TIME AND FREQUENCY DOMAIN ROOM COMPENSATION APPLIED TO WAVE FIELD SYNTHESIS

L. Fuster, J.J. López, A. González, P. Faus

Audio and Communications Signal Processing Group (GTAC),
Institute of Communications and Multimedia Applications
Universidad Politécnica de Valencia - UPV, Valencia, Spain. http://www.gtac.upv.es
lfuster@dcom.upv.es

ABSTRACT

In sound rendering systems using loudspeakers, the listening room adds echoes not considered by the reproduction system, thus deteriorating the rendered audio signal. Specifically, Wave Field Synthesis is a 3D audio reproduction system, which allows synthesizing a realistic sound field in a wide area by using arrays of loudspeakers. This paper proposes a room compensation approach based on a multichannel inverse filter bank calculated to compensate the room effects at selected points within the listening area. Time domain and frequency domain algorithms are proposed to accurately compute the bank of inverse filters. A comparative study between these algorithms by means of laboratory experiments is presented.

1. INTRODUCTION

Nowadays, one of the most promising audio reproduction system is the Wave Field Synthesis (WFS) [1], where sound field is synthesized in an wide area by means of arrays of loudspeakers. This system allows WFS to reproduce an acoustic field inside a volume from the signal recorded on a given surface.

Some of the main problems to implement WFS are related to the interaction of the array with the listening room. The listening room introduces new echoes that are not included in the signal to be reproduced, altering the synthesized sound-field and reducing the spatial effect, and thus, the promised potentiality of this system.

Recently, some developments have been carried out in order to reduce room reflections using a bank of filters before the arrays, by means of a wave domain approach [2], [3]. In contrast to these previous works, we propose a new approach to WFS compensation based on direct solutions of the multichannel inverse filtering problem obtained in time and frequency domain to compensate the room effects at certain control points within the listening area.

Next section briefly introduces the theory related to WFS and multichannel inversion obtained in both time and frequency domains. Section 3 explains the developed experiment in our laboratory and shows some results of the time and frequency employed algorithms. Finally, the conclusions of both methods are presented.

2. THEORY

Firstly, a brief explanation of the Wave Field Synthesis (WFS) system will be given. In section 2.2, an introduction to Multiple-Input-Multiple-Output systems will be commented. Afterwards, the room compensation algorithm would be particularized to WFS. Section 2.4 explains the matrix construction and Toeplitz solver for filter inversion in time domain algorithms. At last, the fast and approximated frequency domain method will be described in section 2.5.

2.1. Wave Field Synthesis

Wave Field Synthesis is a method of sound reproduction, based on fundamental acoustic principles [1],[4]. It enables the generation of sound fields with natural temporal and spatial properties within a volume or area bounded by secondary sources (arrays of loudspeakers). This method offers a large listening area with uniform and high reproduction quality.

The theoretical basis of WFS is given by the Huygens' principle. According to this, the propagation of a wave front can be described by recursively adding the contribution of a number of secondary point sources distributed along the wave front. This principle can be used to synthesize acoustic wave fronts of an arbitrary shape.

A synthesis operator for each loudspeaker can be derived. The general 3D solution can be transformed into the 2-D solution, which is sufficient for reconstructing the original sound field in the plane of listening [5],[6],[7]. For that purpose a linear array of loudspeakers is employed to generate the sound field of virtual sources.

The field rendered by a source at a point R within the area surrounded by the loudspeakers can be expressed as equation (1).

$$P(r_R) = \sum_{n=1}^{N} Q_n(w) \frac{e^{-jk\triangle r}}{\triangle r} \tag{1}$$

where $\frac{e^{-jk\triangle r}}{\triangle r}$ represents the free field propagation between the secondary sources (loudspeakers) and the point R within the listening area. $Q_n(w)$ corresponds to the expression of the nth loudspeaker driving signal for a rendering system of N loudspeakers. These driving signals are dependent on the virtual source, loudspeakers and listening area positions [5].

Because of the separation between the loudspeakers, a spatial aliasing frequency exists. This frequency is given by equation 2.

$$f_{al} = \frac{c}{2\Delta x \sin \Theta_{max}} \tag{2}$$

where c is the speed of sound, Δx represents the separation between loudspeakers and Θ_{max} corresponds to the maximum angle of incidence of the synthesized wave field relative to the loudspeaker array.

