

ITERATIVE ALGORITHMS FOR MULTICHANNEL EQUALIZATION IN SOUND REPRODUCTION SYSTEMS

M. Guillaume, Y. Grenier, and G. Richard

École Nationale Supérieure des Télécommunications
Département TSI. 46 Rue Barrault
75634 Paris Cedex 13, France

ABSTRACT

A fast iterative algorithm, with computation based on the fast Fourier transform (FFT), is presented. It can be used to control a soundfield at several control points with a loudspeaker array from multiple reference signals. It designs an equalizer able to invert long FIR filters and which achieves better performance than traditional FFT-based deconvolution methods with an equal number of coefficients in the inverse filters.

1. INTRODUCTION

Multichannel equalization is used in sound reproduction systems such as transaural systems, ambisonics, or WFS, in order to improve the spatio-temporal quality of sound reproduction. Miyoshi and al. first introduced the MINT theorem [1] in order to have a full control of sound at several points. But, like acoustic echo control [2], multichannel equalization is an application of very high order filters, and the resolution of the subjacent problem prevents a direct inversion as recommended by Miyoshi due to the size of the resultant matrix. Kirkeby and al. proposed a fast deconvolution method [3], able to inverse long electroacoustic responses. They use a regularization parameter, weighting the effort term of the solution, in order to limit the temporal extent of inverse filters so that linear convolutions could be correctly approximated with circular ones.

In this paper, iterative algorithms are presented for the purpose of multichannel equalization. They are able to invert long multiple electroacoustic impulse responses, and achieve better results than traditional methods with the same number of coefficients in the inverse filters. The global aim is to minimize a cost function expressed in the time domain but computations of the update equalizer are more efficiently done in the frequency domain, using circular convolutions (FFT) to calculate linear ones. In section 2, the multichannel equalization problem and the notations used are introduced. In section 3, two different algorithms are presented, and in section 4, results are presented in the case of the cross-talk cancellation problem and compared with other methods. Finally, section 5 suggests some conclusions.

2. CONTEXT AND OBJECTIVES

2.1. Problem description and variable definitions

Figure 1 illustrates the discrete-time multichannel equalization problem encountered in generalized transaural reproduction systems

corresponding author : mathieu.guillaume@enst.fr

[4]. The objective of the equalization is to determine the loudspeaker array inputs $\underline{y}(n)$ such that the sound pressure is controlled at multiple positions in space. In order to achieve this objective, I desired signals $\underline{s}(n)$ are computed from K recorded signals $\underline{x}(n)$ described by the extrapolation relation :

$$\begin{pmatrix} s_1(n) \\ \vdots \\ s_I(n) \end{pmatrix} = \begin{bmatrix} a_{11}(n) & \dots & a_{1K}(n) \\ \vdots & \ddots & \vdots \\ a_{I1}(n) & \dots & a_{IK}(n) \end{bmatrix} * \begin{pmatrix} x_1(n) \\ \vdots \\ x_K(n) \end{pmatrix} \quad (1)$$

When $K = 1$, $x(n)$ represents the temporal signature of a point source, whereas the extrapolation operator A characterizes the propagation in an enclosed space from the source to the I control points.

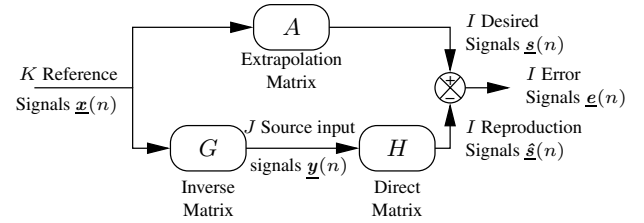


Fig. 1. Multichannel discrete-time equalization problem

The multiple input multiple output (MIMO) system H describes the electroacoustic transfer functions. It relates the J transducers signals to the I reproduced signals at the control points :

$$\begin{pmatrix} \hat{s}_1(n) \\ \vdots \\ \hat{s}_I(n) \end{pmatrix} = \begin{bmatrix} h_{11}(n) & \dots & h_{1J}(n) \\ \vdots & \ddots & \vdots \\ h_{I1}(n) & \dots & h_{IJ}(n) \end{bmatrix} * \begin{pmatrix} y_1(n) \\ \vdots \\ y_J(n) \end{pmatrix} \quad (2)$$

Thus $h_{ij}(n)$ is the electroacoustic impulse response between the transducer j and the control point i .

