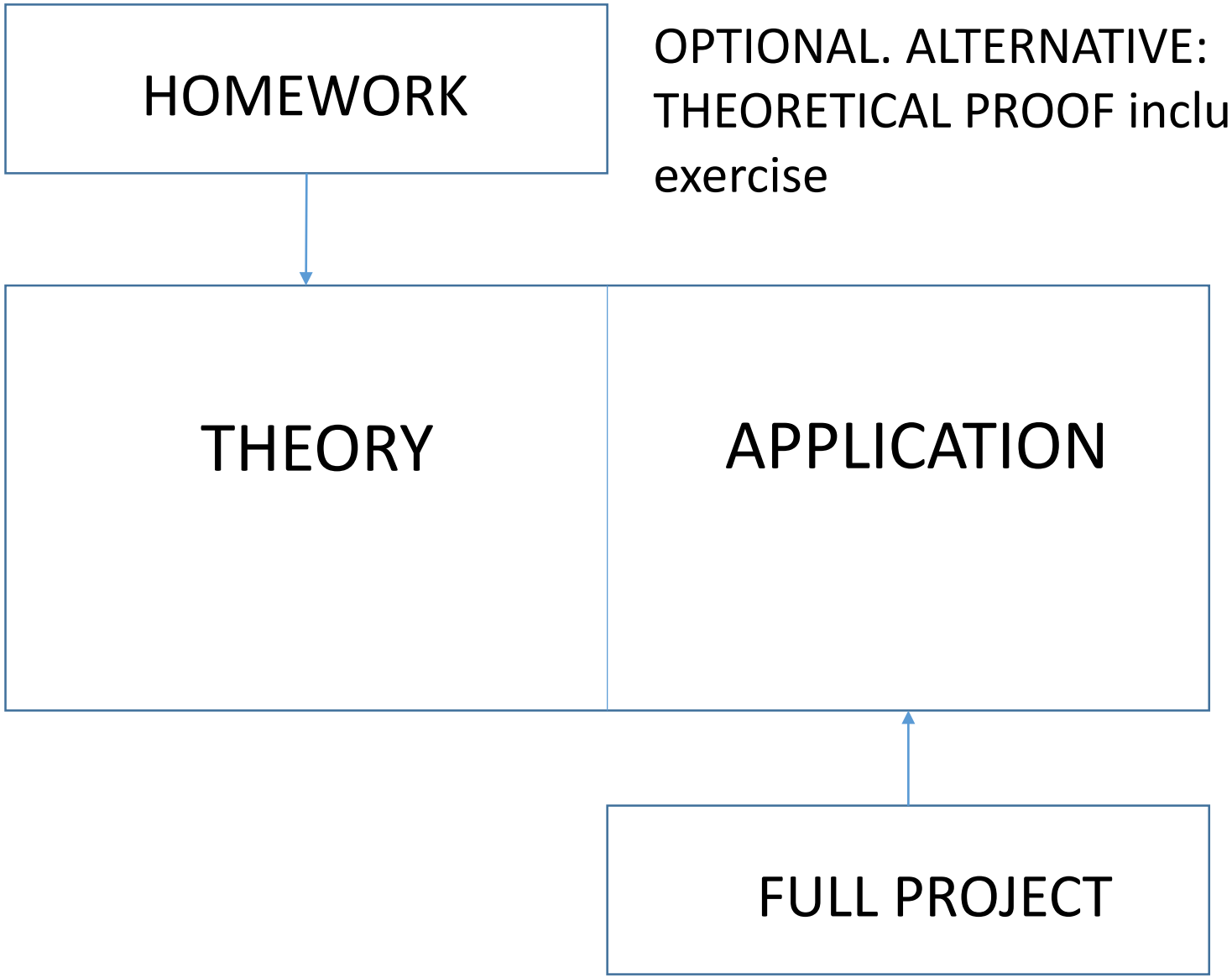
STANDARD EXAM: 2 parts → up to 30 cum laude



OPTIONAL. ALTERNATIVE:
THEORETICAL PROOF including difficult exercise

↪ no implementation SW

STANDARD EXAM: 2 parts → up to 30 cum laude



STANDARD EXAM → up to 30 cum laude

- individual HOMEWORK on Theory:

  a long, **difficult** exercise about (usually) 2D reconstruction, camera calibration and localization + a little bit of image feature extraction

