VOICE OPERATED MULTILINGUAL INFORMATION DISPLAY/RETRIEVAL SYSTEM

Abdul Mobin, Anil Kumar, and S.S.Agrawal

CENTRAL ELECTRONICS ENGINEERING RESEARCH INSTITUTE CENTRE, NEW DELHI, INDIA.

ABSTRACT

This paper describes the design and development of a voice operated multilingual information display and retrieval system. The words or commands spoken by a speaker can be displayed on the multilingual terminal for varification and data entry applications. Previously stored information can also be accessed through a spoken word command and displayed on a CRT terminal.

The system consists of an Isolated Word Recognition Unit, an interface unit and multilingual CRT terminal. The design aspect and the functioning of the system are described in detail The code generated by the recognition unit corresponding to the recognized word is accepted by the interface card through the Z-80PIO Having compared this code with the code stored in the interface card,the information to be displayedis sent to the CRT terminal using serial communication. The system is based on the Z-80 microprocessor and utilises 64 KB of RAM/EPROM for storage of information and software programmes.ASCII data of the information to be displayed corresponding to the recognized word is sent by the interface card and displayed on the terminal. The terminal use a software called GIST and has the capability of displaying some Indian and Roman scripts. The information can also be displayed in different font and graphics forms.

1. INTRODUCTION

Speech is the most natural and efficient means for humans to communicate with each other as well as with machines. Among the human-machine interaction system speech recognition is an attempt to have machines responding to spoken commands. Attempts are being made to have automatic speech recognition systemof various complexities.However, the success so far is limited to handle small vocabulary and to great extent they are speaker dependent systems. This paper describes the development of a voice operated multilingual information display/retrieval system and the flexibilities ofit's practical applications. Both the word recognition system as well as the interface card for displaying the information are Z-80 A microprocessor based hardware modules.

Interesting applications are to be anticipated with automatic speech recognition systems in the areas of industrial controls, transport,communication and military technologies. As one of the interesting applications, the automatic spoken word recognition system can be successfully used for the display and retrieval of stored informations on a CRT by voice commands. The CRT used for this purpose is equipped with a GIST card which accepts serially the ASCII codes of the information to be displayed. To do this a controlling unit has been designed and developed at CEERI. The unit is capable of supplying the serial ASC II codes of information desired to be displayed on the CRT.

With GIST card incorporated in the above system it is possible to display the information in any of the several Indian scripts and Roman script for English either simultaneously or in one of the selected scripts on the screen. Since India is a multilingual country, the objective of developing the Voice Operated Information Display System is to have interaction with computers using multilingual scripts.

2. DESCRIPTION OF THE SYSTEM

The system consists of two major parts, one is a word recognition unit and the other is an interface unit for informtion display. These two units are suitably interfaced as shown in the figure 1.

2.1 Spoken Word Recognizer

The recognition processor is a single board microcomputer based on a Z-80 A CPU. It consists of 64 Kbytes of RAM/EPROM, Parallel interface adapter(PIA), Counter Timer Circuit(CTC) and two Asynchronuous Adapters (ACIA). One ACIA interfaces to the video terminal controller unit.

The system is speaker dependent but it can be easily trained for a new speaker.The speaker has to speak through a close talking microphone. This speech signal is fed to signal conditioning circuitry which consists of pre-amplifier, equalizer, automatic volume control etc. The preprocessed signal is then passed through an audio spectrum analyzer chip ASA-16.It has 16 contiguous band pass filters covering the frequency range of 200Hz to 7KHz. Each bandpass filter is follwed by a rectifier and a low pass filter with cut off frequency of 25 Hz. This arrangement gives the slowly varying signal proportional to sound pressure level in that channel. The ASA-16 also has 16-channel multiplexer which provides spectral data for each of the 16 frequency bands. The multiplexed signal is allowed to pass through the 8-bit analog-to-digital conveter and this digitised data is stored in the memory.

The underlying principle behind the isolated word recognizer is pattern matching. In the first instance, patterns of all the words corresponding to a vocabulary are obtained and their templates stored in the system memory. Next, the pattern of the test word is generated and coma

pared with the templates recorded earlier. In case of a correct match there is a display of the correctly recognized word on the front pannel of the machine. Corresponding to the matched word the system generates a specific code. The code is accepted by the interface unit through I/O device (Z-80 PIO). This code is further processed for another comparison with prestored codes. Having compared with the correct code this interface unit sends the ASCII data of the stored information to CRT terminal through serial communication.

