AprilTag detection for building measurement

Kepa Iturralde[a], Juncheng Shen[a], Thomas Bock[a]

[a] Chair of Building Realization and Robotics, School of Engineering and Design, Technical University of Munich, Germany

E-mail: kepa.iturralde@br2.ar.tum.de

Abstract

Building façade renovation with prefabricated panels requires the installation of supporting connectors. Obtaining the exact locations and angles of the connectors is currently a time-consuming activity. The new method presented in this paper uses AprilTag markers as targets on the wall and measures them based on a photogrammetry concept. The 3D coordinates of the targets are estimated by our algorithm, and then we can obtain the location and direction of each connector. Current results are promising but still need some improvement. Our goal is to measure the exact position and angle of each marker, and the error should not exceed 1mm at about 10m away from the building. This objective is not reached in this paper, but latest advances show very close results.

Keywords –

AprilTags, detection, measure

1 Introduction

When placing prefabricated modules with solar panels on existing buildings, it is important to know the location of the connectors or anchors so the modules fit them with high accuracy. Measuring the location of connectors with Total Station is a time consuming process. Therefore, visual or photograph based methods can be a solution to capture the location of the connectors in an automated manner.

Visual fiducials are artificial landmarks designed to be easy to recognize and distinguish from one another. Although related to other 2D barcode systems such as QR codes [1], they have significantly goals and applications. With a QR code, a human is typically involved in aligning the camera with the tag and photographs it at fairly high-resolution obtaining hundreds of bytes, such as a web address. In contrast, a visual fiducial has a small information payload (perhaps 12 bits) but is designed to be automatically detected and localized even when it is at very low resolution, unevenly lit, oddly rotated, or tucked away in the corner of an otherwise cluttered image. Unlike 2D barcode systems in which the position of the barcode in the image is unimportant, visual fiducial systems provide camera-relative position and orientation of a tag. Fiducial systems also are designed to detect multiple markers in a single image [2]. Visual fiducial systems are perhaps best known for their application to augmented reality, which spurred the development of several popular systems including ARToolkit [3] and ARTag [4]. Real-world objects can be augmented with visual fiducials, allowing virtually generated imagery to be super-imposed. Similarly, visual fiducials can be used for basic motion capture [5]. AprilTag is a kind of fiducial marker as a black-and-white square tag with an encoded binary payload. Usually, those fiducial markers are placed in the space, and a camera takes photos of them. The photos can tell the pose information relative to the camera, while the ids of different tags can be recognized [6]. One paper compared the performance of various fiducial markers including ARTag, ArUco, Stag and AprilTag [7] AprilTag performs well in terms of position, orientation, and detection rate. Therefore, we chose AprilTag as the positioning marker for our project. This research is part of the ENSNARE project [8].

2 Research Gaps and Approach

When it comes to detection of visual fiducials, it is important to have a technique that is not only efficient but also consumes less time than current target measurement with, for instance, Total Stations. The Research Gap found in this field is that there are no existing methods to detect AprilTags using stereo vision. The rest of the paper provides in detail solutions to help tackle these research gaps.

The AprilTag detection process can be described as the following steps as also described in Figure 1:

- Camera and markers preparation
- Image pre-processing
- Geometry extraction
- Marker localization

