A 600 BPS MELP VOCODER FOR USE ON HF CHANNELS

Mark W. Chamberlain
Harris Corporation, RF Communications Division
1680 University Avenue
Rochester, New York 14610

ABSTRACT

The U.S. government has developed and adopted a new Military Standard vocoder (MIL-STD-3005) algorithm called Mixed Excitation Linear Prediction (MELP) which operates at 2.4Kbps. The vocoder has good voice quality under benign error channels. However, when the vocoder is subjected to a HF channel with typical power output of a ManPack Radio (MPR), the vocoder speech quality is severely degraded. Harris has found that a 600 bps vocoder provides significant increase in secure voice availability relative to the 2.4Kbps vocoder. This paper describes a 600 bps MELP vocoder algorithm that takes advantage of the inherent inter-frame redundancy of the MELP parameters. Data is presented showing the advantage in both Diagnostic Acceptability Measure (DAM) and Diagnostic Rhyme Test (DRT) with respect to SNR on a typical HF Channel when using the vocoder with a MIL-STD-188-110B [1] waveform.

INTRODUCTION

A need exists for a low rate speech vocoder with the same or better speech quality and intelligibility as the current performance of 2.4Kbps Linear Predictive Coding (LPC10e) based systems. A MELP speech vocoder at 600 bps would take advantage of robust lower bit-rate waveforms than the current 2.4Kbps LPC10e standard and benefit from better speech quality of the MELP vocoder parametric model. Tactical ManPack Radios (MPR) require lower bit-rate waveforms to ensure 24-hour connectivity using digital voice. Once HF users receive reliable good quality digital voice, wide acceptance will provide for better security by all users. HF user will also benefit from the inherent digital squelch of digital voice and the elimination of atmospheric noise in the receive audio.

The LPC10e vocoder has been widely used as part of NATO's and the US DoD's encrypted voice systems in use on HF channels. The 2.4Kbps system allows for communication on narrow-band HF channels with only limited success. The typical 3 kHz channel requires a relatively high SNR to allow reliable secure communications at the standard 2.4Kbps bit rate. The use of MIL-STD-188-110B waveforms at 2400bps would still require a 3 kHz SNR of more than +12 dB to provide a usable communication link over a typical fading channel.

When HF channels do allow a 2400 bps channel to be relatively error free, the voice quality of LPC10e is still marginal. Speech intelligibility and acceptability of LPC10e is limited to the amount of background noise level at the microphone. The intelligibility is further degraded by the low-end frequency response of the military H-250 handset. The MELP speech model has an integrated noise pre-processor as described in [2] that improves the vocoder's sensitivity to both background noise and low-end frequency roll-off. The 600 bps MELP vocoder would benefit from the noise pre-processor and the improved low-end frequency insensitivity of the MELP model.

The proposed 600 bps system discussed in this paper consists of a conventional MELP vocoder front end, a block buffer for accumulating multiple frames of MELP parameters, and individual block Vector Quantizers for MELP parameters. The low-rate implementation of MELP uses a 25 ms frame length and the block buffer of four frames, for block duration of 100ms. The MELP parameters are coded as shown in Table 1. This yields a total of sixty bits per block of duration 100 ms, or 600 bits per second.

SPEECH PARAMETERS	BITS
Aperiodic Flag	0
Band-Pass Voicing	4
Energy	11
Fourier Magnitudes	0
Pitch	7
Spectrum	(10+10+9+9)

Table 1 - MELP 600 VOCODER

Details of the individual parameter coding methods are covered below, followed by a comparison of bit-error performance of a Vector Quantized 600 bps LPC10e based vocoder contrasted against the proposed MELP 600 bps vocoder. We will discuss Diagnostic Rhyme Test (DRT) and the Diagnostic Acceptability Measure (DAM) results for MELP 2400 and 600 for several different conditions,

0-7803-7225-5/01/$17.00 (c) 2001 IEEE

and compare them with the results for LPC10e based systems under similar conditions. DRT and DAM results represent testing perform by Harris and the National Security Agency (NSA). Harris performed tests shall be identified by a superscript value 1 and NSA data shall be identified by a superscript value 2.

