International Journal of Advanced Information Science and Technology (IJAIST) ISSN: 2319:2682
Vol.4, No.2, February 2015 DOI:10.15693/ijaist/2015.v4i2.463-467

Handheld Communication Aid For Deaf And Dumb People Based On Asr Technique

Ramnath.K V*[1]and Gopalakrishnan.S[2]

[1]PG Scholar, [2]Assistant Professor

Department of Electronics and Communication Engineering,

PSNA college of Engineering and Technology,

Dindigul -624619, India.

Abstract

Deaf and dumb often communicate via sign language, a kind of representation of words through hand and finger positions. But it has got serious limitations because it is not easy to understand by a normal listener on the opposite and to make things worse, not many in the world know sign language at all. This project aims to build a handheld device that would help to make a frequent communication between both deaf & dumb people and normal people using every day spoken language such as English. Although people who affected by dysarthria may also communicate with anyone. The project can be divided into four modules. The first one is Speech to Image Translating module. It involves advanced Speech Recognition unit (android OS smart phone) and a color display. The second one is text to Voice translating module. It involves touchscreen based text recognition using 65K Color Touchscreen TFT Display. The third module is the ability to send SMS to mobile phones. The fourth module is the language learning ability. In this mode the deaf and dumb can use this device to learn letters, numbers and words.

***Index terms*- Automatic speech recognition (ASR), Audio decoder, Voice output communication aid, AMR Translator.**

I. INTRODUCTION

Speech is one of the oldest and most natural means of information exchange between humanbeings. We as human speak and listen to each other in human-human interface. For centuries people have tried to develop machines that can understand and produce speech as humans do so naturally. Obviously such an interface would yield great benefits. Attempts have been made to develop vocally interactive computers to realize voice/speech recognition. In this case a computer can recognize text and give out a speech output. Voice/speech recognition is a field of computer science that deals with designing computer systems that recognize spoken words. It is a technology that allows a computer to identify the words that person speaks into a microphone or telephone a person speaks into a microphone or telephone. We see there are nearly half of the people suffering from hearing loss (deaf) and speech loss (dumb) that might have occurred since birth or during their lifetime later. Those with sensory disabilities such as deafness, dumbness, or deaf-dumb often rely on their working senses to communicate and have access to the world. In some cases, their impaired senses can be partially repaired to be functional to some degree. The drive for human communication is so strong that unexpected human audio, visual, and tactile communication strategies have evolved throughout time.

Nowadays, the statical techniques prevail over ASR applications. Common speech recognition systems these days can recognize thousands of words. The last decade has witnessed dramatic improvement in speech recognition technology, to the extent that high performance algorithms and systems are becoming available. In some cases, the transition from laboratory demonstration to commercial deployment has already begun. The reason for the evolution of ASR, hence improved is that it has a lot of applications in many

aspects of our daily life, for exampletelephone applications, applications for the physically handicapped and illiterates and many others in the area of computer science. Speech recognition is considered as an input as well as an output during the Human Computer Interaction (HCI) design. HCI involves the design implementation and evaluation of interactive systems in the context of the users' task and work.

II. METHODOLOGY

Speech recognition can be defined as the process of converting an acoustic signal captured by a microphone or a telephone, to a set of words. Automatic speech recognition (ASR) is one of the fastest developing ields in the framework of speech science and engineering. As

the new generation of computing technology, it comes as the next major innovation in man-machine iteration, after functionality of text to speech (TTS), supporting interactive voice response (IVR) system.

Automatic speech recognition

The design of user interfaces for speech-based applications is dominated by the underlying ASR technology. More often than not, design decisions are based more on the kind of recognition the technology can support rather than on the best dialogue for the user. When isolated words are all the recognizer can handle, then the success of the application will depend on the ability of designers to construct dialogues that lead the user to respond using single words. A simple ASR technique is shown in Fig.1.

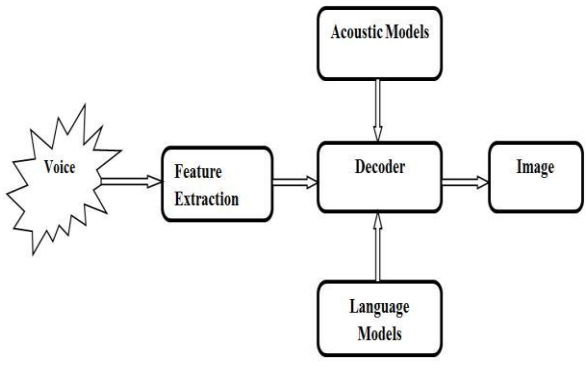


Fig.1- Automatic speech recognition

Word spotting and the ability to support more complex grammars opens up additional flexibility in the design, but can make the design more difficult by allowing a more diverse set of responses from the user. Some current systems allow a limited form of natural language input, but only within a very specific domain at any particular point in the interaction. No systems allow unconstrained natural language interaction, and it's important to note that most human-human transactions over the phone do not permit unconstrained natural language either.

