International Journal Of Engineering Technology Science And Research
IJETSR
Www.Ijetsr.Com
ISSN 2394 – 3386
Volume 5, Issue 1
January 2018

Design of Secure Speech Coders for Military Communications

Akella Amarendra Babu, Aadem Kavitha, V Sudha Rani, E. Soumya, B. Saratha
St. Martin's Engineering College, Secunderabad, India

ABSTRACT

Military communications are provided in battlefield under the harsh acoustic noise environment such as in tanks, helicopters, artillery shelling, speeding low flying aircrafts and consequent disruption of communication links. Typically, there are three technical requirements for military communications - information security, real time systems and variable data rates. Information security requires primarily confidentiality and authentication. Real time systems ensure that there is no unacceptable time delay while processing the speech data for duplex communication. Military communications are network centric. During military operations, the available data rates on the communication channels will vary due to destruction of communication links. Therefore, communications need to be provided at varying available data rates.

In this paper, we incorporate the above three technical requirements of military communications in the algorithms at the speech coding stage. It eliminates the need to have additional equipment in the voice communication systems and keeps the overall size of the communication systems compact and ensures mobility.

Keywords

Mixed excitation Linear Prediction, Code excited Linear Prediction, Noise pre-processing, Adaptive multi rate coder, Variable bit rate coder.

INTRODUCTION

Confidentiality of information and authentication of the source of information are the two important security services that encompass the information security facility. Confidentiality ensures that the communication system is accessible only for authorized parties. Authentication ensures that the speaker is correctly identified, with an assurance that the identity is not false.

The MELPe or enhanced-MELP (Mixed Excitation Linear Prediction) is a United States Department of Defense speech coding standard used mainly in military applications and satellite communications, secure voice, and secure radio devices. LPC coder uses a fully parametric model and produces intelligible speech at 2.4 kbps. However, it generates annoying artifacts such as buzzes, thumps and tonal noises. MELP utilizes additional parameters to capture the underlying signal dynamics. In 2002, the US DoD adopted MELPe as NATO standard, known as STANAG-4591, enabling the same quality as the old 2400 bit/s MELP at half the rate. MELP speech coding algorithm is described in the subsequent paragraphs in this paper.

Authentication provides protection against active attacks like masquerade. A masquerade takes place when one person pretends to be a different person. Although human speech can be recognized by the recipient, it is foolproof only when speech is synthesized using high quality speech coders, typically coders like ITU-T G.711 PCM, ITU-T G.726 ADPCM, ETSI GSM 6.10 RPE-LTP etc. Although vocoders like MELP, CELP etc., attempts to retain the naturalness of the synthetic speech, human auditory system cannot authenticate the speaker when subjected to pressures of battlefield conditions. Therefore, authentication is a security pre-requisite for military communications. Speaker identification algorithms or secure hash functions may be used for authentication.

Today, military communications are **Network Centric**, which use a wide variety of networks and communication bearers. Network Centric Defence communications need to be secure, interoperable and bearer-independent. Bearer-independent communications will operate over any type of communication network e.g. fixed-lines or mobile, commercial or military, terrestrial or satellite.

In Network Centric Communication scenario, links may go down due to battlefield conditions. Overall available bit rate will change continuously. Speech communications are connection oriented virtual circuits. The bit rate allotted to a particular speech circuit depends on the overall availability of the links.

Adaptive Multi Rate Codec (AMR) uses link adaptation to select the bit rates based on link conditions and selects one of the various coding techniques, such as MELP, MELPe, CELP, ACELP etc.,. The usage of AMR requires optimized link adaptation that selects the best codec mode to meet the local radio channel and capacity requirements. If the radio conditions are bad, source coding is reduced and channel coding is increased. This improves the quality and robustness of the network connection while sacrificing some voice clarity. In the particular case of AMR this improvement is somewhere around S/N = 4-6 dB for usable communication. The new intelligent system allows the network operator to prioritize capacity or quality per base station.

A speaker is silent roughly 63% time in a two way conversation. In a conventional speech coder, the same number of bits is allocated to all speech frames. Furthermore, due to dynamic nature of the speech signal, the number of bits required to represent the frames faithfully varies with time. By changing the bit rate as a function of the signal properties, it is possible to yield an average bit-rate that is substantially less than the fixed bit rate of the conventional coder. This is the principle of the source controlled variable rate, where the coding algorithm responds to the time varying local character of the speech signal to determine the bit rate. **FS 1015 LPC coder and the FS MELP** pertain to this family. For these coders, parameters of the speech production model are encoded using different number of bits, depending on whether the frame is voiced or unvoiced. **TIA IS96 variable bit rate (VBR) CELP** is another coder in which the control mechanism is based on the background noise estimate and the energy of the signal.

