Motivational Model

Vision and Speech Interface for Robotic Arm

Clear
Intuitive

Reliable
Confident

Unstable
Buggy

Tech Support

Operator

Detect objects visually

Visual detection of human anatomy (e.g. gesture, gaze)

Object Identification

Track distance of object relative to robot

Perceive what the robot sees through a GUI

Easy to understand

Intuitive

Inconsistent
Ambiguous

Operator

Interact with the environment aurally

Perceive sound

Generate Sound

Convert speech to text

Interpret Text

Convert text to speech

Transparent

Elegant

Incomprehensible

Operator Tech Support

Integrate visual and audio inputs

Create new robot action

Match command with action

See specific logs for each operation

Input item information

Operator Tech Support

Responsive
Robust

Timely
Reliable

Cumbersome

Move robotic arm

Stop execution of commands

Execute actions

Move back into initial state