Optical character recognition

Optical character recognition usually abbreviated OCR, is the mechanical or electronic translation of images of and written, type written or printed text. optical character recognition belongs to the family of techniques performing automatic identification. Text extraction may employ binarization or directly process the original image it consist a survey of existing techniques for page layout an analysis. Mathematical morphology is a topological and geometrical based approach for image analysis. It Provides powerful tools for extracting geometrical structures and representing shapes in many applications.

III.PROPOSED WORK

This proposed work composed of four major modules. First, speech to image translation module. Second, text to voice translation module. Third, sending sms module and finally Language learning ability module. These modules are implemented using ASR technique and Optical character recognition to establish a frequent ommunication between deaf and dumb people to normal people without any need of sign language. A handheld communication device was developed, which constitutes of several components including LPC1313 microcontroller, ARM cortex microprocessor (uses version three), etc. All these components combined with software tools to provide information in text or image. This processes were conducted using LPCX presso IDE compiler, AMR Voice translator and Flash magic. The Block diagram of embedded device is shown in Fig 2 and its working functionality is discussed below.

A. Speech to image translation module

In the speech to image translational module should convert the speech signal should modified into binary values. we can assign the binary values into images. The process is to recognize the words spoken by a normal person and to convert this voice input to an image or text and to display it on the screen of the device. The device has a large storage space, a FAT-32 MicroSDmemory card, in order to store all the images needed. We can store the images into the memory card.

B. Text to voice translation module

When we can go to this module ,to click touch screen "swipe enable" or "swipe disable ". It involves touchscreen based gesture recognition using 65K Color Touchscreen TFT Display. The process is to understand and decode the swipe gesture made on the touchscreen and then to speak out this word/alphabet/numeral in a virtual human voice through an MP3 audio decoder. The user will be able to form sentences using this process quite quickly and easily. The color display would help this process by rendering an onscreen swipe keypad layout for the user to input their gestures.

C. Send SMS to mobile phones

In this module send the SMS to all mobile users. The deaf or dumb need to communicate over long distances and hence the device has an inbuilt GSM module to send SMS. In the GSM unit, we can send the MS,Based on the touchscreen display the user can enter his text

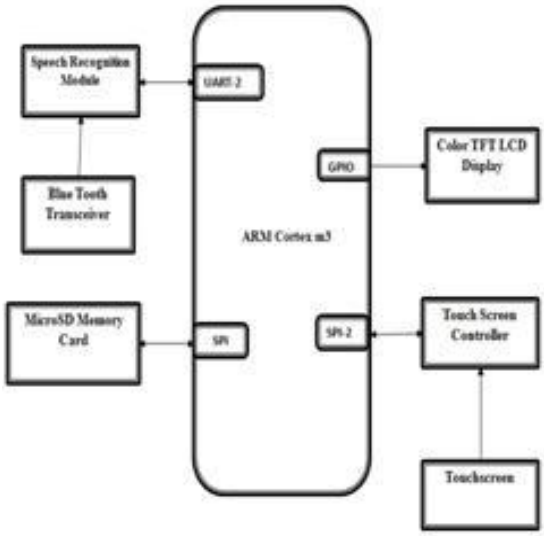
and the mobile number just like in a normal mobile phone to send SMS to others.

D. Language learning ability:

In this mode the deaf and dumb can use this device to learn letters, numbers and words. They will be displayed as pictures in the color display.

Embedded processor

The ARM Cortex™-M3 processor is the industry-leading 32-bit processor for highly deterministic real-time applications and has been specifically developed to enable partners to develop high-performance low-cost platforms for a broad range of devices including microcontrollers, automotive body systems, industrial control systems and wireless networking and sensors. The processor delivers outstanding computational performance and exceptional system response to events while meeting the challenges of low dynamic and static power constraints. The processor is highly configurable enabling a wide range of implementations from those requiring memory protection and powerful trace technology through to extremely cost sensitive devices requiring minimal area.



