

AN INTEGRATED VOICE CODEC AND ECHO CANCELLER IMPLEMENTED IN A SINGLE DSP PROCESSOR

P.J. Wilson, J.M. Puetz, A.V. McCree, and D.T. Wang.

M/A-COM Telecommunications Division
3033 Science Park Road
San Diego, CA 92121

ABSTRACT

A new Linear Predictive Echo Canceller (LPEC) is presented which modifies the classical echo canceller structure to pre-whiten the input signal and hence have better convergence properties in the presence of highly correlated input signals such as speech. This structure is integrated with a 16 KBPS Residual Excited Linear Predictive (RELPE) Vocoder [1] to allow significant execution speed savings by sharing common processing modules. The integrated LPEC has been implemented in the Vocoder Demonstration System (VDS-20) using a single TMS32020 processor and will be available at the conference for demonstrations.

I. INTRODUCTION

Echo cancellation in the telephone public switched network is currently only required for long-haul trunks where significant transmission delays (>40 ms) are present. With the advent of digital voice codecs and speech compression technology, this requirement must be re-evaluated for some applications. At high transmission rates (64 Kbps PCM & 32 Kbps ADPCM), the delay introduced by the voice codec is nominal, normally less than a few samples, whereas at lower bit rates (<16 Kbps) significant processing delays are possible, thereby requiring some form of echo cancellation, even for geographically short transmission distances.

Medium band (9.6-16 Kbps) voice coding algorithms use speech specific or hybrid coding techniques that require the transmission of side information normally in the form of the vocal tract response; these algorithms can introduce significant delays (>40 ms single direction). Practical telephony communication systems must use existing 2-wire phone equipment; the typical 2-4 wire conversion hardware provides poor transmission line impedance matching, typically 6-12 dB echo return loss (ERL). When poor ERL is coupled with voice codec processing delays, performance results are unsatisfactory unless echo cancellation is provided within the system. A cost-effective system solution requires an integrated voice codec and echo canceller.

The design and physical implementation constraints of the voice codec and echo canceller restrict the choice of algorithms to those which may be integrated into a single DSP processor operating full-duplex in real-time. Significant trade-offs may be accomplished by using the inherent delay introduced by the voice codec to simplify the implementation of the echo canceller; traditional echo cancellers must operate at the sampling rate. In addition, the algorithms can share common processing modules to allow faster operation and more complex algorithms to be introduced.

The paper is structured as follows. Section II discusses the voice codec algorithm, the Residual Excited Linear Predictive (RELPE) Vocoder. Section III discusses the echo canceller design and section IV the integration of these two units in a single DSP processor. Section V briefly discusses the hardware implementation, the Vocoder Demonstration System (VDS-20). Section VI presents a performance analysis and Section VII concludes the paper, presenting ongoing areas of research.

II. RELPE VOCODER

The selection of a suitable voice codec algorithm has been structured to meet the following requirements:

- Good subjective quality speech at medium bit rates (9.6-16 KBPS).
- Ability to pass call progress and DTMF tones.
- Small board area with low power consumption.

The Residual Excited Linear Predictive (RELPE) Vocoder algorithm [1] implemented in a single TMS320 processor, has been chosen for its good quality voice transmission and its flexibility for alternate transmission rates.

The RELPE Vocoder algorithm divides into two functions, Analysis and Synthesis, which correspond to the transmitter and receiver respectively. The Analysis algorithm operates by generating the Prediction Error signal (or Residual) obtained by inverse filtering the speech data using the LPC coefficients. A downsampled baseband Residual is quantized using Pitch Predictive ADPCM (PPADPCM) and is transmitted to the Receiver together with the quantized LPC Reflection coefficients. At