LPC SPEECH MODEL

LPC10e has become popular because it preserves nearly all of the intelligibility information, and because the parameters can be closely related to human speech production of the vocal tract. LPC10e as defined in [3] represents the speech spectrum in the time domain rather than in the frequency domain. The LPC10e Analysis process (transmit side) produces predictor coefficients that model the human vocal tract filter as a linear combination of the previous speech samples. These predictor coefficients are transformed into reflection coefficients to allow for better quantization, interpolation, and stability evaluation and correction. The synthesized output speech from LPC10e is a gain scaled convolution of these predictor coefficients with either a canned glottal pulse repeated at the estimated pitch rate for voiced speech segments, or convolution with random noise representing unvoiced speech.

The LPC10e speech model then consists of two half frame voicing decisions, an estimate of the current 22.5 ms frames pitch rate, the RMS energy of the frame, and the short-time spectrum represented by a 10^{th} order prediction filter. A small portion of the more important bits of a frame are then coded with a simple hamming code to allow for some degree of tolerance to bit errors. During unvoiced frames, more bits are free and are used to protect more of the frame from channel errors.

The simple LPC10e model does generate a high degree of intelligibility. However, the speech can sound very synthetic and often contains buzzing speech. Vector Quantizing of this model to lower rates then would still contain the same synthetic sounding speech. The synthetic speech usually only degrades as the rate is reduced. A vocoder that is based on the MELP speech model may offer better sounding quality speech than one based on LPC10e. The remaining portion of the paper investigates the vector quantization of the MELP model.

MELP SPEECH MODEL

MELP was developed by the U.S. government DoD Digital Voice Processing Consortium (DDVPC) [4] as the next standard for narrow band secure voice coding. The new speech model represents a dramatic improvement in speech quality and intelligibility at the 2.4Kbps data rate. The algorithm performs well in harsh acoustic noise such as HMMWV’s, helicopters and tanks. The buzzy sounding speech of LPC10e model has been reduced to an acceptable level. The MELP model represents the next generation of speech processing in bandwidth constrained channels.

The MELP model as defined in MIL-STD-3005 [5] is based on the traditional LPC10e parametric model, but also includes five additional features. These are mixed-excitation, aperiodic pulses, pulse dispersion, adaptive spectral enhancement, and Fourier magnitudes scaling of the voiced excitation.

The mixed-excitation is implemented using a five band-mixing model. The model can simulate frequency dependent voicing strengths using a fixed filter bank. The primary effect of this multi-band mixed excitation is to reduce the buzz usually associated with LPC10e vocoders. Speech is often a composite of both voiced and unvoiced signals. MELP performs a better approximation of the composite signal than LPC10e’s Boolean voiced/unvoiced decision.

The MELP vocoder can synthesize voiced speech using either periodic or aperiodic pulses. Aperiodic pulses are most often used during transition regions between voiced and unvoiced segments of the speech signal. This feature allows the synthesizer to reproduce erratic glottal pulses without introducing tonal noise.

Pulse dispersion is implemented using a fixed pulse dispersion filter based on a spectrally flattened triangle pulse. The filter is implemented as a fixed finite impulse response (FIR) filter. The filter has the effect of spreading the excitation energy within a pitch period. The pulse dispersion filter aims to produce a better match between original and synthetic speech in regions without a formant by having the signal decay more slowly between pitch pulses. The filter reduces the harsh quality of the synthetic speech.

The adaptive spectral enhancement filter is based on the poles of the LPC vocal tract filter and is used to enhance the formant structure in the synthetic speech. The filter improves the match between synthetic and natural bandpass waveforms, and introduces a more natural quality to the output speech.

The first ten Fourier magnitudes are obtained by locating the peaks in the FFT of the LPC residual signal. The information embodied in these coefficients improves the accuracy of the speech production model at the perceptually important lower frequencies. The magnitudes are used to scale the voiced excitation to restore some of the energy lost in the 10^{th} order LPC process. This increases the perceived quality of the coded speech, particularly for males and in the presence of background noise.