  EVALUATION BASED on (i) SUBMISSION and (ii) REPORT

  OR, as an **alternative**,

  THEORETICAL PROOF including **difficult** exercises

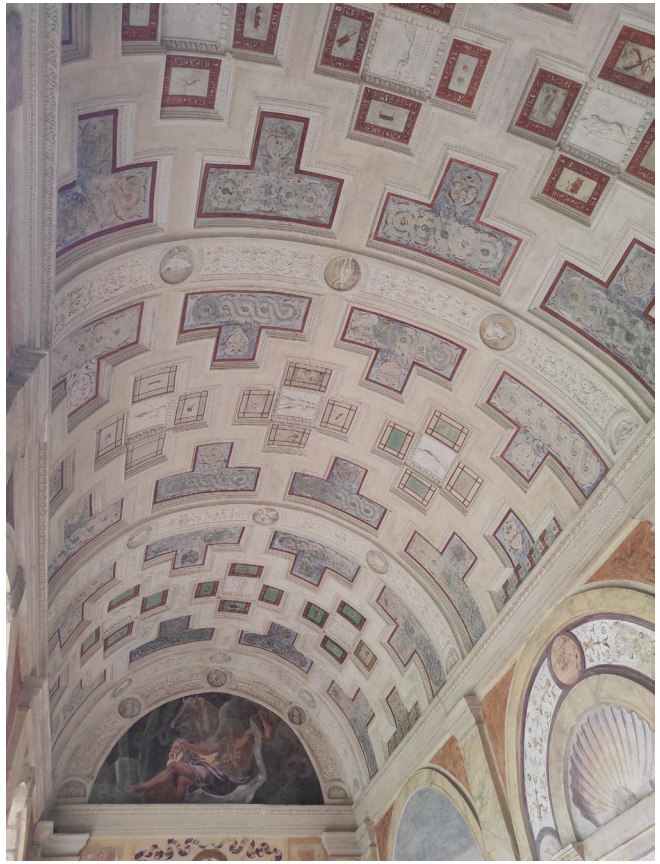
- either individual or team FULL PROJECT:

  an individual or team work involving design, implementation and experimental activity

  EVALUATION BASED on

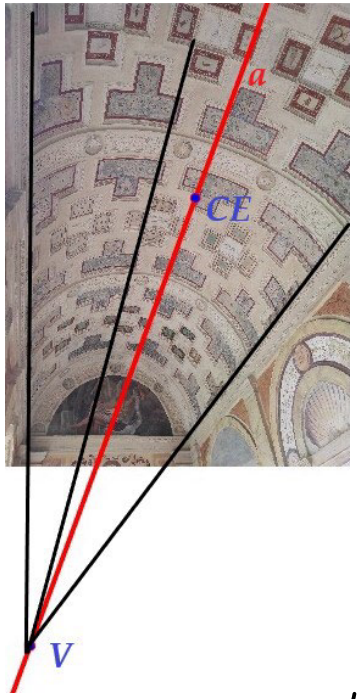
  (i) SUBMISSION, (ii) REPORT and (iii) ORAL DISCUSSION