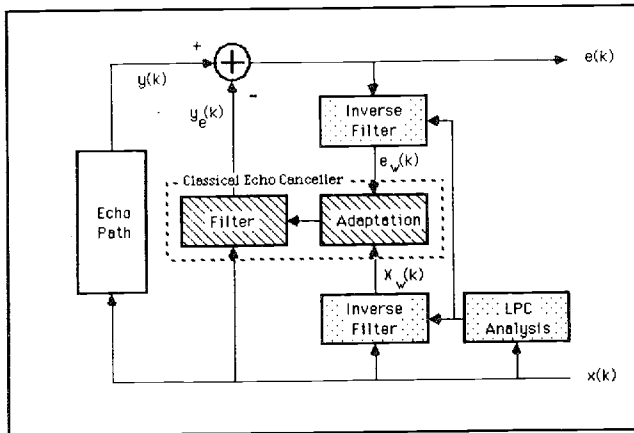


Figure 1. Linear Predictive Echo Canceller Structure

the receiver, the Synthesis algorithm generates a full-band Residual signal by non-linear distortion techniques. This signal is used as the excitation for an all-pole filter to re-synthesize speech. The 16 KBPS RELP Vocoder operates with a baseband of 2 kHz, 8-level residual quantizer, and a frame size of 22.5 ms, thereby introducing a single direction transmission delay of 45 ms.

III. ECHO CANCELLATION

Echoes in a communications network are caused when an impedance mismatch in a 4-wire circuit allows the coupling of transmit and receive paths. For a local subscriber loop, the echo dispersion is typically of short duration (<3 ms). An echo canceller synthesizes a replica of the echo signal and subtracts it from the returned signal. The echo replica, $y_e(k)$, is generated by a discrete convolution of the far end speech, $x(k)$, and echo path impulse response, $h_1(k)$.

$$y_e(k) = \sum_{i=0}^{N-1} h_1(k)x(k-i)$$

In the classical echo canceller [3], this impulse response is modeled as an adaptive transversal filter whose coefficients are updated iteratively.

This technique performs well for a white noise source, but for speech, which is typically highly correlated, performance limitations result in slow convergence and sometimes suboptimal convergence to local rather than global minima.

A modification to the classical structure provides better convergence properties for correlated input signals by implementing pre-whitening. One technique is to apply Linear Prediction [4]; the coefficients obtained from linear prediction analysis are used to inverse filter and decorrelate, or pre-whiten, the speech. This produces a significant performance improvement over the classical echo canceller. Figure 1 presents a block diagram of the Linear Predictive Echo Canceller (LPEC). It should be noted that the LPEC is by definition a block process and hence compatible with a block processing voice codec.

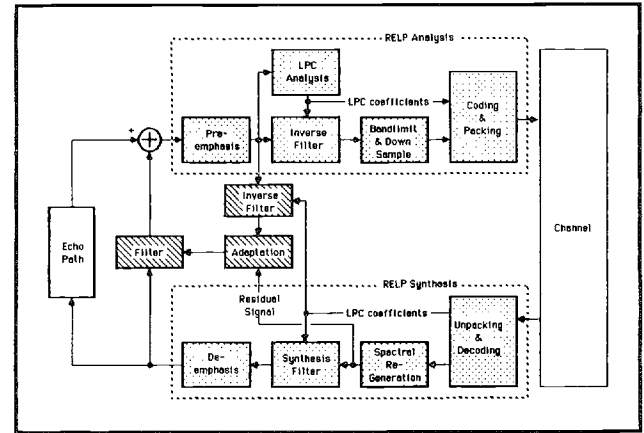


Figure 2. Integrated RELP Vocoder and Linear Predictive Echo Canceller

In addition, the design of an echo canceller must address several system implementation issues:

- Residual echo suppression
- Doubletalk
- Operation with signalling tones

A dual-path model [5] can be used to solve the last two issues. In this way, the echo canceller does not diverge either during a doubletalk condition or in the presence of tones, but the coefficients are frozen to maintain echo suppression.

