ECE8903 Special Problem: Multimodal Speech Capture System

Ankit Kaul
Georgia Institute of Technology, Atlanta, GA.

***Abstract*—This document describes the developments in the Multimodal Speech Capture System (MSCS), which is being designed in the GT Bionics lab, under the guidance of Prof. Mayasam Ghovanloo. Improvements in the stationary version of the MSCS, and the design of a peripheral PCB for enhancing the accuracy of sensor data have been addressed in this document. This project work was performed over Fall 2016.**

Keywords—Multimodal Speech Capture System, FPGA programming, embedded systems, PCB Design

I. Introduction

Our abilities to speak and articulate are associated to Broca's area in the left hemisphere of our brain. A critical language area in the posterior superior temporal lobe, called as Wernicke's area, is involved in language processing. Damage to a discrete part of the brain in the left frontal lobe (Broca's area) of the language-dominant hemisphere has been shown to significantly affect use of spontaneous speech and motor speech control [1].

An ischemic stroke or transient ischemic attack (TIA) occurs if an artery that supplies oxygen-rich blood to the brain becomes blocked. Due to an impedance in blood flow to the brain, the supply of oxygen to the brain cells reduces and brain cells start dying. When brain cells die, the functioning of the body parts that they control is impaired or lost [2]. A stroke can affect any part of the brain, and when it affects the Broca's and Wernicke's areas, it can potentially cause speech and language related disorders.

One of the perennial health issues which plague the human race is damage to certain parts of the brain which results in paralysis of bodily functions. Speech impairment is one such ailment which is very common consequence of stroke and also seen among patients dealing with either trauma, multiple sclerosis, Parkinson's, or other neurological conditions, which lead to cognitive aspects of speech or physical damage to oral motor functions. It is incumbent upon us, the technologists, to develop systems that can help patients, who are dealing with speech paralysis, to communicate and regain their vocal capabilities. A multi-modal Tongue Drive Systems (mTDS) is currently under development at the GT Bionics lab, which has led to development of a new technology called Multimodal Speech Capture System (MSCS) aimed at helping Speech Language Pathologists to train and rehabilitate their patients more effectively.

The initial prototype of the MSCS was a stationary setup which is being transformed into a portable/wearable head-mountable variant, which will transmit the patient's speech data, including tongue motion, voice, and lip movements to a PC. The scope of this project was to proceed with the study and development of the firmware and hardware needed to build a functional prototype of this mobile/wearable version of the MSCS.

The MSCS involves tracking three important speech modes: tongue movement, voice, and lip reading [3]. A permanent magnetic tracer is affixed near the tip of the tongue of a human subject wearing this device. A magnetometer array collects the changes in the magnetic field near the user's mouth, which will be used to track the subject's tongue movements after signal processing in the PC, and help define the tongue's spatial coordinates within the oral space [4]. Two microphones and a webcam are affixed to the 3D-printed chassis of the head-mount which help acquire the requisite audio and video inputs, respectively, from the subject. The speech and visual input signals are then routed through a Mojo board, which is equipped with a Spartan-6 field programmable gate array (FPGA) and USB interface, and communicates these signals to a PC via its USB connection. The PC runs magnetic localization, speech recognition, and image processing algorithms in real time to generate an audiovisual feedback that would enable better control of the articulators and certain aspects of speech facilitated by this novel Human Machine Interface.

II. Hardware

A. Magnet

A disc magnet is used for the purpose of tracking the tongue. It is produced by K&J Magnetics and its part number is D201-N52. It has a diameter of 1/8" and is 1/32" thick and has a surface field of 3309 Gauss.

B. Magnetic Sensors: LSM303D and LSM9DS0

The LSM303D is an ultra-compact high-performance 3D magnetometer and 3D accelerometer. However, the accelerometer function is not used for this purpose. The sensor supports both SPI and I^2C protocol for communication. The detailed specifications for the magnetic sensor can be found here. This magnetic sensor is the key sensor used in the localization of the magnet. 4 magnetic sensors are currently mounted on a single PCB board. There are in all 6 boards currently employed with 2 boards near each cheek and 2 below the chin.

