FROM ALGORITHMS TO SYSTEMS –

IT'S A ROCKY ROAD

Eberhard Hänsler

Signal Theory
Darmstadt University of Technology
haensler@nesi.e-technik.th-darmstadt.de

ABSTRACT

This paper emphasizes the efforts which are necessary to move from algorithms for adaptive compensation of acoustical echoes towards complete hands–free communication systems. It is explained that an adaptive echo compensator is just one – and not even the most important – building block of such a system.

1. REVIEW

While preparing this contribution, a look back to the program of the first International Workshop on Acoustic Echo Control held in Berlin in 1989 seemed appropriate. Astonishingly enough, topics discussed there are still of interest today. Nevertheless, excellent progress has been made since then and answers to a number of questions have been found. The following are a few comments on the most important ones:

1.1 NLMS versus RLS

The superior convergence performance of the Recursive Least Squares (RLS) Algorithm as compared to the Normalized Least Mean Square (NLMS) Algorithm made the RLS Algorithm a desirable tool for the adaptation of a filter for acoustic echo compensation. This statement seemed valid in spite of stability problems and increased numerical complexity. As for numerical stability, these problems are solved [1]. The same is true with respect to complexity. The "Fast Newton Algorithm" [2] exhibits a complexity comparable to the NLMS Algorithm [3]. On the other hand, RLS Algorithms do not offer substantial advantages with respect to tracking system changes. However, for adaptive filters used to compensate acoustical echoes fast tracking is more important than fast initial convergence. Therefore, the use of the NLMS Algorithm in acoustic echo compensators is widely accepted now.

1.2 FIR versus IIR

The echo compensating filter has to model the impulse response of the loudspeaker–enclosure–microphone system (LEMS). Depending on the properties of the enclosure, this function may be several hundreds of milliseconds long resulting in up to several thousands of non–zero sample values. At first glance, using a recursive (IIR) filter seemed, therefore, appropriate. Besides, matching a filter to the (time varying) transfer characteristic of a realistic enclosure such as an office or the interior of a passenger car requires a large number of adjustable parameters. Several investigations have shown that IIR filters need approximately the same number of parameters as FIR filters to achieve a comparable matching [4, 5, 6, 7]. Also considering that the compensation filter has to be adapted, the use of an FIR filter is the only realistic approach.

1.3 Time Domain versus Frequency Domain

The use of frequency domain techniques for acoustic echo compensation can offer a considerable reduction in signal processing complexity. Taking into account the extremely high degree of the (time domain) adaptive filter, complexity reduction is highly desirable. On the other hand, the introduction of time delay is inherent to the use of FFT's and block processing methods. It depends on the application whether the user accepts or the ITU–T–Recommendations allow additional front end delays. Over the last few years the tradeoff between delay and signal processing demands became well understood. Solutions are available that allow to tailor a system with respect to signal delay and processing complexity for specific applications [9, 10].

2. FROM ALGORITHMS ....

Typically about one hundred papers per year are written on acoustic echo and noise control. The majority deals with algorithms for adapting the echo compensating filter. A number of methods have been discussed to overcome the well known weakness of the NLMS Algorithm in case of correlated input signals. Various forms of fixed or adaptive decorrelation filters have been proposed [11, 12]. Individual step sizes matched to the special properties of LEMSs for the coefficients updates can be used to improve convergence [13, 14]. The combination of acoustic echo compensation with noise reduction [15] and the extension to stereo systems [16] are current topics in the literature. Noise reduction is desirable especially in mobile telephones.

Obviously, the area of adaptive algorithms is very well covered whereas the research (and development (?)) community seems to hesitate dealing with the nitty–gritties of complete hands–free telephone systems.

3. .... TO SYSTEMS –

The ITU–T Recommendations [8] put very stringent conditions on hands–free telephone systems. For "ordinary" telephones the echo attenuation has to be at least 45 dB in the case of single talk. In double talk situations this value can be reduced by 15 dB. Beyond that, only a negligible delay may be introduced into the signal path by the hands–free facility.

Despite of all the results gained by the analysis and simulation of "real systems", no adaptive filter can assure to provide the echo attenuation required over the entire duration of a call in all situations that may occur. Therefore, the adaptive echo compensating filter is only one component of a hands–free telephone system and it is not even the most important one! The system has to guarantee the required echo attenuation even in situations where the adaptive filter exhibits extremely poor performance. As a matter of fact, a hands–free system has to be designed to operate without echo compensation. Its comfort to users strongly depends on the "auxiliary" functions necessary to provide reliable operation.

Figure 1 shows a hands–free telephone system that can be connected to an analog (two–wire) line. Its main units are an echo compensator for (electrical) hybrid echoes, an adaptive loss control, and an adaptive compensator for acoustical echoes. No hybrid and therefore no compensator for hybrid echoes are necessary in case of digital (four–wire) transmission.

