Gesture-Based Interaction in Augmented Reality Learning Tools

Olivia Grant
Dept. of Engineering
Some State University
Some City, A State, USA
OGrant@ssu.edu

Eric Roberts
Dept. of Engineering
Some State University
Some City, A State, USA
EricRoberts@ssu.edu

***Abstract*—This paper explores the integration of gesture recognition into augmented reality (AR) platforms aimed at enhancing STEM education. We introduce a prototype AR learning tool that enables students to manipulate virtual scientific models using natural hand gestures. Our system leverages computer vision techniques for gesture detection and integrates with Unity-based AR environments via lightweight APIs. Usability testing with a pilot group of undergraduate students revealed increased engagement, faster task completion, and improved spatial understanding compared to traditional touchscreen controls. These findings suggest gesture-based AR systems can serve as effective interactive learning environments in both in-person and remote educational settings.**

I. Introduction

Augmented reality (AR) has emerged as a promising medium for educational tools, offering students immersive, hands-on experiences with abstract and complex subjects. From molecular biology to circuit design, AR allows users to visualize and manipulate content in 3D space, fostering a deeper understanding of spatial relationships and system dynamics.

However, traditional input methods—such as screen taps, sliders, or handheld controllers—can constrain the natural expressiveness and interactivity of AR systems. This is especially evident in learning environments, where intuitive control is essential to maintain user focus and immersion.

To address this, we propose the use of gesture-based interaction in AR learning tools. Specifically, we developed a prototype that supports common manipulations (rotate, zoom, select, group) through natural hand movements recognized by a monocular RGB camera. The system is lightweight, platform-agnostic, and intended for deployment in classrooms or remote setups using standard consumer-grade hardware.

In this paper, we describe the architecture of our system, the computer vision model used for gesture classification, and the integration pipeline with Unity's AR Foundation framework. We also present the results of a user study evaluating the usability and learning efficacy of the tool in comparison with touch-based interfaces.

II. Related Work

Prior research in AR education has focused heavily on content development and learning outcomes, with less emphasis on input modalities. Systems like ARChem [1] and GeoLearnAR [2] demonstrated the value of interactive 3D content, but typically rely on screen-based or device-orientation inputs.

Gesture interaction has been explored in HCI more broadly, with frameworks such as Leap Motion and OpenPose enabling hand tracking and gesture control across domains. In education-specific contexts, however, most gesture systems are limited to high-end hardware or VR ecosystems, making them inaccessible in standard classroom settings.

Our contribution is distinct in that it combines affordable vision-based gesture recognition with AR educational content in a platform-independent framework. We focus specifically on gestures relevant to task flow in interactive learning—object manipulation, selection, and scene navigation—and evaluate their pedagogical effectiveness through controlled studies.