

METHODS FOR THREE-DIMENSIONAL OBJECT RECOGNITION AND POSITIONING IN AUGMENTED REALITY SYSTEMS

Artikova Muazzam Axmedovna

Tashkent University of Information Technologies (TUIT), Candidate of Technical
Sciences, Associate Professor

Email: muazzamxon@mail.ru

Arislanova Malika Abdulla qizi

Tashkent University of Information Technologies (TUIT), Assistant Lecturer

Email: arislانovamalika1990@gmail.com

Abstract

Augmented Reality (AR) systems combine digital information with the physical environment and allow users to interact with virtual objects in real time. One of the essential technological problems in AR is the accurate recognition and positioning of three-dimensional objects. The quality of interaction largely depends on the ability of the system to determine the location, orientation, scale, and spatial boundaries of both real and virtual objects. This article examines the main methods used for three-dimensional object recognition and positioning in augmented reality systems. Particular attention is given to feature-based recognition, marker-based tracking, depth sensing, point-cloud processing, coordinate transformations, and camera pose estimation. The advantages and limitations of these approaches are discussed from the perspective of accuracy and real-time performance. The analysis shows that combining computer vision techniques with geometric modeling and spatial tracking can improve the stability and precision of AR applications. The considered methods can be used in education, industrial visualization, architecture, navigation, medicine, engineering, and interactive digital environments.

Keywords: augmented reality, 3D object recognition, object positioning, computer vision, spatial tracking, depth sensing, camera pose estimation, geometric modeling, virtual objects, object detection.

1. Introduction

Augmented Reality has developed rapidly due to advances in computer vision, mobile computing, sensors, and three-dimensional graphics. Unlike conventional graphical systems, AR applications must establish a relationship between computer-generated objects and the physical environment. For this reason, recognizing real-world objects and determining their spatial position are fundamental tasks.

An AR system should be able to identify relevant objects, estimate their position and orientation, and maintain this information as the user or camera moves. If these operations are performed inaccurately, virtual objects may appear to move incorrectly, lose their position, or fail to interact naturally with physical objects.

Three-dimensional object recognition involves identifying an object from visual or sensor information and determining its geometric characteristics. Object positioning, in turn, involves estimating its coordinates and orientation within a predefined or dynamically constructed coordinate system. These processes are closely connected with camera calibration, feature extraction, tracking, and geometric transformation.

The purpose of this article is to analyze major methods for recognizing and positioning three-dimensional objects in AR systems and to identify the main factors affecting their accuracy and computational efficiency.

2. Literature Review

Early AR systems frequently used artificial markers for object recognition and camera tracking. Kato and Billinghurst introduced ARToolKit, which demonstrated how visual markers could be detected by a camera and used as reference points for estimating the position of virtual objects. Marker-based systems remain useful because they are relatively simple and can provide reliable tracking under controlled conditions.

However, marker-based approaches require specially prepared physical objects or surfaces. Modern AR systems increasingly use markerless recognition based on natural visual features. Feature-based methods identify distinctive points, edges,

textures, or patterns in camera images. These features can then be matched between consecutive frames to estimate movement and spatial relationships.

Three-dimensional object recognition can also be supported by depth sensors. Depth information provides an additional spatial dimension that is not directly available in conventional RGB images. Devices equipped with depth cameras can generate point clouds in which each point contains three-dimensional spatial information. This makes it possible to estimate surfaces, distances, and object boundaries.

Processing these points allows an AR system to construct an approximate geometric representation of the surrounding environment.

Several methods can be used to determine the position of a three-dimensional object. The simplest approach is reference-based positioning, where an object is located relative to a known marker or predefined coordinate system. Feature-based positioning does not require artificial markers. Instead, the system identifies natural features in the environment and tracks their movement between successive frames. This method provides greater flexibility but can be affected by poor lighting, repetitive textures, motion blur, or significant changes in viewpoint. Depth-based positioning provides direct spatial information and can therefore improve the estimation of object distance from the camera. However, depth sensors may have limited operating ranges and can be affected by reflective, transparent, or very dark surfaces.