The LSM9DS0 is a very similar to the LSM303D except for fact that it has an on board 3D gyroscope along with a 3D magnetometer and a 3D accelerometer. This IC is mounted on the Mojo cape for the low density sensor PCB version of the MSCS. The accelerometer and gyroscope functions are currently not used for this project. This sensor supports both SPI and I^2C protocol for communication as well. The detailed specifications for the magnetic sensor can be found here. 2

magnetic sensors are currently mounted on a single cape PCB board.

C. *Mojo Board*

The Mojo board is an Embedded Micro development kit that comprises of a Spartan 6 XC6SLX9 FPGA chip and an Atmel microcontroller (ATmega32U4). The FPGA chip acts as the controller of the sensor boards and the receiver of the sensor data. The FPGA chip packetizes the received data and sends it to the Atmel micro-controller which acts as an interface between the FPGA board and the computer. The Mojo board currently supports 84 digital IO pins which could support 42 sensors (with 1 data pin and 1 chip select signal per sensor) [5].

D. *Microphone and Webcam*

In order to do multi-modal classification, a pair of microphones and a webcam are added to the overall system. The microphone focuses on the lips of the user and records the lip movements while the pair of microphones record stereo sound.

Figure 1 shows the overall block diagram of the MSCS.

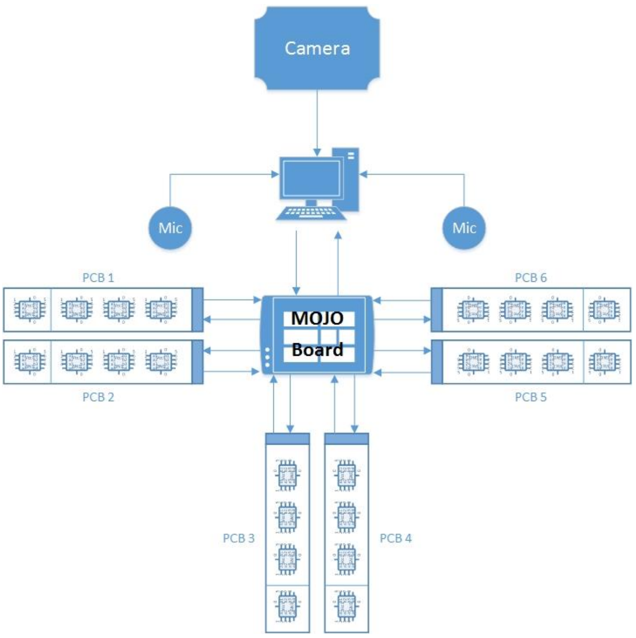

Fig. 1: TTS Hardware Block Diagram

III. SOFTWARE

A. *FPGA Code*

As mentioned earlier, the Mojo board comprises of a Spartan 6 FPGA board which is used to generate the various signals required to control the magnetic sensors and to receive data from the magnetic sensors. It is comprised of the following modules [3]:

1. SPI wrapper: This module acts as a standalone for the PCB. It also acts as an interface between the magnetic sensors and the ComController, which generates the control signals. This SPI wrapper helps in modularizing the code and hence additional PCBs can be added by simply instantiating more wrappers.
2. ComController: The ComController module interfaces the SPI Wrappers to the USB controller. The ComController is responsible for generating the polling signals and co-ordinating different PCB boards. It collates the received data to be sent out to the PC.
3. USB Controller: The USB controller interfaces between the ComController and the PC. The USB is in-charge of sending the packets of sensor data off to the PC.
4. CLK Modules: The delay module and the clock scaler help in generating different clocks. The LSM303D operates at a lower frequency compared to the USB controller. Thus the on-board FPGA clock should be scaled down for it to be usable in the SPI protocol.

