© 2014 IJIRT | Volume 1 Issue 5 | ISSN : 2349-6002

SPEECH RECOGNITION MICROPROCESSOR

Sonarka Maini, Saurabh Jakhmola
Information Technology
Dronacharya College Of Engineering

***Abstract*- This study focused on advantages or applications of Speech Recognition Microprocessor.In computer science and electrical engineering, speech recognition is the translation of spoken words into text. It is also known as "automatic speech recognition" , "computer speech recognition" or just "speech to text" .Speech recognition applications include voice user interfaces such as voice dialling , routing, domotic appliance control, search , simple data entry , preparation of structured documents, speech-to-text processing and aircraft.**

The aim of this master's thesis is to develop an interactive and speaker independent speech recognition system. The system shall be able to identify predetermined keywords from incoming live speech and in response, play audio files with related information. Moreover, the system shall be able to provide a response even if no keyword was identified. For this project, the system was implemented using PocketSphinx, a speech recognition library.

***Index Terms*- Speech Recognition, Microprocessor, PocketSphinx**

I. INTRODUCTION

Automatic speech recognition systems have caught the attention of researchers since the middle of the 20th century. From the initial attempts to identify isolated words using a very limited vocabulary, to the latest advancements processing continuous speech composed by thousands of words, the ASR technology has grown progressively.

Voice recognition systems, which today is one of gradually started to gain importance. In this study, using a PIC18F452 microcontroller, two registered words, compared with a voice command input. Registered with the sound of the word overlap as a result of the exit sign is active. Recognition of the audio signals zero crossing for the count and time was examined. In general, the voice recognition process, sound recording and determination of the expression; processing of sound, making comparison and integration of rest; finally, the expression was created with the realization of the corresponding functions. Developed in this system, without sounds recording of audio signal processing are performed simultaneously. In this system, not a waste of time for this situation may be many advantages to the system. Work is relatively simple and economically cheap.

The study purposes of the system are Mobile applications represent a very important area where ASR systems can be used, for example in GPS navigation systems for cars, where the user can provide navigation commands by voice, or a speech-based song selector for a music player. However, a large amount of existing portable devices lack the processing power needed to execute a high-end speech recognizer. For this reason, there is a large interest in designing and developing flexible ASR systems that can be run on both strong and resource-constrained devices.

Speech recognition applications

II. HISTORY

The technology of Automatic Speech Recognition (ASR) and Transcription has progressed greatly over the past few years. Ever since research of this technology began in 1936, the largest barriers to the speed and accuracy of speech recognition was computer speed and power (or lack there of). With the average the CPU now at and above a Pentium III and RAM levels at 500 MB and up, accuracy levels have reached 95% and better with transcription speeds at over 160 words per minutes.

As mentioned above, the study of automatic speech recognition and transcription began in the 1936 with ATT&T's Bell Labs. At that time, most research was funded and performed by Universities and the U.S. Government (primarily by the Military and DARPA - Defense Advanced Research Project Agency. It was not until the early 1980's when the technology reached the commercial market.

Like most emerging technologies, there were several competing research "camps", each working independently to develop speech recognition. Please view the Speech Recognition Timeline to get a full view of its development.

The first company to launch a commercial product was Covox in 1982. Covox brought digital sound (via The Voice Master, Sound Master and The Speech Thing) to the Commodore 64, Atari 400/800, and finally to the IBM PC in the mid '80s. Along with (or bundled) this introduction of sound to computers came Speech Recognition.

