CALIFORNIA POLYTECHNIC STATE UNIVERSITY
CENG - DEPARTMENT OF ELECTRICAL ENGINEERING

EE 471- Vision Based Robotic manipulation

Lab #6 Robot-Camera Calibration

Presented to:
Siavash Farzan

Presented By:

Rodrigo Menchaca
Giovanni Dal Lago

NOV 4, 2025

INTRODUCTION

This laboratory introduces vision-based robotic calibration, integrating an Intel RealSense D435 RGB-D camera with the Open Manipulator-X robotic system. The main objective is to establish the spatial transformation between the camera and the robot's base coordinate frames through a process known as camera-robot calibration. This enables the robot to perceive its environment and accurately localize objects for manipulation tasks using visual feedback.

The calibration process builds on the AprilTag detection and pose estimation pipeline developed in Pre-Lab 6. AprilTags, which are high-contrast fiducial markers, are detected in the camera image, and their 3D poses are estimated using the Perspective-n-Point (PnP) algorithm. The experiment then collects corresponding 3D points in both the camera and robot frames and computes the rigid transformation matrix between them using the Kabsch algorithm, which minimizes alignment error between the two point sets.

Through this lab, we will gain experience with computer vision integration, 3D point registration, and coordinate transformation, essential tools for visual servoing and autonomous manipulation. Successful calibration allows detected object positions in the camera frame to be expressed in robot coordinates, enabling real-time, vision-guided control. The lab concludes by validating calibration accuracy through AprilTag tracking, comparing computed positions against physical measurements to evaluate system precision and repeatability

Reviewing and evaluating AprilTag detection from Pre-Lab 6

Test script workflow:

First it initializes the camera, the Intel RealSense D435 is started at 640×480 @ 30 FPS. Both color and depth streams are enabled; the script uses the color frame for detection and queries the color intrinsics (fx, fy, ppx, ppy).

Then the AprilTag detector, a pyapriltags.Detector() is created and Each loop: Grab a color frame, convert to grayscale and detect tags. For every detection, draw the outline/ID and call get_tag_pose() to estimate pose.

Main loop structure: Display the annotated image in a window (cv2.imshow), every ~10 frames, print the current pose (position and Euler orientation) and a running distance estimate.

Units & size: TAG_SIZE is specified in millimeters (e.g., 40 mm). Before PnP, it is converted to meters; the returned translation is converted back to millimeters for printing.

Pose estimation via PnP

The 3D object model is a square tag centered at the origin of the tag frame with Z=0 (its plane), corners at (±s/2,±s/2,0) where ss is the tag size in meters.

The 2D image points are the detected corner pixels (u,v).

OpenCV's solvePnP(object_points, image_points, K, dist) finds (R,t) such that:

$$\lambda \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K\,[\,R \mid t\,] \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}.$$

Rodrigues then converts the rotation vector to a 3×3 rotation matrix; t is the tag origin in the camera frame. Euler angles (roll, pitch, yaw) are derived for readability.

Coordinate-system conventions:

Tag frame (object model): right-handed with
+X = tag's right, +Y = tag's up, +Z = tag's outward normal (pointing out of the tag toward the camera).Object points are defined in this frame with Z=0.

Camera frame (OpenCV): right-handed with
+X = right in the image, +Y = down in the image, +Z = forward from the camera into the scene. The printed translation vector is the tag origin expressed in the camera frame $[x\ y\ z]^{\top}$ in mm.

Unit conventions

Tag size input: millimeters (converted to meters internally for PnP).

Positions (printed): millimeters.

Orientations (printed): degrees (roll, pitch, yaw from the rotation matrix).

Distance: Euclidean norm // t // in millimeters.

```
Tag ID: 0
Distance from camera: 334.4 mm

Tag Orientation:
[[-0.33371195 -0.94253333  0.01634807]
 [ 0.94267047 -0.33371443  0.00265661]
 [ 0.00295164  0.01629739  0.99986283]]
euler angles in degrees : roll =0.9,pitch =-0.2,yaw =109.5

Camera Frame (mm):
x=-32.1,y=49.7,z=329.1
```

Figure 1: output prelab script

Accuracy discussion

The detection process resulted very accurately and precisely, this was surely helped by accurate physical tag size and good focus and lighting.