Each of these central units itself comprises of a number of subsystems not shown. The echo compensators require highly sophisticated control procedures based on detectors for far– and/or near–end talker activities, environmental noise, changes of the enclosure properties, etc. [18]. Signals generated by the near–end talker and by near–end noise sources can misadjust the compensators for acoustical echoes within a few sampling intervals. Therefore, the step size control of the filter adaptation has to react rapidly. Since fast estimates of system states such as single or double talk or system change are likely to be unreliable, at least a two stage control is needed: Assuming the worst case in a first step and correcting the control parameters after reliable estimates are available.

The purpose of the adaptive loss control is to guarantee the overall echo attenuation according to customer or ITU–T requirements. In order to reduce the loss inserted to the smallest value necessary, the attenuation effected by (user adjustable) volume controls, the acoustical echo path, and – not to forget – the adaptive echo compensator has to be estimated.

A number of peripheral units have to support the above mentioned components. High pass filters are needed to remove voltage offsets of (cheap) A/D converters. Volume controls – some of them automatic – have to make sure that A/D converters do not over– or underflow. Furthermore, the acoustic echo compensator is based on a linear model of the electro–acoustic transfer system. Limiters at appropriate locations are necessary to prevent nonlinearities caused by overflows. Finally, a center-clipper may be desirable to remove residual echoes from the output signal.

Most important for the development of such a system is the proper tuning and coupling of all subsystems. Experiences gained during the development and implementation of several systems proved that the use of a high performance adaptive echo compensator is a necessary but not a sufficient condition for a hands–free system offering pleasant double talk properties.

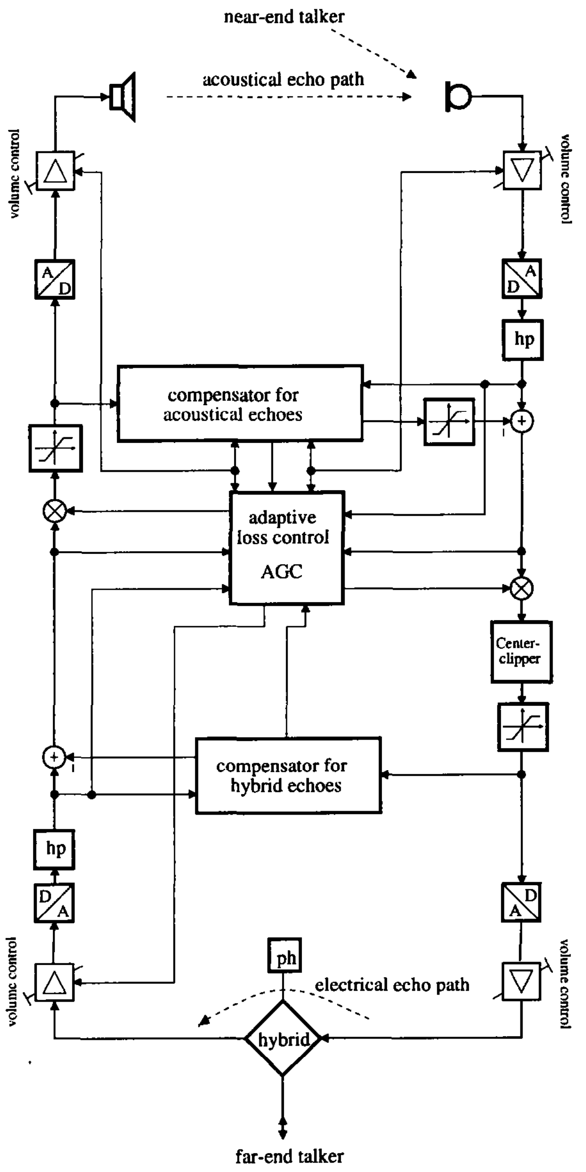

Figure 1: Hands–free telephone system [17]

4. IT'S A ROCKY ROAD

Introducing the hands–free facility to standard consumer telephones means that only low cost hardware should be used. It seems a common agreement that only 16 bit fixed–point arithmetic can be applied. Even with 24 bit signal processors already on the market, the increased chip area leading to higher costs seems to be prohibitive. In addition, only a very restricted number of processing cycles is available for the hands–free function. Similarly severe limitations apply to the amount of usable RAM. Therefore, implementing the hands–free function for consumer mass products is like cooking a first class gourmet dinner using only low–priced ingredients.

A word length of 16 bit seems sufficient for handling speech samples. However, the adaptive filter for compensating acoustic echoes suffers severe performance degradations [19]. This is due to the fact that the impulse response of a LEMS decays exponentially. Furthermore, updates of higher order coefficients may be lost permanently over time. Also products of signal samples by filter coefficients may be omitted during accumulation of the filter output signal. Applying double precision does not cure this problem since it increases the number of processing cycles and RAM cells necessary. Instead, splitting the compensating filter into segments and scaling up the filter coefficients properly can overcome the problem caused by 16 bit word length.

