CHAPTER – 2

Sound / Audio System:

Sound:

Introduction:

Sound is a physical phenomenon produced by the vibration of the matter, such as a violating string or a block of word. As the matters, pressure vibrations are created in the air surrounding it. This alternation of high and low pressure is propagating through air in a wave like motion. When a wave reaches the human ear, a sound is heard.

Basic Sound Concepts:

Sound is produced by the vibration of matter. During the vibration, pressure vibrations are created in the air surrounding it. The pattern of the oscillation is called a waveform.

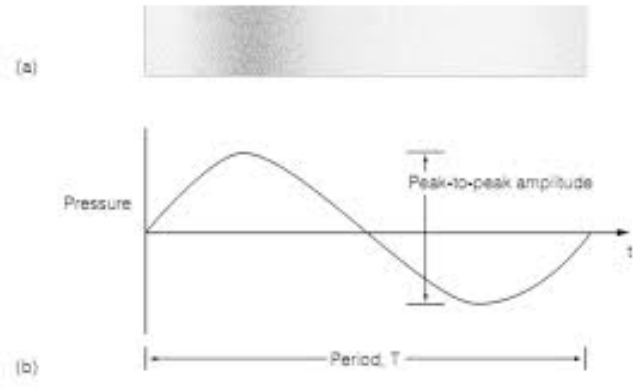


Figure: - Oscillation of an air pressure wave

The waveform repeats the same shape at regular intervals and this pattern is called a period.

Frequency:

The frequency of sound is reciprocal value of the period; it represents the number of periods in a second and measured in hertz (Hz) or cycles per second (cps).

The frequency range is divided into:

Infra – sound – From 0 to 20 Hz
Human hearing frequency range – From 20 Hz to 20 KHz
Ultrasound – From 20 KHz to 1GHz
Hyper sound – From 1GHz to 10 THz

Multimedia systems typically make use of sound only within the frequency range of human hearing.

Amplitude:

A sound also has amplitude, a property subjectively heard as loud less. The amplitude of a sound is the measure of the displacement of the air pressure wave from its mean or quiescent state.

Computer Representation of Sound:

The smooth, continuous curve of a sound waveform is not directly represented in a computer. A computer measures the amplitude of the waveform at regular intervals to produce a series of numbers. Each of these measurements is a sample.

Figure: - Sample waveform

Sampling Rate:

The rate at which a continuous waveform is sampled is called sampling rate. Sampling rates are measured in Hz.

CD standard sampling rate: 44100 Hz

Quantization:

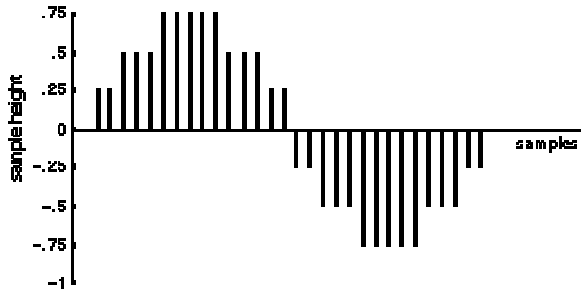


Figure: - Three- bit quantization

Just as a waveform is sampled at discrete times, the value of the sample is also discrete. The resolution or quantization of sample value depends on the number of bits used in measuring the height of the waveform.

- An 8-bit quantization yields 256 possible values.
- 16-bit CD-quality quantization results in over 65536 values.

Sound Hardware:

Before sound can be processed, a computer needs input/output devices. Microphone jacks and built-in speakers are devices connected to an ADC and DAC, respective for the input and output.

Music:

The relationship between music and computers has become more and more important, especially considering the development of MIDI (Musical Instrument Digital Interface) and its important contributions in the music industry today.

- The MIDI Interface between electronic musical instruments and computers is a small piece of equipment that plugs directly into computer's serial port and allows the transmission of music signals. MIDI is considered to be the most compact interface that allows full-scale output.