The MIMO system G is the equalizer. It computes the input signals of the loudspeaker array from the K reference signals :

$$\begin{pmatrix} y_1(n) \\ \vdots \\ y_J(n) \end{pmatrix} = \begin{bmatrix} g_{11}(n) & \dots & g_{1K}(n) \\ \vdots & \ddots & \vdots \\ g_{J1}(n) & \dots & g_{JK}(n) \end{bmatrix} * \begin{pmatrix} x_1(n) \\ \vdots \\ x_K(n) \end{pmatrix} \quad (3)$$

Thus $g_{jk}(n)$ is the impulse response of the equalizer between the reference signal k and the loudspeaker input signal j .

Using equation 3, 2 and 1, it is easily shown that the equalizer has to achieve $HG = A$, independently of the K reference signals

[4]. The problem can then be decomposed into K independent subproblems [5] :

$$\forall k \in [1, \dots, K], \quad \begin{bmatrix} h_{11}(n) & \dots & h_{1J}(n) \\ \vdots & \ddots & \vdots \\ h_{I1}(n) & \dots & h_{IJ}(n) \end{bmatrix} * \begin{pmatrix} g_{1k}(n) \\ \vdots \\ g_{Jk}(n) \end{pmatrix} = \begin{pmatrix} a_{1k}(n) \\ \vdots \\ a_{Ik}(n) \end{pmatrix} \quad (4)$$

2.2. Numerical resolution

Each above subproblem is of infinite dimension and some approximations must be made to solve it numerically. Each impulse response $h_{ij}(n)$, $g_{jk}(n)$, and $a_{ik}(n)$ will be modeled by a FIR filter [2] of L_h , L_g , and $N = L_h + L_g - 1$ (the length of the discrete convolution of h_{ij} and g_{jk}) coefficients. In the following, we use the notations :

$$\begin{aligned} \underline{g}_{jk} &= [g_{jk}(0), \dots, g_{jk}(L_g - 1)]^T \\ \underline{a}_{ik} &= [a_{ik}(0), \dots, a_{ik}(N - 1)]^T \\ \underline{h}_{ij} &= [h_{ij}(0), \dots, h_{ij}(L_h - 1)]^T \\ \mathcal{H}_{ij} &= \begin{bmatrix} h_{ij}(0) & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ h_{ij}(L_h - 1) & & & 0 \\ 0 & \ddots & & h_{ij}(0) \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & h_{ij}(L_h - 1) \end{bmatrix} \\ \mathcal{H} &= \begin{bmatrix} \mathcal{H}_{11} & \dots & \mathcal{H}_{1J} \\ \vdots & \ddots & \vdots \\ \mathcal{H}_{I1} & \dots & \mathcal{H}_{IJ} \end{bmatrix} \\ \underline{g}_k &= [\underline{g}_{1k}^T, \dots, \underline{g}_{Jk}^T]^T ; \underline{a}_k = [\underline{a}_{1k}^T, \dots, \underline{a}_{Ik}^T]^T \end{aligned} \quad (5)$$

Thus, $\mathcal{H}_{ij} \cdot \underline{g}_{jk}$ computes the discrete convolution between the two finite sequences h_{ij} and g_{jk} . The resulting problem is of finite dimension and can be solved numerically. With the above notations, it can be synthesized by the relation :

$$\mathcal{H} \underline{g}_k = \underline{a}_k, \quad \forall k \in [1, \dots, K] \quad (6)$$

Relation 6 describes a system of IN equations with JL_g unknowns. It is well known [5] that the solution which minimizes the quadratic error :

$$\underline{g}_{k_{opt}} = \mathcal{H}^\dagger \underline{a}_k \quad (7)$$

where $\mathcal{H}^\dagger$ is the pseudo-inverse of the matrix $\mathcal{H}$. This solution is also of minimal effort in the case in which the problem is sub-determined, that is when $IN < JL_g$. Unfortunately, the practical dimensions of $\mathcal{H}$ prevent a direct pseudo-inversion. Iterative algorithms represent an alternative solution to a direct pseudo-inversion since they converge to the optimal solution without needing the inversion of $\mathcal{H}$.

3. ITERATIVE ALGORITHMS

At each iteration of the algorithm, the equalizer $\underline{g}_k$ is updated in order to decrease a cost function. Generally, the mean-squared-error (MSE) is used [6] :

$$J_k(\underline{g}_k) = \|\underline{a}_k - \mathcal{H} \underline{g}_k\|_2^2 \quad (8)$$

In this paper, two algorithms are studied : the steepest-descent and the approximative Gauss-Newton algorithms [6]. In the following, the notation $\text{diag}(\underline{m})$ transforms the vector $\underline{m}$ into a square matrix with its elements on the diagonal, and the notation $\text{diag}_n(\mathcal{M})$ applied to a matrix is a block matrix with n matrix $\mathcal{M}$ on its diagonal.

The complexity of these algorithms can largely be decreased if the required computations are done in the frequency domain [7]. For each step of these algorithms, only linear convolutions of two finite discrete sequences have to be computed. This can be done more efficiently in the frequency domain using FFT algorithms. For instance, consider the convolution of $h_{ij}(n)$ and $g_{jk}(n)$, which is accomplished in the time domain by the matrix operation $\mathcal{H}_{ij} \underline{g}_{jk}$. This can also be done in the frequency domain by zero-padding the two sequences in order to have N elements, and by computing the circular convolution of the two sequences by FFT. The matrix operation can be written as :

$$\mathcal{H}_{ij} \underline{g}_{jk} = \mathcal{F}^{-1} \text{diag}(\mathcal{F} \mathcal{M}_h \underline{h}_{ij}) \mathcal{F} \mathcal{M}_g \underline{g}_{jk}$$

where $\mathcal{F} [N \times N]$ denotes the discrete Fourier transform matrix, $\mathcal{M}_h$ the zero-padding matrix for the sequences $\underline{h}_{ij}$, and $\mathcal{M}_g$ the zero-padding matrix for the sequences $\underline{g}_{jk}$ [7]. The following matrix identity is deduced :

$$\mathcal{H}_{ij} = \mathcal{F}^{-1} \text{diag}(\mathcal{F} \mathcal{M}_h \underline{h}_{ij}) \mathcal{F} \mathcal{M}_g$$

This identity can be generalized to the matrix $\mathcal{H}$:

$$\mathcal{H} = \mathcal{F}_I^{-1} \mathbf{H} \mathcal{F}_J \mathcal{M} \quad (9)$$

with $\mathcal{F}_I = \text{diag}_I(\mathcal{F})$, $(\mathbf{H})_{i,j} = \text{diag}(\mathcal{F} \mathcal{M}_h \underline{h}_{ij})$, $\mathcal{F}_J = \text{diag}_J(\mathcal{M})$, and $\mathcal{M} = \text{diag}_J(\mathcal{M}_g)$.

3.1. Steepest Descent Algorithm

In this algorithm, the equalizer $\underline{g}_k$ is updated in the opposite direction of the gradient of the cost function[6] :

$$\begin{aligned} \underline{g}_k(m+1) &= \underline{g}_k(m) - \frac{\mu}{2} \nabla_{\underline{g}_k} J_k(\underline{g}_k) \\ \text{with :} & \\ \nabla_{\underline{g}_k} J_k(\underline{g}_k) &= -2 \mathcal{H}^T \underline{e}_k(m) = \mathcal{H}^T (\underline{a}_k - \mathcal{H} \underline{g}_k(m)) \end{aligned} \quad (10)$$

The iteration step must fulfill the condition $\mu < 2/\lambda_{max}$ where λ_{max} is the largest eigenvalue of the autocorrelation matrix $\mathcal{H}^T \mathcal{H}$ [6]. As this value is unknown, one has to proceed by successive trials in order to assure convergence of this algorithm. This procedure can be automated using an optimal step version of the algorithm, in which the step is chosen in order to minimize the cost function in the next iteration [6]. This operation is synthesized by the following equation :

$$\begin{aligned} \mu(m) &= \left\{ \mu / \nabla_{\mu} J_k(\underline{g}_k) = 0 \right\} \iff \mu(m) = \frac{\underline{e}_k^T(m) \underline{e}_k(m)}{\underline{e}_k^T(m) \underline{e}_k(m)} \\ \text{with : } \underline{e}_k(m) &= \mathcal{H} \mathcal{H}^T \underline{e}_k(m) \end{aligned} \quad (11)$$

The details of the computation are summarized in Table 1.

Initialization	
<i>Frequency Direct Matrix</i>	
$(\mathbf{H})_{i,j} = \text{diag}(\mathcal{F}\mathcal{M}_h \mathbf{h}_{i,j})$	
<i>Reference Vector</i>	
$\mathbf{a}_k = \mathcal{F}_I \mathbf{a}_k$	
Iteration	
<i>Estimated vector</i>	
$\hat{\mathbf{a}}_k(m) = \mathbf{H} \mathcal{F}_J \mathcal{M} \mathbf{g}_k(m)$	
<i>Error vector</i>	
$\mathbf{e}_k(m) = \mathbf{a}_k(m) - \hat{\mathbf{a}}_k(m)$	
<i>Gradient vector</i>	
$\mathbf{d}_k(m) = \mathcal{M}^T \mathcal{F}_J^{-1} \mathbf{H}^H \mathbf{e}_k(m)$	
Normal Version	Optimal Step Version
<i>Update</i>	<i>Optimal Step Vector</i>
	$\mathbf{c}_k(m) = \mathbf{H} \mathcal{F}_J \mathcal{M} \mathbf{d}_k(m)$
	<i>Iteration Step</i>
	$\mu(m) = \frac{\mathbf{e}_k^H(m) \mathbf{c}_k(m)}{\mathbf{c}_k^H(m) \mathbf{c}_k(m)}$
<i>Update</i>	<i>Update</i>
	$\mathbf{g}_k(m+1) = \mathbf{g}_k(m) + \mu(m) \mathbf{d}_k(m)$

Table 1. Steepest Descent Algorithms

3.2. Approximated Gauss-Newton Algorithm

The opposite direction of the gradient of the cost function is not the optimal direction for updating the equalizer. To increase speed convergence, the best direction for updating is obtained by multiplying the gradient vector with the inverse of the Hessian matrix $(\mathcal{H}^T \mathcal{H})^{-1}$ [6]. Admittedly, if the computation of the inverse Hessian matrix was feasible, it would be better to solve directly the problem without an iterative algorithm. But the inverse Hessian matrix can only be approximated, using the frequential Hessian matrix $\mathbf{H}^H \mathbf{H}$. Hence, the structure of the matrix $(\mathbf{H}^H \mathbf{H} + \lambda \text{Id})$, constituted of blocks of diagonal matrix, enables fast algorithms for its inversion. The parameter λ is used for regularization. Finally, the approximated temporal inverse Hessian matrix $\mathbf{W}$ used here is :

$$(\mathcal{H}^T \mathcal{H})^{-1} \approx \mathbf{W} = \mathcal{M}^T \mathcal{F}_J^{-1} (\mathbf{H}^H \mathbf{H} + \lambda \text{Id})^{-1} \mathcal{F}_J \mathcal{M} \quad (12)$$

The resulting algorithm is the approximated Gauss-Newton algorithm, with update equation :

$$\mathbf{g}_k(m+1) = \mathbf{g}_k(m) - \frac{\mu}{2} \mathbf{W} \nabla_{\mathbf{g}_k} J_k(\mathbf{g}_k) \quad (13)$$

The corresponding optimal step version of this algorithm is still obtained by using equation 11 with :

$$\mathbf{c}_k(m) = \mathcal{H} \mathbf{W} \mathcal{H}^T \mathbf{e}_k(m)$$

The successive steps of this algorithm are summarized in Table 2.

It can be noted that a better candidate than the null equalizer can be chosen for the equalizer initialization for the two algorithms :

$$\mathbf{g}_{k0} = \mathcal{M}^T \mathcal{F}_J^{-1} (\mathbf{H}^H \mathbf{H} + \lambda \text{Id})^{-1} \mathbf{H}^H \mathbf{a}_k \quad (14)$$

This initial equalizer is the one used in traditionnal FFT-based methods using regularization [3].

