ECHO CANCELLATION IN TELEPHONY

BY ENGR. M. AJMAL KHAN
EE DEPARTMENT, KFUPM
Dammam, Saudi Arabia

ABSTRACT

Both wireless and long distance landline networks must contend with a variety of different factors that can adversely degrade the quality of voice communications. Network delays can cause echo and other distortions. In the wireless network, acoustic echo, background noise, and other additional disruptions must be dealt with in order to assure voice quality. Voice quality has traditionally been influenced by impairments resulting from the analog nature of older equipment installed in terrestrial networks and the relatively low grade sound quality resulting from speech compression techniques used in digital wireless networks. Although these impairments are gradually being eliminated by the introduction of new technologies, electrical and acoustic echo still remain possible sources to degrade voice signal quality.

This paper discusses the properties of acoustic echo and electrical echo (hybrid echo) in order to explain why different techniques are needed to control and eliminate each type of signal degradation. In addition, for some non-voice applications, superior echo canceller technology is considered to improve overall network quality.

ACOUSTIC AND ELECTRIC ECHO

Mainly two kinds of echoes namely the electric echo and the acoustic echoes exist in communication systems. The electric echo is also called hybrid echo or line echo. This echo can be found in the public-switched telephone network (PSTN), mobile, and IP phone systems. The electric echo is created at the hybrid connections which are created at the two-wire / four-wire PSTN conversion points as shown in Figure-1.

The electric echo can be generated from both the near end and the far end electric devices. The near end echo, hybrid echo, has been around almost since the advent of the telephone itself. Due to the economic reason, we use two-wire system to perform full duplex functions that actually require the performance of a four-wire system. The principle is to use different kinds of "hybrid" to balance and separate the sending and receiving signals. The real hybrid circuits can not be 100% ideal because of the leakage, and the parasitic or parametric deviations. Therefore, part of the signal takes the wrong path from both the near end hybrid and the far end hybrid and thus becomes echo. In the old telephone system, the echo is 28ms or less. In the modern telephone system, the electric echo will be longer. The electric echo in GSM could be up to 80ms. The electric echo in IP telephone could be up to 120ms or even longer.

Acoustic echo is generated from either hands-free telephone or a telephone with poor voice coupling between the earphone and the microphone. In a hands-free telephone, we have to let remote voice go through the

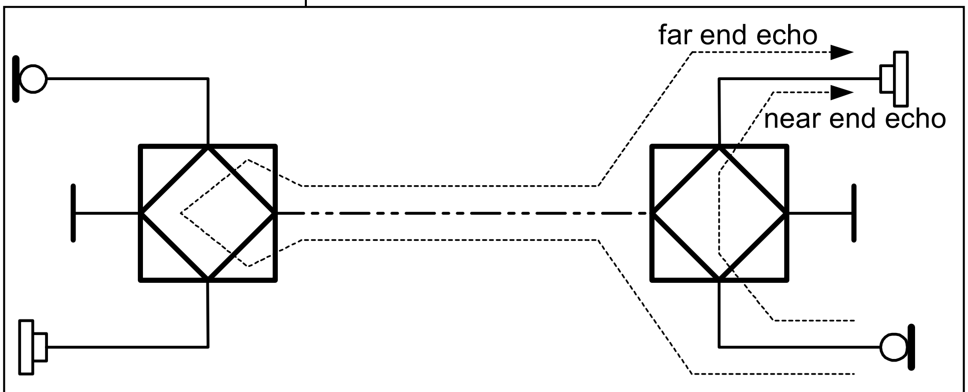

Figure 1: Hybrid Connections and the Resulting Electric Echo

loudspeaker and become part of the microphone signal. When subscriber B is calling subscriber A, and A uses a hands-free telephone, the voice from subscriber B is sent to the loudspeaker. The microphone of subscriber A picks up both the voice from the subscriber A and the loudspeaker voice from subscriber B. Thus, the subscriber B receives an acoustic echo. There are two different components making up this acoustic echo. The first is the direct coupling between the loudspeaker and the microphone, and the second is the undesired remote speech reflected from roof, windows, and walls, etc. The echo from the second component could be as long as 200ms.

