A Multimodal Interface for Human-Computer Interaction Using Voice Recognition and Gesture-Based Mouse Control

Prof. Minal Solanki[1], Akhilesh Hadke[2], Nakshatra Patrange[3]

Assistant Professor, Computer Application[1]
MCA, Computer Application[2,3]
K. D. K College of Engineering, Nagpur, Maharashtra, India
minal.solanki@kdkce.edu.in[1], akhileshhadke.mca23@kdkce.edu.in[2], nakshtrapatrange.mca23@kdkce.edu.in[3]

Abstract: *The integration of voice assistants with hand-gesture-based mouse control marks a major advancement toward the development of more intuitive, accessible, and user-friendly interfaces. By merging the natural ease of voice commands with the precision and responsiveness of hand gestures, this approach offers users a seamless, hands-free method for navigating and interacting with digital environments. Voice assistants serve as a natural language gateway, enabling users to perform tasks through simple spoken commands, while hand gesture recognition provides fine-grained control over cursor movement and interface navigation—eliminating the need for traditional hardware like keyboards and mouse. This dual-mode interaction system not only streamlines multitasking and enhances user efficiency but also significantly improves accessibility for individuals with physical impairments, offering a more inclusive way to engage with technology. By integrating these complementary technologies, users benefit from a smoother, more natural control experience that aligns more closely with human communication patterns. This study investigates the combined use of voice recognition and gesture tracking to create a cohesive, multimodal user interface. It examines the technical challenges involved, such as accuracy, latency, and system integration, while also emphasizing the substantial user experience improvements. Furthermore, the study explores diverse applications of this integrated approach across multiple domains—including accessibility support, gaming environments, professional workplaces, and immersive AR/VR settings—highlighting its potential to redefine the future of human-computer interaction*

Keywords: Voice Assistants, Hand Gesture Recognition, Touchless Interaction, Multimodal Interaction, Human-Computer Interaction (HCI)