

VOICE RECOGNITION SYSTEM

Nikhil Gupta

Siddharth Rawat

Devanshu Arya

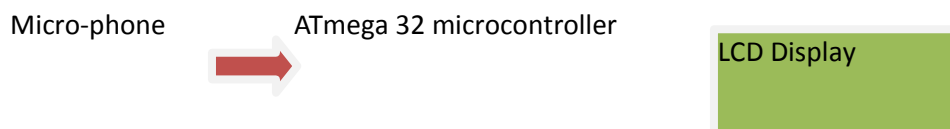
1. PROBLEM STATEMENT

- *To recognize a word spoken by the user*
- *To display the correctly recognized word on a character LCD.*

2. OVERVIEW

- The function of the voice recognition system is to :
 - Create a dictionary of 3-5 words.
 - Compare a test word spoken by the user with the words stored in the dictionary and display the word if a match is found.

3. Basic Block Diagram



4. STRUCTURE

- *The microphone analog circuit goes to the ADC of the MCU. The digitized sampling of the word is passed through the digital **Chebyshev filters** (flash programmed onto the MCU). Once the analysis is done, a **fingerprint** of the word is created and stored in the memory.*

- *While recording the stored fingerprints from the dictionary, against which the fingerprint of the test word is compared using methods such as Euclidean distance and correlation.*
- *At the time of testing, once the comparisons between the test word and the dictionary are made, the LCD which is connected to the MCU displays the word spoken if it has a match in the dictionary.*
- *The software tradeoff in this project is between the number of filters we can implement and the maximum number of cycles we have to adhere to. The more filters there are, the more accurate the speech recognition will be. However, because each filter takes about 320 cycles and we could not implement more than 2000 cycles, we had to trade off the accuracy of the system and limit the number of filters to 5.*

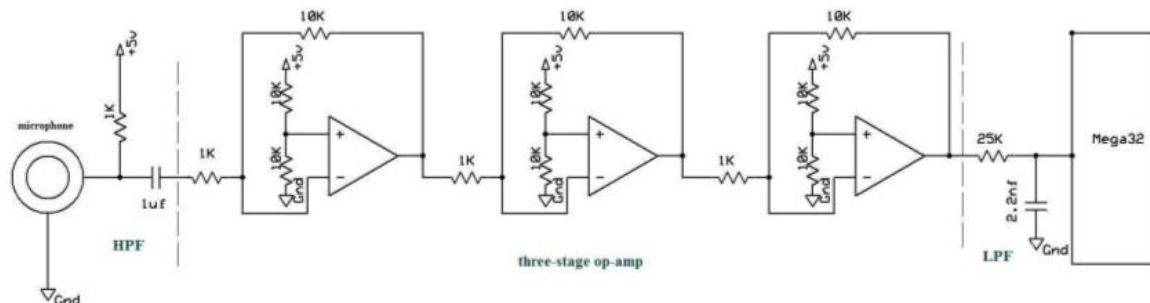
5. ALGORITHMS & TECHNIQUES

- *What we needed to know in this project is how to calculate the frequency to sample speech based on the **Nyquist Rate Theorem**. According to this theorem, the sampling rate must be atleast twice the maximum frequency of the sample input.*
- *We had to learn about **Chebyshev filters** to determine the cutoff frequencies to build the digital filters for human voice.*
- *As for the analysis part of the speech, we needed to know how to calculate **euclidean** distance and **correlation**.*
- ***Tor's speech recognition algorithm**: A complex algorithm for voice recognition, which uses the above-mentioned concepts and algorithms.*

6. SPEECH RECOGNITION ALGORITHM

- *According to the algorithm used in our project, every 250th sample of the input word is analyzed by the digital filters. Each of the five filters then returns some numerical value which is stored in an array.*
- *The process is repeated till the 2000th sample is analyzed, resulting in an array of 40 elements. This is the fingerprint of the spoken word.*

7. CIRCUIT DIAGRAM FOR MICROPHONE CIRCUIT



8. ANALOG CIRCUIT

- The analog circuit consisted of a microphone, a high-pass filter and a series of op-amps.
- The microphone output voltage was of the order of a few millivolts. We had to use two op-amps to get a voltage output of about 5 volts.
- The microphone was directly connected to a capacitor(high-pass filter) to remove signals with frequency less than 50Hz.
- The output of the second op-amp was directly connected to the ADC of the ATmega32.

9. PRACTICAL TESTING

- We started by testing for a single word. However, we faced several problems and had to make several changes to the code. We also tried different words, and finally achieved 100 per cent accuracy for two words. The words used in this case were BYE and ANGRY.
- We then proceeded to perfect the voice recognition system for atleast four words. We attempted several different combination of words such as HELLO, LEFT, CLOSE, ON, OFF, AHEAD and many more, but achieved success with BYE, GO, RIGHT and DOWN.

10. DICTIONARY SIZE AND CORRELATION

- An observation that we made while testing was there was a relation between the number of words of the dictionary and the value of minimum correlation.
- The optimum value of minimum correlation for single word testing was 0.1. This value reduced to 0.0001 when the dictionary size was increased to four words.

11. TASKS COMPLETED

- Voice recognition achieved for two words with 100 per cent accuracy.
- Voice recognition for four words with accuracy close to 80 per cent.

12. FUTURE PROSPECT

- A complete voice recorder and recognizer
- Voice-controlled robotic car
- Voice-controlled security system

13. REFERENCES

- *Tor's Speech Recognition Reference Code:*
<http://www.eecg.toronto.edu/~aamodt/ece341/speech-recognition/sr.c>
- *Bruce's Digital Filters:* _
[Codehttp://instruct1.cit.cornell.edu/courses/ee476/Math/IIRfilterOpt.c](http://instruct1.cit.cornell.edu/courses/ee476/Math/IIRfilterOpt.c)
- http://instruct1.cit.cornell.edu/courses/ee476/FinalProjects/s2005/ik48_ua22/mysite/index.htm
- http://instruct1.cit.cornell.edu/courses/ee476/FinalProjects/s2005/ik48_ua22/mysite/Clabwdf.c
- *electronics.howstuffworks.com/gadgets/high-tech-gadgets/speech-recognition.htm*