

## Wideband Speech Coding & Its Application

Apeksha B. landge.  
M.E. [student]  
Aditya Engineering College Beed

Prof. Amir Lodhi.  
Guide & HOD,  
Aditya Engineering College Beed

### ABSTRACT:

Increasing the bandwidth of sound signals from the telephone bandwidth of 200-3400 Hz to the wider bandwidth of 50-7000 Hz results in increased intelligibility and naturalness creates a feeling of transparent communication. Emerging end-to-end digital communication systems enable the use of wideband speech coding in numerous and diverse applications. In recognition of the need for high-quality wideband speech code several standardization activities have been conducted recently, resulting in the selection of a new wideband speech codec, AMR-WB, at bit rates from 6.6 to 23.85 k bit/s by both 3GPP and ITU-T.

The adoption of AMR-WB by the two bodies is of significant importance because for the first time the same codec will be adopted for wireless as well as wire line services. This will eliminate the need for transcoding and ease the implementation of wideband voice applications and services across a wide range of communication systems and equipment.

This paper presents a summary of wideband speech coding standards for wideband telephony applications. The quality advantages and applications of wideband speech coding are first presented, and then the issue of telephony over packet networks is discussed. Several wideband

speech coding standards are discussed, and special emphasis is given to the AMR-WB

standard recently selected by 3GPP and ITU-T.

*Keywords:* ITU-T, AMR-WB

### Introduction:

In general, speech coding is a procedure to represent a digitized speech signal using as few bits as possible, maintaining at the same time a reasonable level of speech quality. A not so popular name having the same meaning is speech compression. Speech coding has matured to the point where it now constitutes an important application area of signal processing. Due to the increasing demand for speech communication,

speech coding technology has received augmenting levels of interest from the research, standardization, and business communities.

Many signal processing problems including speech coding can be formulated as a well-specified computational problem; hence, a particular coding scheme can be defined as an algorithm. In general, an algorithm is specified with a set of instructions, providing the computational steps needed to perform a task. With these instructions, a computer or processor can execute them so as to complete the coding task. The instructions can also be translated to the structure of a digital circuit, carrying out the computation.

### Desirable Properties of a Speech Coder:

The main goal of speech coding is either to maximize the perceived quality at a particular bit-rate, or to minimize the bit-rate for a particular perceptual quality. The appropriate bit-rate at which speech should be transmitted or stored depends on the cost of transmission or storage, the cost of coding (compressing) the digital speech signal, and the speech quality requirements. In almost all speech coders, the

reconstructed signal differs from the original one. The bit-rate is reduced by representing the speech signal (or parameters of a speech production model) with reduced precision and by removing inherent redundancy from the signal, resulting therefore in a lossy coding scheme. Desirable properties of a speech coder include:

- High Speech Quality. The decoded speech should have a quality acceptable for the target application.
- Robustness Across Different Speakers / Languages. The underlying technique of the speech coder should be general enough to model different speakers (adult male, adult female, and children) and different languages adequately.
- Robustness in the Presence of Channel Errors. This is crucial for digital communication systems where channel errors will have a negative impact on speech quality.
- Good Performance on Nonspeech Signals (i.e., telephone signaling). In a typical telecommunication system.

- **Low Memory Size and Low Computational Complexity.** In order for the speech coder to be practicable, costs associated with its implementation must be low.
- **Low Coding Delay.** In the process of speech encoding and decoding, delay is inevitably introduced, which is the time shift between the input speech of the encoder with respect to the output speech of the decoder. An excessive delay creates problems with real-time two-way conversations, where the parties tend to “talk over” each other. Thorough discussion on coding delay is given next.
- **Low Bit-Rate.** The lower the bit-rate of the encoded bit-stream, the less bandwidth is required for transmission, leading to a more efficient system. This requirement is in constant conflict with other good properties.

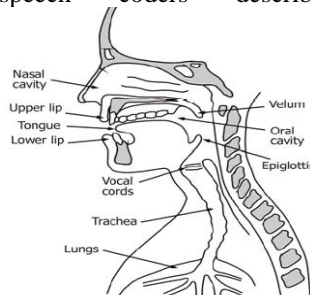
### Introduction of speech coding:

Speech coding is the process of obtaining a compact representation of voice signals for efficient transmission over band-limited wired and wireless channels and/or Today, speech coders have become essential component in telecommunications and in the multimedia infrastructure. Commercial systems that rely on efficient speech coding include cellular communication, voice over internet protocol (VOIP), videoconferencing, electronics toys, archiving, and digital simultaneous voice and data(DSVD), as well as numerous PC-based games and multimedia applications

Speech coding is the art of creating a minimum redundant representation of the speech signal that can be efficiently transmitted or stored in digital media, and decoding the signal with the best possible perceptual quality.