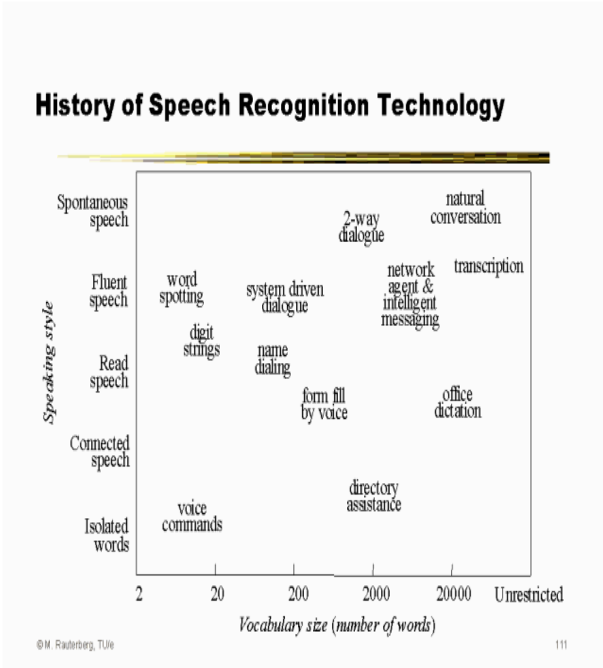
Another company that was founded in 1982 and whose eventual product has become the overwhelming leader in the speech recognition market was Dragon Systems. Scansoft, Inc. now owns and manufactures this product, Dragon Naturally Speakin

Dragon Systems was founded in 1982 by James and Janet Baker to commercialize speech recognition technology. As graduate students at Rockefeller University in 1970, they became interested in speech recognition while observing waveforms of speech on an oscilloscope. At the time, systems were in place for recognizing a few hundred words of discrete speech, provided the system was trained on the speaker and the speaker paused between words. There were not yet techniques that could sort through naturally spoken sentences. James Baker saw the waveforms--and the problem of natural speech recognition--as an interesting pattern-recognition problem.

Rockefeller had neither experts in speech understanding nor suitable computing power, and so the Bakers moved to Carnegie Mellon University (CMU), a prime contractor for DARPA's Speech Understanding Research program. There they began to work on natural speech recognition capabilities. Their approach differed from that of other speech researchers, most of whom were attempting to recognize spoken language by providing contextual information, such as the speaker's identity, what the speaker knew, and what the speaker might be trying to say, in addition to rules of English. The Bakers' approach was based purely on statistical relationships, such as the probability that any two or three words would appear one after another in spoken English. They created a phonetic dictionary with the sounds of different word groups and then set to work on an algorithm to decipher a string of spoken words based on phonetic sound matches and the probability that someone would speak the words in that order. Their approach soon began outperforming competing systems.

In 2000, Lernout & Hauspie acquired Dragon Systems. In 2001, Scansoft, Inc. acquired all rights to Lernout & Hauspie's speech recognition products including Dragon Naturally Speaking. In 2003, Scansoft, Inc. acquires Speechworks.

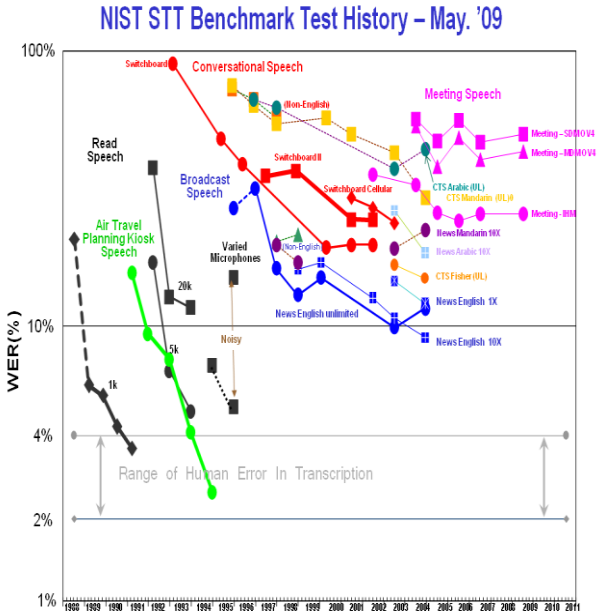
Scansoft, Inc. is presently the world leader in the technology of Speech Recognition in the commercial market.