Computer Music-MIDI:

- Sound waves, whether occurred natural or man-made are often very complex. i.e. they consist of many frequencies.
- Digital sound is relatively. Straight forward to record complex sound. However, it is quite difficult to generate complete sound.
- There is a better way to generate complex sound. This is known as MIDI.
- Common MIDI devices includes electronic music synthesises, modules and MIDI devices in common sound cards.

MIDI Hardware:

An electronic musical Instrument or computer which has MIDI interface should has one or more MIDI ports.

- The MIDI ports or musical instruments are usually labelled with
  IN – for receiving MIDI data
  OUT – for outputting data that are generated by instruments
  THRU – for passing MIDI data to the next instrument

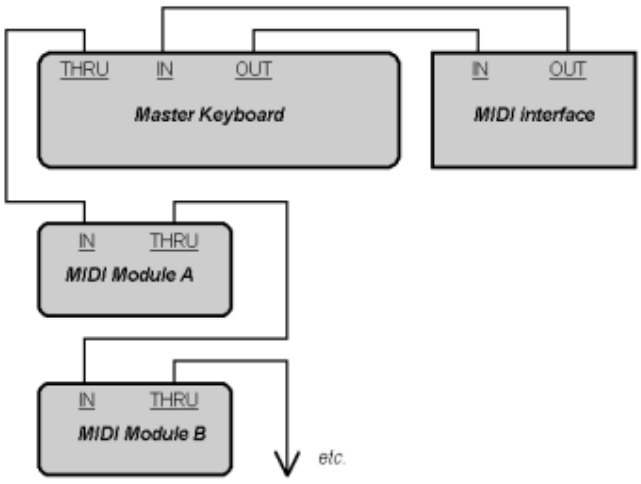


Figure: - Midi Hardware

MIDI Devices:

Through the MIDI interface, a computer can control output of individual instruments. On the other hand, the computer can receive, store or process coded musical data through the same interface. The data are generated with a keyboard and reproduced through sound generator. A sequencer can store data. Further, it may also modify the musical data. In a multimedia system, the sequencer is a computer application.

The heart of any MIDI system is the MIDI synthesizer device. A typical synthesizer looks like a simple piano keyboard.

Most synthesizers have the following components:

- Sound Generators
- Microprocessor
- Keyboard
- Control Panel
- Auxiliary controller
- Memory

MIDI Messages:

MIDI messages transmit information between MIDI devices and determine what kinds of musical events can be passed from device to device. The format of MIDI messages consists of the status byte, which describes the kind of message and data bytes. MIDI messages are divided into two groups.

(i) Channel Messages:

Channel messages go only to specific devices. There are two types of channel messages.

- Channel voice message send actual performance data between MIDI devices, describing keyboard action, controller action and control panel changes. Example, Note on, Note off, channel pressure, control changes etc.
- Channel mode message determine the way that a receiving MIDI device responds to channel voice messages. Example, Local control, All notes off, Omni mode off etc.

(ii) System Messages:

System messages go to all devices in a MIDI system because no channel numbers are specified. There are three types of system messages.

- System real time message are very short and simple, consisting of only one byte. They carry extra data with them. Example, System Reset, Timing clock (MIDI clock) etc.
- System common messages are commands that prepare sequencers and synthesizers to play a song. Examples, Song select, Tune Request etc.
- System exclusive messages allow MIDI manufactures to create customized MIDI messages to send between MIDI devices.

MIDI Devices:

Sound Generator:

- Sound generators do the actual work of synthesizing sound, the purpose of the rest of the synthesizes is to control the sound generators.
- The principal purpose of the generator is to produce an audio signal that becomes sound when fed into a loud speaker by varying the voltage oscillation of audio signal; a sound generator changes the quality of the sound.

Microprocessor:

- The microprocessor communicates with the keyboard to know what the musician is playing and with the control panel to know what note is musician playing and with the control panel to know what command the musician wants to send to the microprocessor.
- The microprocessor specifies note and sound, commands to the sound generators.

OR

Microprocessor sends and receives messages.