LPC1313 ARM cortex microcontroller

The LPC1311/13/42/43 are ARM Cortex-M3 based microcontrollers for embedded applications featuring a high level of integration and low power consumption. The ARM Cortex-M3 is a next generation core that offers system enhancements such as enhanced debug features and a higher level of support block integration. The LPC1311/13/42/43 operate at CPU frequencies of up to 72 MHz. The ARM Cortex-M3 CPU incorporates a 3-stage pipeline and uses a Harvard architecture with separate local instruction and data buses as well as a third bus for peripherals. The ARM Cortex-M3 CPU also includes an internal prefetch unit that supports speculative branching. The peripheral complement of the LPC1311/13/42/43 includes up to 32 kB of flash memory, up to 8 kB of data memory, USB Device (LPC1342/43 only), one Fast-mode Plus I2C-bus interface, one UART, four general purpose timers, and up to 42 general purpose I/O pins.

Memory Storage device

Secure Digital (SD) is a non-volatile memory card format for use in portable devices, such as mobile phones, digital cameras, GPS navigation devices ,and tablet computers. The Secure Digital standard was introduced in 1999 as an evolutionary improvement

over MultiMediaCards (MMC). The Secure Digital standard is maintained by the SD Card Association (SDA). SD technologies have been implemented in more than 400 brands across dozens of product categories and more than 8,000 models. Electrically passive adapters allow the use of a smaller card in a host device built to hold a larger card.

Protocols Used

In asynchronous transmitting, teletype-style UARTs (Universal Asynchronous Receiver Transmitter)send a "start" bit, five to eight data bits, least-significant-bit first, an optional "parity" bit, and then one, one and ahalf, or two "stop" bits. The start bit is the opposite polarity of the data-line's idle state. The stop bit is the data-line's idle state, and provides a delay before the next character can start. (This is called asynchronous start-stop transmission). In mechanical teletypes, the "stop" bit was often stretched to two bit times to give the mechanism more time to finish printing a character.A stretched "stop" bit also helps resynchronization.A code format of UART is shown in Fig 3.

SPI is a serial bus standard established by Motorola and It is a synchronous serial data link that operates in full duplex (signals carrying data go in both directions simultaneously).Devices communicate using a master/slave relationship, in which the master initiates the data frame.

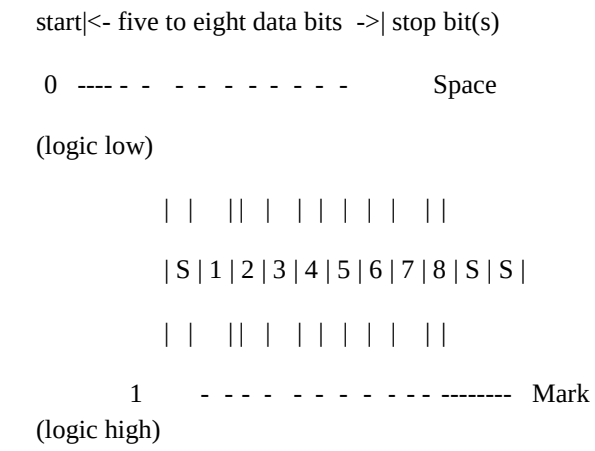


Fig. 3- Asynchronous Code Format of UART

When the master generates a clock and selects a slave device, data may be transferred in either or both directions simultaneously. In fact, as far as SPI is concerned, data are always transferred in

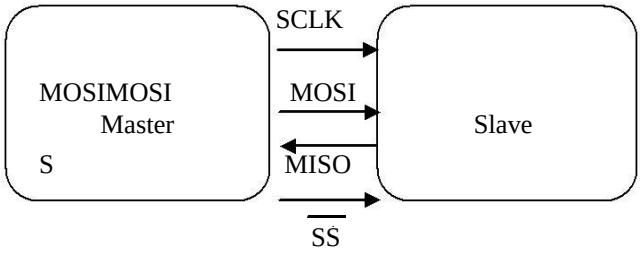


Fig.4-Single Slave of SPI

both directions. It is up to the master and slave devices to know whether a received byte is meaningful or not. So a device must discard the received byte in a "transmit only" frame or generate a dummy byte for a "receive only" frame. A master slave action can be held by either single slave or multi slave. A single slave action of SPI is shown in Fig.4. SPI specifies four signals: clock (SCLK); master data output, slave data input (MOSI); master data input, slave data output (MISO); and slave select (ÇSS).

IV .RESULTS AND DISCUSSION

The speech to image translation module sucessfully translates spoken words into corresponding image stored in the database. A voice signal is recognized using ASR. Initially, a cell phone which supports AMR voice software application should paired with embedded kit using bluetooth. Once after a word is spoke, that word is recognized by ASR and sent to processor. Here, processor acts as master and database as slave. A database checks the received signal and its corresponding word is sent back to the processor and that word is displayed as text or image in TFT display.

To convert text into voice, a keypad is connected with TFT display. A word typed in that keypad is recognized by microcontroller which sends commands through bus to check the word with corresponding words stored in the database. Once the word was matched then it will delivered via speaker.