### Speech Production and Modeling:

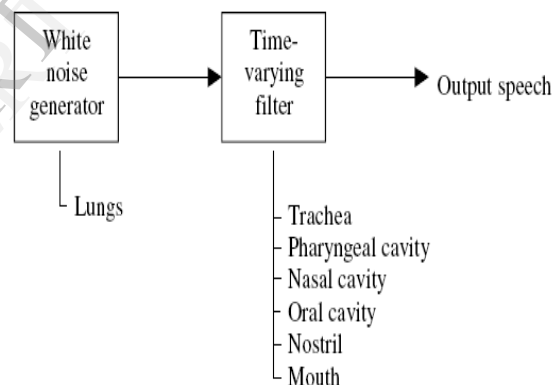
In this section, the origin and types of speech signals are explained, followed by the modeling of the speech production mechanism. Principles of parametric speech coding illustrated using a simple example, with the general structure of speech coders described at the end.



**Figure: Diagram of human speech production system**

A simplified structural view is shown in Figure Speech is basically generated as an acoustic wave that is radiated from the nostrils and the mouth when air is expelled from the lungs with the resulting flow of air perturbed by the constrictions inside the body. It is useful to interpret speech production in terms of acoustic filtering. The three main cavities of the speech production system are nasal, oral, and pharyngeal forming the main acoustic filter. The filter is excited by the air from the lungs and is loaded at its main output by a radiation impedance associated with the lips. The vocal tract refers to the pharyngeal and oral cavities grouped together. The nasal tract begins at the velum and ends at the nostrils of the nose. When the velum is lowered, the nasal tract is acoustically coupled to vocal tract to produce the nasal sounds of speech.

The human speech production system can be modeled using a rather simple structure: the lungs—generating the air or energy to excite the vocal tract—are represented by a white noise source. The acoustic path inside the body with all its components is associated with a time-varying filter. The concept is illustrated in Figure4 This simple model is indeed the core structure of many speech coding algorithms. By using a system identification



**Figure: Simple model of speech coding algorithm.**

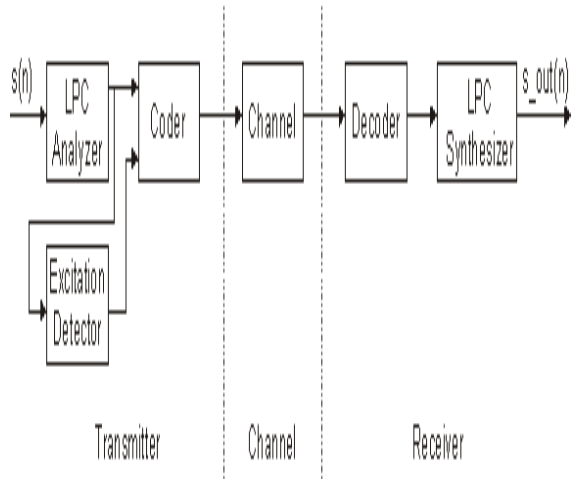
### Speech Recognition Principle:

Speech recognition is performed by identifying a sound based on its frequency content. In order to achieve this, the frequency content of several samples of the same sound must be averaged in a training phase (i.e. the sound's “reference fingerprint” must be generated). Then, the frequency content of a sound input can be compared to the a fore mentioned fingerprint by treating them as vectors and computing the distance between them. If a sound is close enough to the reference, then it is considered to be a match. A MATLAB implementation of this process was created in order to better illustrate it, and experiment with the settings.

### Voice-excited LPC Vocoder :

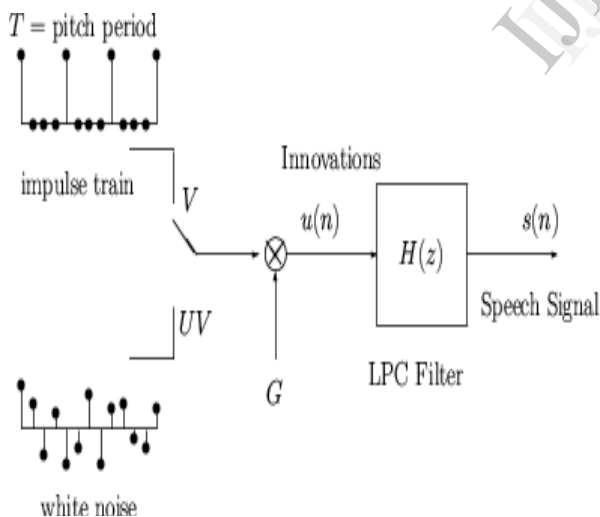
As the test of the sound quality of a plain LPC-10 vocoder showed, the weakest part in this

methodology is the voice excitation. It is known from the report that one solution to improve the quality of the sound is the use of voice-excited LPC vocoders. Systems of this type have been studied by Atalet and Weinstein. Figure shows a block diagram of a voice-excited LPC vocoder. The main difference to a plain LPC-10 vocoder, as shown in Figure is the excitation detector, which will be explained in the sequel.