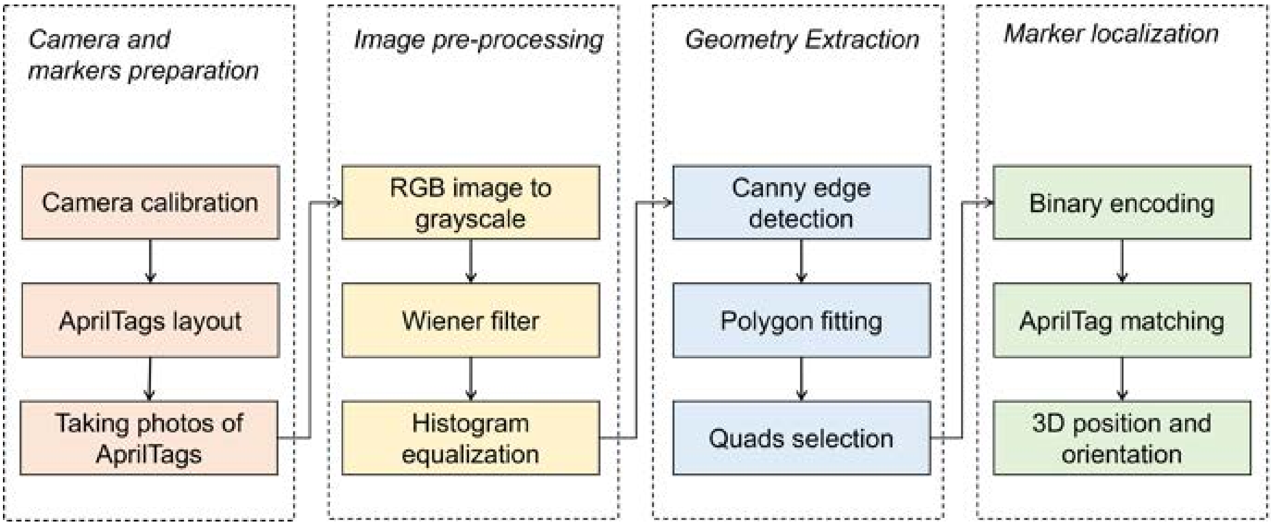

Figure 1. The process flow chart for AprilTag detection

Four steps are employed in the AprilTag detection:

- The color photos of targets are converted to grey.
- Noises are reduced by using an adaptive low pass Wiener filter.
- Contrast is enhanced using Histogram equalization [9]] to ensure the images are of the ideal brightness.
- Canny edge detection algorithm is used to get the edges, fit the edges into several closed polygons and filter out the unqualified polygons, leaving only the quads.

The candidate quads are binary encoded and if they match successfully with the AprilTag library, their IDs, and the pixel positions of the 4 corners can be output. Finally, SolvePnP [10] (perspective-n-point) method is used to solve the 3D position and orientation of each marker according to the 4 corners, marker size, and the camera calibration matrix.

The recognition distance d of Apriltag is related to many factors [11]:

$$d = \frac{f \times h_r \times I_p}{h_p \times S_s \times c} \quad (1)$$

where f is the focal length of the camera in mm, h_r is the real height of the AprilTag in mm, I_p is the height of image sensor in pixels, h_p is the AprilTag height in pixels and S_s is the image sensor's height in mm. c is a constant to change the unit scale. After a comprehensive reference to the formula and considering the need to measure outdoors, we decided to use a full-frame digital camera Sony A7R4 with different lens for target detection.

We can obtain translation matrix t and rotation matrix R of each Apriltag based on four corner coordinates in the image. Written as a transformation matrix in the form of

$$T = \begin{bmatrix} R_{11} & R_{12} & R_{13} & t_1 \\ R_{21} & R_{22} & R_{23} & t_2 \\ R_{31} & R_{32} & R_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2)$$

Suppose there are two targets, and their transformation matrices are T_A and T_B, then the transformation matrix of the latter in the former coordinate system is

$$T_{B\ in\ A} = {T_A}^{-1} * T_B \quad (3)$$

Initially, the center of the bottom left target (id: 1) is chosen as the origin and convert the coordinates of the rest of the targets to the world coordinate system. The problem is that once there is a small deviation in the position and angle of the world coordinate system, the deviation will increase after the remaining coordinates are converted into it. The solution is that we use 4 AprilTag (1 to 4) as a big origin.

There are 16 corner points in total, and we know the exact position between them, then we can use SolvePnP to get their pose matrix more precisely. Figure 2depicts the method of using AprilTags to form a big origin.