2.2 Interface Unit

This unit is also based on Z-80A CPU which is driven by 4.0 MHz clock. It uses a special purpose integrated circuit designed to ease the problems associated with serial communication i.e. Universal Synchronuous/Asynchronuous Receiver and Transmitter (USART).

The unit provides a three line serial communication port. One line transmits character to the terminal, second line receives character entered at the terminal keyboard and the third line simply provides a common system ground connection. Information is put on these wires one bit at a time and in this way it takes eight seperate transmissions to send a whole byte. The baud rate i.e. speed at which the data is transmited is chosen to be 9600 in our case. However, this rate depends upon the specific application and the equipment involved. Since there is no handshaking in this I/O process, it is the microprocessor's responsibility to send characters not faster than speed at which the termial can receive. Similar method is used to accept the incoming characters at the frequent intervals to avoid missing data being sent by the terminal.

To communicate with the terminal, the transfers to and from interface unit are conveyed by using the American standard EIA-RS232 C, because the data are no longer activated by current intensity but by voltage level i.e. +12V and -12V. Also such levels are not TTL compatible and can not be provided by the microprocessor. Thus a physical interface is interposed using a driver and a receiver circuit. The functional diagram of the complete system is shown in Fig.2

3. Testing/Features of the System

The system is tested with 10 words vocabulary spoken by both male and female speakers. Though it can accomodate 40 words vocabulary. The words are spoken once into a microphone to generate the reference pattern. Corresponding to 10 words vocabulary 10 pages of stored information can be retreived. Transcript provides emulation for the DEC VT-52 and Ansi compatible terminals for display of English as well as Indian scripts. Another mode displays high resolution, print quality characters for Indian scripts which can be used for text processing requirements.

ACKNOWLEDGEMENT

The authors are grateful to Dr.W.S.Khokhle, Director CEERI for kind support and encouragement for the above work. The authors are thankful to Mr.M.Ganesan for various discussions and help. They wish to acknowledge the financial support of Deptt. of Electronics, Govt. of India.

REFERENCES

1. Abdul Mobin, S.S.Agrawal and K.D.Pavate "An Online Spoken Word Recognition System using a microprocessor", Proc. European conf. on Speech Technology, Edinburgh, U.K. Vol. 2, pp.296-299, Sept 2-4, 1987.

2. C.K.Chaung and S.W.chan, "Speech Recognition using Variable Frame Rate Coding", IEEE Proc. of Int CASSP, pp.1033-1036, 1983.

3. R.Pieraccini & R.Billi, "Experimental Comparison among Data Compression Techniques in Isolated Word Recognition" IEEE Proc. of ICASSP, pp. 1025-1028, 1983.

4. B.A. Dautrich, L.R.Rabiner and T.B.Martin, "On the use of Filter Bank Features for Isolated Word Recognition" IEEE Conf. ASSP, pp. 1061-1064, 1983.

5. W.Daxer, Speech Communication, Vol.1, No.1, 21, 1982.

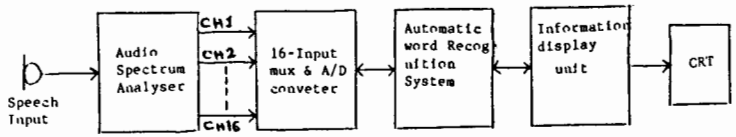

Fig.1. Block Diagram Of the System.

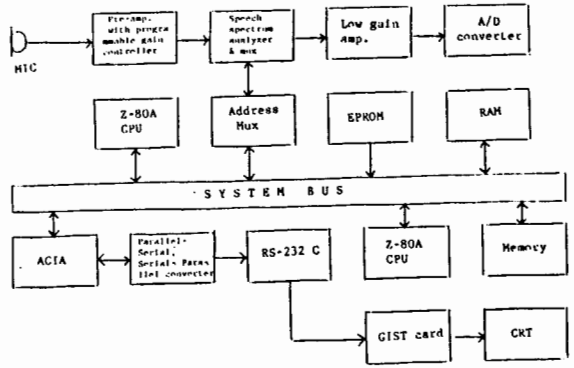

Fig.2. Functional Diagram Of the System.