Another important method is model-based positioning. In this approach, the system compares observed visual information with a previously prepared three-dimensional model. Once corresponding features have been identified, the system estimates the object's position and orientation.

For a rigid object, its spatial transformation can be represented by a homogeneous transformation matrix:

$$T = [R \ t] = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix}$$

This representation makes it possible to combine rotation and translation within a single mathematical framework.

4. Challenges and Applications

The accuracy of three-dimensional object recognition depends on several factors. Lighting conditions, camera resolution, object complexity, occlusion, sensor noise, and computational resources can significantly influence system performance. Occlusion is particularly challenging because parts of an object may become temporarily invisible when another physical object is positioned in front of it.

Real-time processing is also essential. AR systems must continuously analyze sensor data and update virtual objects without noticeable delay. Therefore, algorithms should balance recognition accuracy with computational efficiency.

The described methods have broad applications. In education, students can examine three-dimensional models of scientific objects. In architecture and engineering, virtual structures can be positioned within real environments. In industrial applications, AR can support equipment visualization and maintenance. In medicine, three-dimensional models can be used to visualize anatomical structures. Navigation systems can also use spatial recognition to provide contextual information about physical surroundings.

Conclusion

Three-dimensional object recognition and positioning are fundamental processes in the development of reliable augmented reality systems. Marker-based methods provide stable positioning in controlled environments, while markerless feature-based approaches offer greater flexibility. Depth sensing and point-cloud processing provide additional spatial information and can improve the understanding of physical environments.

Camera pose estimation and coordinate transformations form the mathematical foundation for accurately placing virtual objects within a real-world coordinate system. At the same time, SLAM technologies enable continuous localization and environmental mapping without requiring predefined markers.

Future AR systems are likely to combine visual features, depth information, geometric models, and intelligent recognition algorithms. Such integration can improve positioning accuracy, reduce tracking errors, and support more natural

interaction between users and virtual objects. The development of efficient three-dimensional recognition and positioning methods therefore remains an important research direction for advanced AR applications.

References

1. Azuma, R. T. (1997). A survey of augmented reality. *Presence: Teleoperators and Virtual Environments*, 6(4), 355–385.
2. Billinghurst, M., Clark, A., & Lee, G. (2015). A survey of augmented reality. *Foundations and Trends in Human–Computer Interaction*, 8(2–3), 73–272.
3. Hartley, R., & Zisserman, A. (2003). *Multiple View Geometry in Computer Vision* (2nd ed.). Cambridge University Press.
4. Kato, H., & Billinghurst, M. (1999). Marker tracking and HMD calibration for a video-based augmented reality conferencing system. *Proceedings of the 2nd IEEE and ACM International Workshop on Augmented Reality*, 85–94.
5. Lepetit, V., Moreno-Noguer, F., & Fua, P. (2009). EPnP: An accurate $O(n)$ solution to the PnP problem. *International Journal of Computer Vision*, 81, 155–166.
6. Milgram, P., & Kishino, F. (1994). A taxonomy of mixed reality visual displays. *IEICE Transactions on Information and Systems*, E77-D(12), 1321–1329.
7. Mur-Artal, R., Montiel, J. M. M., & Tardós, J. D. (2015). ORB-SLAM: A versatile and accurate monocular SLAM system. *IEEE Transactions on Robotics*, 31(5), 1147–1164.
8. Szeliski, R. (2022). *Computer Vision: Algorithms and Applications* (2nd ed.). Springer.
9. Newcombe, R. A., Lovegrove, S. J., & Davison, A. J. (2011). DTAM: Dense tracking and mapping in real-time. *Proceedings of the IEEE International Conference on Computer Vision*, 2320–2327.
10. Ericson, C. (2005). *Real-Time Collision Detection*. CRC Press.