III. DARPA TEST HISTORY

The Defense Advanced Research Projects Agency (DARPA) is an agency of the United States Department Of Defense responsible for the development of new technologies for use by the military. DARPA has been responsible for funding the development of many technologies which have had a major effect on the world, including computer networking, as well as NLS, which was both the first hypertext system, and an important precursor to the contemporary ubiquitous graphical interface user .

DARPA began as the Advanced Research Projects Agency (ARPA) created in 1958 by President Dwight D Eisenhower for the purpose of forming and executing research and development projects to expand the frontiers of technology and science and able to reach far beyond immediate military requirements. The administration was responding to the Soviet launching of Sputnik 1 in 1957, and ARPA's mission was to ensure U.S millitary technology be more sophisticated than that of the nation's potential enemies. From DARPA's own introduction:



DARPA Benchmark Test History

IV. Dual ALU Speech Recognition Microprocessor

In the near future, speech recognition will become the method of choice for controlling appliances, toys, tools, computers and robotics. There is a huge commercial market waiting for this technology to mature. This paper proposed a low-cost dual-ALU processor for speech recognition. It does not give an emphasis on sophistication but on low-cost solution.

The dual-ALU architecture provides parallel calculation capability. For the consideration of chip size, the area of the second ALU is only half of the first ALU. We use hardware-software codesign method to implement the speech recognition. The feature extraction bases on LPC-cepstrum coefficients, and template matching employs hidden Markov models (HMM). The processor is designed to process the HMM and connect with the ASIC of LPC-cepstrum.

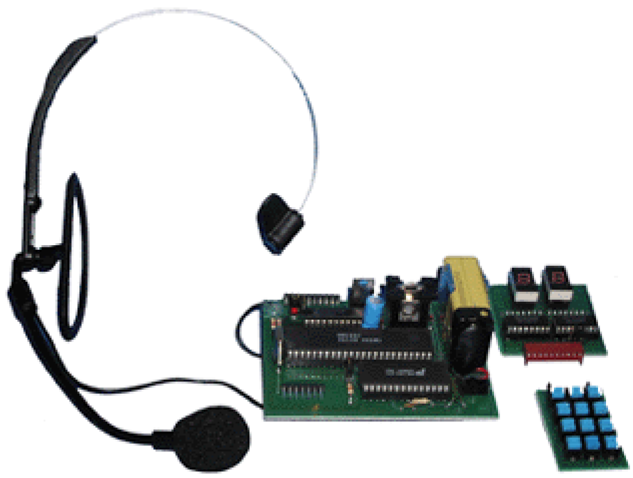
V. SONIC Speech Recognition Microprocessor

The Sonic speech recognition system was developed at the University of Colorado . The recognizer is based on decision-tree state-clustered triphone hidden Markov (CDHMM) models. In such an architecture, the basic building blocks of sounds, also known as phonemes, are characterized by 3 Markovstates. The three states are used to model the statistical distribution at the beginning, middle, and end of the sound unit. Each state is modeled by a mixture of multivariate Gaussian During recognition, features are extracted every 10 ms from the audio signal. The features are represented as a floating-point vector containing 12 Mel Frequency Cepstral Coefficients (MFCCs), energy, and the first and second differences of these parameters. The features encode the short-time spectrum of the audio signal. The resulting feature vector stream has a dimension of 39. Feature computation is extremely efficient consuming less than 1% of the CPU in a typical speech recognizer.

The search network in Sonic is represented using a reentrant static tree-lexicon . Here, phonetic word pronunciations are expanded into sequences of context-dependent HMM states and inserted into a tree-based search structure for efficiency. In this framework, words that share similar initial phoneme sequences also share the same root node in the lexical search tree.

The recognizer implements a two-pass search strategy. The first pass consists of a frame-synchronous, beam-pruned Viterbi token-passing search . Tokens are used to maintain
the state of the recognizer.

CPU requirements for the second-pass search are often negligible compared to the first-pass search.