MELP 2400 PARAMETER ENTROPY

The entropy values shown below give interesting insight into the existing redundancy in the MELP vocoder speech model. MELP's entropy is shown in Table 2 below. The entropy in bits was measured using the TIMIT speech database of phonetically balanced sentences that was developed by the Massachusetts Institute of Technology (MIT), SRI International, and Texas Instruments (TI). TIMIT contains speech from 630 speakers from eight major dialects of American English, each speaking ten phonetically rich sentences. The entropy of successive number of frames was also investigated to determine good choices of block length for block quantization at 600 bps. The block length chosen for each parameter is discussed in the following sections.

SPEECH PARAMETERS	BITS	ENTROPY
Aperiodic Flag	1	0.4497
Band-Pass Voicing	5	2.4126
Energy (G1+G2)	8	6.2673
Fourier Magnitudes	8	7.2294
Pitch	7	5.8916
Spectrum	25	19.2981

Table 2 - MELP 2400 ENTROPY

VECTOR QUANTIZATION

Vector quantization is the process of grouping source outputs together and encoding them as a single block. The block of source values can be viewed as a vector, hence the name vector quantization. The input source vector is then compared to a set of reference vectors called a codebook. The vector that minimizes some suitable distortion measure is selected as the quantized vector. The rate reduction occurs as the result of sending the codebook index instead of the quantized reference vector over the channel.

The vector quantization of speech parameters has been a widely studied topic in current research. At low rate of quantization, efficient quantization of the parameters using as few bits as possible is essential. Using suitable codebook structure, both the memory and computational complexity can be reduced. One attractive codebook structure is the use of a multi-stage codebook as described in [6]. In addition, the codebook structure can be selected to minimize the effects of the codebook index to bit errors. The codebooks presented in this paper are designed using the generalized Lloyd algorithm to minimize average weighted mean-squared error using the TIMIT speech database as training vectors. The generalized Lloyd algorithm consists of iteratively partitioning the training set into decisions regions for a given set of centroids. New centroids are then re-optimized to minimize the distortion over a particular decision region. The generalized Lloyd algorithm is reproduced here from reference [7].

1. Start with an initial set of codebook values $\{Y_i^{(0)}\}_{i=1,M}$ and a set of training vectors $\{X_n\}_{n=1,N}$. Set $k = 0$, $D^{(0)}=0$. Select a threshold ε.
2. The quantization region $\{V_i^{(k)}\}_{i=1,M}$ are given by $V_i^{(k)} = \{X_n\text{:}d(X_n,Y_i) < d(X_n,Y_j) \;\; \forall\, j \neq i\}\;\; i = 1,2,\ldots,M$.
3. Compute the average distortion $D^{(k)}$ between the training vectors and the representative codebook value.
4. If $(D^{(k)}-D^{(k-1)})/D^{(k)} < \varepsilon$, stop; otherwise, continue.
5. $k=k+1$. Find new codebook values $\{Y_i^{(k)}\}_{i=1,M}$ that are the average value of the elements of each quantization regions $V_i^{(k-1)}$. Go to step 2.

APERIODIC QUANTIZATION

The aperiodic pulses are designed to remove the LPC synthesis artifacts of short, isolated tones in the reconstructed speech. This occurs mainly in areas of marginally voiced speech, when reconstructed speech is purely periodic. The aperiodic flag indicates a jittery voiced state is present in the frame of speech. When voicing is jittery, the pulse positions of the excitation are randomized during synthesis based on a uniform distribution around the purely periodic mean position.

Investigation of the run-length of the aperiodic state indicates that the run-length is normally less than three frames across the TIMIT speech database and over several noise conditions tested. Further, if a run of aperiodic voiced frames does occur, it is unlikely that a second run will occur within the same block of four frames. It was decided not to send the Aperiodic bit over the channel since the effects on voice quality was not as significant as quantizing the remaining MELP parameters better.