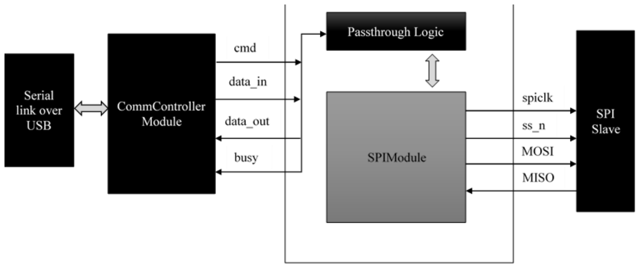

Fig. 2: Spartan 6 Verilog Code Block Diagram

For the purpose of programming the FPGA board, Verilog is being used [5]. The Xilinx ISE can be downloaded from Xilinx website (Take note to download the ISE version and not Vivado): http://tinyurl.com/cayzahj.

IV. CONTRIBUTIONS BY THE AUTHOR

A. *Addition of extra sensor for Magnetic field correction*

The previous TTS version utilized only six low density sensor PCB's for tracking the location of a magnet affixed to a subject's tongue. That system does not account for the effect of Earth's magnetic field on the magnetic location data recorded by the other magnetometers. The absence of a corrective methodology makes the data recorded by the main magnetometers unreliable and skews the tracked location of a user's tongue. With these observations in mind, it was decided to add an extra sensor board which could record the data due to Earth's magnetic field. This would enable us to uphold the integrity of the magnetic data collected for the tongue tracking.

To add just an extra sensor posed a challenge as the magnetometers used in this project have miniscule footprints, LGA-16 in this case. Thus, it was decided to demonstrate the concept with an entire low density PCB, having 4 sensors, and use a single sensor in the eventual design. This was made possible with the hardware modularity that Verilog provides for FPGA programming.

The following steps were followed in order to implement this change:

Editing the Verilog code: The modules ComController.v, top.v, and top.ucsf were modified in order to accommodate for the new sensors. All the variables corresponding to a new PCB, which need to be updated/added in their respective modules, are listed below:

1. ComController.v: This module was designed to be the logical core of the TTA firmware. It drives processes related to sensor initialization, polling, and general system operation.

    a. pcbcount: Stores the number of PCB/sensor modules that the TTS can be currently interfaced to. A change in the value of the variable pcbcount, from 6 to 7, defines the presence of a new slave for the Mojo. This value needs to be updated.

    b. comm_recvd_data_6: A 16-bit array which stores the data received from the communications module. A new variable needs to be added.

    c. polldata_arr: 8-bit array to store data received from sensors, between consecutive polling. New variable needs to be added.

2. top.v : This module is responsible for instantiating other FPGA modules in the TTS firmware, and for defining the logic which supports the functionality of all other operational blocks.

    a. pcbcount: As previously explained, this value needs to be updated to 7.

    b. SPI communication variables (miso, mosi, spiclk, cs0, cs1, cs2, cs3, pwr, and gnd): These variables are defined for each SPI slave as they facilitate SPI communication between the SPI master (Mojo) and SPI slaves (sensor PCBs). A complete set of variables delineated above, needs to be added for every additional PCB/sensor.

    c. pwr and gnd: These variables are similar to the ones mentioned in point b. above, but need to be separately assigned as power and ground pins on the Mojo.

    d. CC_spi_data_out: This 16-bit array stores the device data which needs to be returned to the ComController module. A new variable of this datatype needs to be added for every additional PCB/sensor.

    e. comm_busy: This signal is used to determine if a communication module is busy or not. A new variable needs to be declared for every additional PCB/sensor.

    f. comm_recvd_data: This function takes in the 'CC_spi_data_out' array as an argument and defines the logic to store the data received from communications module. A new such function needs to be added for additional sensors.

    g. SPI Module Instantiation: An SPI module for every slave on the SPI bus needs to be instantiated. This module is defined as an SPIWrapper datatype in the top.v module. It includes the declaration of variables like the mosi, miso, spiclk (SPI Clock), cs (chip select), comm_busy, CC_spi_data, CC_spi_data_out. An entire SPI module needs to be added for every additional PCB/sensor.

