ELECTROACOUSTIC CHARACTERIZATION OF HEARING AIDS: A SYSTEM IDENTIFICATION APPROACH

Todd Schneider and Donald G. Jamieson

Hearing Health Care Research Unit
Dept. of Communicative Disorders
University of Western Ontario
London, Ontario N6G 1H1
Canada

ABSTRACT

The accurate electroacoustic characterization of hearing aids is important for the design, assessment and fitting of these devices. With the prevalence of modern adaptive processing strategies (e.g., level-dependent frequency response, multi-band compression etc.) it has become increasingly important to evaluate hearing aids using test stimuli that are representative of the signals a hearing aid will be expected to process (e.g., speech). Nearly all current hearing aid tests use stationary test signals that can characterize only the steady-state performance of a hearing aid. Our research examines the characteristics of automatic signal processing hearing aids with natural-speech input signals that may cause the hearing aid response to time-vary. We have investigated a number of linear system identification techniques that can be used to develop time-varying models of hearing aids. Using these models, we can begin to characterize performance of hearing aids with real-world signals and explore speech-based transient distortion measures.

1. INTRODUCTION

A number of measurements, such as maximum power output, frequency response and distortion are typically used to characterize hearing aids. The frequency response (gain versus frequency) of a hearing aid is an important measure of performance because it allows predictions concerning the level of acoustic signal provided for a hearing aid user. Current test procedures use pure-tone [1] or stationary, broadband stimuli [2, 3] to measure frequency responses. While some broadband stimuli have speech-like characteristics (e.g., long-term average spectrum or crest-factor), both pure-tone and broadband stimuli lack important characteristics found in natural speech (e.g., non-stationarity). These test signals may, therefore, give results that do not reflect the performance of an automatic signal processing hearing aid with a natural speech input signal.

Because the transient performance of an automatic signal processing hearing aid can be affected by the input signal it receives, another measurement of importance is distortion. Most current distortion tests may not give results that reflect "real-world" performance because they present a single test frequency at a time (e.g., total harmonic distortion [1]) or multiple, pure-tone test frequencies (e.g., intermodulation distortion) that are very poor approximations to natural speech. Newer tests using broadband stimuli that are better approximations to speech have also been developed [4, 5]. However, these tests use stationary, broadband test signals which allow only steady-state distortion measurements.

The problem of linear system identification has been studied extensively in the past (e.g., [6]). Rabiner, Crochiere and Allen [7] compared least squares analysis (LSA), least-mean squares adaptation and short-time spectral analysis methods for finite impulse response (moving average) parameter estimation and found that the LSA provided the most robust estimates of system parameters for both white and band-limited signals across a wide range of signal-to-noise ratios. Williamson, Cummins and Hecox [8] used the LSA method for hearing aid modelling. They applied a low-order MA filter to model the hearing aid and plotted three-dimensional spectrograms showing time, frequency and amplitude. Our research extends this work by applying more powerful MA system identification algorithms. We have also applied ARMA system identification techniques which can give equivalent results to MA models with much lower order models.

This paper describes research done to develop tests that use natural-speech input stimuli for the electroacoustic characterization of hearing aids. Linear system identification techniques are used to develop time-varying models of hearing aids and track characteristics over time. Five system identification techniques are compared to determine which provides the best performance for this application. These techniques provide the means to make speech-based frequency response and distortion measurements (Figure 1).

2. PROBLEM FORMULATION

2.1. System identification

The problem of system identification can be stated as follows: given the output from a device $\{y(n), n = 0 \ldots N-1\}$ that was produced when an input signal $\{x(n), n = 0 \ldots N-1\}$ was applied to the device, generate a model

This work was supported by NSERC, ORTC and Etymonic Design Inc.

0-7803-2431-5/95 $4.00 © 1995 IEEE

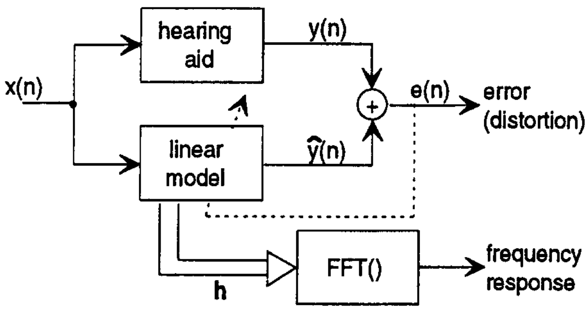

Figure 1: Speech-based hearing aid measurements

that will produce (with minimal error) an estimate $\hat{y}(n)$ of the output. The problem is invariably complicated by noise and distortion that are always present to some degree in the output signal $y(n)$. Although it is possible to derive non-linear models, we will restrict this discussion to linear models. That is, models where $\hat{y}(n)$ is a linear combination of samples of $x(n)$ (moving average or MA models), samples of $x(n)$ and $y(n)$ or samples of $x(n)$ and $\hat{y}(n)$ (auto-regressive moving average or ARMA models).

