National Conference on Innovative Paradigms in Engineering & Technology (NCIPET-2012)
Proceedings published by International Journal of Computer Applications® (IJCA)

An analytical approach for wideband speech coding with the help of filter bank

Ujjwal v. Ramekar
Lecturer
Department of electronics and telecommunication
Sipna's college of engineering and technology,Amravati-444601

ABSTRACT
Interpreting the need of high quality wideband speech coding some standardization events have been conducted. There are various methods and standard for speech coding. these paper present a new technique of speech coding. here interpretation is done with the help of gamma tone filter bank. The o/p of the system is analyzed and process to get the respective train of pulses. the positioning and amplitude and gain of pulses are coded using the process of quantization.

KEYWORDS
Wideband speech signal gamma tone filter bank, speech coding

1. INTRODUCTION
In today's era of technology there is no coding technique which is capable of providing efficient coding speech coding the art of creating the minimally redundant representation of speech signal that can be efficiently transmitted or stored in digital media. & decoded the signal with the best possible perceptual quality in these work analysis of speech is done with the help of gamma tone filter bank. here by applying the concept of filter bank we have done the coding. Speech coding is necessary for cellular phone which has limited data rates for each user it is necessary for voice IP, audio, visualtele conferencing to reduced the bandwidth consumption The organization of paper is as follows section 2 includes the concept and techniques utilized in the proposed work. Section 3 includes parameter of speech. Section 4 includes the synthesis of speech. Section 5 conclude the paper.

2.CONCEPTANDTECHNIQUSUSEDINTHEPRPOSEDWORK

2.1 Speech coding
Speech coding (or digital speech coding) is the process by which a speech signal can be temporally compressed into less bits/second and then decompressed, while preserving its most important content.

• The main objectives of digital speech coding are to lower the bit rate used to represent a speech signal while maintaining an adequate level of perceptual fidelity. In addition, for some applications we need to consider the complexity (computation required to encode/decode) (e.g. telephone line or cell network) and for storage (e.g. MP3)

Basic Components in a Source Coding System

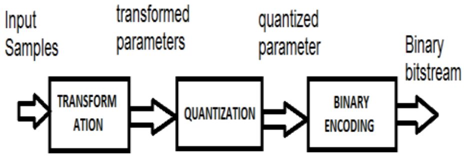

Fig 1:basic coding system

2.1.1Classification of coding algorithms
1. Based on how the speech production and perception models are incorporated

• Open-loop codec (also called vocoders): Extract an excitation signal + a vocal tract system to later re-synthesize the speech, without any attempt to preserve the waveform shape. These more or less correspond to the Model-based codec.

• Closed-loop codec: Apply the source/system model of speech within a feedback loop to iteratively reduce the difference between the original and quantized signals. Based on how the waveform of the speech signal is preserved with the process Based on the nature of the encoded parameters

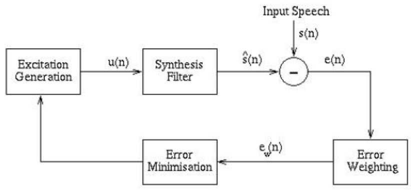

Fig 2:close loop algorithm

2.2Gammatone filter bank
Gamma tone filter can be implemented using the FIR or IIR filter or frequency domain technique here fir can be employed to implement linear phase filter. It has length of 2N-1& obtained by convolving sampled gamma tone impulse response g(n) Of length n=100

$$g(n) = a(nT)^{N-1} e^{-2\pi bERB(f_c)nT} \cos(2\pi f_c nT + \varphi)$$

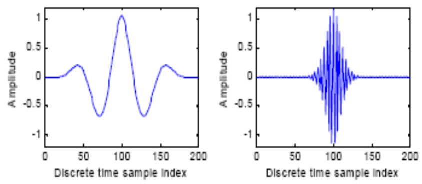

Fig3:Impulse response

3. PARAMETER OF SPEECH

The interpretation of I/P signal is done with the help of linear phase critical phase band filter, peak location ,simultaneous and temporal masking followed by quantization & coding and decoding ,renormalizations

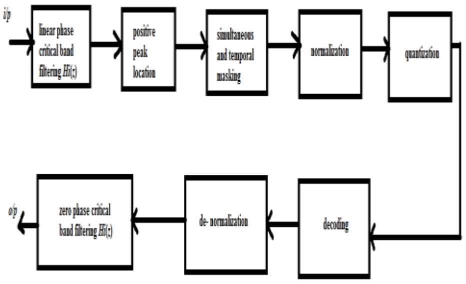

Fig4: of speech and audio signal parameter

3.1Filter bank

Filter bank is group of group of signal onto which certain processing is done with the help of certain processing is done with the help of certain techniques that decompose signal into various frequency sub-bands. This decomposition is necessary because sub-band processing has many advantages over processing of domain

