ACOUSTIC ECHO AND NOISE CONTROL: WHERE DO WE COME FROM – WHERE DO WE GO?

Eberhard Hänsler

signaltheorie

Darmstadt University of Technology
Merckstrasse 25, Darmstadt, Germany
Eberhard.Haensler@nt.tu-darmstadt.de

ABSTRACT

In this paper we describe a short history of acoustic echo and noise control systems and of the International Workshop on Acoustic Echo and Noise Control. We also try to look into the future.

1. WHERE DO WE COME FROM?

Acoustic echo and noise control units as used in hands-free communication systems are comprised of three subunits: A loss control circuit, an adaptive filter parallel to the loudspeaker–enclosure–microphone system (LEMS) and also an adaptive filter within the path of the output signal (see Fig.1). Their functions are obvious: The loss control circuit attenuates the input and/or output signal such that the communication loop remains stable. In addition, echoes caused by the acoustic transfer path from loudspeaker to microphone do not impede a communication between two users. The adaptive filter arranged in parallel to the LEMS is able to cancel echoes according to the degree to which it is matched to the LEMS. The filter within the path of the output signal is used to attenuate remaining echoes and background noise. In the early days of acoustic echo and noise control a socalled center clipper – a nonlinear circuit – took the place of this filter.

Of these units, the loss control circuit has the longest history in hands-free communication systems. In its most simple form it reduces the usually full-duplex communication system to a half-duplex one by simply switching input and output lines on and off alternately. Besides preventing howling and suppressing echoes, any natural conversation was prevented, too. A device for hands-free telephone conversation using voice switching was presented as recently as in 1957 [1]. The introduction of a center clipper in 1974 [2] meant a noticeable improvement. Laboratory experiments applying an adaptive filter for acoustic echo suppression were reported in 1975 [3].

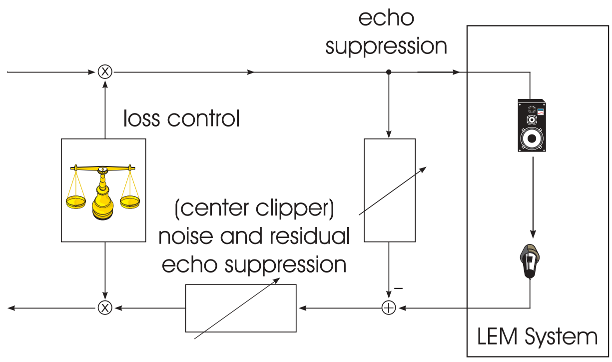

Figure 1: General structure of acoustic echo and noise control systems

The series of biennial international workshops on acoustic echo and noise control started in 1989 in Berlin. Even at that time, hands-free telephone conversations were still made possible by the help of loss control circuits only. The reason for this fact was very simple: the lack of processing power at an affordable cost.

The workshop series was continued in several European countries and made its way to the USA in 1999.

The main interest at the first workshop was focused on adaptive algorithms for acoustic echo cancelling. On the one hand, it was quite clear that the comfort of hands-free communication units would have to be improved. On the other hand, the already foreseeable increase of available processing power would enable the implementation of high-order adaptive filters in the near future. First laboratory implementations of filters using the Least Mean Square Algorithm clearly indicated that this algorithm without any additional support could not solve the problem. Therefore, complexity reduction and stability improvement of the Recursive Least Square Algorithm was one of the main discussion issues. It seems worthwhile to mention that the class of Fast Newton Filters was introduced at the first workshop.

In addition, an astonishing variety of topics that are still under discussion today were presented at the 1989 workshop: Echo cancelling in the frequency domain, subband systems, step-size control, adaptive antennas, noise reduction based on microphone arrays and post filtering, nonlinear techniques for echo cancelling, problems in measuring, etc. It seems strange that it took three more workshops before the problem of stereophonic echo cancelling appeared on the agenda in 1995. Knowing just the titles of the early workshops one may ask whether Ben Joseph Akiba with his "We have seen it all, nothing ever changes" was right. The answer is clear: "Definitely not!" Tremendous progress has been made on all those topics. This is documented in a multitude of papers. An attempt to reference even only a few of them would blow the scope of this contribution. Therefore, only one overview paper and two books published recently are cited here [4, 5, 6]. Each one contains a large number of references to all topics in acoustic echo and noise control.

2. WHERE ARE WE NOW?

Powerful and affordable acoustic echo and noise control units are available now. Their performance is satisfactory, especially if compared to solutions in other voice processing areas like speech recognition or speech to text translation. The fact that echo and noise control systems have not yet entered the market on a large scale seems not to be a technical but a marketing problem: A customer who buys a high quality echo and noise suppression system pays for the comfort of his communication partner. If she/he decides to economize by using a very simple system the partner at the far end is typically too polite to tell her/him about the poor speech quality the system is producing. "Be nice to your communication partner" might be an effective advertising slogan.

