# COMPUTER VISION - PROJECT 3

*Computer Vision 2025-26*
*F. Lincetto, P. Zanuttigh, M. Pavan, L. Marx*

***Topic***: Multi-view Image Sequence Estimation
***Goal***: Given an unordered set of images captured sequentially, estimate the correct sequence order by detecting similarities between pairs.

***Scenario***
In this project, you will work with real images captured by a set of heterogeneous imaging sensors including not only cameras but also multi-spectral, near-infrared and polarized sensors. All these sensors were mounted on a rig and used to capture frame of an object-centric scene. The rig was moved around an object and for each point of view a frame per sensor was taken. Therefore, the frames were taken sequentially. You will be provided with a couple of scenes, an easy one and a more challenging one. For the project you can start by focusing on color (RGB) data.

You can download the scenes at:

- https://lttm.dei.unipd.it/downloads/prj2025/easy.zip
- https://lttm.dei.unipd.it/downloads/prj2025/challenging.zip
- The data is also part of a recently published project (*MultimodalStudio*), you can get from https://lttm.github.io/MultimodalStudio/ the full dataset.

***Your Task:***

Write a Python program which includes the following steps:

1. Color Image Loading
   Start by loading the unordered set of RGB images of the easy scene upper loop (the images are captured at two different heights from the ground; there is also a complete lower loop but focus on the upper one for the moment). All the images were captured with the same exposition and the same white balancing. You may want to pre-process the images by reducing noise or enhancing the image appearance.
2. Feature Extraction
   For each image, extract meaningful features. You can use SIFT, ORB, or other feature descriptors of your choice.

3. Feature Matching
   Exploiting the per-image feature previously extracted, compute matches between different image pairs
4. Sequence Estimation
   Based on the feature matches between different image pairs, try to define a set of criteria to determine which is the most likely next frame in the sequence. Mind that the sequence is a close loop. You may want to consider which is the pair with the most matches but also mind that it is not guaranteed that the pair with most matches is made of two consecutive images.
5. Repeat with the Challenging Scene
   Repeat the assignment with RGB images of the challenging scene, upper loop.

***Optional Tasks***

You can also try one or more of these optional tasks. They are slightly more challenging, but they can grant you a higher score if properly solved.

- Repeat the task with Near-Infrared images
  Near-Infrared frames capture information carried by the visible and near-infrared spectrum, up to 1000 nm. Therefore, Near-Infrared frames may contain different features, thus few adjustments may be needed.
- Repeat the task with Multi-spectral images
  These Multi-spectral frames have 9 channels and a lower resolution. Therefore, the extracted features may be more difficult to match. You can try to employ variants to deal with this kind of frames.
- Repeat the task by including also the lower loop
  Until now we worked only with frames belonging to the upper loop. Let's try to include also the lower loop. The lower loop is captured before the upper loop, thus images of the lower loop come before images of the upper loop in the sequence. Mind that a frame of the upper loop and a frame of the lower loop may have many features in common, but they can't be subsequent frames (expect in one case!). For this assignment, assume that you don't know which frames belong each loop, otherwise computing the upper loop, then separately the lower loop and finally concatenating them would be easy.

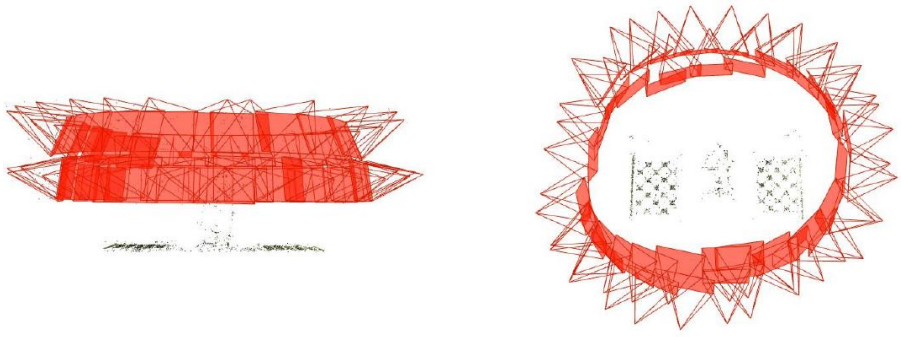

Figure 1. Overview of the RGB camera poses in a sample scene.