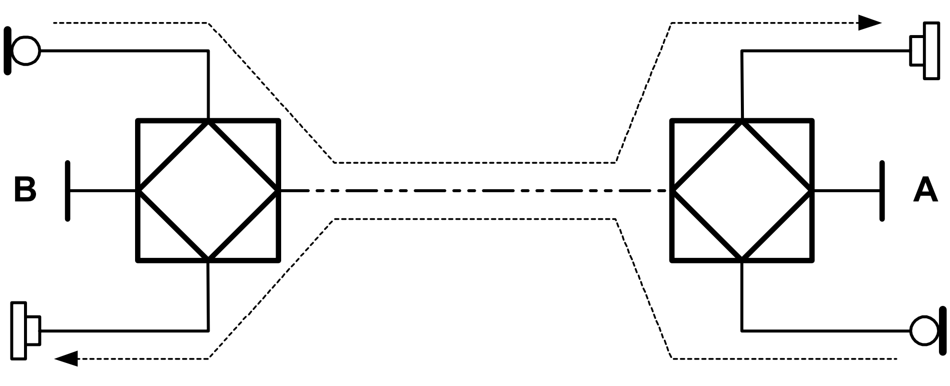

Figure 2: Acoustic Echo

Acoustic echo can be present in both wireline and wireless applications. Most wireline configurations that are exposed to echo conditions are equipped with echo cancellers; therefore wireline acoustic echo is controlled via standard echo cancellation algorithms and associated Non Linear Processing (NLP) techniques.

ECHO CANCELLATION IN TELEPHONES

A concept commonly called "echo cancellation" for echo control was invented at Bell Laboratories in 1964. An echo canceller is a device that analyzes the received speech signal, generates an algorithmic model of the "estimated echo" signal, and subtracts the "estimated echo" from the signal(s) that is returned via the long distance circuit. Figure 3 illustrates a typical Echo Canceller configuration for a long distance wireline application.

As shown in Figure 3, Echo Canceller 1 (EC-1) forms a replica of the expected echo signal that will be returned to subscriber "A". This is done by sampling "A's" speech signal (R_{in}), and passing the sample through a filter that matches the transfer characteristics of the "tail-end circuit" (i.e. the filter emulates the impedance and signal delay characteristics of the network equipment and hybrid connected to subscriber "B"). Subscriber "A's" echo replica is inverted, and then summed with the actual signal that appears at the S_{in} port of EC-1. The S_{in} signal may contain a combination of "A's" echo and "B's" speech (e.g. during double-talking), but only "A's" echo signal will be "cancelled" (i.e. "B's" speech will pass through EC-1 unimpaired and appear on signal S_{out}). In this example a second Echo Canceller (EC-2) is installed at the other end of the long distance circuit (i.e. adjacent to subscriber "A"). It should be understood that EC-2 performs the equivalent function as EC-1. That is, EC-2 samples "B's" speech signal, generates a replica of "B's" echo, and sums the inverted replica of "B's echo" with the actual signal that appears at the S_{in} port of EC-2, thereby canceling "B's" echo while allowing "A's" speech signal to pass through EC-2 unimpaired to appear at the S_{out} port.

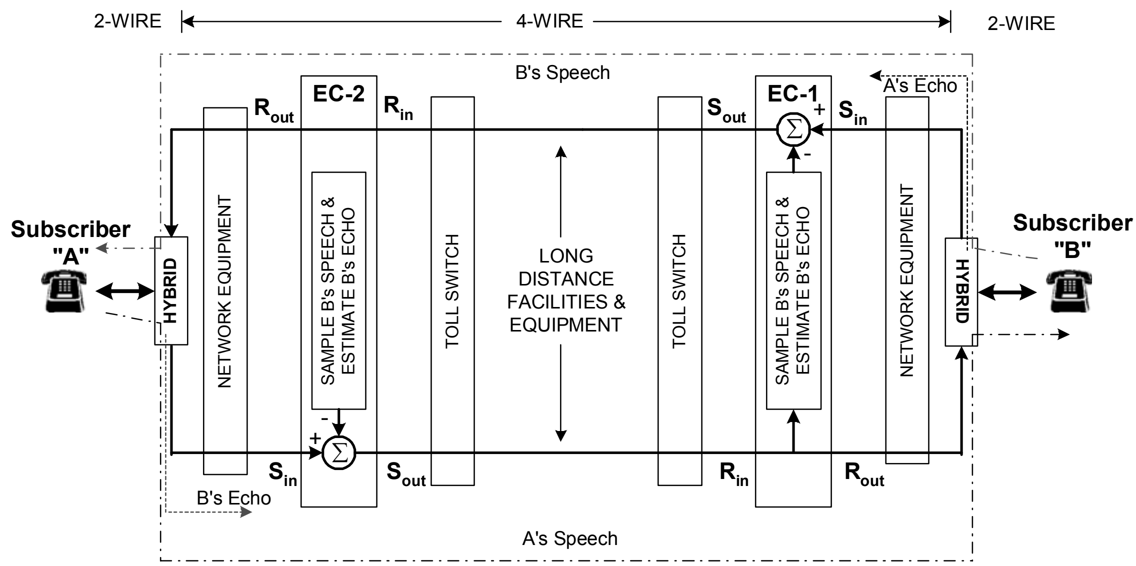

Figure 3: Echo Canceller

NON-LINEAR PROCESSOR (NLP)