Keyboards:

The keyboard affords the musician's direct control of the synthesizer. Pressing keys on the keyboard signals the microprocessor what notes to play and how long to play them.

Control Panel:

The control panel controls those functions that are not directly connected with note and durations.

Auxiliary Controller:

Auxiliary controllers are available to give more control over the note played on the keyboard.

Memory:

Synthesizer memory is used to store patches for the sound generators and settings on the control panel.

Speech:

- Speech can 'perceived', understood and generated by humans and also by machine. A human adjusts himself/herself very efficiently to different speakers and their speech habits. The speech can understand by human.
- The human speech signal comprises a subjective lowest spectral component known as pitch, which is not proportional to frequency. The human ear is most sensitive in the range from 600 Hz to 6000 Hz.

Speech signal have two properties which can be used in speech processing.

(i) Voiced speech signals: show during certain time intervals almost periodic behaviour.

(ii) The spectrum of audio signals shows characteristic Maxine; which are mostly 3-5 frequency bands. These Maxine called formats, occur between of resonance of the vocal tract.

- A machine can also support speech generation and recognition with computers, one can synthetically generate speech, where the generated signals do not sound quite natural but can be easily understood.

Speech Generation:

- An important requirement for speech generation is real-time signal generation. With such requirement met, n speech output system could transform text into speech automatically without any lengthy pre-processing.
- Generated speech must be understandable and must sound natural.
- The requirement of under stable is a fundamental assumption and the natural sound of speech increases user acceptance.

Basic Notations:

- The lowest periodic spectral component of the speech signal is called the fundamental frequency. It is presented in a voiced sound.
- A phone is the smallest speech unit, such as the m of mat and b of bat in English.
- Allophones mark the variants of phone. For example, the aspirated p of pit and the unaspirated p of spit are allophones of the English phenome P.
- The morph marks the smallest unit which carries a meaning itself.
- A voiced sound is generated through the vocal cords; m, v and l are examples of voiced sound. The pronunciation of a voiced depends strongly on each speaker.
- During the generation of unvoiced sound, the vocal cords are opened f and s is unvoiced sound.
- Vowels – a speech created by the relatively free passage of breath through the larynx and oral charity. Example, a, e, I, o and u
- Consonants – a speech sound produced by a partial or complete obstruction of the air stream by any of the various contradictions of the speech organs. Example, m from mother, ch from chew.

Speech Analysis:

Speech analysis/input deals with the research areas which are as follows:

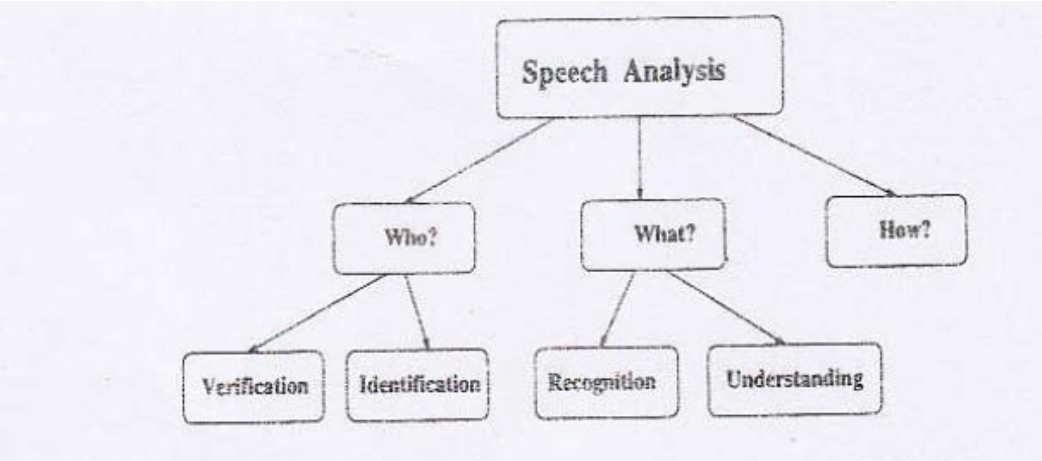


Figure 3.9: *Research areas of speech analysis.*

(1) Who?

