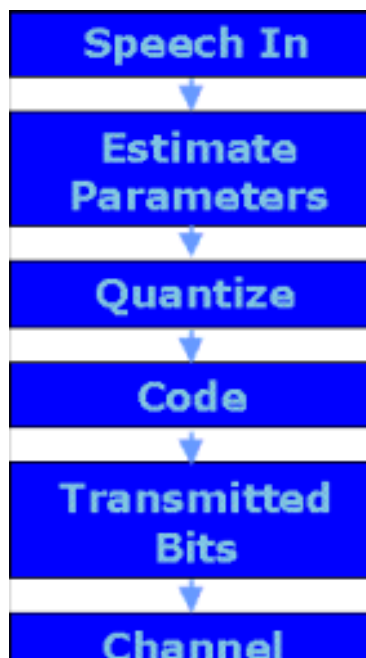




## Application of Vocoders to Wireless Communications

The need for increased utilization of available wireless communication spectrum has fueled the development of voice coding technology. From simple waveform coding techniques operating at 64 kbps, the advance of speech coding algorithms has produced communication quality systems at 2 kbps and below. This allows up to 32 communications channels to operate in the bandwidth formerly occupied by one channel. In addition to the increased capacity, new digital communication systems offer enhanced security due to the ease of implementing encryption algorithms.

Vocoders generally operate by modeling a segment (or frame) of the speech waveform on the order of 20 ms. The speech model parameters are estimated, quantized, coded, and transmitted over the communication channel (Figure 1). At the receiver, the transmitted values are decoded, reconstructed, and used to synthesize speech.



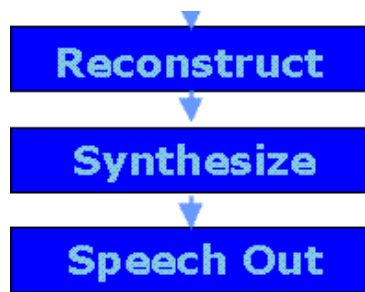


Figure 1

### Channel Degradations Drive Vocoder Design

Vocoders which have been designed for extremely low bit error rates, such as those encountered in land line communications, often experience serious degradation when applied to the much higher bit error rates found in wireless communications. Consequently, it is important to consider robustness to channel degradations during the vocoder algorithm design process. The various techniques which can be used to improve vocoder robustness to bit errors are discussed in the following sections.

### Coding Gain Vs. Error Persistence Trade-off

Due to the slowly varying nature of the speech production mechanism, the speech model parameters contain much frame to frame redundancy. The vocoder can exploit this redundancy by predicting the current frame parameters from the reconstructed parameters of previous frames. In a simple differential encoding technique, the difference between the estimated parameter for the current frame and the reconstructed value from the previous frame is quantized, coded, and transmitted. However, if the parameters of the current frame are highly dependent on parameters transmitted in previous frames, then errors in the parameters caused by channel errors can have long persistence times. A trade-off between coding gain and error persistence is achieved by reducing the dependence on previous reconstructed values using a multiplicative constant (prediction coefficient) between 0 and 1.

### Bit Prioritization Improves Robustness

Some simple high rate waveform coding techniques can be designed to be equally sensitive to bit errors anywhere in the transmission stream. However, vocoders generally have some speech model parameters which are more sensitive to bit errors and others which are less sensitive to bit errors. This can be determined by listening to the effect of bit errors on each parameter. In addition, parameters are often encoded with multiple bits per parameter producing a range of sensitivity from Most Sensitive Bits (MSBs) to Least Sensitive Bits (LSBs). The bits used to encode all of the parameters can then be ordered from most sensitive to least sensitive based on listening tests and bit order within a parameter. The robustness to bit errors is improved by adding redundant bits through the use of error control coding. Optimizing the error control coding so that more protection against bit errors is provided to the MSBs and less protection is provided to the LSBs achieves the best performance.

### Interleaving Spreads Bursts of Errors

Wireless communication channels with fades of the signal power are prone

to errors occurring in bursts. Burst errors can cause problems by breaking error control codes when the number of errors exceeds the maximum correctable errors for the specific code used. For short bursts, intra-frame interleaving improves performance by spreading the burst of errors over several different code words. For example, if four codewords each containing 23 bits that can correct up to 3 bit errors are used in a frame consecutively, then a burst error of length 4 bits will break a single code. However, if the four codewords are interleaved (i.e. bit 1 codeword 1, bit 1 codeword 2, bit 1 codeword 3, bit 1 codeword 4, bit 2 codeword 1, ...) then each codeword contains only one error which is easily corrected. Since intra-frame interleaving only modifies the bit ordering within the current frame, no additional delay is generally needed for implementation. If additional delay can be tolerated, inter-frame (more than just the current frame) interleaving can be used to further increase the performance with longer burst errors.

### **Error Mitigation**

When very high bit error rates occur on the communication channel, error mitigation strategies must be developed for reducing the impact of uncorrected errors. The number of errors in the current frame can be estimated based on the number of corrected and detected errors in the error control words. The estimated number of errors can then be used in a number of error mitigation strategies.