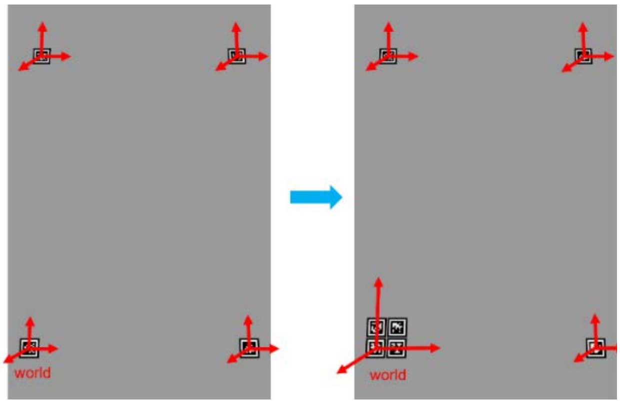

Figure 2. AprilTag to form a big origin

In the real test, one very serious factor that affects accuracy is the size of the AprilTag in the image. When we photographed a 15m building, our 200mm x 200mm AprilTag was very small and unclear in the frame. Therefore, we take photos from different angles, one image contains only 4 adjacent AprilTags and all images cover the entire facade. The advantage of this method is that each target in the photo is as large as possible Figure 3 depicts shooting local targets.

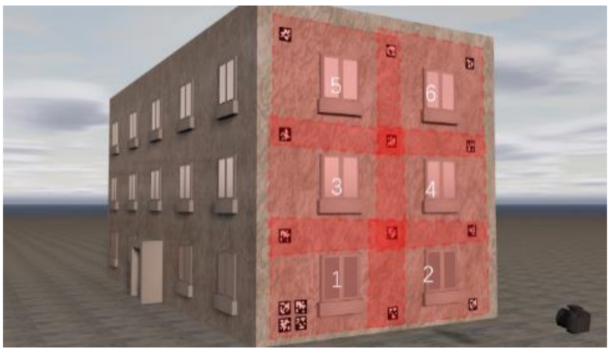

Figure 3. Shooting local targets

The camera needs to be calibrated after each shot because the focal length changes when shooting.

Image pre-processing is performed with the python OpenCV. Noise is reduced from the read images, and all AprilTags are detected from the different images. The transformation matrix, id and 2D coordinates of the corners of each AprilTag are obtained. If the image contains the big origin (id<=4), the exact transformation matrix is got with SolvePnP. When this is not the case, the coordinate system of the rest of the AprilTags is converted to the coordinate system of the bottom left AprilTag of the current image. All coordinate systems are then unified into the world coordinate system. The 3D coordinates, as well as the Euler angles from the rotation matrix, are acquired from the transformation matrix. Figure 4 explains the process of the algorithm employed:

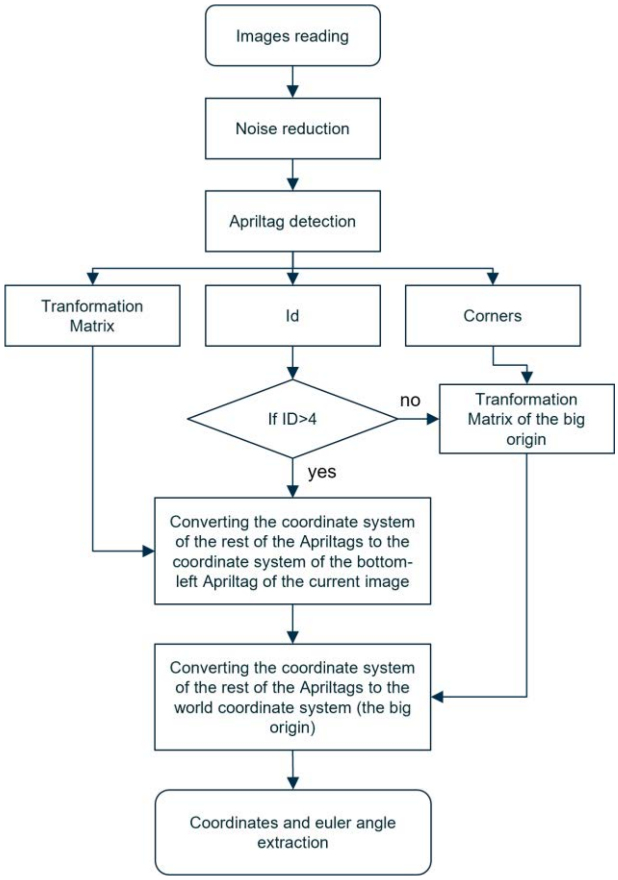

Figure 4. AprilTag position extraction process

After extensive testing, the accuracy of the method is as follows. The accuracy is fair in the case of low floors and decreases with the height of the floor as explained in Table 1.

