NOT MEASUREMENT SENSITIVE

**MIL-STD-3005**
**20 December 1999**

# DEPARTMENT OF DEFENSE TELECOMMUNICATIONS SYSTEMS STANDARD

## ANALOG-TO-DIGITAL CONVERSION OF VOICE BY 2,400 BIT/SECOND MIXED EXCITATION LINEAR PREDICTION (MELP)

**AMSC N/A** **AREA TCSS**

# FOREWARD

1. This standard is approved for use by all Departments and Agencies of the Department of Defense (DoD) and is a replacement for FIPSPUB-137, Telecommunications: Analog to Digital Conversion of Voice by 2,400 Bit/Second Linear Predictive Coding.

2. This standard contains design requirements for analog-to-digital (A-D) conversion of voice by 2,400 bit/second Mixed Excitation Linear Prediction (MELP). Adherence to this standard is required to produce interoperable systems at the defined rate and to meet or exceed the minimum performance requirements.

3. Appendix A of this document contains an example of an interoperable MELP algorithm. This information is provided as guidance only.

4. Appendix B contains guidelines for verification of all new implementations of this standard. New implementations must be verified to guarantee that the standard was correctly implemented. This verification process will determine if the standard is interoperable with other MELP implementations and will verify that the performance of the implementation meets or exceeds the performance of the MELP reference coder.

5. Beneficial comments (recommendations, additions, deletions) and any pertinent data which may be of use in improving this document should be addressed to: R224, National Security Agency, 9800 Savage Road STE 6516, Ft. Meade, Maryland 20755-6516 by using the Standardization Document Improvement Proposal (DD Form 1426) appearing at the end of this document or by letter.

CONTENTS

<u>APPENDIX</u>

<u>FIGURE</u>



<u>TABLE</u>

# 1. SCOPE

1.1 Scope. This standard establishes interoperability and performance requirements for analog-to-digital (A-D) conversion of voice by 2,400 bit/second Mixed Excitation Linear Prediction (MELP). The requirements presented in this document must be met in order for systems to be interoperable at 2,400 bit/second. Minimum performance requirements are also provided, but may be exceeded. The performance requirements are provided in Appendix B.

# 2. APPLICABLE DOCUMENTS

2.1 General. Documents listed in this section are required in order for the document user to fully understand the guidance being provided by this standard.

2.2 Government documents.

2.2.1 Specifications, standards, and handbooks. The following specifications, standards, and handbooks form a part of this document to the extent herein. Unless otherwise specified, the issues of these documents are those listed in the issue of the Department of Defense Index of Specifications and Standards (DoDISS) and supplement thereto, cited in the solicitation.

STANDARDS

FEDERAL

| | |
|---|---|
| FED-STD-1016 | Telecommunications: Analog to Digital Conversion of Radio Voice by 4,800 Bit/Second Code Excited Linear Prediction (CELP) |
| FED-STD-1037 | Glossary of Telecommunications Terms |
| FIPSPUB-137 | Telecommunications: Analog to Digital Conversion of Voice by 2,400 Bit/Second Linear Predictive Coding |

MILITARY

| | |
|---|---|
| MIL-STD-188-113 | Interoperability and Performance Standards for Analog-to-Digital Conversion Techniques |

(Unless otherwise indicated, copies of the above specifications, standards, and handbooks are available from the Standardization Document Order Desk, 700 Robbins Avenue, Building 4D, Philadelphia, PA 19111-5094.)

(Copies of the Federal Information Processing Standards (FIPS) are available to Department of Defense activities from the Standardization Document Order Desk, 700 Robbins Avenue, Building 4D, Philadelphia, PA 19111-5094. Others must request copies of FIPS from the National Technical Information Service, 5285 Port Royal Road, Springfield, VA 22161-2171.)

2.2.2 Other Government documents, drawings, and publication. The following other Government documents, drawings, and publications form a part of this document to the extent specified herein. Unless otherwise specified, the issues are those cited in the solicitation.

| | |
|---|---|
| DoDISS | Department of Defense Index of Specifications and Standards |

(Copies of the DoDISS are available on a yearly subscription basis either from the Government Printing Office or the DoDSSP Subscription Services, 700 Robbins Avenue, Building 4D, Philadelphia, PA 19111-5094.)

2.3 Other publications. The following documents form a part of this standard to the extent specified herein. Unless otherwise specified, the issues of the documents which are DoD adopted should be those listed in the issue of the DoDISS specified in the solicitation. The issues of the documents which have not been adopted should be those in effect on the date of the cited DoDISS.

NORTH ATLANTIC TREATY ORGANIZATION (NATO)

STANDARDIZATION AGREEMENT (STANAG's)

| | |
|---|---|
| STANAG 4198 | Parameters and Coding Characteristics That Must be Common to Assure Interoperability of 2400 BPS Linear Predictive Encoded Digital Speech |
| STANAG 4209 | The NATO Multi-Channel Tactical Digital Gateway -- Standards for Analogue to Digital Conversion of Speech Samples |

(Application for copies should be addressed to the Naval Publications and Forms Center, 5801 Tabor Avenue, Philadelphia, PA 19120-5099.)

(Non-Government standards are generally available for reference from libraries. They are also distributed among non-Government standards bodies and using Federal agencies.)