It has two phase

I)architecture exploratic

II)logic design

3.1.1Summary of filter bank

n many application arena the decomposition and subsequent processing in frequency domain benefits in various way of performance

1)for wireless communication, accurate channel selection

2)in adaptive equalization fast convergence & lower complexity

3)multi resolution image compress & transform

Following fig shows ,filter bank is the collection of various filter with some common i/p & o/p

It consist of m-analysis filter Hk(I) which separate the i/p signal x(n)

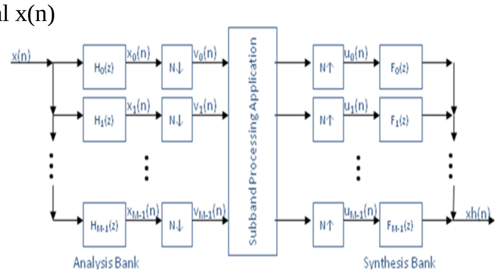

Fig 5:generalized block diagram of filter bank

The ultimate aim of choosing various filter bank parameter is to minimized the error. Below figure shows the frequency response of filter bank and performance of filter bank

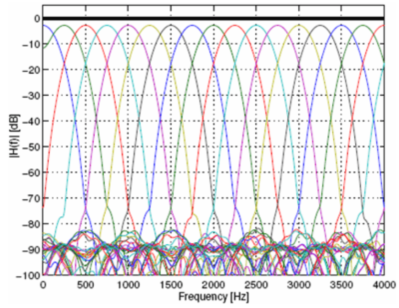

Fig 6: Gamma tone filter bank

Gamma tone filter bank is the group of collection of no of filter Through which we can analyze the speech it can be operated in various so main like frequency as well as time Gamma tone filter bank in time domain It is used to describe shape of the impulse response function of the auditory system is estimated by the reverse correlation function of neural firing times

The gamma tone filter is in time domain as

gt(t) oc tn 1 exp(-27rbt) cos(27r Jot + ¢»

Gamma tone filter bank in frequency domain

If *GT(J)* represents the Gamma Tone filter in the frequency domain (frequency response function) then

GT(J) ex [1 + *j(J - fo)/bj-n* + [1 + *j(J + fo)/brn* (-00 < *f* < 00)

3.2Masking of speech

Masking is the effect when the perception of one sound is on the anther sound is done .the masking is the phenomenon which is used to find out auditory systems ability to separate the component of complex sound the ultimate aim of Appling masking is to produce a more accurate perceptual parameter of firing pulses trail experiment gives that simultaneous masking removes an average of around 10%of pulse while application of temporary post making removes an average of 55%of pulses

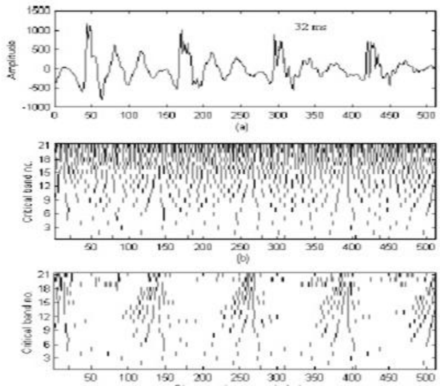

Fig 7:discrete time index

4. SPEECH SYNTHESIS

Speech synthesis is the artificial production of speech of human. Pc is used for the purpose is called as speech synthesizer And it can be implemented with the help of hardware synthesizing techniques The most important parameter for quality of speech synthesis are naturalness and intelligibility. Naturalness describe how closely the o/p sound. The two primary & important techniques for the synthesis of sound are concatenative & format synthesis

1) concatenative synthesis

It is based on the stringing together of segment of various previously recorded speech. This synthesis produced the most natural sound

2) format speech

The format synthesis don't take human speech sample into consideration. Here they used artificial speech & parameter such as fundamental frequency voicing & noise level are varied over time to creat a waveform of artificial speech. This method is also known as rules-based synthesis The synthesis can be done with the help of filter bank which gain simplicity & less complex city The systematic block diagram of speech synthesis with the help of gamma tone filter bank is as follow

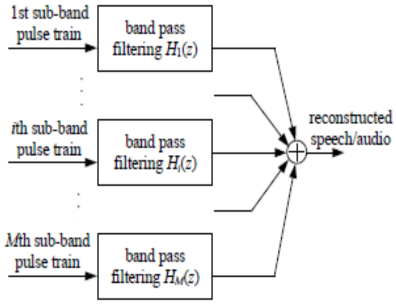

Fig 8:synthesisof speech with the help of filter bank

It consist of three band pass filter to each band pass filter a train of pulse is provided. The o/p of these filter is given to next stage to reconstruct speech

5. CONCLUSION

Speech coding is necessary in today's technical era. Here an attempt is made to design speech coding with the help of gamma tone filter bank. new speech paradigm has been designed which employs a filter bank to derive the accurate compact time and frequency parameter.