3. const.ucf : This file defines the mapping of the system's SPI variables to the digital I/O pins of the Mojo board. It is generated by the constraints manager, a part of Xilinx's ISE suite. The following changes were made to the const.ucf file:

Table 1: New Port declarations in the const.ucf file

NET	Mojo Port Number
pwr_6	p30
gnd_6	p29
miso_6	p33
mosi_6	p32
spiclk_6	p35
cs0_6	p34
cs1_6	p41
cs2_6	p40
cs3_6	p51

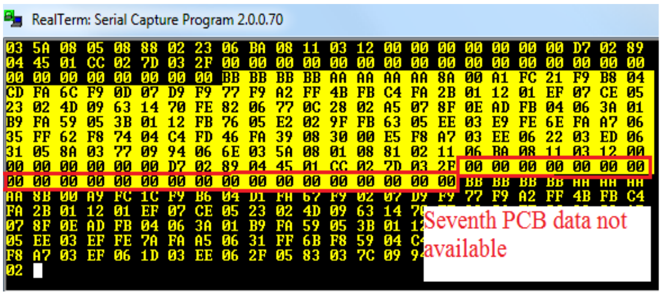

Fig. 3: Modified TTS code without the seventh PCB

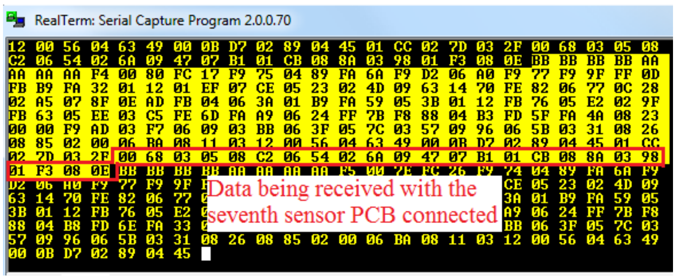

Fig. 4: Modified TTS code with the seventh PCB connected

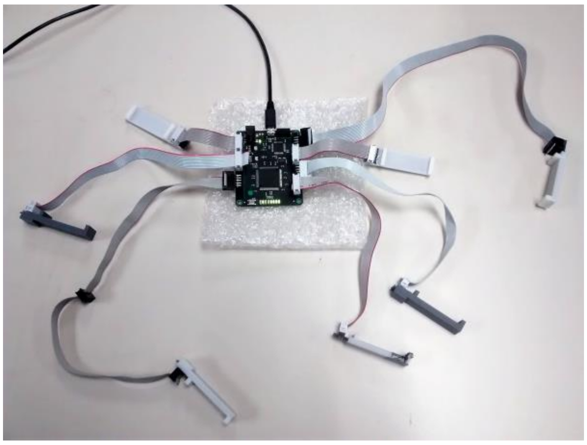

Fig. 5: Modified MSCS firmware setup with seven sensor PCBs

B. Design of a PCB cape for breaking out Mojo IO ports

An important challenge with using the Mojo board in the MSCS was that the connectors used to interface the mojo with the PC were bulky and would occupy considerable area near the Mojo's connector output pin set. In particular, these connectors would occupy more than required lateral space and thus would make some pins of the mojo inaccessible. The following image describes this issue.

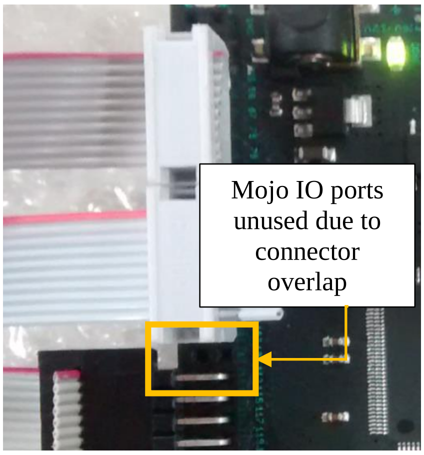

Fig. 6: Effective utilization of Mojo Digital IO pins using orthogonal connectors

This problem was addressed initially by using orthogonal pin connectors at the mojo's IO port connection. These pins were used in conjunction with the 10-pin 2x5 Socket-Socket 1.27mm IDC (SWD) connectors in an alternate fashion, which is shown in figures 7 and 8. This configuration helped solve the problem of ineffective utilization the Mojo's IO ports.