**Figure :** Block diagram of a voice-excited LPC vocoder.

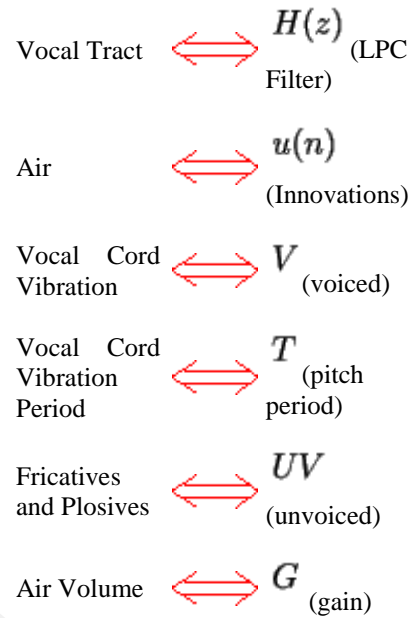
**Mathematical model of LPC Analysis**



**Figure: Mathematical model of LPC Analysis.**  
The above model is often called the LPC Model.

- The model says that the digital speech signal is the output of a digital filter (called the LPC filter) whose input is either a train of impulses or a white noise sequence.

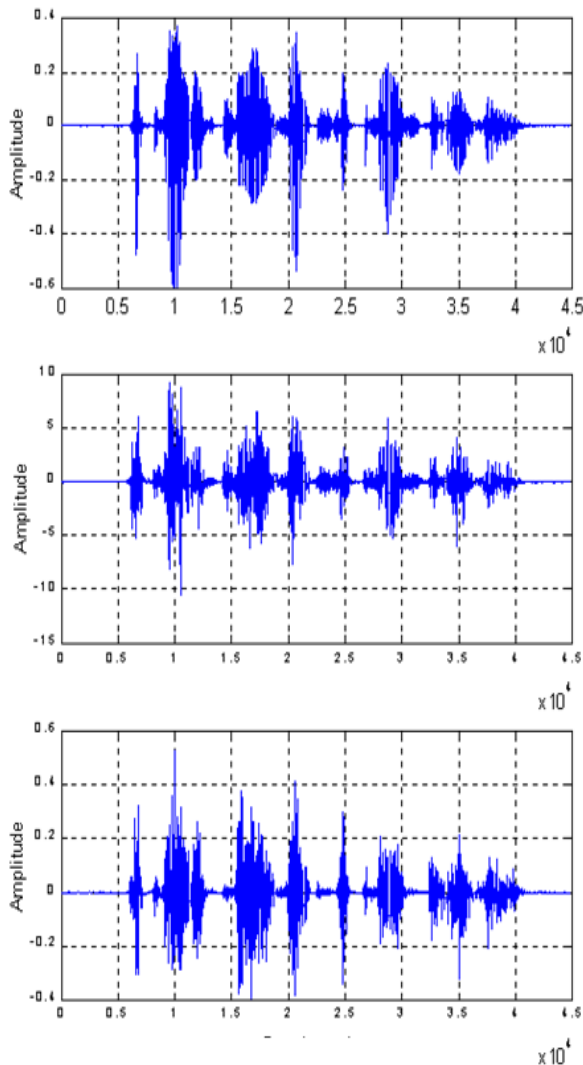
The *relationship* between the physical and the mathematical models:



Which is equivalent to saying that the input-output relationship of the filter is given by the linear difference equation.

**Experimental result:**

The LPC method to transmit speech sounds has



Islamic Azad University South Tehran Branch  
Tehran, Iran world Applied Sciences Journal 59-66,  
2010, ISSN 1818 – 4952

[2] An HSBE LPC Low bit wideband speech coding  
algorithm IET International conference by ying Na,  
zhao xiao-hui Dong 0534, 9989 07 august 2009

[3] T Lalith et al speech recognition using neural  
network IEEE international conference on signal  
processing system 2009, pp-248-252

[4] Richard v.cox. “speech coding” AT&T  
labs(research lab) 2000 CRC press [http://www  
engnetbase.COM](http://www.engnetbase.COM)>

some very good aspects, as well as some drawbacks.  
The huge advantage of vocoders is a very low bit rate  
compared to what is achieved for sound transmission.  
On the other hand, the speech quality achieved is  
quite poor.

*Waveform of the sentence "A pot of tea helps to  
pass the evening":*

#### CONCLUSION:

The results achieved from the voice excited  
LPC are intelligible. On the other hand, the plain LPC  
results are much poorer and barely intelligible. This  
first implementation gives an idea on how a vocoder  
works, but the result is far below what can be  
achieved using other techniques. Nonetheless the  
voice-excited LPC used gives understandable results  
and is not optimized. The tradeoffs between quality  
on one side and bandwidth and complexity on the  
other side clearly appear here. If we want a better  
quality, the

#### REFERANCES:

[1] Hybrid NQ and Neural models for ISF  
Quantization in wide band speech Mansor Sheikhan  
& Sahar company Departmental of Electrical Engg.