Last year homework

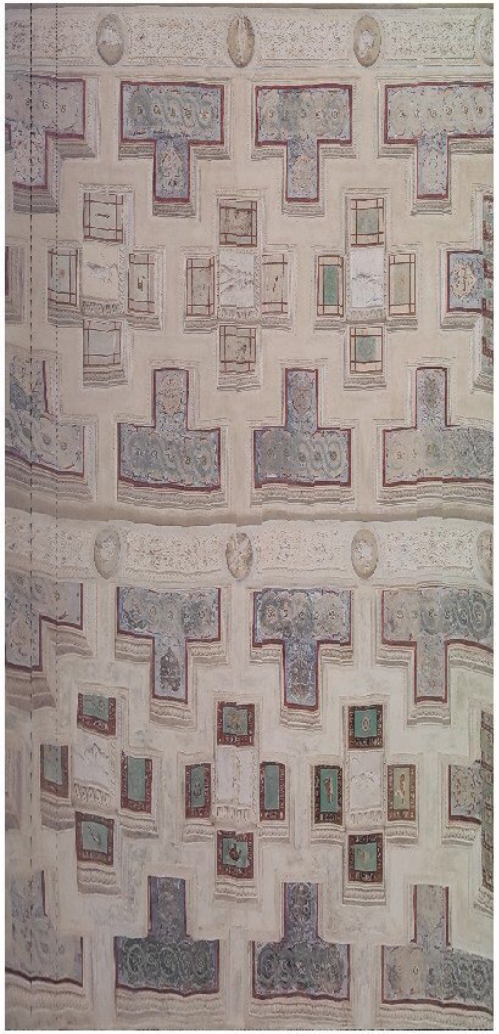


Palazzo Te - Mantova
cylindric vault

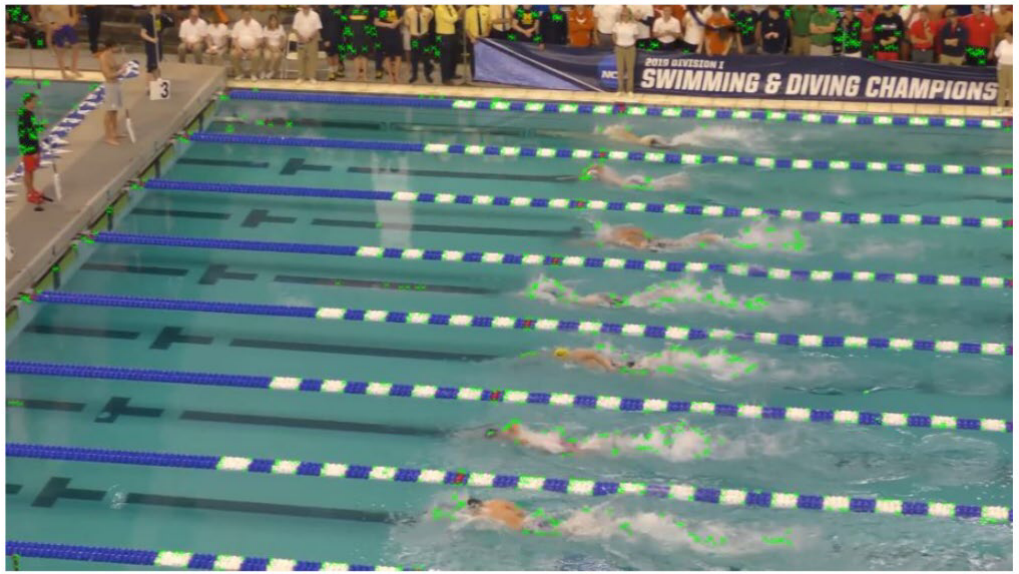
Knowing it is cylindrical...
→ you know something more. you reconstract