2.2. MIMO Systems

A system with multiple linearly related inputs and outputs is commonly denoted as a MIMO system (Multiple-Input-Multiple-Output). The relation between each input and each output is described by a linear time invariant (LTI) system, and so by its corresponding impulse response. Therefore, a MIMO system of L inputs and M output is comprised of MxL impulse responses.

The inverse filtering problem in practical multichannel audio reproduction systems, basically consists in designing a matrix $\mathbf{H}$ of digital finite duration filters (each column of $\mathbf{H}$ represents a different vector of filters for each signal to be rendered), whose convolutions with the signal transmission channels (matrix $\mathbf{C}$), or electroacoustic system matrix, best approximates a desired response (matrix $\mathbf{A}$). Figure 1 shows a diagram of a typical multichannel inversion problem.

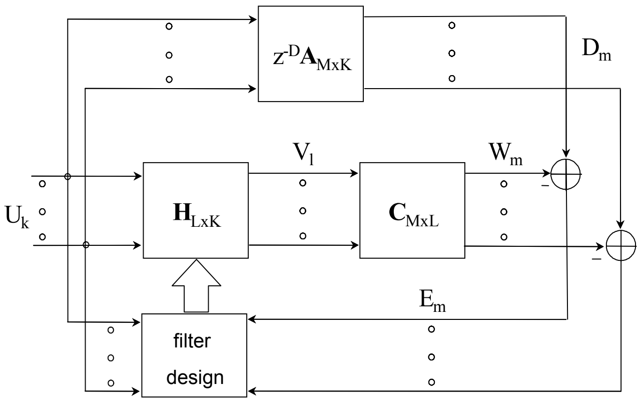

Figure 1: *Multichannel inverse filtering problem.*

Figure 1 illustrates a multichannel deconvolution problem where matrix $\mathbf{C}$ represents the actual transmission channels and matrix $\mathbf{H}$ is the bank of inverse filters used for deconvolution. The K program signals to be rendered are represented in figure 1 as U_k, the L signals that feed the transducers are denoted as V_l and the K desired signals at the reception points are D_m. The difference between the received signals, represented by W_m, and the desired signals are named error signals and denoted by E_m. Driving signals pass through inverse filters prior to feed the transmission channels. This configuration is typical in multichannel sound reproduction systems where inverse filters are usually calculated by the least squares method in time domain [8].

The calculation of the inverse filters can be carried out in a setup stage because the main room reflections can be considered invariant for each specific room.

Different methods have been proposed to obtain the bank of correction filters. Some compute an approximate solution in time domain [9] and others compute an approximate solution in frequency domain using Fast Fourier Transforms (FFT) [10].

2.3. Application to room compensation in WFS rendering systems

Figure 2 represents the block diagram of the multichannel inverse filtering for WFS. The filter matrix $\mathbf{H}$ is calculated using the transmission channels responses, which are measured a priori (matrix $\mathbf{C}$), and the desired signals at the listening or reproduction points. The input signals to the matrix $\mathbf{A}$ and the matrix $\mathbf{H}$ are not the original sound, but the excitation signals for the secondary sources provided by the WFS rendering algorithm. The filter matrix $\mathbf{H}$ must be composed of $L \times L$ inverse filters in the case of a WFS system.

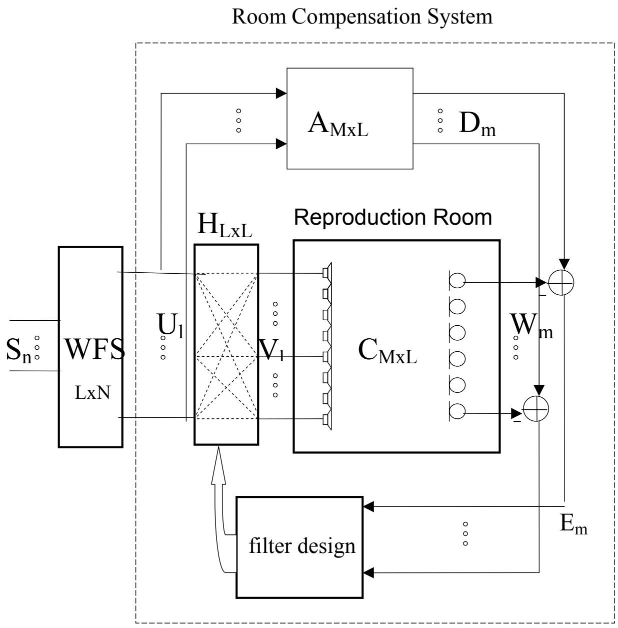

Figure 2: *Room correction system for WFS.*

Using this approach the room compensation system is independent of the WFS rendering algorithm at the previous stage. The number of inputs to the MIMO WFS compensation system will be equal to the number of loudspeakers.