GSM module is installed in embedded processor to send SMS between two paired devices.

The deaf and dumb can use this device to learn letters, numbers and words through language learning ability. They will be displayed as pictures in the color display.

V.CONCLUSION

There are various speech recognition techniques and translation techniques existing for people affected by dysathria to communicate with converser in this survey were studied. So, in this survey, the Automatic Speech Recognition (ASR) effectively recognize voice with greater tolerance to variability of speech. Automatic speech recognition (ASR) is a high complexity algorithm used to implement on a mobile device. In this survey, it has been concluded that there are several techniques discovered for the development of portable, voice output communication aid controllable by automatic speech recognition. The device can be configured to enable the user to create either simple or complex messages using a combination of a relatively small set of input words. Most of the papers mainly focused on VIVOCA for dysarthria people. Since I concluded, that a VOCA for dysarthria people. Since I concluded, that a Voice Input Voice Output handheld communication aid using a new speech recognition technique is proposed in my work.

The new form of communication mechanism will be introduced in my future work. This proposed work is used for normal people to communicate with deaf and dumb people through the help of ASR. But communication involves response from both the sides. Since, in my future work, a new text to voice translation mechanism will be proposing with the ability to send SMS and language learning capability. Through this technique deaf and dumb people and also dysarthria people can communicate their thoughts to the world.

REFERENCES

[1] L. J. Ferrier, H. C. Shane, H. F. Ballard, T. Carpenter, and A. Benoit, ―Dysarthric speakers' intelligibility and speech characteristics in relation to computer speech recognition,‖ *Augmentative AlternativeCommun.*, vol. 11, no. 3, pp. 165–175, 1995.

[2] M. S. Hawley et al., ―A speech-controlled environmental control system for people with severe dysarthria,‖ Med. Eng. Phys., vol. 29, no. 5, pp. 586–593, 2007.

[3] M. S. Hawley, S. P. Cunningham, F. Cardinaux, A. Coy, S. Seghal, and P. Enderby, ―Challenges in developing a voice input voice output communication aid for people with severe dysarthria,‖ in Proc. AAATE―Challenges Assistive Technol., 2007, pp. 363–367.

[4] V. López-Ludeña, R. San-Segundo, R. Martín, D. Sánchez and A. García, ―Evaluating a Speech Communication System for Deaf People,‖ IEEE latinamerica transactions, vol. 9, no. 4, july 2011

[5] J. Murphy, ―I prefer contact this close: Perceptions of AAC by people with motor neurone disease and their communication partners,‖ *Augmentative Alternative Commun.*, vol. 20, pp. 259– 271, 2004.

[6] M. Parker, S. P. Cunningham, P. Enderby, M. S. Hawley, and P. D. Green, ―Automatic speech recognition and training for severely dysarthric users of assistive technology: The STARDUST project,‖ Clin. Linguistics Phonetics, vol. 20, no. 2–3, pp. 149–156, 2006.

[7] L. R. Rabiner and B. H. Juang, ―An introduction to hidden Markov models,‖ IEEE ASSP Mag., Jun. 1986.

[8] U. Sandlerand Y.Sonnenblick,―A system for recognition and transla- tion of the speech of handicapped individuals,‖ in 9th Mediterranean Electrotech. Conf. (MELECON'98), 1998, pp. 16–19.

[9] H. V. Sharma and M. Hasegawa-Johnson, ―State-transition interpola- tion and MAP adaptation for HMM-based dysarthric speech recogni- tion,‖ in NAACL HLT Workshop Speech Language Process. Assistive Technol., 2010, pp. 72–79.

[10] N. Thomas-Stonell,A.L.Kotler, H.A. Leeper, and P. C. Doyle, ―Computerized speech recognition: Influence of intelligibility and perceptual consistency on recognition accuracy,‖ *Augmentative Alternative Commun.*, vol. 14, no. 1, pp. 51–56, 1998.

[11] J. Todman, N. Alm, J. Higginbotham, and P. File, ―Whole utterance approaches in AAC,‖ *Augmentative Alternative Commun.*, vol. 24, no. 3, pp. 235–254, 2008.

[12] B.Wisenburn and D. J. Higginbotham, ―An AAC application using speaking partner speech recognitiontoautomaticallyproducecontextuallyrelevantutterances:Objectiveresults,‖AugmentativeAlternative Commun., vol. 24, no. 2, pp. 100–109, 2008.

[13] B.Wisenburn and D. J.Higginbotham, ―Participant evaluations of rate and communication efficacy of an AAC application using natural language processing,‖ Augmentative Alternative Commun., vol. 25, no. 2, pp. 78–89, 2009.

[14] S. Young et al., Cambridge University Engineering Department The HTK book, 2006.