SOFTWARE FOR THE SIMULATION, PERFORMANCE ANALISYS AND REAL-TIME IMPLEMENTATION OF WAVE FIELD SYNTHESIS SYSTEMS FOR 3D-AUDIO

Sergio Bleda

Dept. Physics & Signal Theory
Univ. de Alicante
sergio@dfists.ua.es

José J. López

Dept. of Communications
Univ. Politécnica de Valencia
jjlopez@dcom.upv.es

Basilio Pueo

Dept. Physics & Signal Theory
Univ. de Alicante
basilio@dfists.ua.es

ABSTRACT

Wave Field Synthesis is a method for 3D sound reproduction, based on the precise construction of the desired wave field by using an array of loudspeakers. The main purpose of this work is to present a set of software tools that brings to the audio community a feasible an easy way to start working with wave field synthesis systems. First in the paper, an introduction to different 3D sound techniques and an overview of WFS theory and foundations are given. Next, a series of software tools specially developed to simulate, analyze and implement WFS systems are presented. The first software module helps the user in the design of the array of loudspeakers to be employed in the reproduction by computing the equations for each speaker signal excitation. Another tool simulates the wave field generated by the arrays and analyses both performance and quality of the acoustic field. Finally a user friendly tool for realtime convolution capable of producing the excitation signals for the array of loudspeakers is presented. Also, different experiments that have been carried out with this software in order to evaluate the precision and behaviour of different WFS configurations are presented and interpreted.

1. INTRODUCTION

The final aim of the 3D sound systems is the reconstruction of the acoustic sensations that a particular listener would perceive in a real environment. At present time, there is a strong trend towards increasing the realism of the sound reproduction systems. Both spatial sensation as re-creation of acoustic environments are sought. The simplicity of this concept hides a series of important physical and technological complications. In fact, they represent a matter of investigation and constant development in sound engineering

Multichannel sound systems, well established in the cinema industry, try to re-create these types of acoustic sensations. From the first sound rendering systems up to the present 5.1, 6.1 and 7.1 systems, a great evolution of signal processing techniques has taken place. These signal processing techniques have been used mainly in digital compression of multichannel sound, giving rise to different standards, which became the property of different companies. However, these systems, although they are suitable for cinema do not provide a true periphonic space sensation appropriate for the accurate reproduction of music, and a good localization of instruments in the sound space.

On the other hand, the systems based on the HRTF have also experienced a considerable evolution. The delivery of immersive binaural signals can be carried out through headphones or loudspeakers. Using the former, the "inside-the-head-effect" is suffered and with the latter, a reduced sweet spot (optimal listener position) is obtained, [1].

Alternatively, Ambisonics or Virtual Surround Panning are more advanced than the typical surround systems, and are adequate for less restricted areas [2][3] than the binaural systems with cross-talk cancellers. The Solution to improve the listeners area size in these systems is to rise the number of used loudspeakers, with the implied complexity and difficulty, as well as to improve the transmission formats.

Nevertheless, the most promising system nowadays is the Wave Field Synthesis (WFS), where sound field is synthesized in an extended area by means of arrays of loudspeakers.

2. WAVE FIELD SYNTHESIS FUNDAMENTALS

2.1. Principles

Wave Field Synthesis (WFS) is a sound reproduction technique that, by analogy to holography and in base to the Huygens principle, reproduces an acoustic field inside a volume from the stored signals recorded in a given surface. Huygens principle tells that the wave front radiated by a source behaves like a distribution of sources that are in the wave front, named secondary sources.

Wave field synthesis was first proposed with application to 3D sound by Berkhout [4][5]. The synthetic wave front is created by loudspeaker arrays that substitute the individual loudspeakers. We would be in the ideal situation, when we achieve an area completely surrounded by loudspeakers feed with signals that create a volumetric velocity proportional to the particle velocity normal component of the original wave front. The main advantage of these systems is the great extension of the useful listening area; since all the loudspeakers surrounded volume compose an accurate wave field reproduction zone.

In practice it is not necessary to surround completely the listener by a surface in the three dimensions, it is enough with a linear loudspeaker array located in front of the listener. In figure 1, a typical WFS configuration is presented, where a virtual sound source is synthesized by a loudspeaker array in the location of the listener, as well as in the whole room.