However, this was neither an efficient nor an aesthetically appealing design. At this point, it was decided that a PCB should be designed which would work as a cape for the Mojo. This cape would help in an effective breakout of the sensor signals, from the Mojo to the Low Density sensor PCBs.

Fig. 7: Effective utilization of Mojo Digital IO pins using orthogonal connectors

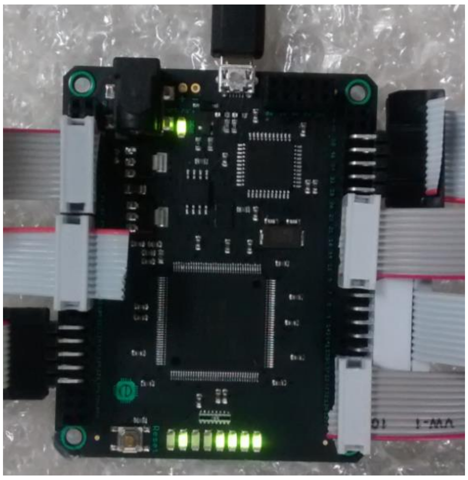

Fig. 8: Connection of seven connectors including 3 orthogonal connectors

The schematic of a cape PCB designed for the Mojo board is shown in figure 9. One of the most important features of this cape is that all the sensor breakout connectors (2x5) lie within the area delineated by two 2x50 header connectors, which are connected to the Mojo. This allows for the cape to be designed with the exact dimensions as that of the Mojo. This helps in ensuring effective usage of volume inside the MSCS module, and improves the modularity of the entire system.

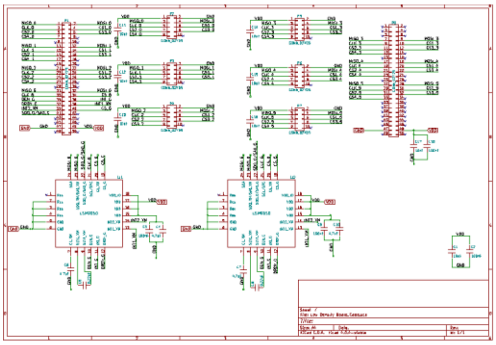

Fig. 9: Schematic of the cape PCB for the low density MSCS version

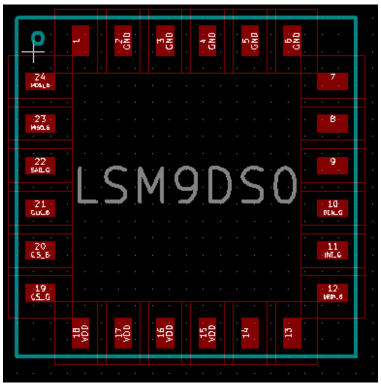

Fig. 10: A custom LGA-24 footprint for the LSM9DS0 prepared in KiCad

The LSM9DS0 was chosen to replace the legacy LSM303D because this IC can also allow for orientation measurement (gyroscope) and movement detection (accelerometer) in more advanced and portable versions of the MSCS (the mTTS). Two footprints of the same IC have been used on the cape to provide redundancy in the event of damage to one of the footprints due to unexpected reasons like shorting/damage caused to IC pads due to overheating during soldering. These ICs are surface mount devices and have considerably small footprint in terms of area, which is 16mm^2 for their LGA-24 footprint. This footprint did not exist in KiCad's open-source library and had to be custom drawn using the KiCad footprint editor. The schematic symbol was also drawn in KiCad using the schematic library editor.

The cape PCB's design is extremely efficient as it provides flexibility for the connections/breakout to be adjusted based on the position of the sensor on the MSCS. This flexibility is provided using 6 identical 2x5 connectors for all the 6 sensor PCB connections. 10-pin ICD headers were used to breakout Mojo signals on the cape. 10-pin ICD ribbon cables will be used as connectors between the headers and the sensor boards on board the MSCS structure. The cables will run along the length of the structure and the orientation that suits best for the space constraint of the model could be selected by the system integrator.