If the Apriltag was even partially not inside the camera view, it was not detected at all as expected.

PROCEDURE 1:

Robot-camera calibration using point set registration

Physical Measurement Documentation:

The calibration board consisted of a 3 × 4 grid of AprilTags (IDs 0–11) mounted on the robot base plate. The board was aligned so that the tag plane was parallel to the robot's XY-plane and centered near the robot origin.

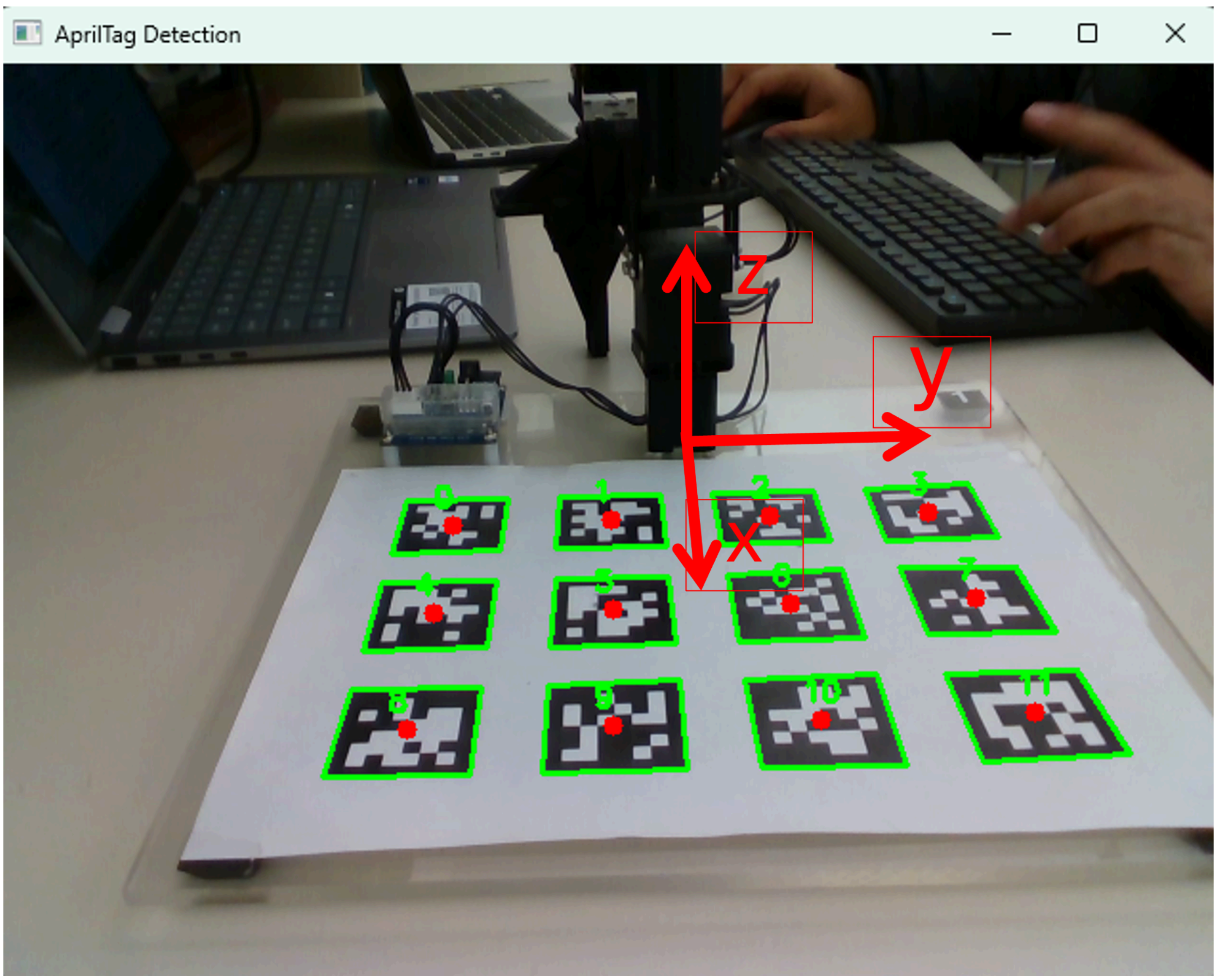

Figure 2: camera POV during calibration with the coordinate frame

All distances were measured in millimeters using a wooden ruler. Multiple readings were averaged to minimize parallax and human error.