Mixed Excitation Linear Prediction

A block diagram of the **MELP** model of speech production is shown in Figure 1, which is an attempt to improve upon the LPC model. MELP decoder utilizes a sophisticated interpolation technique to smooth out inter frame transitions. A randomly generated period jitter is used to perturb the value of the pitch period so as to generate an aperiodic impulse train. The MELP coder extends the number of classes into three: unvoiced, voiced, and jittery voiced. The latter state corresponds to the case when the excitation is aperiodic but not completely random, which is often encountered in voicing transitions. This jittery voiced state is controlled in the MELP model by the pitch jitter parameter and is essentially a random number. A period jitter uniformly distributed up to +/- 25% of the pitch period produced good results. The short isolated tones, often encountered in LPC coded speech due to misclassification of voicing state, are reduced to a minimum.

Shape of the excitation pulse for periodic excitation is extracted from the input speech signal and transmitted as information on the frame.

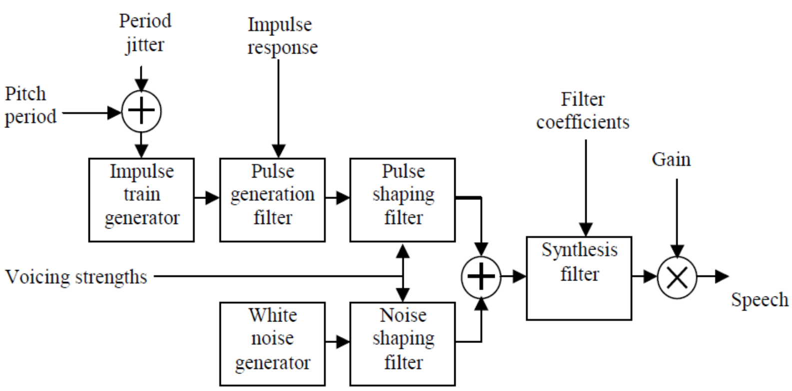

Figure 1 The MELP model of speech production

The shape of the pulse contains important information and is captured by the MELP coder through Fourier magnitudes of the prediction error. These quantities are used to generate the impulse response of the pulse generation filter (Figure 1), responsible for the synthesis of periodic excitation.

Periodic excitation and noise excitation are first filtered using the pulse shaping filter and noise shaping filter, respectively; with the filters' outputs added together to form the total excitation, known as the mixed excitation, since portions of the noise and pulse train are mixed together.

In Figure 1, the frequency responses of the shaping filters are controlled by a set of parameters called voicing strengths, which measure the amount of ''voicedness.'' The responses of these filters are variable with time, with their parameters estimated from the input speech signal, and transmitted as information on the frame.

Shaping Filters

The MELP speech production model makes use of two shaping filters (Figure 1) to combine pulse excitation with noise excitation so as to form the mixed excitation signal. Responses of these filters are controlled by a set of parameters called voicing strengths; these parameters are estimated from the input signal. By varying the voicing strengths with time, a pair of time-varying filters results. These filters decide the amount of pulse and the amount of noise in the excitation, at various frequency bands.

In FS MELP, each shaping filter is composed of five filters, called the synthesis filters, since they are used to synthesize the mixed excitation signal during decoding. Each synthesis filter controls one particular frequency band, with pass bands defined by 0–500, 500–1000, 1000–2000, 2000–3000, and 3000–4000 Hz. The synthesis filters connected in parallel define the frequency responses of the shaping filters. Figure 2 shows the block diagram of the pulse shaping filter, exhibiting the mechanism by which the frequency response is controlled. VS 1 to 5 are the voiced strengths.

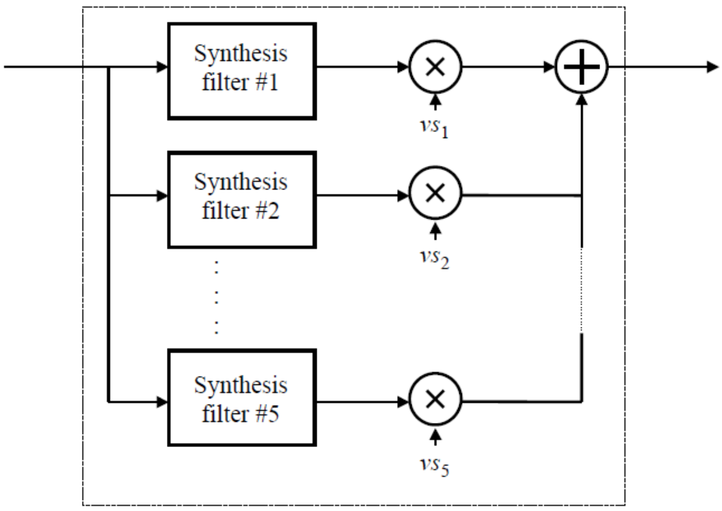

Figure 2: Block diagram of the pulse shaping filter

Thus, the two filters complement each other in the sense that if the gain of one filter is high, then the gain of the other is proportionately lower, with the total gain of the two filters remaining constant at all times.