Initialization	
<i>Frequency Direct Matrix</i>	
$(\mathbf{H})_{i,j} = \text{diag}(\mathcal{F}\mathcal{M}_h \mathbf{h}_{i,j})$	
<i>Reference Vector</i>	
$\mathbf{a}_k = \mathcal{F}_I \mathbf{a}_k$	
<i>Frequency Inverse Autocorrelation Matrix</i>	
$\mathbf{W} = (\mathbf{H}^H \mathbf{H} + \lambda \text{Id})^{-1}$	
Iteration	
<i>Estimated vector</i>	
$\hat{\mathbf{a}}_k(m) = \mathbf{H} \mathcal{F}_J \mathcal{M} \mathbf{g}_k(m)$	
<i>Error vector</i>	
$\mathbf{e}_k(m) = \mathbf{a}_k(m) - \hat{\mathbf{a}}_k(m)$	
<i>Gradient vector</i>	
$\mathbf{d}_k(m) = \mathcal{M}^T \mathcal{F}_J^{-1} \mathbf{H}^H \mathbf{e}_k(m)$	
<i>Modified Gradient vector</i>	
$\mathbf{b}_k(m) = \mathcal{M}^T \mathcal{F}_J^{-1} \mathbf{W} \mathcal{F}_J \mathcal{M} \mathbf{d}_k(m)$	
Normal Version	Optimal Step Version
<i>Update</i>	<i>Optimal Step Vector</i>
	$\mathbf{c}_k(m) = \mathbf{H} \mathcal{F}_J \mathcal{M} \mathbf{b}_k(m)$
	<i>Iteration Step</i>
	$\mu(m) = \frac{\mathbf{e}_k^H(m) \mathbf{c}_k(m)}{\mathbf{c}_k^H(m) \mathbf{c}_k(m)}$
<i>Update</i>	<i>Update</i>
	$\mathbf{g}_k(m+1) = \mathbf{g}_k(m) + \mu(m) \mathbf{b}_k(m)$

Table 2. Approximated Gauss-Newton Algorithms

4. RESULTS

In this section, results are presented in the case of typical transaural systems ($I = J = K = 2$). The equalization problem is known as cross-talk cancellation. The equalizer aims to compensate for both the room and the the loudspeakers responses, and also to cancel the cross-talk from the left loudspeaker to the right ear and vice-versa.

The direct transfert matrix H has been measured by sweep techniques [8]. The corresponding responses are represented on figure 2. The responses have been truncated to $L_h = 8192$ samples. The length of each inverse filter has been fixed to $L_g = 8193$ coefficients so that the convolution of the equalizer and the direct transfer matrix is of length 16384 samples.

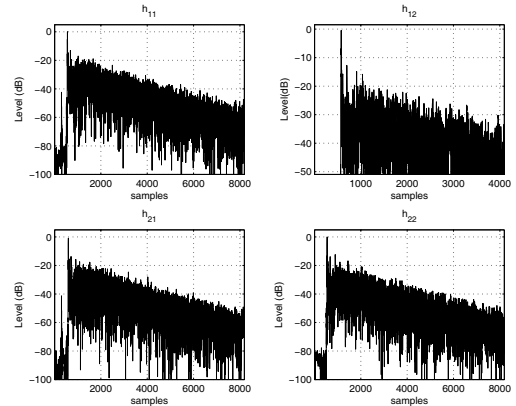


Fig. 2. Direct Transfert Matrix

Results of traditional FFT-based methods, synthesized by the equalizer of equation 14 are depicted on figure 3. The best results have been obtained with $\lambda = 10^{-2}$ for the regularization parameter. Nevertheless pre-echos and post-echos are visible (presence of secondary peaks before and after the main peak, approximately -35dB level). These artifacts are present because regularization isn't sufficient to limit the temporal extent of the inverse filters and then cannot cope with the time aliasing due to circular convolutions.