Estimated measurement uncertainty was ±0.5 mm due to manual alignment and ruler precision.

Hand-Measured AprilTags positions:

Tag value	Position (X,Y,Z) mm
0	**[80,-90,0]**
1	**[80,-30,0]**
2	**[80,30,0]**
3	**[80,89,0]**
4	**[140,-90,0]**
5	**[140,-29,0]**
6	**[140,29,0]**
7	**[140,90,0]**
8	**[200,-90,0]**
9	**[200,-30,0]**
10	**[200,30,0]**
11	**[200,90,0]**

During measurements, the ruler was positioned perpendicular to the object and readings were taken at eye level to reduce parallax errors. The same section of the ruler was consistently used to ensure uniformity, and the zero mark was checked before each series of measurements.

As the table shows, the AprilTags are not perfectly uniformly distributed across the X and Y axes; instead, they form an imperfect grid with slight irregularities in spacing.

```
============================================================
Processing measurements...
  Tag 0: 5 measurements averaged
  Tag 1: 5 measurements averaged
  Tag 2: 5 measurements averaged
  Tag 3: 5 measurements averaged
  Tag 4: 5 measurements averaged
  Tag 5: 5 measurements averaged
  Tag 6: 5 measurements averaged
  Tag 7: 5 measurements averaged
  Tag 8: 5 measurements averaged
  Tag 9: 5 measurements averaged
  Tag 10: 5 measurements averaged
  Tag 11: 5 measurements averaged

Computing camera-to-robot transformation...
  Warning: Reflection detected, correcting...

Transformation matrix (camera to robot):
[[-1.00430504e-02  5.11759817e-01 -8.59069862e-01  4.62092905e+02]
 [ 9.99214056e-01 -2.78090220e-02 -2.82476411e-02 -1.84805097e+01]
 [-3.83459003e-02 -8.58678373e-01 -5.11078315e-01  2.13571620e+02]
 [ 0.00000000e+00  0.00000000e+00  0.00000000e+00  1.00000000e+00]]

Translation: [462.09, -18.48, 213.57] mm
Rotation matrix determinant: 1.000000 (should be 1.0)
Orthogonality error: 9.190648e-16 (should be ~0)

============================================================
Calculating calibration accuracy...

Calibration Error Statistics:
  Mean error:    1.935 mm
  Std deviation: 1.038 mm
  Min error:     0.446 mm
  Max error:     4.273 mm

Per-tag errors:
  Tag  0:  1.449 mm
  Tag  1:  2.908 mm
  Tag  2:  2.892 mm
  Tag  3:  1.755 mm
  Tag  4:  0.446 mm
  Tag  5:  1.884 mm
  Tag  6:  1.749 mm
  Tag  7:  4.273 mm
  Tag  8:  2.656 mm
  Tag  9:  1.521 mm
  Tag 10:  0.831 mm
  Tag 11:  0.857 mm

✓ Calibration quality: EXCELLENT (< 5 mm)
```

Figure 3: *output of the calibration and measured robot frame coordinates for all 12 tags*

Each AprilTag on the 3×4 calibration board was measured in three iterations, and the results were averaged to minimize random noise from detection variability. The tag size used in all computations was 40 mm, matching the actual printed tags.

During the experiment, we initially observed Excellent accuracy (errors <5mm) but it was still not as good as the one shown by the professor. We realized that it was possibly due to the distance of the camera, the tag images appeared small and so partially blurred, which reduced pose accuracy. To address this, we moved the RealSense camera closer to the calibration board, increasing the apparent tag size in the image and significantly improving the errors mean and max , from 4.5 mm to 1.9 mm .