1.2Kbps / 2.4 Kbps MELP Speech Coders

The 2.4Kbps MELP algorithm divides the 8Kbps sampled speech signal into 22.5ms frames for analysis, whereas The 1.2Kbps MELP algorithm divides the 8Kbps sampled speech signal into groups of three 22.5ms frames into a 67.5ms super frame for analysis. Depending upon the type of speech present in the signal, inter-frame redundancy can be exploited to efficiently quantize the parameters.

Communication System

Figure 3 and Figure 4 give block diagram of Encoder and Decoder incorporating the features described above.

Figure 3 Block diagram of encoder

Figure 4 Block diagram of Decoder

Authenticator: Speaker identification algorithms or secure hash functions may be used for the authentication.

VBR / AMR Selector: Adaptive Multi Rate Codec selects a suitable coding algorithm depending upon the link conditions and Variable Bit Rate selector selects appropriate bit rate based on the background noise estimate and the energy of the signal.

Encoder / Decoder: MELPe or CELP coding algorithm takes original speech signal and produce the coded speech as the output. Noise pre-processing algorithm may be combined with the speech coding algorithm to eliminate the effect of harsh acoustic noise environment. The Decoder takes coded speech and produces synthetic speech.

Comparson of Military Vs Civil Vocoders

Figure 5 compares the performance of the NATO Stanag 4591 voice coder with that of comparable civil voice coders. Despite lower requirements for throughput, the NATO coder provides better speech intelligibility. As a result, this voice coder is being considered for adoption in civil standards. Such civil standards will then be interoperable with military, and development of compliant equipment benefits from economies of scale.

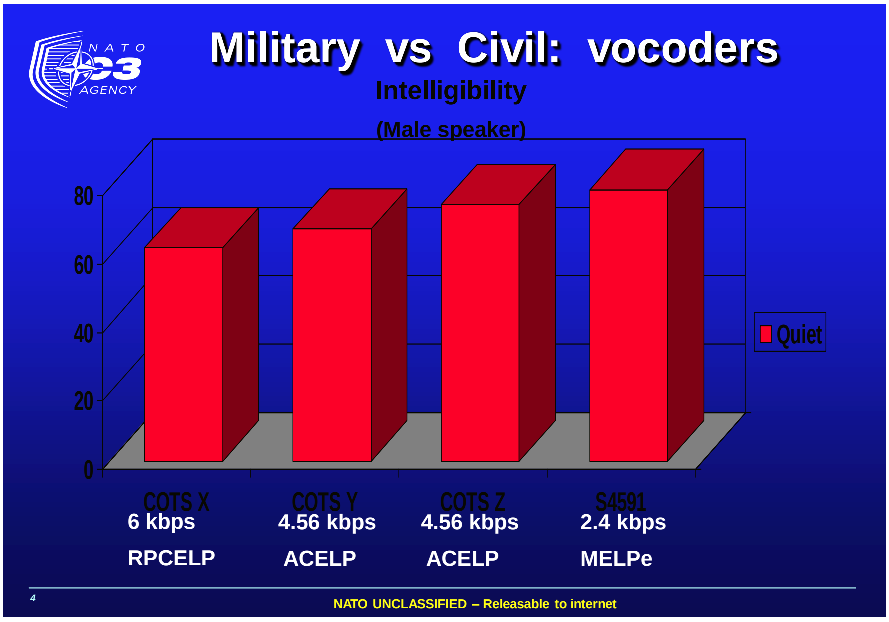

Figure 5: Comparison of performance of civil coders with military vocoders.

Conclusions

In this paper, we have presented Encoder and Decoder block diagrams for the Military Communication System incorporating the information security features and produce intelligible and clear speech communications, even under harsh acoustic noise environment.

References

[1] Wai C. Chu, (2003), Speech Coding Algorithms, Wiley Interscience.
[2] John S. Collura, Diane F. Brandt, Douglas J. Rahikka, National Security Agency, 9800 Savage Rd, STE 6516, Ft. Meade, MD 20755-6516, USA.
[3] Martin Rainer, Cox Richard V., "New Speech Enhancement Techniques for Low Bit Rate Speech Coding." IEEE Speech Coding Workshop-99, Porvoo Finland.
[4] Malah David, Cox Richard V., Accardi Anthony J., "Tracking Speech Presence Uncertainty to Improve Speech Enhancement in Non-Stationary Noise Environments." IEEE ICASSP-99 Conference, Phoenix AZ.
[5] Dr Michael Street , Standards for Network Centric Defence , CIS Division, NATO C3 Agency