Chapter 6
The Voice-Enabled Virtual Mouse for Contactless Computing

P. TamilSelvi
http://orcid.org/0009-0006-9713-9181
Sathyabama Institute of Science and Technology, India

S. Keziah
http://orcid.org/0009-0001-4331-9543
Sathyabama Institute of Science and Technology, India

K. G. Sarah Princes
Sathyabama Institute of Science and Technology, India

Chola Vendhan
Sathyabama Institute of Science and Technology, India

ABSTRACT

This chapter analyzes the design, system architecture, and performance challenges of multimodal virtual mouse systems for contactless computing. By integrating voice-controlled commands with visual-gesture recognition, these systems enable fully hands-free interaction on desktop environments. The chapter focuses on real-time gesture processing, voice

DOI: 10.4018/979-8-3373-9785-6.ch006

command context-aware integration, and coordinated multimodal execution to ensure responsive and reliable user interaction using consumer-grade hardware. It also explores the applications, performance enhancement techniques, and emerging trends that influence the efficiency and capabilities of voice-enabled virtual mouse platforms. The work aims to develop effective, accessible, and robust contactless computing solutions that address the limitations of single-modality systems, emphasizing adaptive and inclusive design for modern desktop settings.

1. INTRODUCTION

The accelerated growth of digital technologies across work, education, and general society has prompted a significant shift towards more natural, intuitive, and contactless methods of human-computer interaction (Yadav et al., 2023). Traditional input devices like keyboards and mice require direct physical interaction, which poses challenges in hygiene-sensitive environments and shared workspaces (Ranaweera et al., 2022). This has led to increased interest in alternative, touch-free input systems such as gesture-based and voice-controlled interfaces, which are powered by advances in sensing technologies and artificial intelligence (Hasan et al., 2022). Contactless interaction systems, including gesture recognition and voice control, are revolutionizing human-computer communication by providing more inclusive, efficient, and adaptive ways of interacting with digital systems (Khan et al., 2024).

Gesture-based interaction leverages computer vision to track hand movements, enabling users to control cursors, select objects, and navigate without physical devices (Rida et al., 2023). These systems offer a natural mapping of hand actions to computer responses, reducing user learning time and enhancing spatial accuracy (Srinivasan et al., 2024). Real-time gesture recognition, however, requires overcoming challenges like varying lighting conditions and background complexity (Priesnitz et al., 2022), ensuring stability and smooth interaction over long durations (Chen et al., 2022). Complementing gesture control, voice-enabled systems provide another layer of interaction, facilitating command execution through speech. Speech recognition has become more reliable with advances in