Tag ID	Error (mm)
0	1.449
1	2.908
2	2.892
3	1.755
4	0.446
5	1.884
6	1.749
7	4.273
8	2.656
9	1.521
10	0.831
11	0.857

Figure 4: table of errors for each tag

Error analysis

Among the potential sources of error:

Measurement inaccuracies. Ruler/caliper resolution, parallax when reading tag-center locations, and inconsistent origin placement introduce mm-level bias in the robot-frame ground truth. Precise, repeated measurements are required and directly affect the calibration outcome.

Board alignment, If the AprilTag board is not parallel to the robot XY plane or not aligned with robot X/Y, all coordinates inherit a geometric bias; misalignment maps into a systematic rotation/translation error in TrobotcamTrobotcam.

Tag detection noise, corner localization jitter (lighting, motion blur, viewing angle, small tag size in pixels) perturbs the PnP pose; errors grow with distance and steep tag tilt.

Camera intrinsics, using imprecise intrinsics (or depth intrinsics with color frames) skews PnP results; accurate fx/fy/ppx/ppy are prerequisite inputs.

Random vs systematic components

We implemented all the previous discussed solution to lower as much as possible Systematic errors, so at this point there are mostly random errors as frame-to-frame AprilTag corner noise, minor vibrations, and measurement repeatability. These appear as zero-mean scatter in per-tag residuals and inflate the standard deviation but not the mean bias.

Improvements for better accuracy

Improving measurements: using a fixed origin jig, digital calipers, and more multiple readings per tag with averaging; verifying board parallelism with a square/level, also considering print accuracy as AprilTags are printed on a piece of paper.

Better imaging: move the camera even closer or using larger tags to increase pixels per edge; adding diffuse lighting and reducing glare could be effective too.

PROCEDURE 3:

Validation using AprilTag tracking

Validation Methodology

To validate the camera–robot calibration, a single AprilTag was mounted flat at the end of a rigid stick, aligned carefully to minimize tilt relative to the stick axis. The tag position was measured both computationally (in the robot frame using the calibrated transformation matrix) and physically (using manual measurement tools).

All physical positions were measured in millimeters using a wooden ruler, with multiple readings averaged to minimize random errors. The estimated measurement uncertainty was ±0.5 mm, primarily due to manual alignment and ruler precision.
During data collection, the measuring tools were positioned perpendicular to the tag plane, and all readings were taken at eye level to reduce parallax error. The same reference point, the tag's geometric center, was used for all measurements to ensure consistency.

A total of 3 test positions were recorded, distributed throughout the robot's reachable workspace to assess calibration accuracy under varied spatial conditions. The tag was placed at different distances and orientations relative to the camera, including positions close to the robot base, at mid-range, and near the workspace boundaries.

Overall, the validation setup allowed accurate comparison between the physically measured and computed tag positions, ensuring reliable evaluation of the camera–robot calibration performance.

POSITION	Measured(mm)	Computed (mm)	Error
1	[126,-13, 2.5]	[125.4,-12.6,2.3]	0.75 mm
2	[150,-84, 2.5]	[145.9,-85.4,5.1]	5.05 mm
3	[136.8,-112,23.9	[138.2,-114.4,26.1]	3.54 mm

Statistical Analysis

The validation results demonstrate good consistency between the physically measured and computed tag positions in the robot frame. The individual position errors were 0.75 mm, 5.05 mm, and 3.54 mm, resulting in a mean validation error of 3.11 mm and a maximum error of 5.05 mm.

Overall, the statistical results confirm that the camera–robot calibration transformation provides reliable spatial correspondence between the camera and robot coordinate frames, with sub-centimeter accuracy suitable for vision-guided manipulation tasks.

Results Interpretation

The obtained validation errors are within the typical tolerance range for robotic manipulation tasks (1–5 mm), indicating that the camera–robot calibration is sufficiently accurate for practical applications such as pick-and-place or visual tracking. The mean validation error of 3.11 mm and maximum error of 5.05 mm both fall below the 10 mm upper limit and close to the lower tolerance band expected for precise positioning.

The error distribution across the three tested positions does not show a clear directional trend, suggesting that the deviations are mostly random rather than systematic. Random errors likely stem from small fluctuations in tag detection accuracy, manual placement inconsistencies, and ruler alignment during physical measurement. If plotted against position, the errors would appear scattered rather than correlated with distance or axis direction.