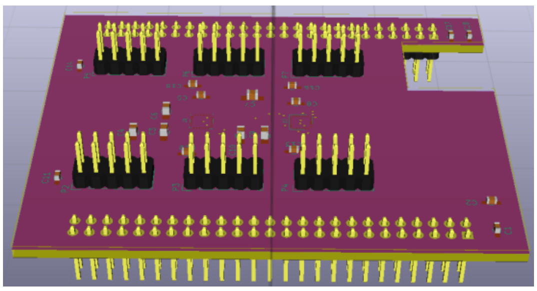

Fig. 11: A side view of the cape PCB

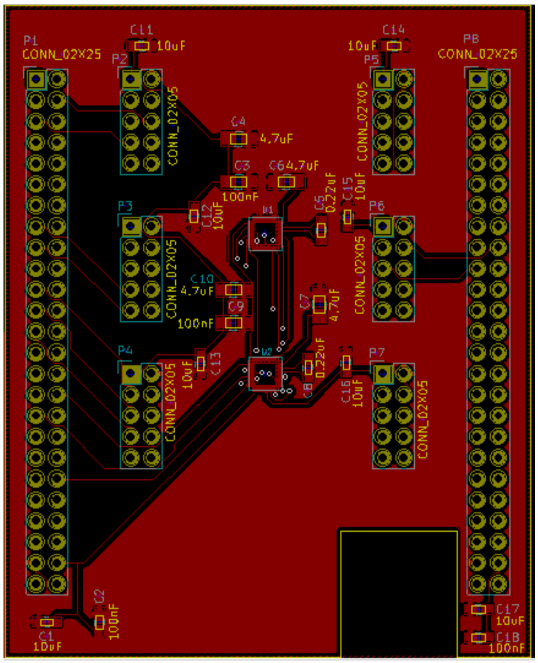

Fig. 12: Top Layer of the cape PCB for the low density MSCS version

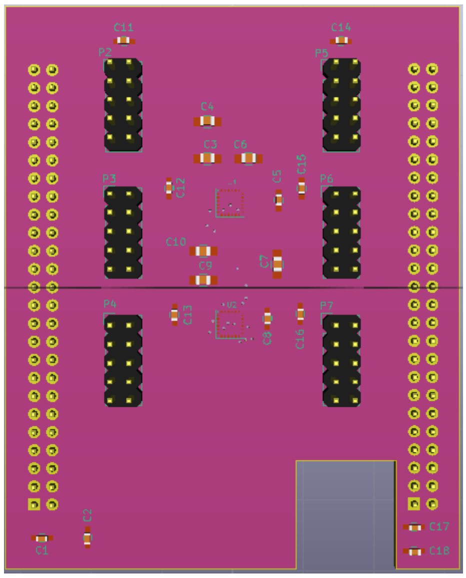

Fig. 13: A 3D view of the Top Layer of the cape PCB

Fig. 14: A 3D view of the Bottom Layer of the cape PCB

Figure 14 shows the position of the two 2x50 header connectors which exactly fit on top of the Mojo board. The dimensions of this cape are 3.1inches by 2.5inches, which are the exact dimensions of the Mojo board.

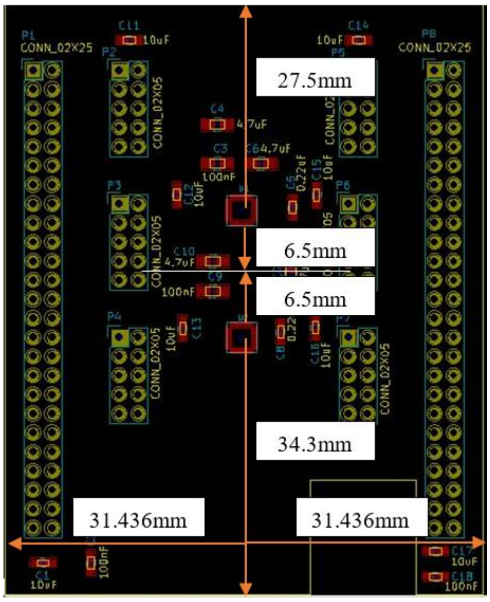

Fig. 15: Top Layer of the cape PCB for the low density MSCS version

The Mojo cape PCB has two LSM9DS0 ICs mounted on its front layer. The magnetometer will allow for measuring Earth's magnetic field, which will be used for a better calibration of the device. This will ensure that the integrity of the data recorded by other magnetometers is upheld. The position of the ICs on the cape are shown in figure 15. These positions allow both the ICs to be fairly located at the center on the top surface of the cape, thus allowing for a symmetric configuration for data measurement and calibration.