The theory used to analyze echo cancellation is based on an assumption that the echo signal returned (S_{in}) can be represented as a linear expression of subscriber "A's" speech signal (R_{in}). In reality, subscriber "A's" speech signal may encounter Coder/Decoder (CODEC) equipment, a hybrid, loaded and non-loaded network sections, and several local switches before it reaches the S_{in} port of the Echo Canceller, as represented by the box labeled "network equipment" in Figure 4. As a result, the S_{in} signal is delayed, distorted, dispersed, decoded, re-coded, and may have had significant noise added to it. Fortunately, the non-linear and stochastic components in the S_{in} signal are relatively small, and a well-designed echo canceller can reduce the echo signal level by about 30 dB. However, when an echo signal is delayed more than several hundred milliseconds, it can be perceived at levels of -50dB or -60 dB, depending on subscriber "A's" ambient noise level. Therefore a Non-Linear Processor (NLP) is incorporated into most echo cancellers to handle the condition called "residual echo".

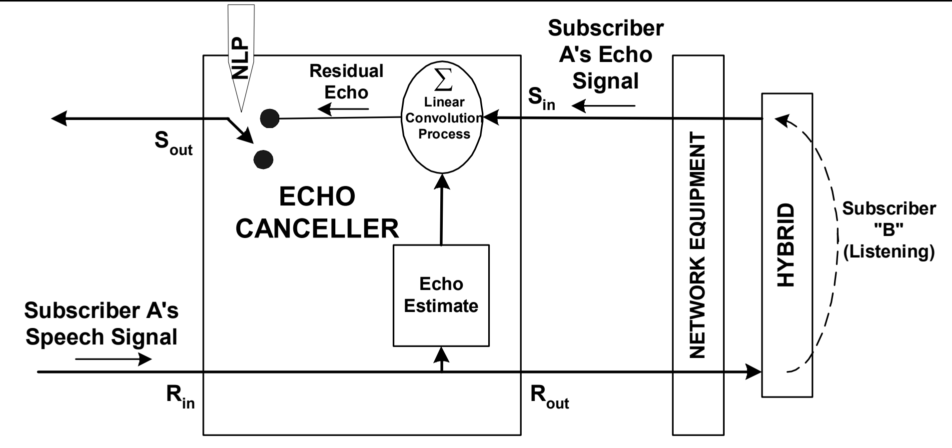

Figure 4: Non Linear Processor

ECHO CANCELLATION IN CELLULAR TELEPHONY

Typical cellular applications, such as Global System Mobile (GSM), Time Division Multiple Access (TDMA), and Code Division Multiple Access (CDMA) require echo control, even though subscribers may be physically located only a few hundred meters apart. This is because cellular networks utilize extensive processing to achieve signal compression, channel coding, and TDMA/CDMA frame interleaving. These functions are performed in the Base Station Controller (BSC), and typically introduce about 100 ms of processing time (delay) for each direction of transmission.

A typical mobile-to-PSTN (public switched telephone network) call configuration is illustrated in Figure 5. In this example, mobile subscriber "A" is speaking, and as shown "A's" echo (returned by the PSTN hybrid) is controlled by the Echo Canceller (EC) located on the PSTN side of the Mobile Switching Center (MSC). If not properly cancelled, "A's" echo would arrive at the mobile unit with a delay of approximately 200 ms, which would severely degrade voice quality.

It should be understood that in this example the PSTN subscriber "B" does not hear an echo because a hybrid does not exist in the mobile network equipment (i.e. there is no hybrid to reflect subscriber "B's" voice back from the mobile network connections). In wireless applications, the echo canceller equipment should ideally be located close to the PSTN hybrids (the origin of the echo), but because the number of Mobile-to-PSTN connections is lower than the number of hybrids in the PSTN, the Mobile Switching Center (MSC) location is a more economical solution. In addition, a call connected to a roaming wireless subscriber may be routed through multiple MSCs, and the propagation time between MSCs adds to the overall tail-end delay. Therefore, international telecommunications union (ITU) standards

