A.n.i B

An expressive conversational robot with intent recognition and app support

Project by Venkata Tej Kiran Reddy Polamreddy

Presentations: https://www.youtube.com/@TejKiran/videos

Table of Contents

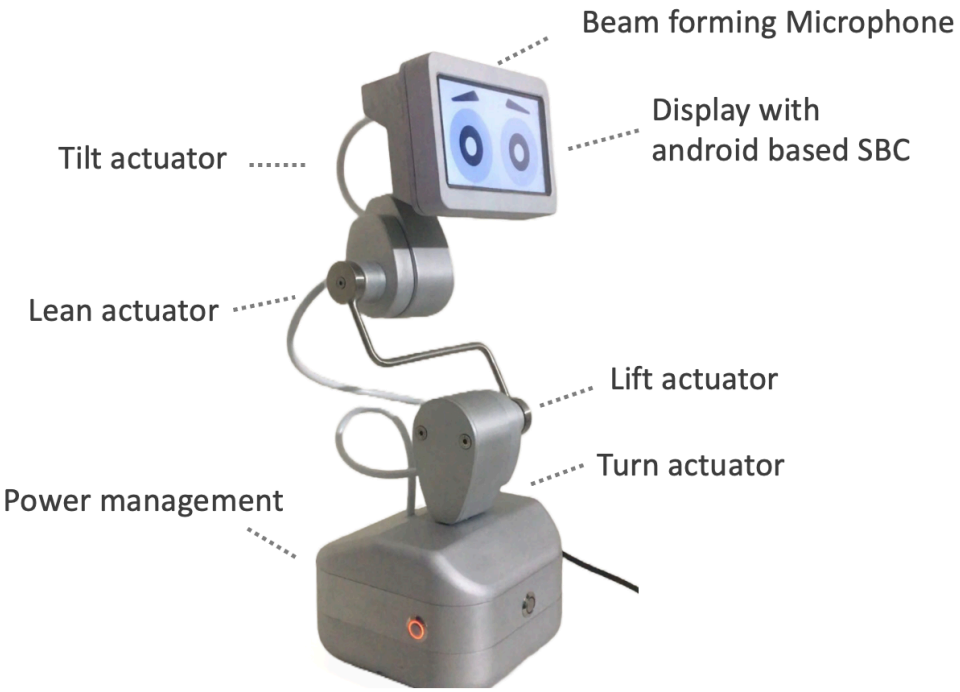
Overview

An interactive robot with NN based voice pipeline and motion support for custom app development.

Robot's Hardware

a) Robot’s Body

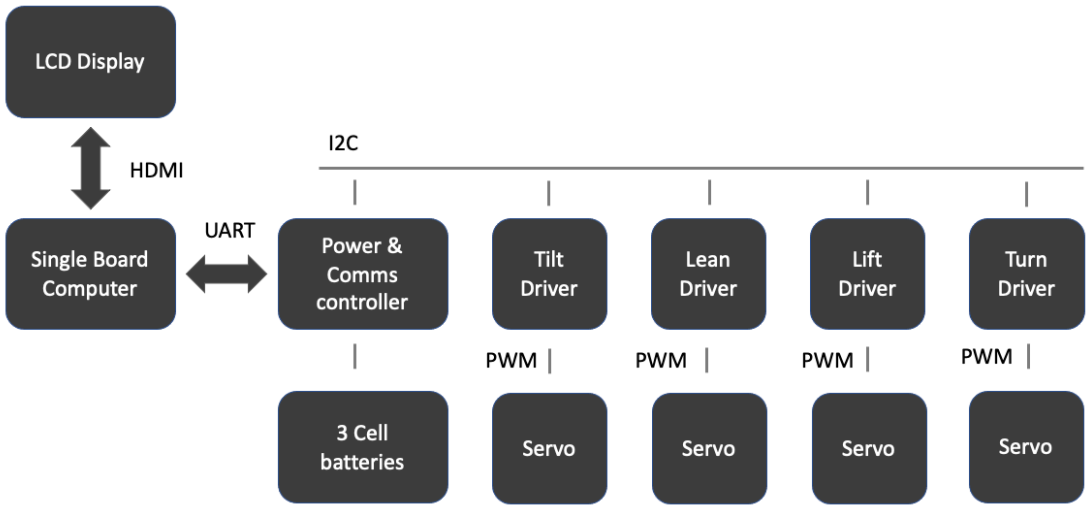
The Hardware consists of 4 different actuators connected serially with a computer & a display module mounted at the top that shows a pair of animated eyes as shown in figure below



The physical hardware is made from aluminum sheets, and stainless steel plates reworked using lathe machines and power tools.

b) Robot's Electronics

Robot's electronic communication wiring diagram is as shown below



Servo drivers are customized with in-house developed PCBs to achieve the precise time-based [3] motion control of the RC Servos. Each driver consists of a microcontroller and a buck converter. The controller runs software based on the Arduino stack [1] with custom servo management logic and communication protocol. The servo drivers convert the commands received over the two-wire communication line into servo motion.

Also, a power management module was designed to handle Lithium battery management and charging tasks. The state machine for power management is taken care by the same controller which handles the communication.

c) Robot's Software stack

The robot's software can be divided into 3 modules

- Power and motion control software for each RC Servo driver
- Device power management, user request handling and motion parameter forwarding software
- Android OS and application software stack developed to run on the Single Board Computer

Arduino-based software stack developed for RC Servo drivers has the logic to receive motion parameters(eg: total time of motion, easing function of motion, delta angle etc.) over I2C interface and can perform servo motion in the desired manner. The software also handles the calibration, power, and signal management of the connected servo motor.