For linear modelling, $\hat{y}(n)$ is computed using one of the following methods:
(i) moving average (MA)

$$\hat{y}(n) = \sum_{j=0}^{J-1} a_j x(n-j) \quad (1)$$

(ii) equation-error ARMA

$$\hat{y}(n) = \sum_{j=0}^{J-1} a_j x(n-j) + \sum_{k=1}^{K-1} b_k y(n-k) \quad (2)$$

(iii) output-error ARMA

$$\hat{y}(n) = \sum_{j=0}^{J-1} a_j x(n-j) + \sum_{k=1}^{K-1} b_k \hat{y}(n-k) \quad (3)$$

For the MA formulation, define

$$\mathbf{h}^t = [a_0 a_1 \ldots a_{J-1}]$$

and

$$\mathbf{z}^t(n) = [x(n)x(n-1)\ldots x(n-J+1)]$$

For the equation-error ARMA formulation, define

$$\mathbf{h}^t = [a_0 a_1 \ldots a_{J-1} b_1 b_2 \ldots b_{K-1}]$$

and

$$\mathbf{z}^t(n) = [x(n)x(n-1)\ldots x(n-J+1) \\ y(n-1)y(n-2)\ldots y(n-K+1)]$$

We use the least squares method to solve this problem: minimize E with respect to the model coefficients ($\mathbf{h}$) where

$$E = \sum_{n=0}^{N-1} (y(n) - \hat{y}(n))^2 \quad (4)$$

It can be shown that for the MA and equation-error formulations, the solution is

$$\mathbf{h} = \mathbf{R}_{zz}^{-1} \mathbf{r}_{yz} \quad (5)$$

where

$$\mathbf{R}_{zz} = \sum_{n=0}^{N-1} \mathbf{z}(n)\mathbf{z}^t(n) \quad (6)$$

and

$$\mathbf{r}_{yz} = \sum_{n=0}^{N-1} y(n)\mathbf{z}(n) \quad (7)$$

The output-error ARMA formulation is much more difficult to solve because the output signal estimate $\hat{y}(n)$ depends on delayed estimates $\hat{y}(n-k)$; specialized techniques are required [9].

Both the MA and equation-error ARMA solutions are biased by output noise and distortion. However, both these methods require less computation and are much easier to implement than the output-error ARMA method. Also, ARMA methods can give equivalent results to MA models with much lower orders when modelling systems (like hearing aids) that are inherently ARMA.

2.2. Hearing aid modelling

We want to extract the "underlying" linear response of a hearing aid that may be distorting and/or have relatively high-levels of output noise. To accomplish this goal, some assumptions regarding the characteristics of automatic signal processing hearing aids must be made:

1. Hearing aid output noise is uncorrelated with the input or output signal.
2. A distorting hearing aid has a slowly-varying "underlying" linear response.
3. On average, the linear response of a hearing aid will dominate its response.

Two sets of test signals (input/output), one clean and one distorted, were used for all analysis. These signals were collected using our experimental hearing aid test system [10]. The clean-signal set is the nonsense word 'asil' put through a low-distortion hearing aid with the input and output signals sampled at a rate of 25kHz using 16-bit resolution. The distorted-signal set is the nonsense word 'abil' put through a distorting hearing aid and sampled as above.

The linear system identification methods are compared using mean-square output error (MSOE), plots of the mean frequency response (computed over all analysis blocks) and informal listening tests. If two methods extract the same "underlying" linear system for a given set of input and output signals, their MSOE should be roughly equivalent. We can also expect that the average frequency response for a hearing aid will be relatively smooth, if a model represents the "underlying" linear system. Perhaps the most reliable tests of all are listening tests. These tests allow the detection of incorrect frequency shaping (caused by biased coefficients), background noise (caused by rapidly changing filter coefficients) and other anomolies that are difficult to detect by other means.

Signal	*MA Sol'n Method*	
	Block	RLS
clean	2.14×10^5	4.02×10^4
distorted	2.00×10^7	5.84×10^6

Table 1: Mean-square output error for clean and distorted test-signal sets

3. IMPLEMENTATION AND RESULTS

3.1. MA models

The time-varying analysis necessary to develop MA models of hearing aids may be implemented on a block-wise or sample-by-sample basis. Both strategies were implemented to compare their performance.

In a block-wise analysis the LSA matrix equations (Eqn. 5) are not, in general, Toeplitz. Thus, the Levinson method cannot be used. Instead, Marple's efficient least-squares method [11] was implemented. A "slow" ($\mathcal{O}(N^2)$) recursive least squares (RLS) method [12] with exponential weighting ($\lambda = 0.9985$, 10% memory at 512 sample delay) was also implemented. (The slow method was used because it was easy to extend to higher dimensions for ARMA modelling.)