2.4. Matrix Construction and Toeplitz Solver

A typical problem in signal processing consists in the computation of inverse FIR filter banks in order to equalize a given MIMO system. This is commonly called multichannel deconvolution. The algorithm used in this paper can cope with time domain deconvolution problems where a huge set of linear equations must be solved to design good filters under the least squares error criterion.

For our specific case, the block diagram of figure 2 can be described by means of the following set of equations in time domain: $\mathbf{A} = \mathbf{CH}$, where $\mathbf{C}$ is composed of $M \times L$ blocks $\mathbf{C}_{ij}$ and $\mathbf{H}$ is composed of $L \times L$ vectors $\mathbf{h}_{jk}$, [11]. Each block $\mathbf{C}_{ij}$ is a Toeplitz matrix that carries out convolutions with the ijth channel. Each product between matrix $\mathbf{C}_{ij}$ and vector $\mathbf{h}_{jk}$ performs the convolution between ijth transmission channel and jkth inverse filter. Bank of filters is usually calculated by the least squares method solving $\mathbf{C}^T\mathbf{CH} = \mathbf{C}^T\mathbf{A}$. The symmetric matrix $\mathbf{C}^T\mathbf{C}$ is given by,

$$\mathbf{C}^T\mathbf{C} = \begin{bmatrix} \sum_{m=1}^{M} \mathbf{R}_{m1\,m1} & \dots & \sum_{m=1}^{M} \mathbf{R}_{mL\,m1} \\ \sum_{m=1}^{M} \mathbf{R}_{m1\,m2} & \dots & \sum_{m=1}^{M} \mathbf{R}_{mL\,m2} \\ \vdots & & \vdots \\ \sum_{m=1}^{M} \mathbf{R}_{m1\,mL} & \dots & \sum_{m=1}^{M} \mathbf{R}_{mL\,mL} \end{bmatrix} \tag{3}$$

where $\mathbf{R}_{ij\,kl} = \mathbf{C}_{ij}^T \mathbf{C}_{kl}$ is a Toeplitz matrix.

Therefore matrix $\mathbf{C}^T\mathbf{C}$ is composed of $L \times L$ blocks of $n_h \times n_h$ elements. Each block has a Toeplitz structure. In the case of WFS, the multichannel system has L input signals from the WFS matrix. There will be L different sets of equations with different right-hand side, the columns of $\mathbf{C}^T\mathbf{A}$, but sharing the same main matrix, $\mathbf{C}^T\mathbf{C}$.

Furthermore, the main matrix has a Toeplitz-block structure. This structure itself does not allow the employment of efficient solution techniques. However, performing simple rows and columns permutations, a Block Toeplitz matrix can be achieved from the Toeplitz-block one. Thus a generalization of the *fast* methods used in the scalar Toeplitz case can be used for solving the Block Toeplitz case.

While the solution of a general linear equations set of order n requires $O(n^3)$ operations, there exist several methods for taking advantage of the structure of a Toeplitz matrix. We refer to the well known Levinson and Trench as *fast* Toeplitz solvers because they require $O(n^2)$ arithmetic operations for the solution of an $n \times n$ Toeplitz set of equations.

In order to save computation time, we solve the generic equations set $\mathbf{Rh} = [1, 0, \ldots, 0]^T$ using Durbin's algorithm [12]. This algorithm exploits the simplified form of the right-hand side and provides a further computational cost reduction compared to Levinson algorithm. Durbin algorithm provides first column of the inverse matrix with a computational saving of 50% compared to the general case. In order to solve the set of equations for each program signal, the Gohberg-Semencul formula [13] can be used. More details about this algorithm can be found in [11].

2.5. Fast Deconvolution using Regularization

As an alternative to the computation of multichannel inverse filter in time domain, a fast deconvolution method was proposed in [10]. Fast deconvolution computes the inverse filters in frequency domain, taking profit of the FFT. The main benefit of the algorithm is the reduction of the computation time.