2.4 Order of precedence. In the event of a conflict between the text of this standard and the references cited herein, the text of this standard should take precedence. Nothing in this document, however, supersedes applicable laws and regulations unless a specific exemption has been obtained.

## 3. DEFINITIONS

3.1 Terms. Definitions of terms used in this standard should be as specified in the current edition of FED-STD-1037. In addition, the following definitions are applicable for the purpose of this standard.

3.1.1 Adaptive spectral enhancement. This feature enhances the formant structure of the synthetic speech by use of an adaptive spectral enhancement filter that is applied to the mixed excitation.

3.1.2 Aperiodic pulses. Aperiodic pulses are used in the excitation model of the synthesizer when the aperiodic flag is set to 1. The aperiodic flag is set to one when the jittery voiced state is encountered during the voicing decision process. This feature is used to reduce the buzzy quality of the synthetic speech signal.

3.1.3 Fourier magnitude modeling. Fourier magnitude modeling involves determining the Fourier magnitudes of the first 10 pitch harmonics of the prediction residual and vector quantizing them with 8 bits for transmission. The use of this technique improves the accuracy of the speech production model at the perceptually important lower frequencies.

3.1.4 Hamming codes. A class of linear codes used for forward error correction. These codes are used only in the unvoiced mode.

3.1.5 Jitter. Random variations introduced into the duration of a signal.

3.1.6 Linear prediction coding. A method for approximating the current speech sample by using a linear combination of past and future speech samples. This method efficiently represents a speech signal and its spectrum characteristics with a very small number of parameters when combined with an appropriate excitation signal.

3.1.7 Mixed excitation. The combination of a periodic function (such as a pulse train) and random noise for use in the excitation model. This combination is applied to sub regions of the frequency domain of the excitation signal.

3.1.8 Prediction coefficients. A set of values that are calculated using a short segment of the input speech signal and provide an estimate of the spectral properties of that signal. These values are determined by performing linear prediction analysis on the input signal. The goal of the analysis is to produce values that minimize the short term mean-squared prediction error over the input segment.

3.1.9 Pulse dispersion. Uses a fixed filter to spread the excitation energy within a pitch period.

3.1.10 Uniform quantizer. A uniform quantizer uses levels and step sizes that are distributed uniformly.

3.1.11 Weighted Euclidean distance. The euclidean distance is a distortion measure between two vectors. In this standard the euclidean distance is determined by summing the squared difference between two vectors for a select number of samples. Normally the euclidean distance is the square root of the measure described in the previous sentence.

3.2 Acronyms used in this standard. The acronyms used in this standard are defined as follows:

A-D - Analog to Digital

DoD - Department of Defense

DoDISS - Department of Defense Index of Specifications and Standards

DoDSSP - Department of Defense Single Stock Point

FEC - Forward Error Correction

LPC - Linear Prediction Coding

LSB - Least Significant Bit

LSF - Line Spectrum Frequency

MELP - Mixed Excitation Linear Predictions

MSB - Most Significant Bit

MSVQ - Multi-Stage Vector Quantizer

STANAG - Standardization Agreement

# 4. GENERAL REQUIREMENTS

Not applicable

# 5. DETAILED REQUIREMENTS

5.1 General. The Mixed Excitation Linear Prediction coder is based on the traditional Linear Prediction Coder (LPC) parametric model, but also includes five additional features. They are mixed excitation, aperiodic pulses, adaptive spectral enhancement, pulse dispersion, and Fourier magnitude modeling. A MELP frame interval is 22.5 ms $\pm 0.01$ percent in duration and contains 180 voice samples (8,000 samples/second).

5.2 Analog specification. The recommended analog requirements for the MELP coder are for a nominal bandwidth ranging from 100 Hz to 3800 Hz. Although the MELP coder will operate with a more band limited signal, performance degradation will result. To ensure proper operation of the MELP coder, the A-D conversion process should produce peak values of (or near) -32768 and 32767. Additionally, the coder should have unity gain, which means that the output speech level should match that of the input speech.

5.3 Parameter quantization and encoding. The MELP parameters which are quantized and transmitted are the final pitch ($P_3$); the bandpass voicing strengths ($Vbp_i$, i = 1, 2, …, 5); the two gain values ($G_1$ and $G_2$); the linear prediction coefficients ($a_i$, i= 1, 2, …, 10); the Fourier magnitudes; and the aperiodic flag. The use of the following quantization procedures is required for interoperability among various implementations.

5.3.1 Pitch and overall voicing. The final pitch ($P_3$), and the low band voicing strength ($Vbp_1$), are quantized jointly using 7 bits, as follows. If $Vbp_1 \leq 0.6$, then the frame is unvoiced and the all-zero code is sent. Otherwise, the log of $P_3$ is quantized with a 99-level uniform scalar quantizer (see 5.3.7) ranging from log20 to log160. The resulting index (range 0 to 98) is then mapped to the transmitted 7-bit codeword using the encode/decode values in table I. All 28 codes with Hamming weight of 1 or 2 are reserved for error protection. This table is also used in decoding the 7-bit pitch code to determine if a frame is voiced, unvoiced, or whether a frame erasure is indicated. A frame is determined unvoiced if the pitch code is all zero or has only one bit set. If two bits are set, then a frame erasure is indicated. Otherwise, the voiced mode is used and the pitch index is determined from the received code according to table I.