THETA: Triangulated Hand-State Estimation for Teleoperation and Automation in Robotic Hand Control

Huang, Alex (School: ASU SCENE)
Karthik, Akshay (School: ASU SCENE)

The teleoperation of robotic hands is limited by the high costs of depth cameras and sensor gloves, commonly used to estimate hand relative joint positions (XYZ). We present a novel, cost-effective approach using three webcams for triangulation-based tracking to approximate relative joint angles (theta) of human fingers. We also introduce a modified DexHand, a low-cost robotic hand from TheRobotStudio, to demonstrate THETA's real-time application. Data collection involved 40 distinct hand gestures using three 640x480p webcams arranged at 120-degree intervals, generating over 48,000 RGB images. Joint angles were manually determined by measuring midpoints of the MCP, PIP, and DIP finger joints. Captured RGB frames were processed using a DeepLabV3 segmentation model with a ResNet-50 backbone for multi-scale hand segmentation. The segmented images were then HSV-filtered and fed into THETA's architecture, consisting of a MobileNetV2-based CNN classifier optimized for hierarchical spatial feature extraction and a 9-channel input tensor encoding multi-perspective hand representations. The classification model maps segmented hand views into discrete joint angles, achieving 97.18% accuracy, 98.72% recall, F1 Score of 0.9274, and a precision of 0.8906. In real-time inference, THETA captures simultaneous frames, segments hand regions, filters them, and compiles a 9-channel tensor for classification. Joint-angle predictions are relayed via serial to an Arduino, enabling the DexHand to replicate hand movements. Future research will increase dataset diversity, integrate wrist tracking, and apply computer vision techniques such as OpenAI-Vision. THETA potentially ensures cost-effective, user-friendly teleoperation for medical, linguistic, and manufacturing applications.

Awards Won: