Proceedings of the Institute of Acoustics

Time versus frequency domain correction of room responses.

P.G. Craven, Algol Applications Ltd., 26 The Nursery,
Sutton Courtenay, Oxon. OX14 4UA
& consultant to B&W Loudspeakers Ltd.

1. INTRODUCTION

A number of papers have recently been published concerning the"deconvolution" of acoustic impulse responses (e.g. Elliott and Nelson[1], Kuriyama and Furukawa [2]). These papers have concentrated on time-domain processing, whereby a transversal filter pre-processes the signal so that the final impulse response should approximate a delta function in some least-squares manner.

Although the least-squares criterion is reasonable when one is trying to achieve zero ouput, as in noise cancellation, it is easy to contrive examples which show an extremely negative correlation with psychoacoustic experience in assessing error in reproduced sound. (E.g. a phase inversion is considered by many people to be hardly detectable, yet one can hardly do more damage in terms of r.m.s. error !)

At B&W we have therefore preferred to use frequency domain processing for generation of speaker and room correction filters, even though the end product is a conventional transversal filters operating in the time domain. We can thus give explicit treatment to the error as a function of frequency, imposing criteria on the maximum allowed boost at any one frequency etc. etc..

It is a standard result that unconstrained deconvolution can be efficiently implemented by division of Fourier Transforms. However practical room responses are non-minimum phase (Neely and Allen [3]), leading to an acausal inverse. As practical implementations allow for only a finite "modelling delay" of m samples, we must constrain the generated filter to have zero response before t=-m. At first sight, this appears to require an explicit time-domain least-squares optimisation; however it is the purpose of this paper to show that such filters can be more efficiently generated by a processing method based largely in the frequency domain.

The mathematics used also provides a useful expression for the final frequency response, regardless of whether the implementation is by time- or frequency-domain processing.

2. CONSTRAINED DECONVOLUTION

In this section we show how to derive a filter which deconvolves a given response in a least-squares sense, subject to the constraint that the filter may extend only a finite number of samples in the negative time direction. (This duration is given by the "modelling delay" in a practical implementation).

2.1 Notation

t is the discrete time variable, taking integer values

Upper-case letters denote sequences

$\otimes$ is the convolution operator for sequences

R^{-1} is the convolutional inverse (assuming this exists) of R

R^{T} is the time-reverse of R

R[0:m] is a truncated version of R, equal to R over the range $t = 0...m$, zero elsewhere

$<\ ,\ >$ denotes scalar product

δ_n denotes a sequence with 1 at t=n and zeroes elsewhere.

A room response R is considered to be standardised by shifting in time to compensate for air delay etc., so that its first non-zero value occurs at t=0. It can be decomposed as

$$R = M \otimes A$$

where M is minimum-phase and A is all-pass, and both M and A are causal (with first non-zero entry at t=0).

2.2 Lemma

The time-shifted all pass responses

$$\{ \delta_n \otimes A \} \quad n = -\infty, \; .. -1, 0, 1, 2, ...\infty$$

form an orthonormal set.

Proof: The autocorrelation of an all-pass is the unit impulse.

I.e. $$<A , \delta_j \otimes A> \; = 1 \text{ if } j=0, \quad = 0 \text{ otherwise.}$$

Hence $$<\delta_i \otimes A , \delta_j \otimes A> \; = 1 \text{ if } i=j, \quad = 0 \text{ otherwise.}$$

As required.

2.3 Theorem

If R is causal and its minimum-phase part M is invertible, the least-squares solution of

$$X \otimes R = \delta_0$$

subject to the constraint that X(t) = 0 when r < -m, is given by

$$X = A[0:m]^T \otimes M^{-1}$$

where A is the all-pass part of R, and M is the minimum-phase part.

Proof: Let $Y = X \otimes M$, so that $X = Y \otimes M^{-1}$.

$$\text{Then } X \otimes R = (Y \otimes M^{-1}) \otimes (M \otimes A) = Y \otimes A.$$

Since M and M^{-1} are both causal, the constraint that X extend at most m samples into negative time translates to the same constraint on Y.
Hence the problem is equivalent to finding the least-squares solution Y, subject to the constraint, of

$$Y \otimes A = \delta_0.$$

Expanding Y as a sum of impulses

$$Y = \sum_{i=-m}^{\infty} y(i) . \delta_i$$

we thus solve

$$\sum_{i=-m}^{\infty} y(i) . (\delta_i \otimes A) = \delta_0$$

Using the orthonormality of the $\{\delta_i \otimes A\}$ (lemma 2.2), the least-squares solution is given by

$$Y(i) = \langle \delta_i \otimes A, \delta_0 \rangle = \begin{cases} A(-i) & \text{if } -m \leq i \leq 0 \\ 0 & \text{otherwise.} \end{cases}$$

Hence $Y = A[0:m]^T$
and therefore $X = A[0:m]^T \otimes M^{-1}$, as required.