The system is designed to hold 6 PCB boards in all, 2 on each of the left, right, and central sides of the MSCS. These boards are low-density boards, with each board containing 6 sensors. Thus, the entire module supports 24 sensors. As shown in an illustration (figure 16) of the current MSCS setup below [6], the blue part is where the mojo and the cape will be mounted, while the six sensor PCBs will be mounted on the front side of the module.

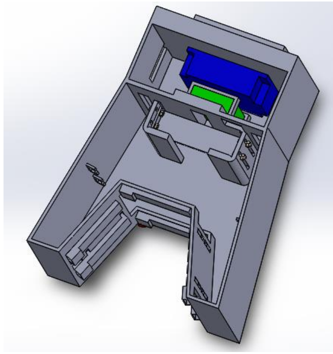

Fig. 16: The low density MSCS setup showing location of the Mojo and cape PCB (blue) [6]

The low density boards are accessed via extremely thin ribbon cables (1.27mm IDC cable) which connect the cape to the sensor boards via a 10-pin 0.5 mm pitch FPC adapter from Adafruit. The cape is connected to the Mojo via a 50-pin 0.5 mm pitch FPC adapter and header pins. The cape board is currently being fabricated, and will be ready to be interfaced with the Mojo, for validation, by the end of Fall 2016. The PCB design was completed using the KiCad software and OSH Park was chosen for PCB fabrication.

V. Conclusion

This project was aimed at working towards an improved functionality of the MSCS by improving its accuracy, reliability, ease of use, and impact on the speech rehabilitation outcome. This project helped the author to gain a deep insight on various challenges faced in a bio-medical device development, the MSCS in this case. The author has strived to describe his understanding of the working of the MSCS hardware, and simultaneously highlight his contribution to the project. The Verilog code which interfaces the Mojo and the low density sensor boards, has been successfully modified to add a new sensor (LSM303D) for better calibration of the device. Attempts have been made to obtain data from the LSM9DS0 instead of the LSM303D but the time constraint has limited the development of Verilog drivers for the LSM9DS0. Furthermore, a cape PCB has been designed for the Mojo board which helps in effective breakout of the Mojo SPI signals to the low density sensor boards on the MSCS. These designs contribute toward increasing the modularity of the system, which is a precursor to a futuristic head-mountable and portable version of the MSCS, the mTTS. While working on this project, the author has acquired valuable skills with respect to system design, Verilog coding, different soldering methods, and 3D printing.

For future work, the head-mountable mTTS needs to be implemented with the high density PCBs, for which a new cape needs to be designed. The next generation mTTS device would also require Verilog drivers to be developed for the LSM9DS0.

VI. REFERENCES

[1] http://memory.ucsf.edu/brain/language/anatomy

[2] https://www.nhlbi.nih.gov/health/health-topics/topics/stroke/causes

[3] Sebkhi N., Desai D., Islam M., Lu J., Wilson K., and Ghovanloo M.. " A Multimodal Speech Capture System for Speech Rehabilitation and Learning." Engineering in Medicine and Biology Society (EMBC), IEEE, 2012.

[4] Cheng, C. , Xueliang H., and Ghovanloo M.. "Towards a magnetic localization system for 3-D tracking of tongue movements in speech-language therapy." Engineering in Medicine and Biology Society, IEEE, 2009.

[5] https://embeddedmicro.com/tutorials/mojo

[6] Testa E. “Special Problem Paper: Tongue Tracking System”, Georgia Institute of Technology, Atlanta, GA, Fall 2016.