In frequency domain, the matrix $\mathbf{H}$ (or control filter matrix) which minimizes the quadratic error between the desired response (matrix $\mathbf{A}$) and the room response (matrix $\mathbf{C}$) is given by:

$$\mathbf{H}_{LSE}(z) = [\mathbf{C}^T(z^{-1})\mathbf{C}(z) + \beta\mathbf{I}]^{-1}\mathbf{C}^T(z^{-1})\mathbf{A}(z) \quad (4)$$

where β is the regularization parameter, which allows that $\mathbf{C}^H\mathbf{C} + \beta\mathbf{I}$ was no singular for $\beta > 0$. Although a high β implies a more biased solution to the original least squares problem.

A detailed explanation of the fast deconvolution method can be found in [10]. This method has proved to be very useful and easy to use, but can suffer from circular convolution effects when the inverse filters are not long enough compared to the duration of the responses of the transmission channels, as will be shown in section 3.

3. EXPERIMENT AND RESULTS

The purpose of the laboratory experiments was to validate the multichannel inversion of MIMO systems computed using Block Toeplitz solvers in time domain and fast deconvolution in frequency domain as a possible practical solution to room compensation for WFS reproduction systems. The section 3.1 explains our developed laboratory experiment. Following, results for this laboratory experiment are shown.

3.1. Experiment Setup

In our prototype an U-shaped WFS array of 32 loudspeakers with a separation of $18cm$ was installed in a real room of dimensions $3 \times 4 \times 2.5m$. The room is slightly acoustical conditioned with reflections in walls, ceiling and floor. Figure 3 shows an array of loudspeakers and the measuring microphones.

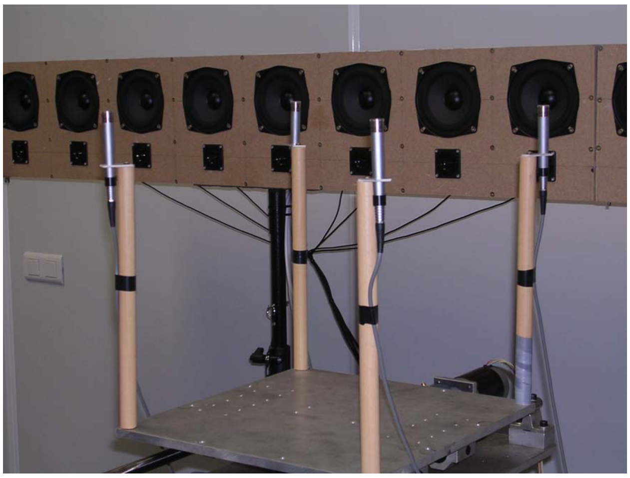

Figure 3: *Setup of WFS array and the measuring microphones mounted on the positioning system.*

The room impulse responses (RIR) between each loudspeaker and 196 listening points separated $5cm$ and situated over a square inside the listening area at the loudspeakers horizontal plane were measured, using several pressure microphones by means of an automatic positioning system. A set of 32×196 RIR were obtained using a Maximum Length Sequences (MLS) measurement method, which has been specially adapted for fast measuring of multichannel systems [14]. The position of the control points has been centered within the loudspeakers arrays.

From these RIR, the matrix $\mathbf{C}$ containing the room responses with the direct signal and the first reflections was built. Concretely, the signals were obtained with a sampling frequency of $48kHz$, and decimated to $8kHz$, because frequencies above $2kHz$ were no considered because of the aliasing frequency. In our experiment the minimum spatial aliasing frequency is $f_{al} \approx 950Hz$. Each RIR was windowed in time domain taking only the first $50ms$.

Then, the bank of inverse filters $\mathbf{H}$ was computed to correct the undesired effects of the room. First, a time domain algorithm was used to obtain an ideal channel up to $2kHz$ with the corresponding delay by means of fast Block Toeplitz solvers. Secondly, the fast deconvolution method was used in frequency domain to calculate these filters. In this case, the regularization parameter was used to constrain the band of interest. In the algorithm, matrix $\mathbf{A}$ is designed to emulate free field conditions.