Circuit Of Speech Recognition Microprocessor

VI. Advantages of Speech Recognition Microprocessor

- Speech is a very natural way to interact, and it is not necessary to sit at a keyboard or work with a remote control.
- No training required for users.
- Flexibility
- Reusable Data

VII. Disadvantages Of Speech Recognition Microprocessor

- Even the best speech recognition systems sometimes make errors. If there is noise or some other sound in the room (e.g. the television or a kettle boiling), the number of errors will increase.
- Speech Recognition works best if the microphone is close to the user (e.g. in a phone, or if the user is wearing a microphone). More distant microphones (e.g. on a table or wall) will tend to increase the number of errors.

VIII. Applications of Speech Recognition Microprocessor

- **In-car systems**

Typically a manual control input, for example by means of a finger control on the steering-wheel, enables the speech recognition system and this is signalled to the driver by an audio prompt. Following the audio prompt, the system has a "listening window" during which it may accept a speech input for recognition.

- **Health care**

In the health care sector, speech recognition can be implemented in front-end or back-end of the medical documentation process. Front-end speech recognition is where the provider dictates into a speech-recognition engine, the recognized words are displayed as they are spoken, and the dictator is responsible for editing and signing off on the document. Back-end or deferred speech recognition is where the provider dictates into a digital dictation system, the voice is routed through a speech-recognition machine and the recognized draft document is routed along with the original voice file to the editor, where the draft is edited and report finalised. Deferred speech recognition is widely used in the industry currently.

- **Military**

High-performance fighter aircraft

Substantial efforts have been devoted in the last decade to the **test** and evaluation of speech recognition in fighter aircraft. Of particular note is the U.S. program in speech recognition for the Advanced Fighter Technology Integration (AFTI) aircraft , and a program in France installing speech recognition systems on Mirage aircraft, and also programs in the UK dealing with a variety of aircraft platforms. In these programs, speech recognizers have been operated successfully in fighter aircraft, with applications including: setting radio frequencies, commanding an autopilot system, setting steer-point coordinates and weapons release parameters, and controlling flight display.

- **Helicopters**

Training air traffic controllers

Training for air traffic controllers (ATC) represents an excellent application for speech recognition systems. Many ATC training systems currently require a person to act as a "pseudo-pilot", engaging in a voice dialog with the trainee controller, which simulates the dialog that the controller would have to conduct with pilots in a real ATC situation. Speech recognition and synthesis techniques offer the potential to eliminate the need for a person to act as pseudo-pilot, thus reducing training and support personnel. In theory, Air controller tasks are also characterized by highly structured speech as the primary output of the controller, hence reducing the difficulty of the speech recognition task should be possible**.**

- **Telephony and other domains**

ASR in the field of telephony is now commonplace and in the field of computer gaming and simulation is becoming more widespread. Despite the high level of integration with word processing in general personal computing. However, ASR in the field of document production has not seen the expected increases in use.

IX. Conclusion

The study is concluded as the importance of Speech recognition microprocessor.It is explained by its architecture and various uses of it.The history is also well explained.Advantages and disadvantages are also explaines with its various applications.

Finally, I would like to state that in the next generation of wireless communication systems, there will be a need for the rapid deployment of independent users. Since network scenarios cannot rely on centralized and organized connectivity and can be conceived as applications of Speech

Recognition Microprocessor. So, it becomes the best solution of different problems of network.

REFERENCE

[1] J. Ahola, Ambient Intelligence, ERCIM (European Research Consortium for Information and Mathematics) NEWS, N. 47, October 2001.

[2] B. Kalaavathi, et. Al, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1470874Raj Kamal, Mobile Computing, Oxford University Press, 2009.

[3] Proceedings IT in Academics, Sinhgad Institute of Management, Pune, India, 2009.

[4] G. Anastasi, M. Conti, E. Gregori, A. Passarella, A power saving architecture for web access from mobile computers, in: Proceedings of the Networking 2002, Lecture Notes in Computer Science, vol. 2345, Springer, Berlin, 2002.

[5] Various sites like google,Wikipedia,yahoo

[6] Various IEEE papers and projects