The Power Management software handles tasks of battery level monitoring, charging control, device power, attention button requests. This software plays the central role of collecting and forwards motion parameters from Single Board computer to Motor Driver modules.

The robot's brain consists of a single board computer that runs on a custom build android operating system(based on an existing repository [2]). Several android services were developed for the robot to handle communication, system management, backend-server connection handling, audio processing, and other tasks. Through these services, a front-end application with animated eyes could generate an expression by synchronising the motion between actuators, graphical eyes shown on display while doing conversation.

Robot's Motion and Emulator app

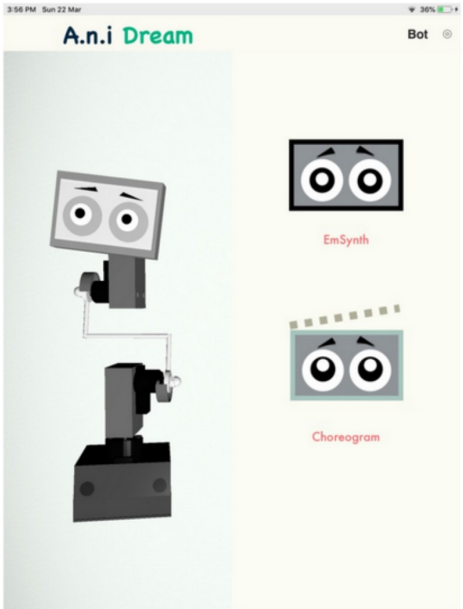
a) Robot Motion Representation

The Robots motion consists of physical actuator motion synchronous with graphical eye animation and is represented as a sequence of keyframes. The complete set of actuators and graphical objects are specified as the motion elements of the robot.

A keyframe is defined as a complete set of motion parameters defined for every element of the robot which needs to be followed or pursued for a given amount of time. Example motion parameters include start position, end position, time of motion ,and easing function etc.

b) Robot Motion Synthesis

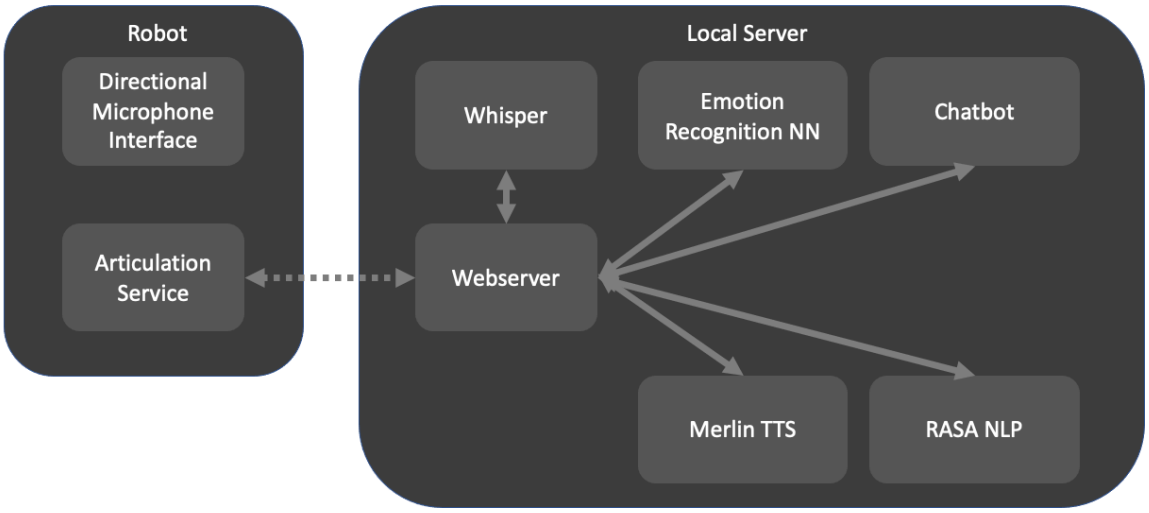
An emulator app called A.n.i Dream(for iPad) is developed for synthesizing these keyframes. The front end of this emulator app is as shown in figure below.



The emulator app provides two key features. The Choreogram feature can be used to create motion sequences for music tracks or poems. The Emsynth feature can be used to create an attitude set for the robot to be used in a conversation.

Robot Interaction

An NLP pipeline prototype was created to enable conversation with the robot. Various modules used to enable this pipeline are shown in figure below.



Through this NLP pipeline, the robot can perform visually emotional conversations by enacting the attitudes created using the above-mentioned Emsynth feature.

First, the robot listens to the user's speech. Then, this audio will be processed through a series of neural networks (listed as Speech-to-Text(google STT), Emotion Recognition, Chatbot[6], Text-to-Speech[4], and Intent recognition[5]) and generates a response with its emotion levels and predicts an Action label(If the user requests for an action). Next, the emotional levels will be mapped to an expression in the attitude set. Finally, the robot speaks back to the response while animating an expression creating a seamless emotional response.

References:

1. https://create.arduino.cc/projecthub
2. https://github.com/brobwind/pie-device-brobwind-rpi3
3. https://github.com/luisllamasbinaburo/Arduino-Easing
4. https://github.com/CSTR-Edinburgh/merlin
5. https://rasa.com/
6. https://github.com/lukalabs/cakechat

Art work Attributions:

https://support.iconfinder.com/en/articles/18231-license-basic