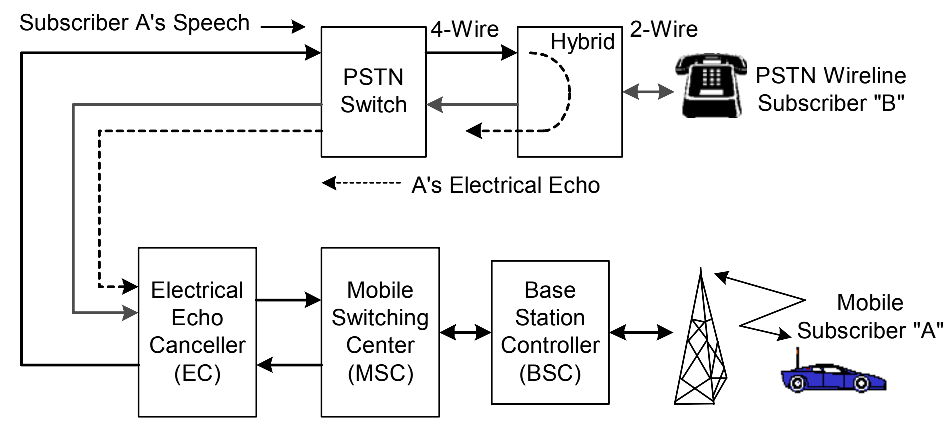

Figure 5: Typical Mobile to PSTN Application

recommend that echo cancellers used for wireless applications be capable of accommodating up to 64 ms tail-end delay.

In general, the voice quality of digital wireless applications is typically lower than PSTN applications. This is a result of wireless voice signal encoding techniques and the higher acoustic noise levels that often surround mobile subscribers. The popular use of hands-free mobile telephone sets in vehicles, coupled with high background noise levels (e.g. engine, wind, traffic sounds), increases the magnitude of this problem for many digital wireless network operators. Similarly, voice quality and performance issues frequently arise because of the wide range of environmental conditions associated with wireless phone usage (e.g. noisy public areas, outdoor conditions, inexpensive wireless phone sets). In attempting to improve speech quality and minimize performance differences between wireless and PSTN calls, echo canceller equipment is often confronted with the task of controlling acoustic echo in addition to traditional electrical echo. There are two fundamental reasons for why acoustic echo is more difficult to control than traditional electrical echo. First, acoustic echo tends to change its characteristics as the subscriber moves around within an acoustic space. For example, acoustic echo generated while using a hands-free speakerphone in a vehicle will change characteristics if the subscriber lifts his hands, tilts his head, or makes similar modest movements. An even a more dramatic change will occur if the subscriber opens a window, or the vehicle travels over a rough surface. This acoustic phenomenon is equivalent to a hybrid that has very significant dispersion characteristics, or when multiple hybrid replacements (e.g. adding or dropping conference call connections) occur within the duration of a single call. These conditions require continuous adaptation to changes that can be fast and/or drastically different. Second, the acoustic echo generated in a wireless environment passes through a non-linear path of speech vocoder equipment. Traditional echo cancellers use a form of the Normalized Least Squares algorithm (based on a variant of a Least Squares estimation) that performs reasonably well when the echo is a linear function of the original speech signal e.g. electrical echo. However, the performance of this type of algorithm deteriorates significantly when the echo is a non-linear function of the original signal e.g. acoustic echo. Therefore, non-linear methods must be used to control acoustic echo. However, additional steps should be taken when implementing non-linear functions to ensure potential side effects are minimized and the acoustic echo is properly distinguished from valid speech that should be allowed to pass through the echo canceller.

CONCLUSIONS

Echo control has gone through tremendous technological developments since its original inception. Continued trends in technology will support ever-increasing levels of integration and density. Echo cancellers are presently designed to process more than 600 channels (64 kbps DS0 or equivalent signals) on a single moderate size printed circuit board. Future echo cancellers will be equipped with specialized functions that further enhance voice quality and provide new network services.

Although higher levels of concentration are emerging, the majority of subscribers will still access telecommunications networks via two-wire local loop connections. Therefore, echo will be generated by the hybrids that terminate these local loops, and consequently echo canceller products will continue to be deployed into the foreseeable future.

*** *Date: 27th October 2002*

REFERENCES

[1] S. Haykin, Adaptive Filter Theory, 3rd ed., Prentice-Hall, 1996.

[2] X. Y. Gao, W. M. Snelgrove, "Adaptive Linearization of a Loudspeaker", ICASSP 1991, Vol. 3, pp. 3589-3592.

[3] M. E. Knappe, R. A. Goubran, "Steady State Performance Limitations of Full-Band Acoustic Echo Cancellers", ICASSP 1994, Adelaide, South Australia, Vol. 2, pp. 73-76.

[4] A. N. Birkett, R. A. Goubran, "Acoustic Echo Cancellation for Hands-free Telephony Using Neural Networks", Neural Network for Signal Processing 1994, IEEE. Workshop Proceedings, Sept. 194, pp. 249-258.

[5] H. F. Olsen, Acoustical Engineering, Toronto, D. Van Nostrand Company, Inc., 1964.

[6] A. N. Birkett, R. A. Goubran, "Limitations of Hands-free Acoustic Echo Cancellers due to Nonlinear Loudspeaker Distortion and Enclosure Vibration Effects", IEEE ASSP Workshop on Appi. Of Signal Processing to Aud. And Acoustics, New Paltz, New York, Oct. 1995.