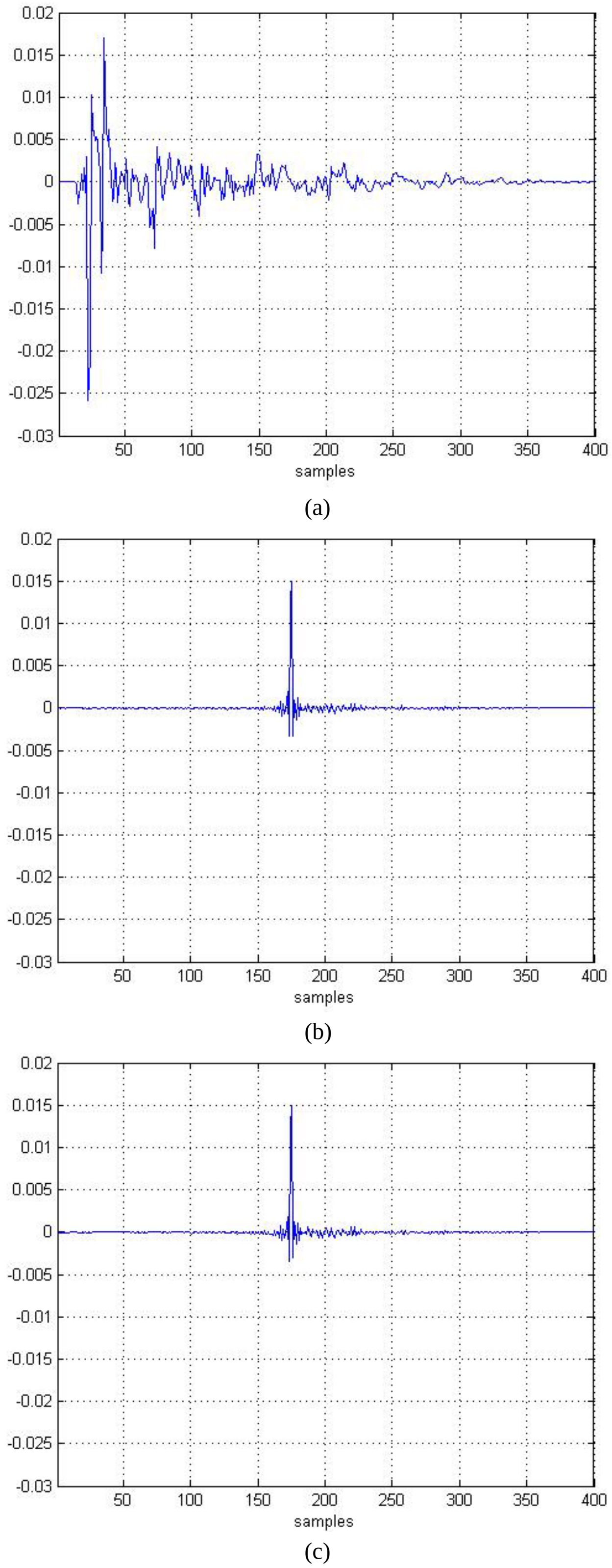

Figure 4: Impulse response between a given loudspeaker and a point within the listening area: (a) before the compensation algorithm, (b) after compensation by the fast Block Toeplitz solvers and (c) after the compensation by the fast deconvolution algorithm.