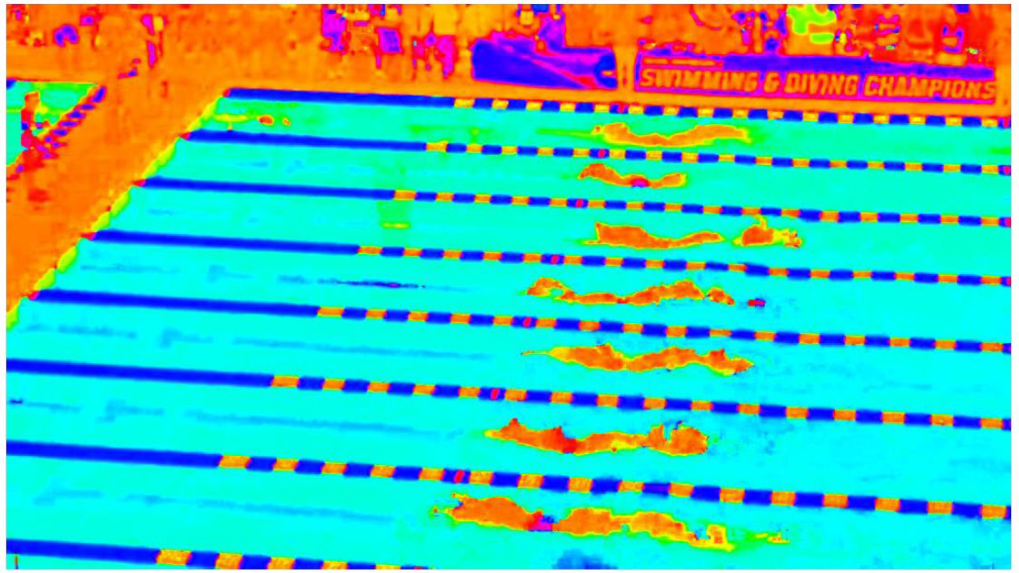


+develop theory to solve it

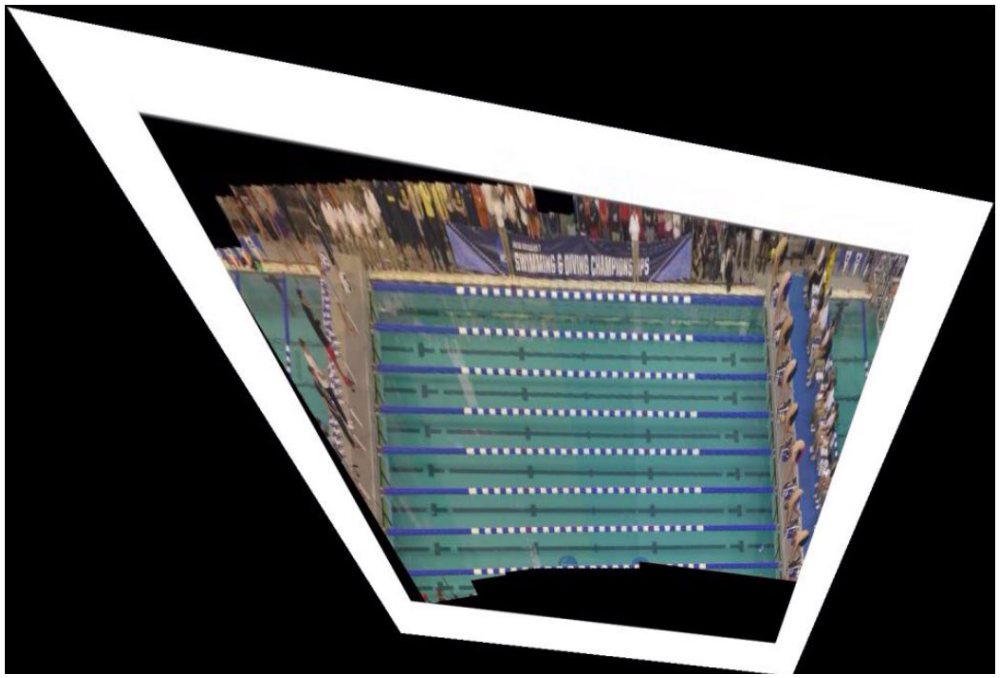


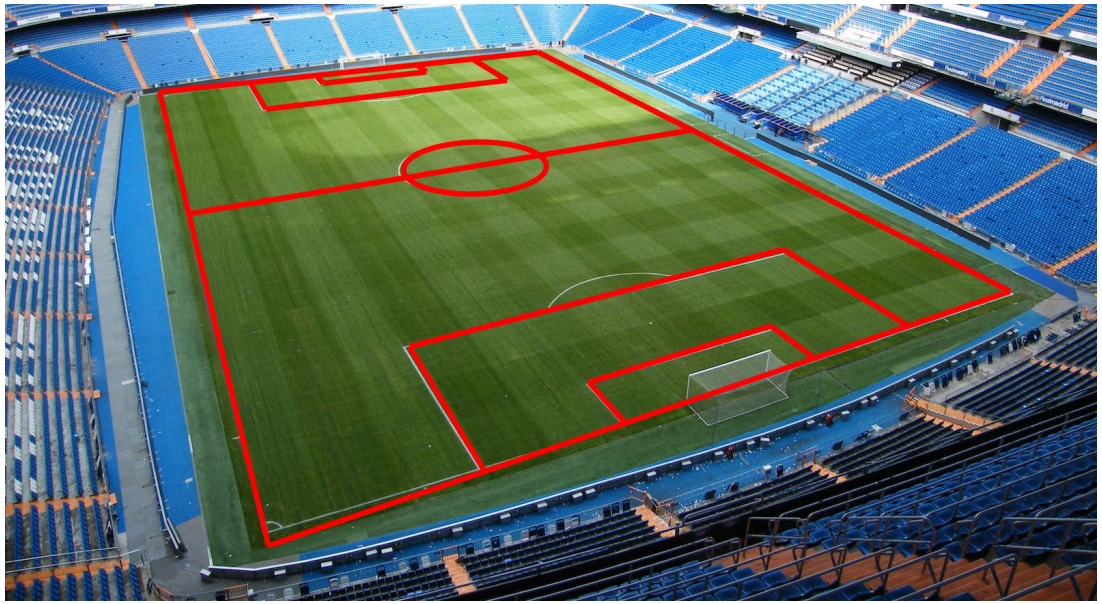
some past projects by your colleagues





identify/measure people swimming trajectory

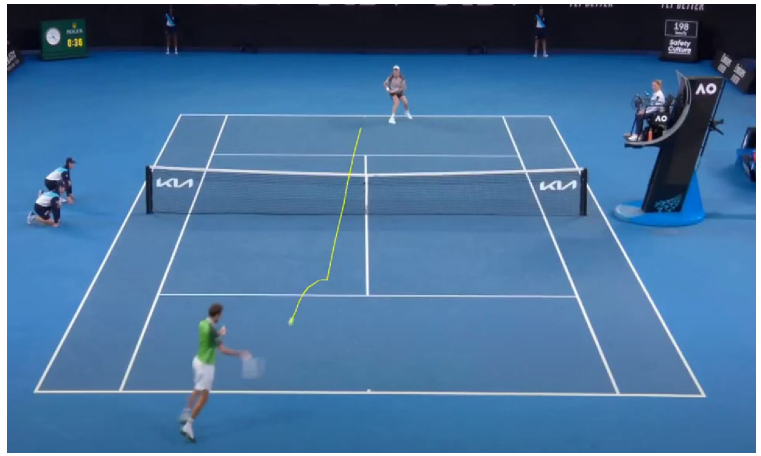


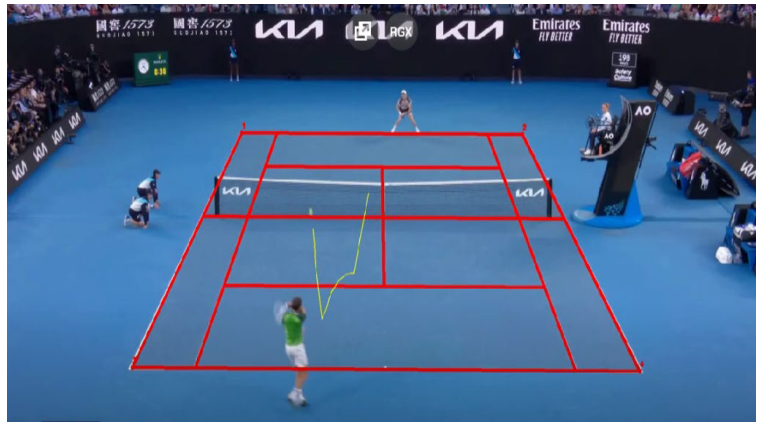


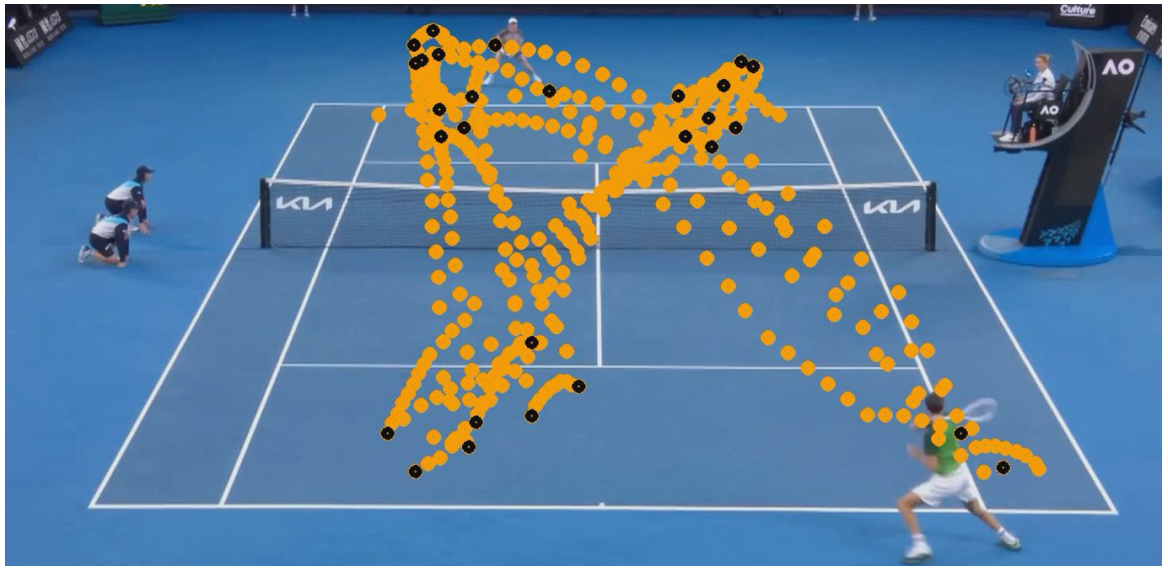
Identify / complete
soccer fieldlines

3D trajectory of tennis ball

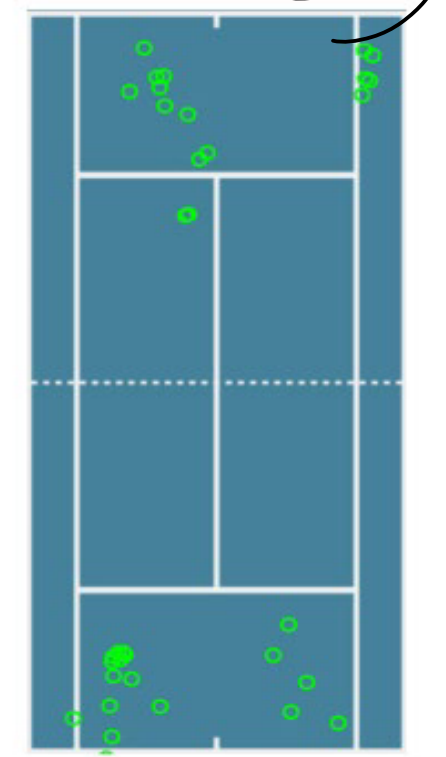
identify in 3D trajectory







correct bounce on ground racket



Object tracking

basketball tracking

Robust tracking

VII. Object tracking

Robust tracking

HOMEWORK

- **HOMEWORK** on Theory:

- A LONG AND DIFFICULT EXERCISE
- INDIVIDUAL (no collaboration)
- NOT MANDATORY (alternative: THEORETICAL PROOF with EXERCISES)
- PUBLISHED END NOVEMBER
- UNIQUE DEADLINE: within the 1° EXAM CALL (about end of January)
- but still VALID IN FUTURE CALLS AND YEARS