- 1) Adaptive smoothing of speech model parameters - Since the speech model parameters generally change slowly with time, reducing the frame to frame variation for frames with high error rates improves performance.
- 2) Frame repeating - If the number of estimated errors is very high, so that there is little chance of decoding very many parameters correctly, it is usually best to just repeat the parameters from the prior frame.
- 3) Muting - If the number of estimated errors continues to be very high, it is generally best to mute the output.

### **IMBE Vocoder**

In order to demonstrate some of these design considerations, relevant details of the IMBE™ (Improved Multi-Band Excitation) Vocoder implementation for APCO Project 25 North American land mobile radio system will be described. The total bit rate is 7200 bps with a frame size of 20 ms. The IMBE™ Vocoder model parameters consist of a pitch or fundamental frequency, a set of Voiced/Unvoiced (V/UV) parameters, and a set of spectral magnitudes. In the IMBE model, the spectrum is divided into a number of frequency bands, the V/UV parameters indicate whether each band is voiced (contains periodic energy) or unvoiced (contains noise like energy). This model provides improved performance for speech in background noise or speech with mixed voicing.

The coding gain vs. error persistence trade-off resulted in selection of a prediction coefficient varying between .7 for low pitched speakers and .4 for high pitched speakers. These values provided good coding gain with low error persistence.

The total of 144 bits per frame is allocated to speech model parameters (87 bits), synchronization (1 bit), and error control codes (56 bits). The speech

model parameter bits are divided into 4 groups from most sensitive to bit errors to least sensitive. The most sensitive are encoded with a [23,12] Golay code with an additional error detecting code. The next group is encoded with three [23,12] Golay codes. The next group is encoded with three [15,11] Hamming codes and no error control codes are applied to the least sensitive group. Intra-frame interleaving is used to spread burst errors over multiple codewords. Adaptive smoothing of speech model parameters, frame repeating, and muting are used to mitigate uncorrected errors.

### Voice Quality Test Results

The results of a voice codec evaluation conducted by the Telecommunications Industry Association (TIA) for APCO Project 25 are shown in Figure 2. This figure shows the Mean Opinion Score (MOS) as a function of average bit error rate and fading characteristics. The MOS is generated by asking listeners to rate the vocoded speech on a scale of 1 (bad) to 5 (excellent). The fading characteristics were generated by assuming a multipath environment with the receiver moving at the indicated speed in Miles Per Hour (mph). For a given vocoder, a higher average bit error rate produces a lower MOS due to higher speech model parameter distortion. For a given vocoder, a slower speed produces a lower MOS due to longer duration fades (long bursts of errors) which are more difficult to correct with error control codes. The four vocoders compared in this evaluation are the IMBE vocoder, Vector Sum Excited Linear Prediction (VSELP), Sinusoidal Transform Coder (STC) and Code Excited Linear Prediction (CELP). The high voice quality scores together with low complexity scores resulted in the selection of IMBE as the APCO Project 25 vocoder standard.

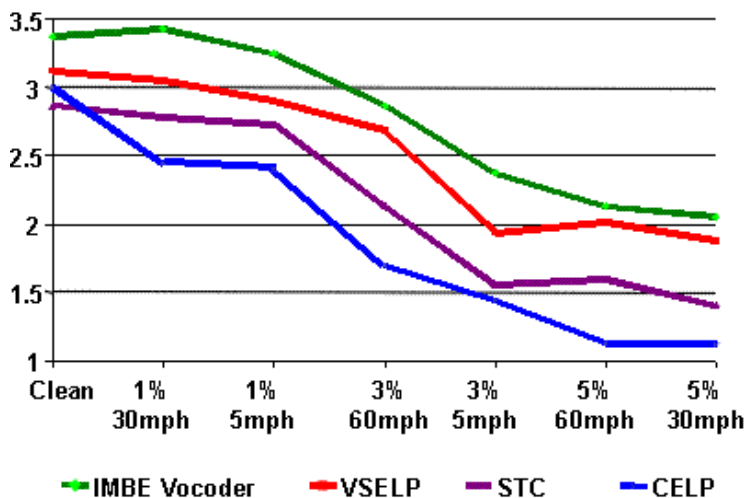


Figure 2. APCO Voice Quality Test

### APCO Project 25 Update

Currently, a number of Public Safety Agencies have adopted Project 25 standards. In the immediate future, an additional number will be seeking compliance. A partial list of these organizations is as follows:

Public Works & Government Services Canada

Aerospace, Marine and Electronic Systems

Security, Communications & Tactical Information Systems

United States Department of Commerce

U. S. Government, Drug Enforcement Agency

National Exchange Police Information of Australia Endorsed  
compliance with APCO 25 Standards

Digital Voice Systems, Inc. 234 Littleton Road Westford, MA 01886 USA  
Tel: 978 392 0002 Fax: 978 392 8866 [info@dvsinc.com](mailto:info@dvsinc.com)

Copyright © 2006 Digital Voice Systems, Inc. All Rights Reserved.