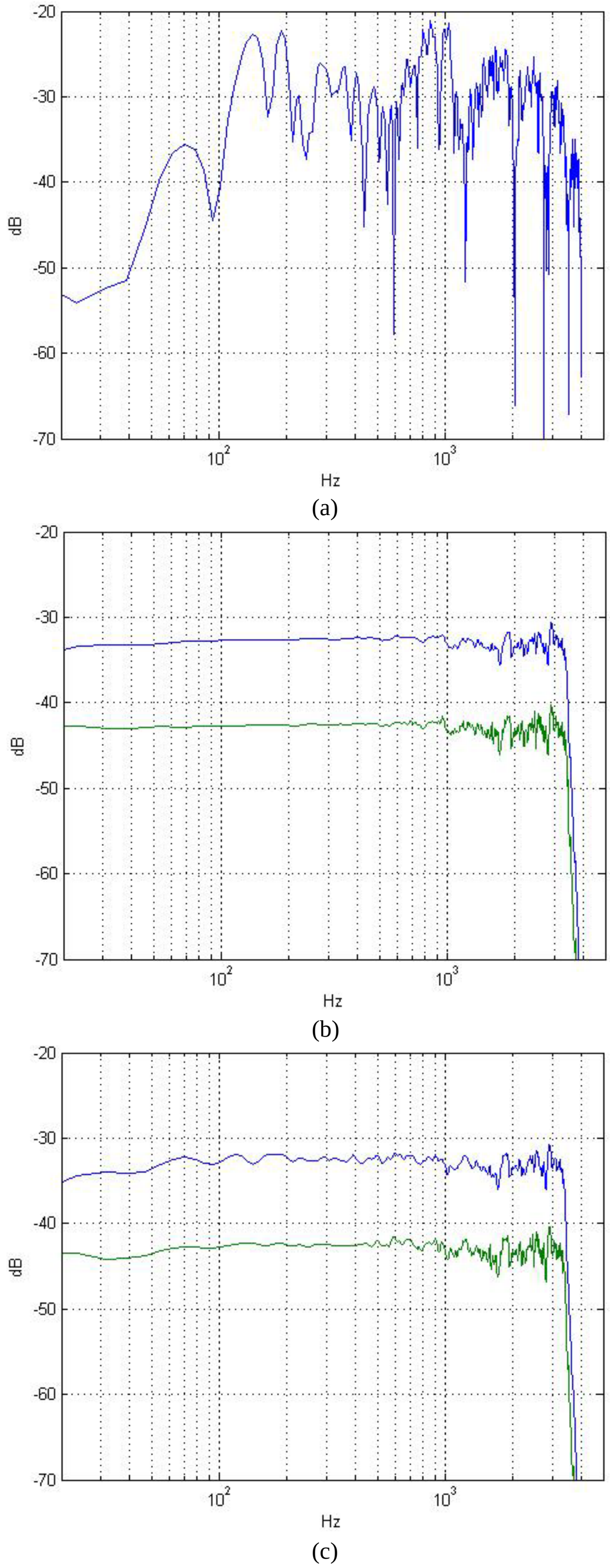

Figure 5: Frequency domain response for filters with length of 512 and 1024 samples: (a) before the compensation algorithm, (b) after compensation by fast Block Toeplitz solvers and (c) after the compensation by fast deconvolution algorithm.

In order to evaluate the system, measures were simulated at the same listening points with the bank of inverse filters working. For every virtual source the driving signal of each loudspeaker was computed with the WFS system. These driving signals were filtered through its corresponding inverse filter bank and added at each loudspeaker before being rendered. The acoustical path was simulated convolving the loudspeaker excitation signal by the previously measured acoustic channels.

3.2. Results

In order to compare the results obtained using Block Toeplitz solvers in time domain and fast deconvolution algorithm in frequency domain, a comparison of the compensation with both algorithms has been carried out. A randomly selected RIR between a loudspeaker and a control point has been taken for this purpose.

Firstly, the impulse response has been plotted before and after applying the bank of filters, figure 4a shows the IR between the loudspeaker and the microphone before the compensation algorithm. Figure 4b after applying Block Toeplitz, and figure 4c after applying fast deconvolution. Both methods show very similar compensation in time domain.

Figure 5a shows the frequency domain response before applying the compensation filter. Figure 5b represents the same response after the bank of filters computed in time domain. The upper plot shows the response with a bank of inverse filters of length of 512 samples and the lower plot for a length of 1024 samples (10 dB apart). Both responses are very similar.

In the same way, figure 5c shows the response for fast deconvolution algorithm using a regularization parameter of 0.01. For a 512 length (upper plot), the filters computed in frequency domain contain peaks and deeps in the whole range of frequencies. These peaks and deeps are due to the circular convolution effects. With a inverse filter length of 1024 samples, the results for the fast deconvolution algorithm has a plainer response than before, but it doesn't achieve the quality shown by the fast Block Toeplitz solvers with 512 samples, figure 5b.

In order to evaluate the acoustic filed rendered by the WFS system, a single source has been simulated within the listening room. Figure 6a represents the field rendered for a single source of $800Hz$ with the WFS array in free field. This response accurately simulates the produced by a real source. Figure 6b shows the same source but into the real room. The field obtained this way does not seem as perfect as the obtained in free field conditions due to the room reflections. The field rendered after applying the inverse filter bank is shown in the lower figures. Figure 6c represents the reproduced field with fast Block Toeplitz solvers. Meanwhile, the fast deconvolution algorithm is represented in figure 6d. In both cases, a very similar field to the original one is achieved.