- Human speech has certain characteristics determined by a speaker. Hence speech analysis can serve to analyze who is speaking i.e. to recognize a speaker for his/her identification and verification.

(2) What?

- Another main task of speech analysis is to analyze what has been said i.e. to recognize and understand the speech signal itself.

(3) How?

Another area of speech analysis tries to research speech patterns with respect to how a certain statement was said.

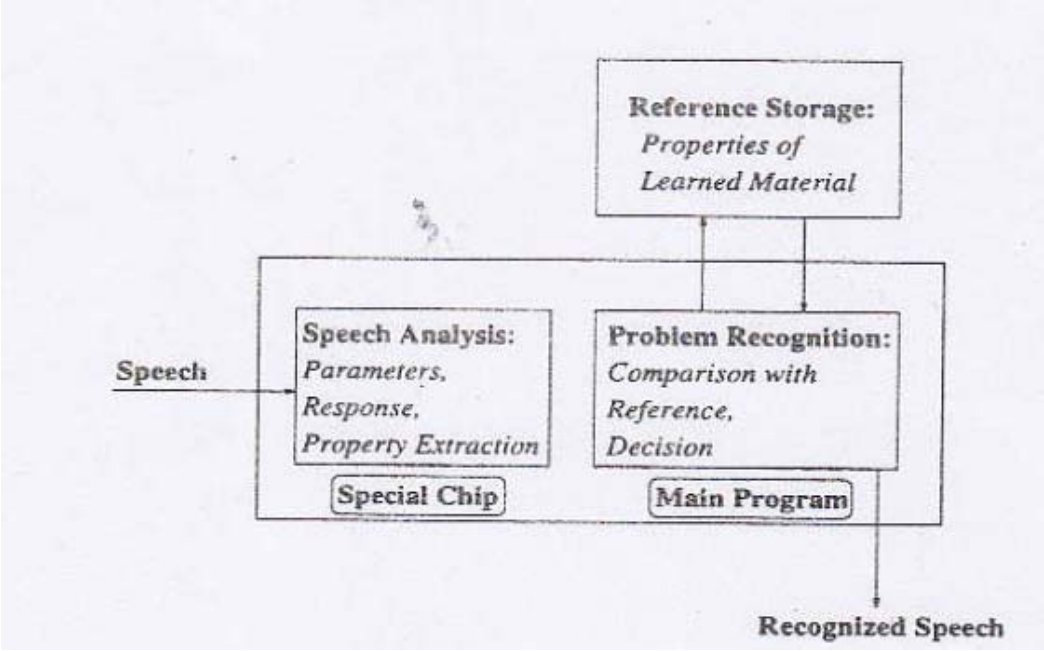
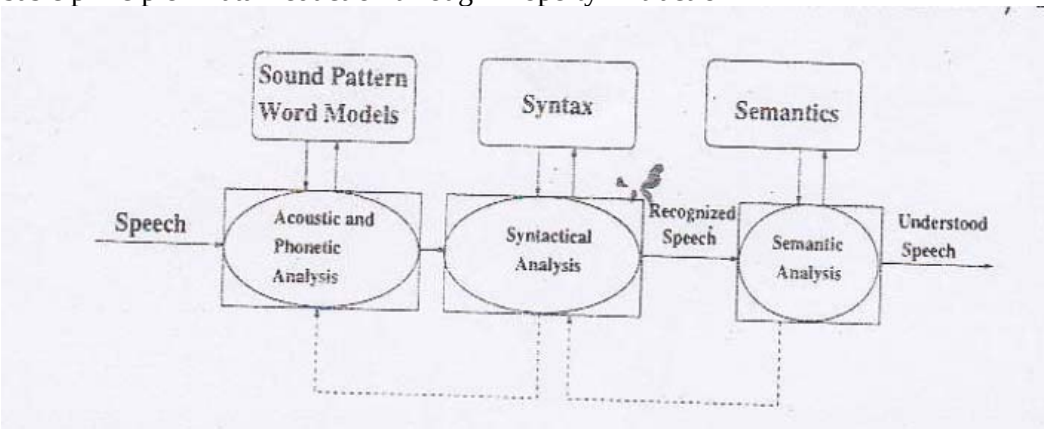


Figure: - Speech recognition system: task division into system components, using the basic principle "Data Reduction through Property Extraction"



Speech Transmission:

- The area of speech transmission deals with efficient coding of the speech signal allow speech / sound transmission at low transmission rates over networks.