Two systems using adaptive echo compensation have been implemented recently based on 16 bit fixed–point hardware: A fullband system [20] and a subband system. The fullband system takes approximately 35 MIPS and 2 k words of RAM. The subband system requires about 20 MIPS and 3 k words of RAM. The reduction of processing cycles necessary for the subband system has to be paid for by increased memory requirements and an overall signal delay of about 40 ms. The fraction of processing cycles used by the adaptive compensator for acoustical echoes is less than 50 % in both systems.

5. OUTLOOK

Voice communication systems with hands–free facilities are on the market. More and more systems providing improved double–talk capability will be offered. While trying to predict future developments two aspects seem important: Being aware of the world wide efforts towards solutions of the acoustical echo compensating problem over more than one decade, it seems unlikely that a real epoch–making new concept will show up. On the other hand, foreseeable progress in integrated circuits will ease the hardware restrictions. Certainly, this will stimulate developers to present increasingly better implementations.

REFERENCES

[1] Gilloire, A.; Moulines, E.; Slock, D.; Duhamel, P.: State of the art in acoustic echo cancellation. Figueiras-Vidal, A. R. (Ed.): Digital Signal Processing in Telecommunications. Springer, London, U.K., 1996, p. 45–91.

[2] Moustakides, G.,Théodoridis, S.: Fast Newton Transversal Filter – A new class of adaptive estimation algorithms. IEEE Trans. on Signal Processing 39 (1991), p. 2184–2193.

[3] Pétillon, T.; Gilloire, A.; Théodoridis, S.: A comparative study of efficient transversal algorithms for acoustic echo cancellation. Proc. EUSIPCO-92 (1992), Brussels, Belgium, p. 119–122.

[4] Hänsler, E.; Schultheiß, U.: On-line identification of a loudspeaker-enclosure-microphone-system for compensating acoustical echoes. Proc. IFAC Identification and System Parameter Estmation 1985, York, U.K., 1985, p. 1841–1846.

[5] v. Zitzewitz, A.: Considerations on acoustic echo cancelling based on realtime experiments. Proc. EUSIPCO-90 (1990), Barcelona, Spain, p. 1987–1990.

[6] Mboup, M.; Bonnet, M.: On the adequatness of IIR adaptive filtering for acoustic echo cancellation. Proc. EUSIPCO-92 (1992), Brussels, Belgium, p. 111–114.

[7] Liavas, A. P.; Regalia, P. A: Acoustic echo cancellation: Do IIR filters offer better modelling capabilities than their FIR counterparts? To be published in IEEE Trans. on Signal Processing, 1997.

[8] International Telecommunication Union: General characteristics of international telephone connections and international telephone circuits – acoustic echo controllers. ITU–T Recommendation G.167, 1993.

[9] Egelmeers, G. P. M.: Real time Realization Concepts of large adaptive Filters. Ph.D. Thesis, Technical University Eindhoven, Eindhoven, The Netherlands, 1995.

[10] Estermann, P. G.: Adaptive Filter im Frequenzbereich: Analyse und Entwurfsstrategie. Ph.D. Thesis, Eigenössische Technische Hochschule Zürich, Zürich, Switzerland, 1996.

[11] Antweiler, Ch.; Schmitz, A.: Acoustic Echo Control Combined with Two Orthogonalizing Techniques. Proc. EUSIPCO-94, Edinburgh, Scotland, U.K., 1994, 1748–1751.

[12] Burger, Th.; Schultheiß, U.: A robust acoustic echo canceller for a hands–free voice–controlled telecommunication terminal. Proc. of the Third European Conf. on Speech Communication and Technology (1993), Berlin, Germany, p. 1809–1812.

[13] Antweiler, Ch.; Grunwald, J.; Quack, H.: Approximation of optimal step size control for acoustic echo cancellation. Proc. ICASSP 97 (1997), München, Germany, p. 295–298.

[14] Makino, S.; Kaneda, Y.; Koizumi, N.: Exponentially weighted stepsize NLMS adaptive filter based on the statistics of a room impulse response. IEEE Trans. on Speech & Audio Processing 1 (1993), p. 101–108.

[15] Martin, R.: Algorithms for hands–free voice communication in noisy environments. Proc. 9. Aachener Kolloquium "Signaltheorie" (1997), Institut für Technische Elektronik der RWTH, Aachen, Germany, p. 233–238.

[16] Sondhi, M. M.; Morgan, D. R.; Hall, J. L.: Stereophonic acoustic echo cancellation – An overview of the fundamental problem. IEEE Signal Processing Letters 2 (1995) 8, p. 148–151.

[17] Burger, Th.: Personal communication.

[18] Breining, Ch.: Control of a hands–free telephone set. To be published in Signal Processing 61 (1997).

[19] Breining, Ch.; Hänsler, E.; Schertler, Th.: Freisprechen – Die Jagd nach akustischen Echos. To be published in Frequenz 51 (1997).

[20] Schertler, Th.; Schmidt, G. U.: Impementation of a low cost acoustic echo canceller. Proc. IWAENC–97, London, U.K., 1997.