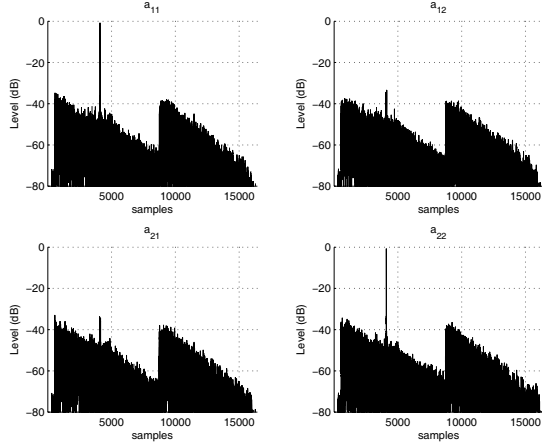


Fig. 3. Results with traditional equalization methods

Better performance could be achieved by using more coefficients in the inverse filters, or by using the iterative algorithms presented in this paper. Time aliasing doesn't occur in the inverse filter because linear convolutions are properly computed in the frequency domain, using FFT algorithms with correctly zero-padded sequences. The corrected system is presented on figure 4, after the convergence of the algorithm. Instead of secondary peaks, a quasi-constant noise reconstruction, approximately -50 dB level, is visible.

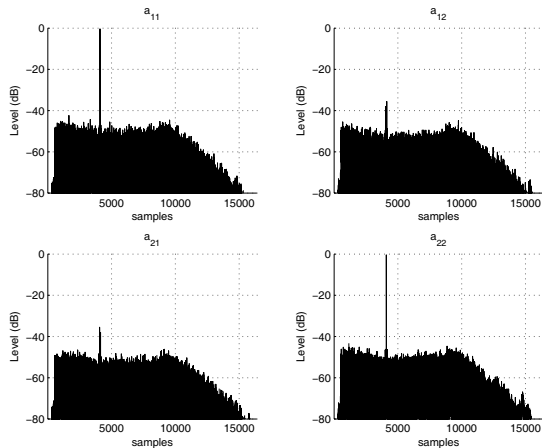


Fig. 4. Results after convergence of the algorithm

Concerning the performance of these algorithms, the MSE is represented on figure 5 for the optimal step gradient algorithm and

for several values of λ for the optimal step approximated Gauss-Newton algorithm. It is seen that the accuracy of the obtained solution is better when the regularization parameter is low. On the other hand, the initial convergence speed is better when the regularization parameter is high. $\lambda = 10^{-2}$ seems to be a good tradeoff between speed convergence and accuracy of the solution. Note that this value corresponds to the one used in the traditional deconvolution method. Moreover, it is seen that approximated Gauss-Newton algorithm outperforms gradient algorithm concerning both speed convergence and accuracy of the solution.

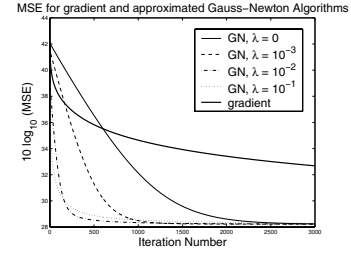


Fig. 5. Convergence properties of optimal step algorithms

5. CONCLUSION

Iterative multichannel equalization algorithms have been presented in this paper. They enable the inversion of long multiple electroacoustic impulse responses. Results have been given in the case of classical transaural applications and demonstrate the efficiency and accuracy of our approach, which is very suitable when optimal performance is required with a moderate number of coefficients in the inverse filters. It is important to emphasize that the proposed algorithms apply for any number of control points and/or loudspeakers.

6. REFERENCES

- [1] M. Miyoshi and Y. Kaneda, "Inverse control of room acoustics using multiple loudspeakers and/or microphones," *ICASSP*, 1986.
- [2] C. Breining and al., "Acoustic echo control : An application of very-high order adaptive filters," *IEEE Signal Processing Magazine*, 1999.
- [3] O. Kirkeby and al., "Fast deconvolution of multichannel systems using regularization," *IEEE Transactions on Speech and Audio Processing*, 1998.
- [4] O. Kirkeby and P.A. Nelson, "Digital filter design for inversion problems in sound reproduction," *Journal of Audio Engineering Society*, 1999.
- [5] J. Bauck and D.H. Cooper, "Generalized transaural stereo and applications," *Journal of the Audio Engineering Society*, 1996.
- [6] G.-O. Glentis and al., "Efficient least squares adaptive algorithms for fir transversal filtering," *IEEE Signal Processing Magazine*, 1999.
- [7] J.J. Shynk, "Frequency-domain and multirate adaptive filtering," *IEEE Signal Processing Magazine*, 1992.
- [8] S. Müller, "Transfer-function measurement with sweeps," *Journal of Audio Engineering Society*, 2001.