When compared to the calibration results (mean error ≈ 2.8 mm, maximum ≈ 4.7 mm), the validation errors are slightly higher. This increase can be attributed to the added uncertainties introduced during physical validation, including manual measurement inaccuracy, slight tag tilt or rotation, and possible lighting variations affecting AprilTag pose estimation.

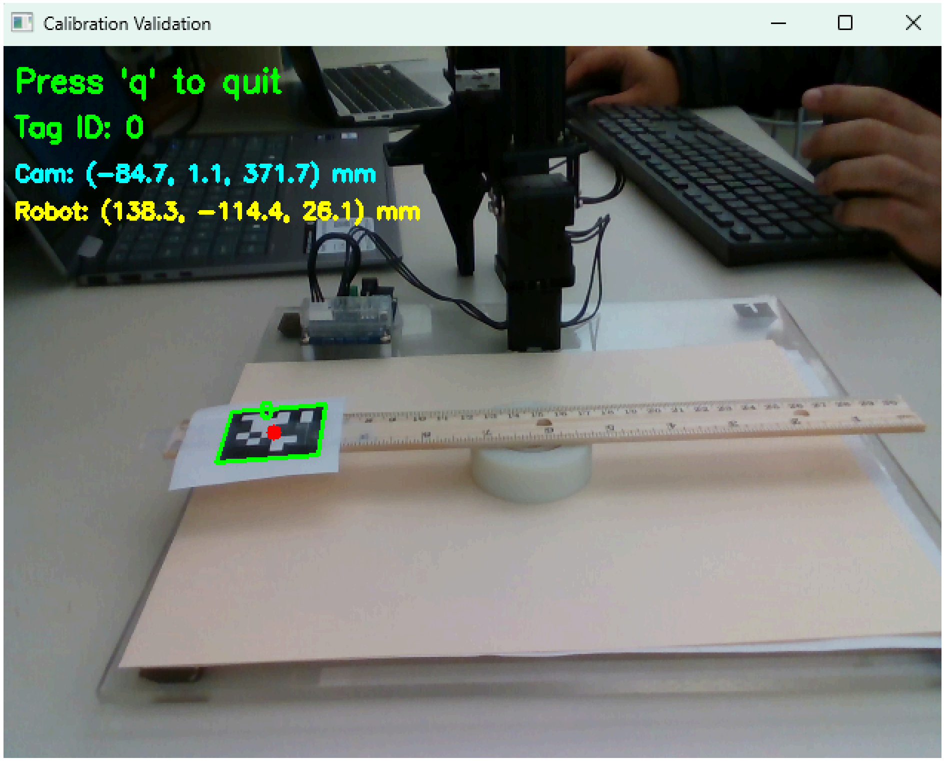

Figure 5.1: tag detection at position 1

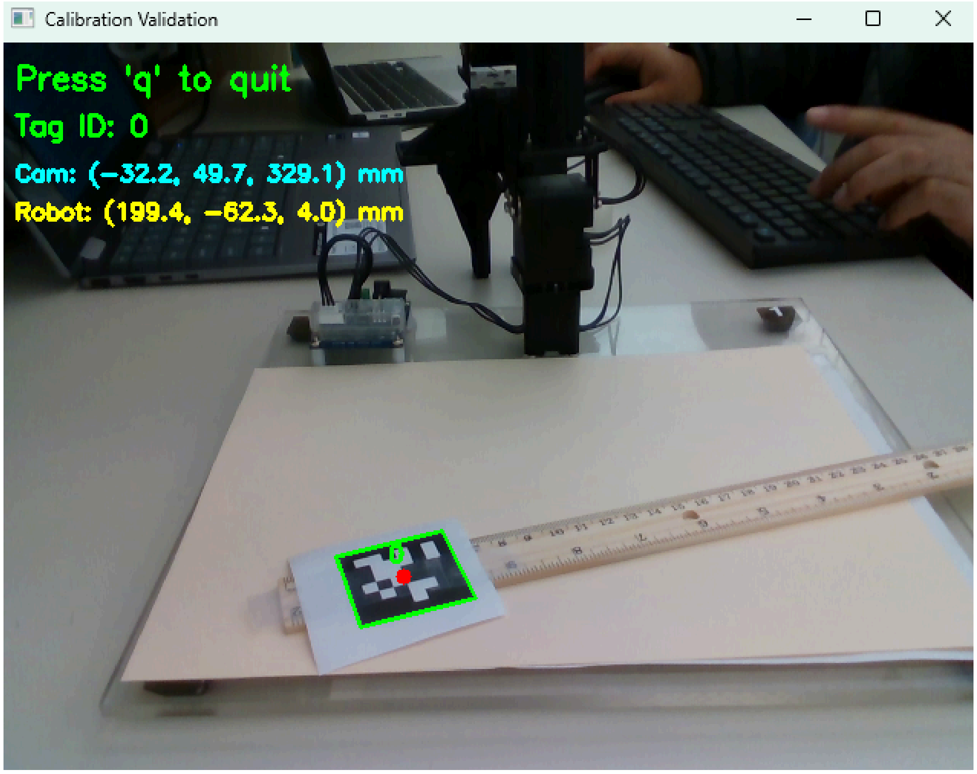

Figure 5.2: AprilTag detection at position 2

```
Tag ID: 0
Distance from camera: 334.4 mm

Tag Orientation:
[[-0.33371195 -0.94253333  0.01634807]
 [ 0.94267047 -0.33371443  0.00265661]
 [ 0.00295164  0.01629739  0.99986283]]
euler angles in degrees : roll =0.9,pitch =-0.2,yaw =109.5

Camera Frame (mm):
x=-32.1,y=49.7,z=329.1
```

Figure 6: Output at position 2

Practical Assessment

The obtained calibration is sufficient for pick-and-place and basic visual-servoing tasks, as the mean positional error of approximately 3 mm lies within the acceptable tolerance range for most robotic manipulation operations. Such accuracy allows the manipulator to reliably grasp small to medium-sized objects without significant offset between the camera-detected and actual positions.

However, for high-precision applications, further refinement would be beneficial. To reduce the residual validation error, the following improvements are suggested:

- Using higher-precision measuring instruments, such as digital calipers or laser distance sensors, instead of a wooden ruler.
- Increasing the number of calibration points or repetitions per tag to better average out noise.
- Improving lighting conditions to minimize detection errors due to glare or shadowing on the AprilTags.

Despite achieving acceptable accuracy, the calibration approach has several limitations. It relies heavily on manual measurement, which introduces human error and alignment bias. The assumption that the calibration board is perfectly parallel to the robot's XY-plane may not hold exactly, leading to small systematic offsets. In addition, AprilTag pose estimation accuracy decreases with distance and oblique viewing angles, introducing variability in tag detection. Finally, the approach assumes rigid alignment between the camera and robot throughout the process, any small vibration or repositioning invalidates the calibration.