- DEVELOPED IN MATLAB
- EVALUATION based on (i) SUBMISSION and (ii) REPORT

EXAM

- **PROJECT**:

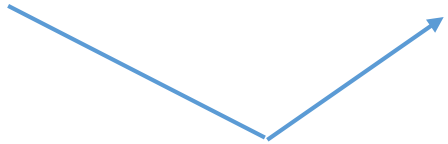
- TEAMS of 1, 2, or 3 STUDENTS (collaboration is welcome)
- MANDATORY IN STANDARD EXAM MODE
- PUBLISHED MID DECEMBER,
- DEVELOPED IN (ALMOST) ANY PROGRAMMING LANGUAGE
- MANY DEADLINES = TWO DAYS BEFORE EACH EXAM CALL
- EVALUATION based on
- (i) SUBMITTED MATERIAL
- (ii) pdf REPORT and ppt PRESENTATION
- (iii) DISCUSSION

STANDARD EXAM: EVALUATION → 30L

CHOOSE BETWEEN TWO POSSIBILITIES :

1° THEORY (Homework* or Theoretical Proof including difficult exercise) 10L/30 + FULL PROJECT 20L/30

2° HOMEWORK* good evaluation (> 7/10) Pass /Fail + FULL PROJECT 30L/30



* the evaluation of the homework is saved for future calls: you don't have to deliver it twice

NON-STANDARD EXAM: EVALUATION → ≤30
NO LAUDE

THREE POSSIBILE «ESCAPE STRATEGIES»:

1° THEORY (Homework* or Theoretical Proof including difficult exercise) **10**/30 + **SHORT** PROJECT **20**/30

NO PROJECT

2° HOMEWORK* **9**/30 + EXTENDED PROOF (including difficult exercises) **18**/30

NO HOMEWORK & NO PROJECT

3° EXTENDED PROOF (including difficult exercise) → max total: **25**/30

* the evaluation of the homework is saved for future calls: you don't have to deliver it twice

Project evaluation:

maximum score (20/20) is NOT the DEFAULT outcome.

Here are some minimal requirements:

- Avoid unnecessary assumption or restrictions
- Awereness of the theory thought in the course: avoid to re-invent (often worst) solutions
- Awereness of the theory thought in the course: capability to autonomously apply theory to derive solutions
- Autonomy in the retrieval and/or construction of the needed data sets
- Experimental results based on real images, not only simulated ones
- Develop a solution with a certain degree of generality, without tailoring it to just a single case (e.g., single video or image)