3. WHERE DO WE GO?

Future research and development in the area of acoustic echo and noise control certainly will be governed by the lack of processing power restrictions. This has a number of consequences:

- It will no longer be necessary to program in assembler language. Even if the efficiency of future compilers will not be remarkably increased, future hardware will be able to run the procedures for echo and noise control in real-time.
- Implementation of even sophisticated procedures on ordinary (office) PCs will be possible. This will make it easier to test modifications of existing procedures or of completely new ideas in real-time and in real environments.
- The performance of future systems will approach theoretical limits given by the environment they have to work in.
- This does not necessarily mean that future systems will be perfectly reliable in all situations. The reliability of estimation procedures used to detect system states like a system change or a double-talk situation depends on the length of the usable data record. Since, however, the working environment is highly time-varying and nonstationary the usage of too long records can cause the loss of the real-time capability.

As a result the performance of future acoustic echo and noise control systems will no longer be limited by the restricted capabilities of affordable hardware. It will depend on the quality of the algorithms implemented.

Up to now the NLMS Algorithm plays the role of the "working horse" for acoustic echo and noise control. The Affine Projection Algorithm offers improved performance at modest additional implementation and processing cost without causing stability problems that are difficult to solve. Rules for step-size control used for the NLMS Algorithm, however, have to be reconsidered. Already under reconsideration are changes of the cost function for the filter adaptation. The minimum mean square error leading to a quadratic error surface will give way to more complex functions that are better matched to the optimization problem.

Hesitancy in dealing with multi-microphone and multi-loudspeaker systems is disappearing, as the number of papers at this workshop illustrates. As indicated, the necessary processing power will be available and with more and more electronic devices being implemented anyhow, the reluctance to wire such systems will also disappear.

Besides increased processing speed and memory size of digital signal processing hardware, the reduction of energy consumption will be an important development. This enables the use of algorithms that demand high processing speed on battery powered systems like mobile telephones or hearing aids.

In general acoustic echo and noise control units will not just remain part of telephones with hands-free capability. They will become features of many consumer products like speech recognition systems, speaker identification systems, hearing aids, video and audio conference systems, computer games, public address systems, car communication systems, voice control units, etc. The combination with other systems may trigger synergy effects so that some functions can be utilized by several units.

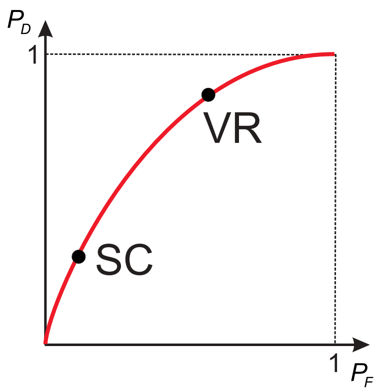

Figure 2: Operating points for step-size control (SC) and voice recognition (VR) on a Detector Operating Characteristic

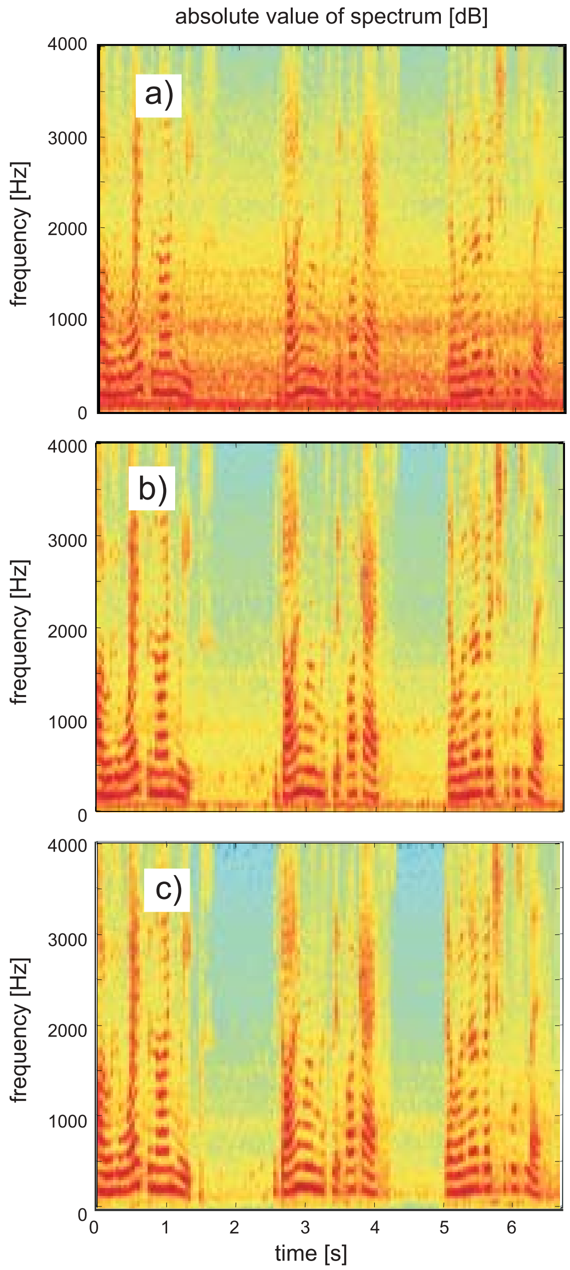

Figure 3: Reducing noise components in between the harmonics of the pitch of voiced speech segments: Time-frequency diagram of a) noisy speech, b) of noisy speech after spectral reduction, and c) after spectral subtraction and "pitch adaptive" filtering