6. ACKNOWLEDGEMENT

I take this opportunity to express my sincere gratitude to Prof. Dr.S.A.LADHAKE principal Sipna's college of engineering and technology, who have contributed in making the work a great success. i express my gratitude to my H.O.D.Prof.A.P.THAKARE & Prof.Dr.A.A.GURJAR, whose constant help & encouragement guided me to complete my work.

7. REFERENCES

[1] Kubin, G., and Kleijn, W. B., "On speech coding in a perceptual domain", in Proc. IEEE ICASSP, 1999, pp. 205208.

[2]Rubin, F., "Arithmetic stream coding using fixed precision registers", IEEE Trans Inf. Theory, vol. IT-25, no. 6, 1979, pp. 672-675.

[3] Zwicker, E., and Terhardt, E. "Analytical expressions for critical-band rate and critical bandwidth as a function of frequency", J. Acoust. Soc. Am., vol. 68, no. 2, November1980, pp. 1523-1525.

[4]Zwicker, E., and Zwicker, U. T., "Audio engineering and psychoacoustics: matching signals to the final receiver, the human auditory system", J. Audio Eng. Soc., vol. 39, no. 3, March 1991, pp. 115-126.

[5]Ambikairajah, E., Davis, A. G., and Wong, W. T. K., "Auditory masking and MPEG-1 audio compression", Electr. & Commun.

[6]Brandenburg, K. B., and Stoll, G., "ISO-MPEG-1 audio: A generic standard for coding of high-quality digital audio", J. Audio Eng. Soc., vol. 42, no. 10, Oct. 1994, pp. 780-792.

[7] Krasner, M. A., "The critical band coder – digital encoding of speech signals based on the perceptual requirements of the auditory system", in Proc. IEEE ICASSP, 1980, pp. 327330

[8]P. 1. M. Johannesma, (1972) "The pre-response stimulus ensemble of neurons in the cochlear nucleus," Symposium on Hearing Theory, 22-23 June, 1972, IPO, Eindhoven

[9] Filterbank Reconstruction of Bandlimited Signals from Nonuniform and Generalized Samples Yonina C. Eldar, Student Member, IEEE and Alan V. Oppenheim, Fellow, IEEE

[10] K.Yao and J. B. Thomas, "On some stability and interpolatory properties of nonuniform sampling expansions," IEEE Trans. Circuit Theory, vol. CT-14, pp. 404–408, Dec. 1967.

[11] "Speech Coding Overview" Jason Woodard,http://wwwmobile.ecs.soton.ac.uk/speech_codecs

[12] SPEECH CODING: FUNDAMENTALS AND APPLICATIONS MARK HASEGAWA-JOHNSON University of Illinois at Urbana–Champaign Urbana, Illinois ABEER ALWAN University of California at Los Angeles Los Angeles, California

[13]Sub band Speech analysis using Gammatone Filter banks and optimal pitch extraction methods for each sub band using average magnitude difference function (AMDF) for LPC Speech Coders in Noisy Environments Suma S.A.; Dr. K.S.Gurumurthy University Visvesvaraya College of Engineering, K.R. Circle Bangalore 01, India

[14] R. G. Amado and J. V. Filho, Pitch detection algorithms based on zero-cross rate and autocorrelation function for musical notes,Proc. Int. Conf. on Audio, Language and Image Processing, (Shanghai, China, 2008) pp. 449-454.

[15] K. Järvinen, J. Vainio, P. Kapanen, T. Honkanen, P. Haavisto, R. Salami, C.Laflamme, J.-P.Adoul, "GSM Enhanced Full Rate Codec," IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Munich, Germany, 20–24 April 1997, pp. 771–774.

[16]J.D.Gordy and R.A.Goubran, A combined LPC-based speech coder and filtered-X LMS algorithm for acoustic echo cancellation," in Proc. IEEE ICASSP, vol.4,pp.125-128, May

[17]Oxenham, A.J. Plack, C.J. Suppression and the upward spread of masking, Journal of the Acoustical Society of America, 104 (6) pp.3500–3510

[18]R. G. Willey, "Recovery of bandlimited signals from unequally spaced

samples," *IEEE Trans. Commun.*, vol. COMM-26, pp. 135–137, Jan.1978.