IV. VOICE CODEC AND ECHO CANCELLER INTEGRATION

Current implementations of digital voice communication systems that require echo cancellation typically use separate units. In our experience, the integration of a stand alone echo canceller with the RELP Vocoder typically degrades subjective voice quality because of:

- Analog noise introduced by multiple A/D-D/A conversions
- Residual echo suppression silencing the channel
- Poor Echo Return Loss Enhancement (ERLE) for voice

The LPEC was designed to overcome these observed problems.

The LPEC (Figure 1) performs Linear Prediction Analysis on the Receive In signal, $x(k)$, to obtain the pre-whitened signal, $x_w(k)$. To obtain the optimum ERLE performance for LPEC, pre-emphasis and a Hamming window must be applied to obtain the LPC coefficients. However, the RELP Vocoder Analysis algorithm [1] performs the same operations as LPEC to provide a spectrally flattened or pre-whitened Residual signal for transmission to the Receiver. Therefore, the RELP Vocoder Synthesis algorithm can provide the required LPC coefficients and pre-whitened Receive In signal, $x_w(k)$, for the LPEC with no extra processing required. Figure 2 presents the structure for an integrated unit. This integrated structure is slightly sub-optimal due to the LPC coefficient and Residual signal quantization introduced by the RELP Vocoder. However, no significant performance penalty has been observed.

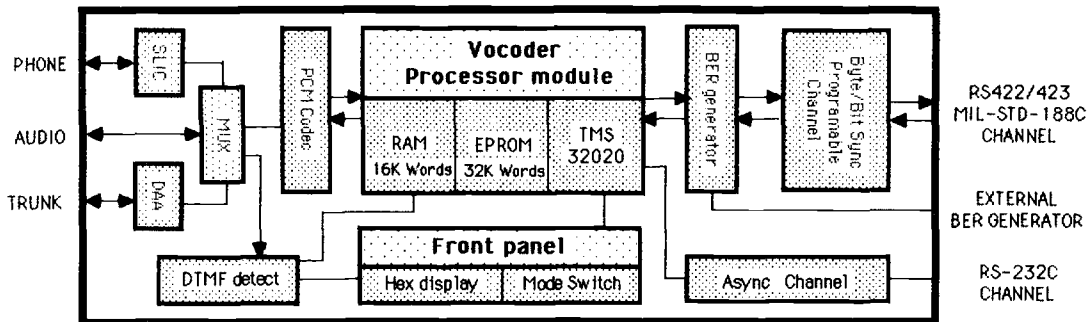


Figure 3. VDS-20 Functional Block Diagram

V. VDS-20

The integrated RELP/LPEC unit is implemented in the Vocoder Demonstration System (VDS-20), a second generation real-time test bed for voice processing algorithms utilizing a single TMS32020 processor. Figure 3 presents a functional block diagram of the VDS-20, which is based heavily on the VDS-10 [2]. The VDS-20 has the following key features:

- Single TMS32020 processor with 32K of external program memory and 16K of external data RAM
- Front panel mode switch for selecting voice algorithms
- Software controlled frame size and data rate
- Flexible serial transmission channels, including RS-422/3 and MIL-STD-188, operating in bit synchronous or packet mode
- Audio/telephone/trunk/4-wire E&M interfaces with level monitoring
- DTMF recognition with external display
- Bit error generation (internal and external)
- System diagnostics with external readout
- UART interface

Three echo cancellers have been implemented in the VDS-20:

- Classical Echo Canceller
- Linear Predictive Echo Canceller
- Integrated LPEC

Each structure employs 32 filter taps (4 ms) and is based on the dual-path model. Table 1 presents the processor resource usage for each of the three structures. The primary performance bound for real-time operation is execution speed; the percentage of real-time usage is listed for each module. Note that the integrated LPEC structure requires only slightly more execution time than the classical echo canceller while the standard LPEC requires significantly more processing. The penalty induced by the integrated unit is an increase in the data memory requirement. Current estimates are that the reduced processing load of the integrated LPEC will allow an increase in the echo path model to 80 taps, yielding cancellation of echoes up to 10 ms in duration with this same structure.