4. CONCLUSIONS

A listening room compensation method for Wave Field Synthesis reproduction systems has been presented and validated in the present work. In contrast to other more computationally efficient methods [3] based on plane wave decomposition, this approach carries out time domain multichannel filtering to compensate for the listening room effects. At first approach, a high number of control points has been considered inside the listening room, in order to assure that compensation is achieved in a wide enough listening area. The number and relative positions of the control points is beyond the scope of the present work and is a subject of current research.

The bank of inverse filters has been calculated using two different methods. On one hand, the solution has been obtained in time domain under the criterion of the minimization of the quadratic error at all the control points, inverting the pseudoinverse using efficient algorithms for the inversion of Block Toeplitz matrices. On the other hand and approximate method called *Fast Deconvolution* has been also employed and compared with the first one. Fast deconvolution is significantly faster than solutions in time domain, because employs FFT, but needs longer filters than the first method for achieving equivalent quality filters.

The comparison of the results obtained in the experiments confirm that time domain computation avoids the negative circular convolution effects that appear in frequency domain inversion methods and allows to achieve the best filter bank for a given impulse response size under the least squares criterion. The drawback of time domain computation is the time employed in the computation of filter bank; some hours for the experiment proposed compared to seconds in the case of fast deconvolution. Anyway, the filter bank is computed only once in the set-up stage of the WFS system. If the array is not moved from its position, the filter bank remains the same.

In general the main advantages of employing room correction using MIMO inversion techniques can be summarised as follows. The use of direct solutions over the whole area of control takes into account not only the wall reflections, but also the floor and ceiling reflections, in contrast with other methods where the simplification of the problem to 2D does not guarantee the acoustic field accuracy within the whole listening area in case of severe ceiling/floor reflections. The control points within the area of interest could be placed everywhere. In this way, the control points can be adapted to the specific listening area that would be controlled. It is also possible to somewhat correct degradations due to loudspeaker's directivity, [15].

5. ACKNOWLEDGEMENTS

This work was supported by Spanish Ministry of Science and Technology (MCYT) under Projects ref. TIC2003-06841-C02-02, ref. TIC2003-08238-C02-01 and by FEDER funds.

6. REFERENCES

[1] A.J.Berkhout, "A holographic approach to acoustic control," *Journal of the Audio Engineering Society*, vol. 36, no. 12, pp. 977–995, December 1988.

[2] Edo Hulsebos, Diemer de Vries, and emmanuelle Bourdillat, "Improved microphone array configurations for auralization of sound fields by wave field synthesis," in *110th AES Convention*, Amsterdam, The Netherlands, May 2001, Audio Engineering Society, p. Prepr.5337.

[3] S. Spors, H. Buchner, and R. Rabenstein, "Efficient active listening room compensation for wave field synthesis," *116th AES Convention*, 2004.

[4] A.J. Berkhout, D. de Vries, and P. Vogel, "Acoustic control by wave field synthesis," *J. of Acoustic Society of America*, vol. Vol. 93(5), pp. 2764–2778, 1993.

[5] E.N.G.Verheijen, *Sound Reproduction by Wave Field Synthesis*, Ph.D. thesis, Delft Univeristy of Technology, 1997.

