Accelerating Computer Vision for Assisted Indoor Climbing

Asad Khan, Laura Madrid, & Lucas Noritomi-Hartwig

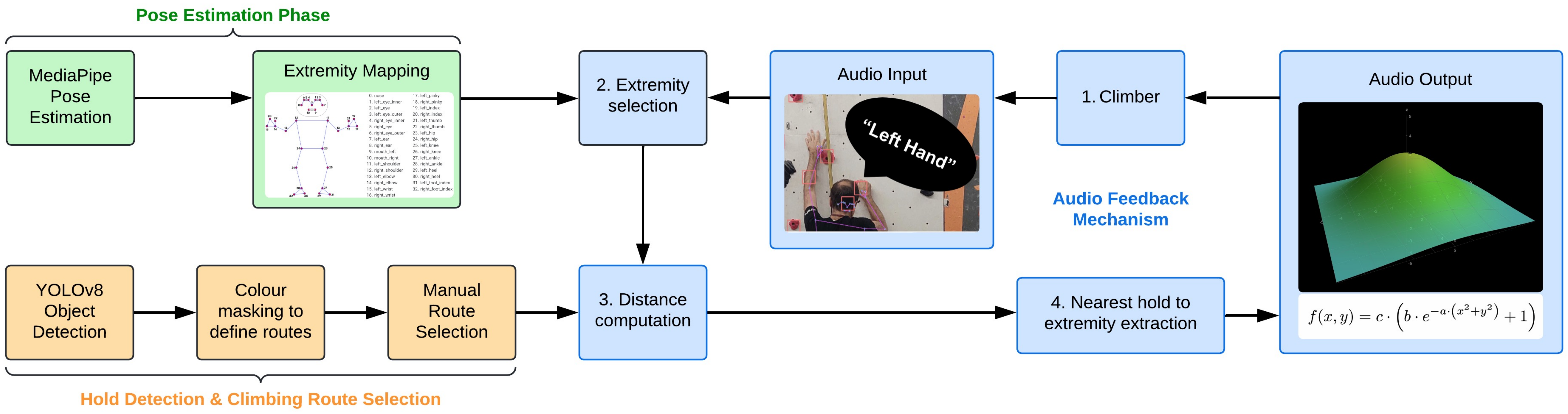
UNIVERSITY OF TORONTO

Abstract

We introduce an indoor rock-climbing assistance tool (IRCAT) designed to help visually impaired individuals navigate and complete indoor top-rope rock climbs. IRCAT implements pose estimation and object detection to capture each instance of the climb. IRCAT uses speech recognition to interpret the climber's verbally chosen extremity to move next during the climb and provides audio feedback, guiding the climber to the nearest hold based on the specified extremity. If no command is given, IRCAT loops through a predetermined order of extremities as the climb progresses. Since the climber cannot rely on sight, IRCAT listens and dynamically adjusts with each new command, helping the climber detect the route through auditory cues. The tool demonstrates its effectiveness in a controlled environment, and with additional development, it has the potential to assist in various rock-climbing situations. Our objective with this project is to investigate existing methods and create a practical solution to support visually impaired rock climbers.

Climbing Phase

Assisted Indoor Rock-Climbing Pipeline



Implementation Details

Pose Estimation Phase:

- MediaPipe Pose for pose estimation, using landmarks to calculate extremity centroids and measure distance to the next rock hold.

Hold Detection & Climbing Route Selection:

- YOLOv8 detects climbing holds, and HSV masks classify them by color to define routes.
- The user interface allows the belayer to select and adjust routes by adding or removing holds.

Audio Feedback Mechanism:

- Audio Input:
  - *speech_recognition* + Google's Web Speech AP listens for extremity: "hand" or "foot", and side: "left" or "right".
  - Given proper command, adjust current extremity.
- Audio Output:
  - The *sounddevice* package plays audio, with pitch based on distance between centroid and target hold - smaller distance means higher pitch, like a metal detector.

NVIDIA Tools:

- Parallelization of the audio feedback mechanism using CUDA enhances the efficiency of computing distance, thereby facilitating a more continuous and responsive audio output signal. This approach reduces computational latency, ensuring higher precision and temporal consistency with the auditory feedback.
- Additionally, the YOLOv8 model used was fine-tuned on:
  - NVIDIA GeForce RTX 3090 GPU (CUDA Version: 12.6)
  - AMD Ryzen 9 5900X 12-Core, with 132 GB of RAM.