In addition, the integrated LPEC introduced no subjective degradation of voice quality to the RELP Vocoder.

Module	Classical	LPEC	Integrated
Input/Output	6 %	6 %	6 %
μ-law Conversion	6 %	6 %	6 %
RELP Analysis	22 %	22 %	22 %
RELP Synthesis	8 %	8 %	8 %
Classical EC	22 %	22 %	22 %
LPC Analysis	-	5 %	-
Inverse Filter	-	2x4 %	4 %
TOTAL	64 %	77 %	68 %
Program memory	3K	3K	3K
Data memory	2K	2K	2.3K

Table 1. VDS-20 Processor Resources

Filter taps	32
Echo dispersion	4 ms
20 db ERLE	250 ms
Steady state ERLE	40 dB
Convergence time	2 s

Table 2. Preliminary LPEC Performance Specifications

VI. ECHO CANCELLER PERFORMANCE

The three echo canceller structures were simulated in FORTRAN with finite arithmetic to model TMS320 operation, i.e., 16-bit data words and a 32-bit accumulator. A stationary echo path impulse response, $H(z)$, was used; this was measured using the VDS-20. A standard PCM speech test pattern (Figure 4) of "One, two, three,..." spoken by an English male, was used as the input to the RELP Vocoder. The RELP Vocoder output was used as input to the three structures under test. The test pattern was chosen deliberately to demonstrate the poor performance of the classical echo canceller. This occurs for several reasons:

- Not continuous speech
- Mainly voiced segments (high correlation)
- Low background noise

The performance of the three structures was characterized by two objective measures, ERLE and NORM, calculated once per speech frame:

$$ERLE = 10 \log_{10} \frac{\sum_{i=1}^{180} v_i^2}{\sum_{i=1}^{180} e_i^2} \quad NORM = 10 \log_{10} \frac{\|H\|^2}{\|H - h_e\|^2}$$

The Echo Return Loss Enhancement (ERLE) is the standard measure for echo canceller performance and gives a measure in dB of the level of the perceived echo below the signal; this enhancement is in addition to any loss provided by hardware impedance matching. NORM is a measure of the modelling of the echo path impulse response. It compares the given stationary impulse response, $H(z)$, to the echo canceller's transversal filter coefficients, $h_0(z)$, for each speech frame. When the input signal is white and stationary, $NORM = ERLE$ [4], so NORM is an indication of the ERLE performance if the adaptive filter coefficients are frozen with a white input signal.

Figures 5 and 6 present the ERLE and NORM measures for the PCM test pattern respectively. It can be seen that the classical echo canceller performs poorly. Clearly the LPEC performs very well, converging to 20 dB in 250 ms, with a final convergence better than 40dB ERLE. The classical echo canceller typically has a steady state ERLE of 25 dB for speech after convergence (not shown in diagram). In addition, the performance of the integrated LPEC is very similar to that of the LPEC. Table 2 presents the preliminary specification for the integrated LPEC.

VII. CONCLUSION

The Linear Predictive Echo Canceller modifies the classical echo canceller structure to pre-whiten the input signal and hence provides better convergence properties in the presence of highly correlated input signals such as speech.

An integrated LPEC structure is presented which shares common processing modules with the RELP Vocoder to allow significant execution speed savings, thereby reducing the LPEC processing requirements to those of the classical structure, while maintaining the performance advantages. An ERLE of 20dB is realized in 250 ms with a steady state ERLE exceeding 40dB. The current implementation handles echo dispersions of 4 ms, which is sufficient for a telephone subscriber local loop. Further work is in progress to extend this specification to 32 ms for more generalized applications.

The integrated LPEC has been implemented in the Vocoder Demonstration System (VDS-20) using a single TMS32020 processor and will be available at the conference for demonstrations.