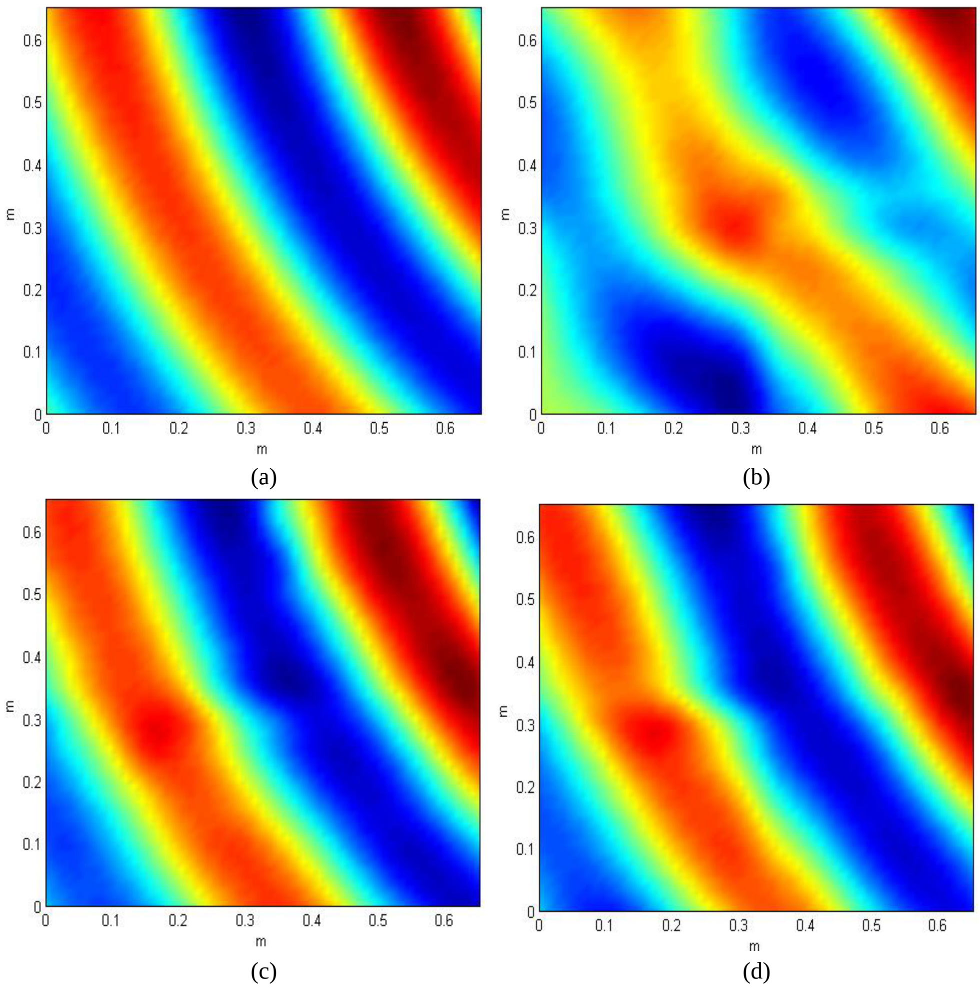

Figure 6: Rendered field by a source signal of $800Hz$: (a) original source reproduced by WFS in free field, (b) measured field in a real room, (c) field after applying the fast Block Toeplitz solvers and (d) after the compensation fast deconvolution algorithm.

[6] P.Vogel, *Application of Wave Field Synthesis in Room Acoustics*, Ph.D. thesis, Delft University of Technology, 1993.

[7] E.W.Start, *Direct Sound Enhancement by Wave Field Synthesis*, Ph.D. thesis, Delft University of Technology, 1997.

[8] P. A. Nelson, F. Orduna-Bustamante, and H. Hamada, "Multichannel signal processing techniques in the reproduction of sound," *Journal of Acoustic Society of America*, vol. 44, pp. 973–989, 1996.

[9] M. Miyoshi and Y. Kaneda, "Inverse filtering of room acoustics," *IEEE Trans. Acoustics, Speech and Signal Processing*, vol. 36, pp. 145–152, 1988.

[10] O. Kirkeby, P.A. Nelson, H. Hamada, and F. Orduna-Bustamante, "Fast deconvolution of multichannel systems using regularization," *IEEE Trans. Acoustics, Speech and Signal Processing*, vol. 6 (2), pp. 189–194, 1998.

[11] J.J. López and A. González, "Two steps levinson algorithm for time domain multichannel deconvolution," *Electronics Letters*, vol. 36 (7), pp. 686–688, 2000.

[12] P. M. Clarkson, "Optimal and adaptive signal processing," *CRC Press*, 1993.

[13] I. C. Gohberg and A. A. Semencul, "On the inversion of finite toeplitz matrices and their continuous analogs," *Mathematics Issted.*, vol. 2, pp. 201–233, 1972.

[14] A. González, P. Zuccarello, M. de Diego, and G. Pinero, "Simultaneous measurement of multichannel acoustic systems," *Journal of Acoustic Society of America*, vol. 52, pp. 26–42, 2004.

[15] D. de Vries, "Sound reinforcement by wavefield synthesis: Adaptation of the synthesis operator to the loudspeaker directivity characteristic," *Journal of the Audio Engineering Society*, vol. 44, no. 12, pp. 1120–1131, December 1996.