2.4 Corollary

If the modelling delay m is zero, the result of theorem 2.3 reduces to

$$X = A(0) . M^{-1}$$

that is, the "constrained causal inverse" of the full room resonse is a scaled version of the unconstrained inverse of its minimun phase part.

3. FINAL FREQUENCY RESPONSE

The final, corrected, response is

$$X \otimes R = (A[0:m]^T \otimes M^{-1}) \otimes (M \otimes A)$$
$$= A[0:m]^T \otimes A \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.1)$$

Three interesting cases arise:

(i) m is infinite.

Then (3.1) reduces to

$$A^T \otimes A = \delta_0 .$$

As expected, the impulse response is perfect and the frequency response is flat.

(ii) m is zero.

Then (3.1) simplifies to

$$A(0).A$$

The impulse response is a scaled version of the all-pass component of the original impulse response, and the frequency response is flat.

(iii) m is non-zero and finite

Since convolution with an all-pass does not affect frequency response, and neither does time reversal, the freqency response of (3.1) is the same as that of the truncated all-pass response A[0:m].

This is easily plotted for various values of m. When m is small, the response will show substantial broad peaks and troughs, the size of which can be estimated to first order as (the square root of) the proportion of the energy of A which lies after the truncation point t=m.

4. COMPUTATIONAL EFFICIENCY

The minimum-phase inverse M^{-1} is obtained in the frequency domain by computing its phase as the Hilbert transform of the amplitude response. The main computational load is in performing Fourier transforms. Starting from a measurement in the time domain, two forward and one inverse transform yield M^{-1} in the Fourier domain.

The all-pass component A is obtained by convolution $A = R \circledast M^{-1}$. As the two sequences R and M^{-1} already exist in the Fourier domain, this involves one more inverse Fourier transform to obtain A as a time-series.

It m is zero, the convolution by $A[0:m]^T$ is trivial. Otherwise one forward and one inverse transform will deliver the final result as a time series.

Assuming that a Fourier transform of n points requires $4n.\log_2 n$ operations (faster algorithms are possible), the 6 Fourier transforms for the complete computation will take 24.n.log[2]n operations.

The usual direct methods for solving the least-squares problem would require of order n^3 operations. Marple [3] has published a clever but extremely complicated algorithm which requires about $131.n^2$ operations. It can be seen that for n of order 1000, one would naively expect the frequency domain approach to be faster by a factor of about 500.

Allowing for the possibility that n may need to be somewhat greater in the Fourier case in order to allow for "wrap round" effects, and that n also needs to be rounded up to the next power of 2, it would perphas be more reasonable to claim an improvement factor in excess of 150.

5. DISCUSSION AND CONCLUSIONS

The results reported here would seem to open the way for room-correction to be discussed more generally in the frequency-domain, rather than exclusively in the time-domain as hitherto.

Practical room equalisers, if implemented using transversal filters, have a restricted extent in both directions, and it has to be admitted that frequency domain methods do not at present provide an elegant means of dealing with this. However, in view of our initial commments about the least-squares criterion, it may be wondered whether the mathematically more satisfying constrained least-squares solution has any psychoacoustic advantage over a straightforward truncation or windowing of the generated filter.

In view of the results on the corrected frequency response, it would seem that modelling delays of just a few milliseconds are to be treated with caution. We have yet to confirm this result by listening tests, however.

6. REFERENCES

[1] S J ELLIOTT & P A NELSON 'Multiple-Point Equalisation in a Room using Adaptive Digital Filters', JAES vol.37 no.11 pp 899-907 (1989)

[2] J KURIYAMA & Y FURUKAWA 'Adaptive Loudspeaker System' JAES vol.37 no.11 pp 919-926 (1989)

[3] S T NEELY & J B ALLEN 'Invertibility of a Room Impulse Response', JASA vol.68 no.1 pp 165-169 (1979)

[4] S L MARPLE 'Efficient Least Squares FIR System Identification' IEEE trans. ASSP, vol.29 no.1 (1981)