REFERENCES

- [1] PJ Wilson et al., "Implementation of the RELP Vocoder using the TMS320", ICASSP-84.
- [2] PJ Wilson et al., "The Transmission of In-band Signaling for Medium Band Voice Codec Implementations", ICASSP-85.
- [3] MM Sondhi & DA Berkley, "Silencing Echoes on the Telephone Network", Proc IEEE, 1980, pp 948-63.
- [4] S Yamamoto et al., "An Adaptive Echo Canceller with Linear Predictor", Trans IECE Japan, 1979, pp 851-857.
- [5] K Ochiai et al., "Echo Canceller with Two Echo Path Models", IEEE Trans COM-25, 1977, pp 589-595.

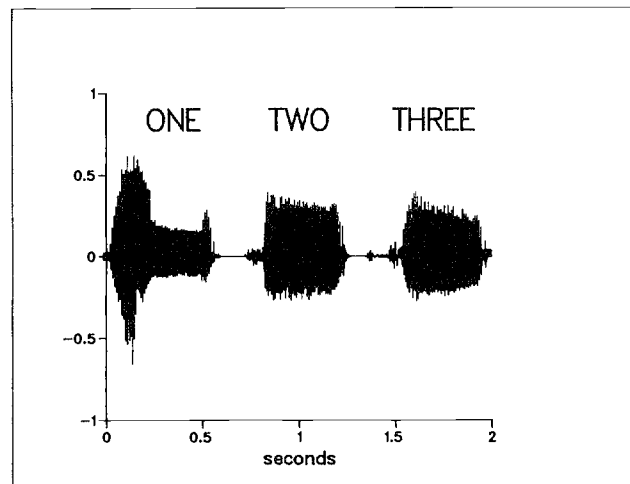


Figure 4. PCM Input Test Pattern

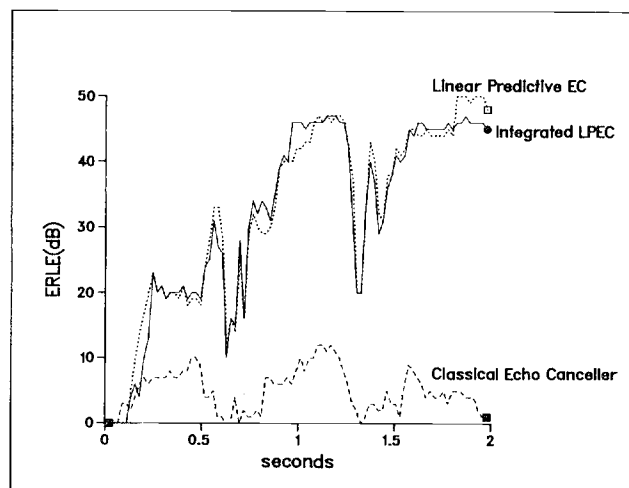


Figure 5. ERLE Measurement For PCM Test Pattern

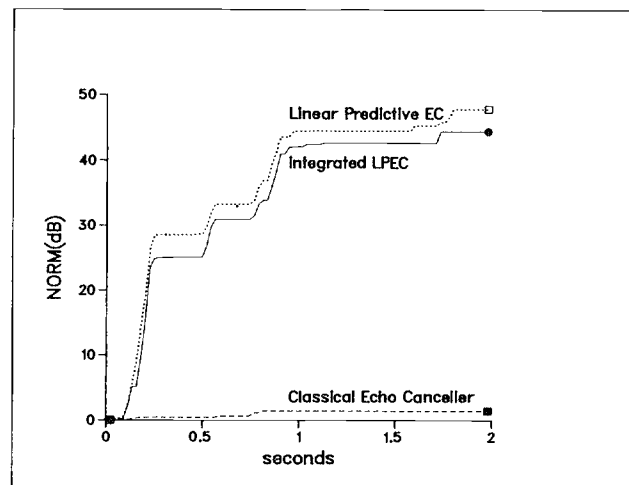


Figure 6. NORM Measurement For PCM Test Pattern