BANDPASS VOICING QUANTIZATION

The band-pass voicing (BPV) strengths control which of the five bands of excitation are voiced or unvoiced in the MELP model. The MELP standard sends the upper four bits individually while the least significant bit is encoded along with the pitch. Table 3 illustrates the probability density function of the five bandpass voicing bits. These five bits can be easily quantized down to only two bits with very little audible distortion. Further reduction can be obtained by taking advantage of the frame-to-frame redundancy of the voicing decisions. The current low-rate coder uses a four-bit codebook to quantize the most probable voicing transitions that occur over a four-frame block. A rate reduction from four frames of five bit band-pass voicing strengths is reduced to only four bits. At four bits, some audible differences are heard in the quantized speech. However, the distortion caused by the band-pass voicing is not offensive.

BPV DECISIONS	PROB
Prob(u,u,u,u,u)	0.15
Prob(v,u,u,u,u)	0.15
Prob(v,v,v,u,u)	0.11
Prob(v,v,v,v,v)	0.41
Prob(remaining)	0.18

Table 3 - MELP 600 BPV MAP

ENERGY QUANTIZATION

MELP's energy parameter exhibits considerable frame-to-frame redundancy, which can be exploited by various block quantization techniques. A sequence of energy values from successive frames can be grouped to form vectors of any dimension. In the MELP 600 bps model, we have chosen a vector length of four frames of two gain values per frame. The energy codebook was created using the K-means vector quantization algorithm and is described in [7]. The codebook were trained using training data scaled by multiple levels to prevent sensitivity to speech input level. During the codebook training process, a new block of four energy values are created for every new frame so that energy transitions are represented in each of the four possible location within the block. The resulting codebook is searched resulting in a codebook vector that minimizes mean squared error.

For MELP 2400, two individual gain values are transmitted every frame period. The first gain value is quantized to five bits using a 32-level uniform quantizer ranging from 10.0 to 77.0 dB. The second gain value is quantized to three bits using an adaptive algorithm that is described in [5]. In the MELP 600 bps model, we have vector quantized both of MELP's gain values across four frames. Using the 2048 element codebook, we reduce the energy bits / frame from 8 bits per frame for MELP 2400 down to 2.909 bits per frame for MELP 600. Quantization values below 2.909 bits per frame for energy were investigated, but the quantization distortion becomes audible in the synthesized output speech and effected intelligibility at the onset and offset of words.

FOURIER MAGNITUDES QUANTIZATION

The excitation information is augmented by including Fourier coefficients of the LPC residual signal. These coefficients or magnitudes account for the spectral shape of the excitation not modeled by the LPC parameters. These Fourier magnitudes are estimated using a FFT on the LPC residual signal. The FFT is sampled at harmonics of the pitch frequency. In the current MIL-STD-3005, the lower ten harmonics are considered more important and are coded using an eight-bit vector quantizer over the 22.5 ms frame.

The Fourier magnitude vector is quantized to one of two vectors. For unvoiced frames, a spectrally flat vector is selected to represent the transmitted Fourier magnitude. For voiced frames, a single vector is used to represent all voiced frames. The voiced frame vector was selected to reduce some of the harshness remaining in the lowrate vocoder. The reduction in rate for the remaining MELP parameters reduce the effect seen at the higher data rates to Fourier magnitudes. No bits are required to perform the above quantization.

PITCH QUANTIZATION

The MELP model estimates the pitch of a frame using energy normalized correlation of 1kHz low-pass filtered speech. The MELP model further refines the pitch by interpolating fractional pitch values as described in [5]. The refined fractional pitch values are then checked for pitch errors resulting from multiples of the actual pitch value. It is this final pitch value that the MELP 600 vocoder uses to vector quantize.

MELP's final pitch value is first median filter (order 3) such that some of the transients are smoothed to allow the low rate representation of the pitch contour to sound more natural. Four successive frames of the smoothed pitch values are vector quantized using a codebook with 128 elements. The codebook was trained using the k-means method as described in [7]. The resulting codebook is searched resulting in the vector that minimizes mean squared error of voiced frames of pitch.

