Speech To Speech Language Translator

Umeaz Kheradia, Abha Kondwilkar

B.E (Electronics & Telecommunication), Rajiv Gandhi Institute Of Technology.

Abstract- In the nascent stage of developing a personalised interpreter, we develop a prototype which uses a speech processing hardware and online translators to provide the user with real time translation. Speech processing hardware works on the principle of 'compare and forward', i.e., a database is already stored in the unit which is used for comparing with the input speech and the result is forwarded for further processing. The need arises from the inability of dictionaries and human translators to suit our needs for better communication. In this situation the prototype proposed will suffice the purpose reasonably well and minimise the communication inefficiencies.

Index Terms- Speech Processing Unit, Recognition, Database, Online Translator, Control System

I.   INTRODUCTION

Language is the asset which differentiates human communication from that of animals. Thus, it is of utmost importance that all individuals can communicate with each other without any hindrances. But due to the diversity in geography and thus in cultures, many languages have been nurtured parallel to each other. Today this cultural gap is bridged by interpreters. But the interpreters have their own limitations.

At present we need 'Personalised Interpreters' which will reduce our dependence on dictionaries and human interpreters. This will reduce the hindrance posed by the language barrier. In this paper, we will mainly focus on the prototype of the above mentioned personalised interpreter.

This prototype will comprise of a microcontroller which will act as a link between the various blocks. The use of speech IC's will be made for the speech to text conversion. To enable internet accessing we shall be using a GSM module. The speaker at the output end will provide the desired translated version of the input.

II.   RELATED WORK

Photocopier as a Translator: A device has been developed by Fuji Xerox which can scan a printed sheet of Japanese text from a newspaper or magazine and give out its translated version in Chinese, English or Korean, while retaining the original layout. [1]

Non-talking dictionaries: They allow immediate word translation and back translation with the help of a predecided database. These dictionaries are bidirectional, trilingual and multilingual.

Speech-to-Speech Electronic Translators: They make use of the latest advancements in the field of speech recognition. They understand what you say in your own language and immediately repeat the same in one or more languages. [2]

Bing Translator: It is a service provided by Microsoft as part of its Bing services to translate texts or entire web pages into different languages. It is a text to text online translator.
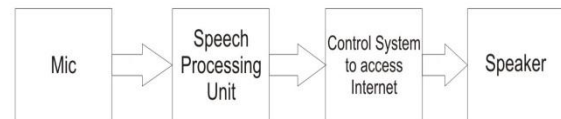
III.   CONCEPT

Figure 1: Basic Block Diagram

The input speech is given through the microphone which then goes to the speech processing unit. This unit processes the input and the word which was spoken is recognised. This word is forwarded to the online translator for translation and the output is given out through the speaker.

IV.   WORKING

The input speech first goes to the speech IC (HM2007) of the speech processing unit. This IC works in two modes: [3]
·   Training: Stores the database.
·   Recognition: Compares with the database.

The IC requires an external memory which is sufficed by an SRAM. Speech IC along with the static RAM forms the fundamental block of the speech processing unit. The database is stored in the SRAM and then HM2007 is used in the recognition mode. As mentioned above, in recognition mode, comparison of the input and the database takes place and a particular eight bit BCD address is given as the result. This address gives the location of the found word or it represents the exception which takes place. The address remains on the D-bus of the speech IC for a very small time which may prove to be insufficient for extracting all the eight bits of the address. Thus, a latch is used which increases the time to read the address.
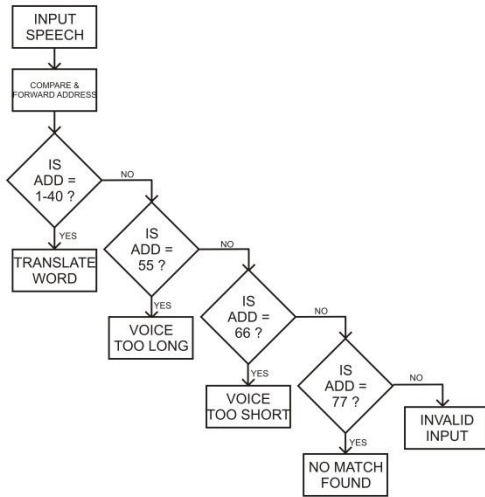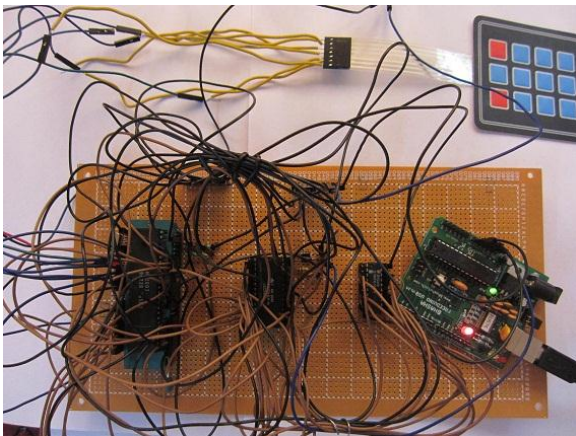
Figure 2: Flow Chart

The address is forwarded to the control system, in this case, a personal computer via an arduino. The arduino converts the parallel address to serial form for serial communication with the personal computer.

A user interface is created which performs the following functions:

- BCD to decimal conversion.
- If address is between 1 and 40, forward it to the online translator.
- Allows to select the language to be translated in.
- Generates a message box if an exception occurs

V. RESULTS AND FUTURE SCOPE

Figure 3: Final Circuit Board

Figure 4: User Interface

The input speech given through the mic is detected by the speech IC HM2007. It is then compared with the stored database. Let the input word be cat. Its address in the database is 6 (binary coded decimal form).Therefore the input is present in the database. Thus, we can see its address in freeduino window on the computer. Then in the user interface created using Visual Basic we select the language in which we desire to see the output. If we select Japanese we can see the translated version, i.e., 'neko' on the screen. As expected that the translated output should be in speech form could not be made possible.

The future scope of this research could be in two different ways:

1. <u>Personalised Interpreter:</u> The prototype can be developed into a standalone device which could serve as the 'Personalized Interpreter'. This would be revolutionary in communication technology.
2. <u>Integrating with Mobile Phone:</u> Using Dictionary as the database: The speech processing unit can also be embedded into presently available mobile phones. The dictionary present in the mobile phones can serve as the database   for the speech processing unit. As the number of words in the dictionary is usually inexhaustible this could prove to be an efficient and excellent translator.

Speech Processing Software: With the advent of high speed processors in present day cell phones,   speech processing in real time is not a dream anymore. Thus a monolithic chunk of code could be   enough for converting speech to text without pre-existing database. This would increase the gamut of the language which needs to be translated.

REFERENCES

[1] http://blog.translatus.com/2007/10/fuji-xerox-introduces-copier-translator.html
[2] http://www.ectaco.com/Electronic-Dictionary
[3] Datasheet of HM2007

AUTHORS

First Author – Umeaz Kheradia: B.E (Electronics & Telecommunication), Rajiv Gandhi Institute of Technology. umeaz_kheradia17@yahoo.com

Second Author – Abha Kondwilkar: B.E (Electronics & Telecommunication), Rajiv Gandhi Institute of Technology. abhassk@yahoo.co.in