Table 1. The accuracy of the algorithm

Target height	Translation deviation	Angle deviation
2.8m	2mm to 3mm	0.05°
6.1m	2mm to 5mm	0.1°
8.9m	7mm to 9mm	0.2°
12.1m	20mm	0.6°
15.1m	50mm	1°

These results are still optimal but open the way for further development.

Figure 5. Experimental setup for AprilTag detection

An experiment was conducted to verify the AprilTag detection algorithm as seen in Figure 5.

3 Conclusion

The algorithm to detect the AprilTags precisely was described in the paper. Experimental validation was also provided. The new detection algorithm includes to adjust the 2D detection algorithm and external environment to solve the problem of instability. Calibration was done and was combined with calibration algorithms to get the pose of markers.

The amount of AprilTags that needs a building in order to be measured, depends on the complexity and size of the building. Besides, if the building is covered by vegetation, traffic signals, or other objects, the approach would need to consider taking more and closer range pictures. This problem is common to other data acquisition techniques.

Currently the method is being improved and the results are already better and the main objective (1mm in 10m) is getting close under certain light conditions.

Acknowledgements

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 958445.

References

[1] C.-H. Chu, D.-N. Yang, and M.-S. Chen, "Image stablization for 2D barcode in handheld devices," in *Proceedings of the 15th ACM international conference on Multimedia*, Sep. 2007, pp. 697–706. doi: 10.1145/1291233.1291394.

[2] E. Olson, "AprilTag: A robust and flexible visual fiducial system," in *2011 IEEE International Conference on Robotics and Automation*, 2011, pp. 3400–3407. doi: 10.1109/ICRA.2011.5979561.

[3] D. Wagner, G. Reitmayr, A. Mulloni, T. Drummond, and D. Schmalstieg, "Pose tracking from natural features on mobile phones," in *2008 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*, 2008, pp. 125–134. doi: 10.1109/ISMAR.2008.4637338.

[4] M. Fiala, "ARTag, a fiducial marker system using digital techniques," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, 2005, vol. 2, pp. 590–596 vol. 2. doi: 10.1109/CVPR.2005.74.

[5] A. C. Sementille, L. E. Lourenço, J. R. F. Brega, and I. Rodello, "A Motion Capture System Using Passive Markers," in *Proceedings of the 2004 ACM SIGGRAPH International Conference on Virtual Reality Continuum and Its Applications in Industry*, 2004, pp. 440–447. doi: 10.1145/1044588.1044684.

[6] J. Wang and E. Olson, "AprilTag 2: Efficient and robust fiducial detection," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2016, pp. 4193–4198. doi: 10.1109/IROS.2016.7759617.

[7] M. Kalaitzakis, S. Carroll, A. Ambrosi, C. Whitehead, and N. Vitzilaios, "Experimental Comparison of Fiducial Markers for Pose Estimation," in *2020 International Conference on Unmanned Aircraft Systems (ICUAS)*, 2020, pp. 781–789. doi: 10.1109/ICUAS48674.2020.9213977.

[8] "ENSNARE," *https://www.ensnare.eu/*, Jan. 24, 2023.

[9] M. Abdullah-Al-Wadud, Md. H. Kabir, M. A. Akber Dewan, and O. Chae, "A Dynamic Histogram Equalization for Image Contrast Enhancement," *IEEE Transactions on Consumer Electronics*, vol. 53, no. 2, pp. 593–600, 2007, doi: 10.1109/TCE.2007.381734.

[10] A. Kriegler and W. Wöber, "Vision-based Docking of a Mobile Robot," Jan. 2020. doi: 10.3217/978-3-85125-752-6.

[11] S. M. Abbas, S. Aslam, K. Berns, and A. Muhammad, "Analysis and Improvements in AprilTag Based State Estimation," *Sensors*, vol. 19, no. 24, 2019, doi: 10.3390/s19245480.