SPECTRUM QUANTIZATION

LPC spectrum of MELP is converted to line spectral frequencies (LSFs) [8] which is one of the more popular compact representations of the LPC spectrum. The LSF's are quantized with a four-stage vector quantization algorithm [9]. The first stage has seven bits, while the remaining three stages use six bits each. The resulting quantized vector is the sum of the vectors from each of the four stages and the average vector. At each stage in the search process, the VQ search locates the "M best" closest matches to the original using a perceptual weighted Euclidean distance [5]. These M best vectors are used in the search for the next stage. The indices of the final best at each of the four stages determine the final quantized LSF.

The low-rate quantization of the spectrum quantizes four frames of LSFs in sequence using a four-stage vector quantization process. The first two stages of codebook uses ten bits, while the remaining two stages uses nine bits each. The search for the best vector uses a similar "M best" technique with perceptual weighting as is used for the MIL-STD-3005 vocoder. Four frames of spectra are quantized to only 38 bits.

The codebook generation process uses both the K-Means and the generalized Lloyd technique. The K-Means codebook is used as the input to the generalized Lloyd process. A sliding window was used on a selective set of training speech to allow spectral transitions across the four-frame block to be properly represented in the final codebook. It is important to note that the process of training the codebook requires significant diligence in selecting the correct balance of input speech content. The selection of training data was created by repeatedly generating codebooks and logging vectors with above average distortion. This process removes low probability transitions and some stationary frames that can be represented with transition frames without increasing the over-all distortion beyond unacceptable levels.

DAM / DRT PERFORMANCE

The Diagnostic Acceptability Measure (DAM) [10] and the Diagnostic Rhyme Test (DRT) [11] are used to compare the performance of the MELP vocoder to the existing LPC based system. Both tests have been used extensively by the US government to quantify voice coder performance. The DAM requires the listeners to judge the detectability of a diversity of elementary and complex perceptual qualities of the signal itself, and of the background environment. While the DRT is a two choice intelligibility test based upon the principle that the intelligibility relevant information in speech is carried by a small number of distinctive features. The DRT was designed to measure how well information as to the state of six binary distinctive features (voicing, nasality, sustension, sibiliation, graveness, and compactness) have been preserved by the communications system under test.

The DRT performance of both MELP based vocoders exceeds the intelligibility of the LPC vocoders for most test conditions. The 600bps MELP DRT is within just 3.5 points of the higher bit-rate MELP system. The rate reduction by vector quantization of MELP has not effected the intelligibility of the model noticeably. The DRT scores for HMMWV demonstrate that the noise pre-processor of the MELP vocoders enables better intelligibility in the presence of acoustic noise.

TEST CONDITION	DRT	DAM
Source Material (QUIET)	95.9[1]	85.8[1]
MELPe 2400 (QUIET)	94.0[1]	69.1[1]
MELPe 600 (QUIET)	90.5[1]	54.9[1]
LPC10e 2400 (QUIET)	89.4[1]	50.0[1]
LPC10e 600 (QUIET)	86.8[1]	47.1[1]
Source Material (HMMWV)	91.0[2]	45.0[2]
MELPe 2400 (HMMWV)	74.4[2]	52.6[2]
MELPe 600 (HMMWV)	65.0[1]	40.3[1]
LPC10e 2400 (HMMWV)	68.7[1]	37.6[1]
LPC10e 600 (HMMWV)	61.9[1]	35.3[1]

Table 4 - VOCODER DRT/DAM TESTS

The DAM performance of the MELP model demonstrates the strength of the new speech model. MELP's speech acceptability at 600 bps is more than 4.9 points better than LPC10e 2400 in the quiet test condition, which is the most noticeable difference between both vocoders. Speaker recognition of MELP 2400 is much better than LPC10e 2400. MELP based vocoders have significantly less synthetic sounding voice with much less buzz. Audio of MELP is perceived to being brighter and having more low-end and high-end energy as compared to LPC10e.