100^{th}-order filters were used for both test cases. The block-wise analysis was done every 1024 samples on non-overlapping blocks. For the RLS method, the frequency response was computed from filter coefficients that are "sampled" every 1024 points. All frequency responses were computed using 512-point complex FFTs. The MSOE for both methods with each signal set is shown in Table 1. The mean frequency responses are shown in Figures 2 and 3.

The RLS method gives lower MSOE. This is a result of sample-by-sample coefficient updating which allows better tracking of rapid changes. The frequency responses show that both methods give similar (i.e., slightly rough) frequency responses for clean signal. However, both methods fail for the distorted signal and give quite different frequency responses.

Listening tests reveal that both methods give excellent performance for the clean signal. The residual is primarily noise with some low-level signal. The output of the RLS model has some "zipper noise" caused by rapidly changing filter coefficients. The block-wise method fails completely on the distorted signal; the output signal from the model has large clicks and sounds very different from hearing aid output. The RLS model output sounds better than the block-wise model. However, it has high-levels of background zipper noise.

The memory parameter of the RLS filter (λ) was varied to determine the effect it had on performance. Larger values of λ reduced the level of zipper noise somewhat and improved the frequency response, but even $\lambda = 1$ did not give satisfactory performance.

3.2. ARMA Models

Three ARMA system identification methods were implemented and compared. We first implemented a block-wise

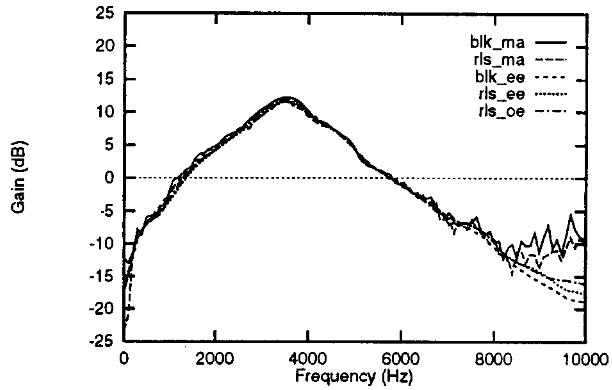

Figure 2: Average frequency responses with clean test-signal set for RLS moving average (rls_ma), block-wise MA (blk_ma), ARMA RLS output-error (rls_oe), ARMA RLS equation-error (rls_ee) and block-wise ARMA equation-error (blk_ee)

Signal	*ARMA Sol'n Method*		
	Block	RLS eqn. err	RLS out. err
clean	4.64×10^4	4.66×10^3	3.67×10^3
distorted	1.72×10^6	8.20×10^3	6.36×10^3

Table 2: Mean-square output error for clean and distorted test-signal sets

equation-error method where the autocorrelation matrix (Eqn. 6) was computed recursively on Hanning-windowed, non-overlapping blocks. The coefficients are directly computed via Cholesky decomposition. At the start of a block, the algorithm uses input and output signal samples from the end of the previous block to initialize the new filter (i.e., the signal is used to interpolate coefficients at block boundaries).

A RLS equation-error formulation was implemented using a two-dimensional version of the "slow" RLS method that was used above for MA modelling. Finally, an output-error ARMA formulation was implemented using a three-dimensional version of the "slow" RLS method that realizes the filtered-error version of this algorithm [9, 13].

16^{th}-order filters were used for all methods. All frequency responses computed as above. The block-wise analysis was done every 1024 samples; the RLS filters used $\lambda = 0.9985$ (as above). The MSOE for all three methods with each signal set is shown in Table 2. The mean frequency response graphs are shown in Figures 2 and 3.

Both of the RLS methods give lower (and similar) MSOE values for both signals. This is a result of their better tracking capability because they update the model coefficients each sample. The RLS output-error formulation gives the smallest MSOE. This is probably because coefficients are not biased by noise and distortion. The small increase in MSOE for RLS methods between the clean- and distorted-

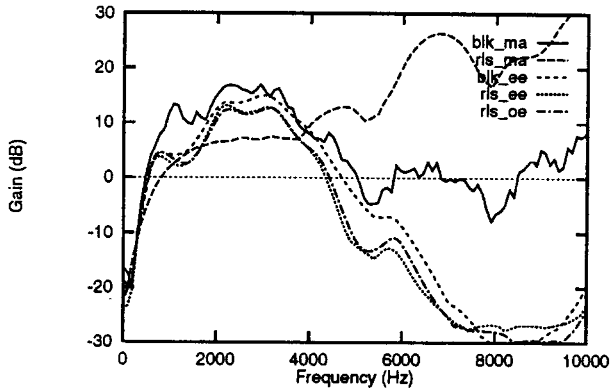

Figure 3: Average frequency responses with distorted test-signal set for RLS moving average (rls_ma), block-wise MA (blk_ma), ARMA RLS output-error (rls_oe), ARMA RLS equation-error (rls_ee) and block-wise ARMA equation-error (blk_ee)

test signals implies that they are probably not extracting the "underlying" linear system very well.