The following example, however, shows that this is not always the case: A voice activity detector in a speech recognition system is tuned for a high conditional detection probability $P_D = P(\text{speech}|\text{speech})$. In contrast, a voice activity detector in an step-size control unit is optimized for a low conditional false alarm probability $P_F = P(\text{speech}|\text{noise})$ (see Fig. 2).

In order to enhance the performance of future echo and noise control systems the estimation of signals or states of (sub-)systems will be complemented by the estimation parameters of models based on a priori knowledge about speech and noise sources and probabilities of user behavior. If, for example, one is able to estimate the pitch frequency of voiced speech segments, the performance of noise reduction by spectral subtraction may be improved by filtering the noisy speech signal in such a way that the noise in between harmonics of the pitch is reduced (see Fig. 3) [7].

Furthermore, wherever it is feasible, one should acquire additional information. In modern cars the exact number of revolutions per minute of the engine is available on a bus system. Noise reduction can use this information to suppress noise caused by the engine (see Fig. 4) [8].

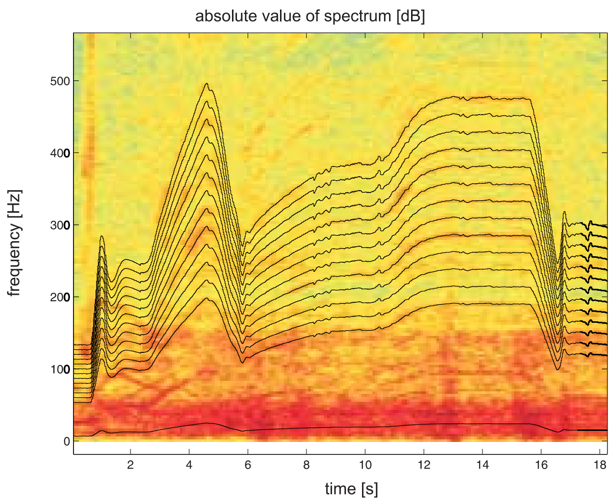

Figure 4: Time-frequency diagram of car noise with harmonics due to the engine marked in black

Today's systems have to be "hand-tuned" according to the specific application they are intended to be used in. With the processing power available "self-configuring" systems will become feasible. When put into a new environment they will automatically determine, for example, the optimal order of the adaptive filters, the number of subbands and the level of thresholds such as the spectral floor for noise reduction, only to mention a few.

4. CONCLUSION

Customer demands are time-variant. Using available systems, customers will certainly ask for better performance. Therefore, the need for new and better ideas will remain. Acoustic echo and noise control will continue to be one of the most interesting problems in digital signal processing.

5. REFERENCES

[1] W. F. Clemency, F. F. Romanow, A.F. Rose: The Bell System Speakerphone. A.I.E.E. Trans., vol. 76, Part I, 148ff, 1957.

[2] D. A. Berkley, O. M. M. Mitchell: Seeking the Ideal in "Hands-Free" Telephony. Bell Laboratories Record, vol. 52, 318–325, 1974.

[3] G. Pays, J. M. Person: Modèle de Laboratoire d'un Poste Téléphonique à Haut–parleur. F.A.S.E. 75, 88–102, Paris, France, 1975.

[4] Ch. Breining, P. Dreiseitel, E. Hänsler, A. Mader, B. Nitsch, H. Puder, Th. Schertler, G. Schmidt, J. Tilp: Acoustic Echo Control: An Application of Very-High-Order Adaptive Filters. IEEE Signal Processing Magazine, vol. 16, 42–69, July 1999.

[5] S. L. Gay, J. Benesty (Ed.): Acoustic Signal Processing for Telecommunication. Kluwer Academic Publishers, Boston, 2000.

[6] J. Benesty, T. Gänsler, D. R. Morgan, M. M. Sondhi, S. L. Gay: Advances in Network and Acoustic Echo Cancellation. Springer, Berlin, 2001.

[7] J. Tilp: Single-channel Noise Reduction with Pitch-adaptive Post-filtering. Proc. EUSIPCO-2000, vol. 1, 171–174, Tampere, Finland, 2000.

[8] H. Puder, F. Steffens: Improved Noise Reduction for Hands-Free Car Phones Utilizing Information on Vehicle and Engine Speeds. Proc. EUSIPCO-2000, vol. 3, 1851–1854, Tampere, Finland, 2000.