SECURE VOICE AVALIBILITY

Secure voice availability is directly related to the bit-error rate performance of the waveform used to transfer the vocoder's data and the tolerance of the vocoder to bit-errors. A 1% bit-error rate causes both MELP and LPC based coders to degrade voice intelligibility and quality as seen in table 5. The useful range therefore is below approximately a 3% bit-error rate for MELP and 1% for LPC based vocoders.

The 1% bit-error rate of the MIL-STD-188-110B waveforms can be seen for both a Gaussian and CCIR Poor channels in figures 1 and figure 2, respectively. The curves indicate a gain of approximately seven dB can be achieved by using the 600 bps waveform over the 2400bps standard. It is in this lower region in SNR that allows HF links to be functional for a longer portion of the day. In fact, many 2400 bps links cannot function below a 1% bit-error rate at any time during the day based on propagation and power levels. Typical ManPack Radios using 10-20W power levels make the choice in vocoder rate even more mission critical.

TEST CONDITION	DRT	DAM
MELPe 2400	91.5[2]	54.7[2]
MELPe 600	85.2[1]	43.1[1]
LPC10e 2400	81.4[2]	N/A
LPC10e 600	79.5[1]	38.3[1]

Table 5 - BER 1% DRT/DAM TESTS

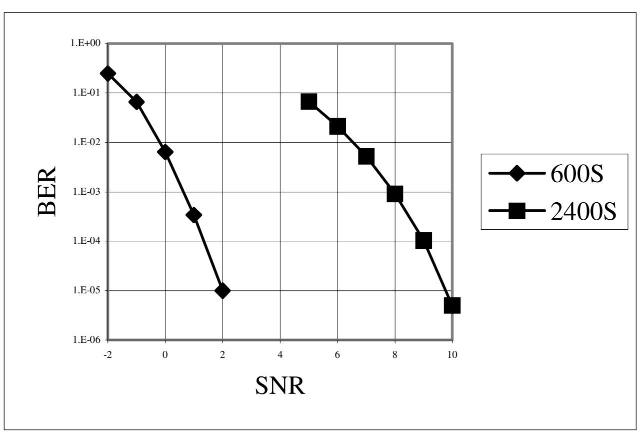

Figure 1 MIL-STD-188-110B AWGN

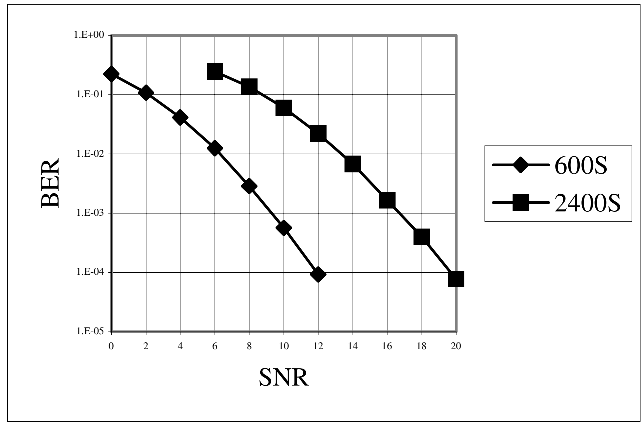

Figure 2 MIL-STD-188-110B CCIR POOR

HARDWARE IMPLEMENTATION

The MELP vocoder discussed in this paper runs real-time on a sixteen bit fixed-point Texas Instrument's TMS320VC5416 digital signal processor. The low-power hardware design resides in the RF-5800H/PRC-150 ManPack Radio and is responsible for running several voice coders and a variety of data related interfaces and protocols. The DSP hardware design runs the on-chip core at 150MHz (zero wait-state) while the off-chip accesses are limited to 50 MHz (two wait-state). The data memory architecture has 64K zero wait-state on chip memory and 256K of two wait-state external memory which is paged in 32 K banks. For program memory, we have an additional 64K zero wait-state on chip memory and 256K of external memory that is fully addressed by the DSP.