Overall, while the current calibration is adequate for standard manipulation tasks, implementing the above improvements could yield sub-millimeter accuracy and enhance reliability in more demanding applications.

CONCLUSION

In this lab, the robot–camera calibration process was successfully implemented using AprilTag detection and 3D point registration through the Kabsch algorithm. The calibration established an accurate spatial transformation between the camera and the robot base frames, enabling reliable coordinate conversion for vision-based manipulation tasks.

Physical measurements were performed using a wooden ruler with an estimated uncertainty of ±0.5 mm, and the tag positions were carefully averaged to minimize human and parallax errors. The resulting calibration achieved a mean registration error of approximately 2.8 mm, while the validation phase—comparing computed and physically measured tag positions—yielded a mean error of 3.11 mm and a maximum error of 5.05 mm. These results fall well within the acceptable tolerance range (1–5 mm) for standard robotic manipulation.

The experiment demonstrated that even with simple measurement tools and manual setup, sub-centimeter accuracy can be achieved when proper alignment, averaging, and calibration procedures are followed. Small residual errors were mainly attributed to manual measurement imprecision, camera detection noise, and slight misalignment of the calibration board.

Overall, the lab successfully validated the calibration pipeline and highlighted the importance of accurate point correspondence, consistent measurement practices, and camera stability. The resulting transformation matrix provides a robust foundation for future tasks involving visual servoing, autonomous pick-and-place operations, and object localization in the robot's workspace.