All of the frequency responses are similar for both test-signal sets. In listening tests for the clean signal, all methods performed very well. The block-wise method has slightly more signal in the residual; the outputs from both RLS models have higher-levels of background zipper noise caused by rapidly changing model coefficients.

Listening tests on the distorted test-signal set revealed that both RLS model outputs sound distorted. The value of λ chosen allows them to update their coefficients so rapidly that they compute a piecewise linear approximation to the non-linear characteristics of the hearing aid. Setting $\lambda = 1$ did not improve performance very much—the model output still sounded distorted. Both RLS methods also have higher levels of background (zipper) noise and a noticable high-frequency emphasis.

The block-wise equation-error model output sounds like a less-distorted version of hearing aid output signal. There are some artifacts present in the model output signal (e.g., a few "thumping" transients and low-level "swishy" background noise). However, this method performs better than either of the RLS methods.

4. CONCLUSIONS

All methods work well on low-noise, undistorted signals. However, there are large differences in performance for the distorted-test signal set. Both MA methods examined here fail when used with distorted signals. Overall, ARMA methods give much better performance than MA methods.

Both ARMA RLS schemes track too well and are unable to extract the underlying linear system because they update their coefficients each sample. Increasing the memory of the RLS filters by setting $\lambda = 1$ improved their performance only slightly.

The block-wise equation error ARMA method gives the best peformance of all methods examined here. Because it estimates the auto-correlation matrix over an entire block, it can, to a certain extent, "average out" the effects of distortion and extract the underlying linear system. Because coefficients are updated only once each block, there is no background zipper noise introduced. However, instantanous updating of coefficients at block boundaries can sometimes cause transients. Using the signal to interpolate partially solves this, but some transients are still introduced.

5. REFERENCES

[1] American National Standards Institute, "Specification of Hearing Aid Characteristics," in *ANSI 3.22*, ANSI, New York, 1987.

[2] American National Standards Institute, "Testing Hearing Aids with a Broadband Noise Stimulus," in *ANSI 3.42*, ANSI, New York, 1992.

[3] T. Schneider and D. G. Jamieson, "A Dual-Channel MLS-based Test System for Hearing Aid Characterization," *J. Audio Engineering Society*, vol. 41, no. 7/8, pp. 583–593, 1993.

[4] J. M. Kates, "On Using Coherence to Measure Distortion in Hearing Aids," *J. Acoustical Society of America*, vol. 91-Pt. 1, no. 4, pp. 2236–2244, 1992.

[5] T. Schneider and D. G. Jamieson, "Using MLS coherence for broadband distortion measuements on hearing aids," *J. Acoustical Society of America*, 1995. (to appear March 1995).

[6] L. Ljung, *System Identification: Theory for the User.* Prentice-Hall, New Jersey, 1987.

[7] L. R. Rabiner, R. E. Crochiere, and J. B. Allen, "FIR System Modelling and Identification in the Presence of Noise and with Band-Limited Inputs," *IEEE Trans. on Acoustics, Speech and Signal Proc.*, vol. 4, no. ASSP-16, pp. 319–333, 1978.

[8] M. J. Williamson, K. L. Cummins, and K. E. Hecox, "Speech Distortion Measures for Hearing Aids," *J. Rehabilitation Research and Development*, vol. 24, no. 4, pp. 277–282, 1987.

[9] J. Shynk, "Adaptive IIR Filtering," *IEEE ASSP Magazine*, pp. 4–21, April 1989.

[10] D. G. Jamieson and T. Schneider, "Electroacoustic Evaluation of Assistive Hearing Devices," *IEEE Engineering in Medicine and Biology Magazine*, pp. 249–254, April/May 1994.

[11] S. L. Marple, "Efficient Least Squares System Identification," *IEEE Trans. on Acoustics, Speech and Signal Proc.*, vol. ASSP-29, no. 1, pp. 62–73, 1981.

[12] S. Orfandis, *Optimum Signal Processing.* MacMillan, New York, 2^{nd} ed., 1988.

[13] K. Kurosawa and S. Tsujii, "An IIR Parallel-Type Adaptive Algorithm Using the Fast Least Squares Method," *IEEE Trans. on Acoustics, Speech and Signal Proc.*, vol. 37, no. 8, pp. 1226–1230, 1989.