- The goal is to provide the receiver with the same speech/sound quality as was generated at the sender side.

Some Techniques for Speech Transmission:
(1) Pulse Code Modulation:

A straight forward technique for digitizing an analog signal is pulse code modulation. It meets the right quality demand stereo audio signals in the data rate used for CD. Its rate is 176400 bytes/s.
(2) Source Encoding:

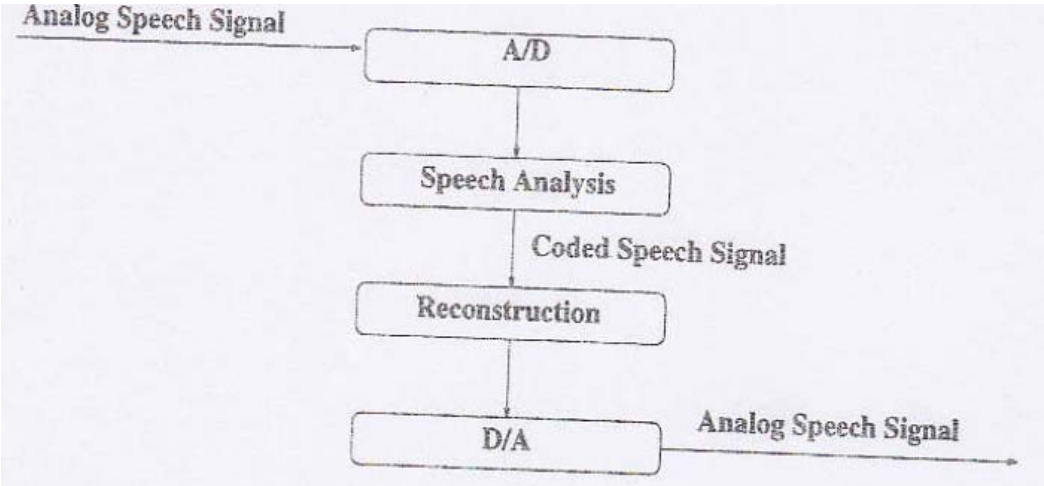


Figure: - Component of a speech transmission system using source encoding

In source encoding transmission depends on the original signal has certain characteristics that can be exploited in compression.
(3) Recognition-Synthesis Method:

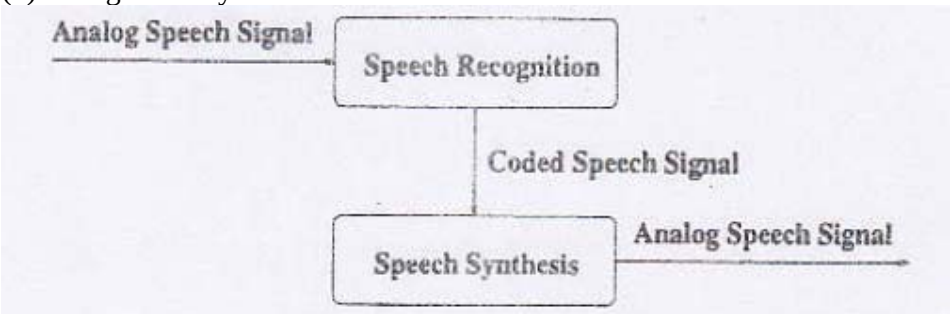


Figure: - Component of a recognition Synthesis for speech transmission

This method conducts a speech analysis and speech synthesis during reconstruction speech elements are characterized by bits and transmitted over multimedia system. The data rate defines the quality.