The 2400 bps MELP source code was developed by NSA, Microsoft, ASPI, Texas Instruments, and ATT. The source code consists of TI's 54X assembly language source code combined with Harris's MELP 600 vocoder. This code has been modified to run on the TMS320VC5416 architecture using the FAR CALLING run-time environment, which allows DSP programs to span more than 64K. The code has been integrated into a C calling environment using TI's C initialize mechanism to initialize MELP's variables and combined with a Harris proprietary DSP operating system.

Run-time loading on the MELP 2400 target system allows for Analysis to run at 24.4 % loaded, the Noise Pre-Processor is 12.44% loaded, and Synthesis to run at 8.88 % loaded. Very little load increase occurs as part of MELP 600 Synthesis since the process is no more than a table lookup. The additional cycles for MELP 600 vocoder is contained in the vector quantization of the spectrum in Analysis.

CONCLUSIONS

The speech quality of the new MIL-STD-3005 vocoder is indeed much better than the old FED-STD-1015 [3] vocoder. This paper has investigated the use of Vector Quantization techniques on the new standard vocoder combined with the use of the 600 bps waveform as is defined in U.S. MIL-STD-188-110B. The results seem to indicate that a 5-7 dB improvement in HF performance is possible on some fading channels. Furthermore, the speech quality of the 600 bps vocoder is better than the existing 2400 bps LPC10e standard for several test conditions. However, on air testing is required to validate the simulation results presented. If the on air tests confirm the results presented in this paper, low-rate coding of MELP should be considered to be added to the MIL-STD-3005 for improved communication and extended availability to ManPack radios on difficult HF links.

ACKNOWLEDGMENTS

The author wishes to acknowledge the contributions of John Collura of the National Security Agency and to all participating members of the U.S. government's DoD Digital Voice Processing Consortium (DDVPC), in their efforts to create the 2.4Kbps Mixed Excitation Linear Prediction (MELP) voice coding algorithm standard.

REFERENCES

(1) MIL-STD-188-110B – "Mil. Std. Interoperability and Performance Standards for Data Modems", Draft Version Revised 7 March 2000
(2) Collura, John S., "Noise Pre-Processing for Tactical Secure Voice Communications." IEEE Speech Coding Workshop-99, Porvoo Finland
(3) "Analog to Digital Conversion of Voice by 2400 bits/second Linear Predictive Coding", Federal Standard 1015, Nov 1984
(4) Supplee Lynn M., Cohn Ronald P., Collura John S., McCree Alan V., "MELP: The New Federal Standard at 2400 bps", IEEE ICASSP-97 Conference, Munich Germany
(5) "Analog-to-Digital Conversion of voice by 2400 bit/second Mixed Excitation Linear Prediction (MELP)", MIL-STD-3005, Dec 1999
(6) Gersho A., Gray R. M., Vector Quantization and Signal Compression, Norwell, MA:Kluwer Academic Publishers, 1991
(7) Linde Y., Buzzo A., Gray R. M., "An Algorithm for Vector Quantization Design", IEEE Transactions on Communications, COM-28:84-95, Jan 1980
(8) Soong F., Juang B., "Line Spectrum Pairs (LSP) and Speech Compression", IEEE Int. Conf. On Acoustics, Speech, and Signal Processing, 1983
(9) Juang B. –H., Gray A. H. Jr., "Multiple Stage Vector Quantization for Speech Coding", In International Conference on Acoustics, Speech, and Signal Processing, volume 1, pages 597-600, Paris France, April 1982
(10) Voiers William D., "Diagnostic Acceptability Measure (DAM): A Method for Measuring the Acceptability of Speech over Communications Systems", Dynastat, Inc.; Austin Texas
(11) Voiers William D., "Diagnostic Evaluation of Speech Intelligibility.", In M.E. Hawley, Ed, Speech Intelligibility and Speaker Recognition (Dowder, Huchinson, and Ross; Stroudsburg, PA 1977)