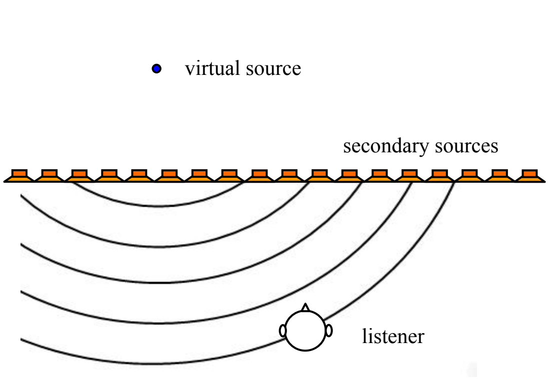


Figure 1: *Simulated wave field of a virtual source behind the loudspeaker array.*

2.2. Formulation

It is known that an arbitrary sound field within a closed volume can be generated with a distribution of monopole and dipole sources on the surface of this volume [5]. For practical purposes, this method has been adapted to make use of linear loudspeakers arrays surrounding the listening area, rather than planes of loudspeakers. Using the geometry of figure 2, the sound field created by the virtual source can be synthesized by the array of loudspeakers at the analysis point [6], according to the next expression,

$$P(\vec{r},\omega) = \sum_{n=1}^{N} \left[Q(r_n,\omega) G(\theta_n,\omega) \frac{e^{-jkr_n^{'}}}{r_n^{'}} \right] \Delta x \qquad (1)$$

where N is the number of loudspeakers in the array, ω is the angular frequency, k is the wave number, $Q(r_n,\omega)$ is the driving signal of the nth loudspeaker, ϕ_n is the angle between the main axis of the nth loudspeaker and its connection line to the analysis point, $G(\theta_n,\omega)$ is the directivity index of the nth loudspeaker, Δx is the spatial interval between the array loudspeakers, r_n is the distance between the virtual source and the nth loudspeaker, $r_n^{'}$ is the distance between the nth loudspeaker and the analysis point and $\vec{r}$ defines de analysis point.

Derivation of the driving signals for a line array of loudspeakers is found on [7] and is done by the equation,

$$Q(r_n,\omega) = S(w) \frac{\cos\theta_n}{G(\phi_n,\omega)} \sqrt{\frac{jk}{2\pi}} \sqrt{\frac{r_0}{s_0 + r_0}} \frac{e^{-jkr_n}}{\sqrt{r_n}} \qquad (2)$$

where $S(\omega)$ is the virtual source radiated signal, θ_n is the angle between the virtual source and the main axis of the nth loudspeaker, and $G(\phi_n,\omega)$ is the directivity index of the virtual source (note that this is not the same index than before). The driving signal is only a delayed version of the virtual source signal multiplied by a factor depending of the distance between virtual source and each element of the loudspeaker array r_n and the angle θ_n between them. The factor $\sqrt{jk}$ introduces an emphasization of 3 dB per octave. Also notice that the driving signal depends on the listener position. This dependence is very weak and it is a consequence of the two-dimensional approximation [6].

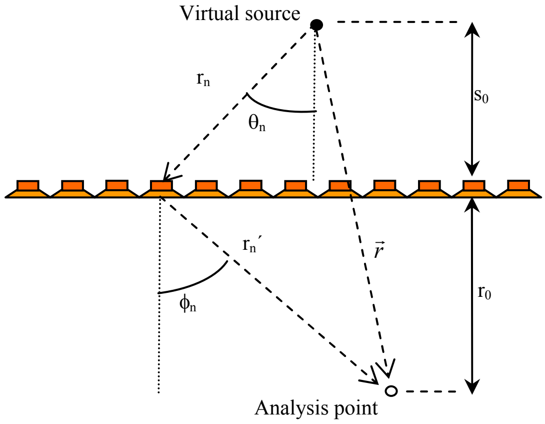


Figure 2: *Geometric description for WFS.*

2.3. Application to 3D audio systems

The knowledge of the room impulse response is needed to obtain a complete processing and representation WFS system. The signals that feed the secondary sources are extrapolated of a enough dense set of measured impulse responses. Figure 3 shows a WFS system diagram that synthesizes the wave field produced by a primary source in the simulated room.

It is not advisable to use the complete impulse response for the whole simulated points. It is a better idea to use the direct sound and the first echoes (10.000 samples at 44.1 kHz approx.), then rest of the impulse response is simulated with a reverberator with the typical parameters of the simulated room.

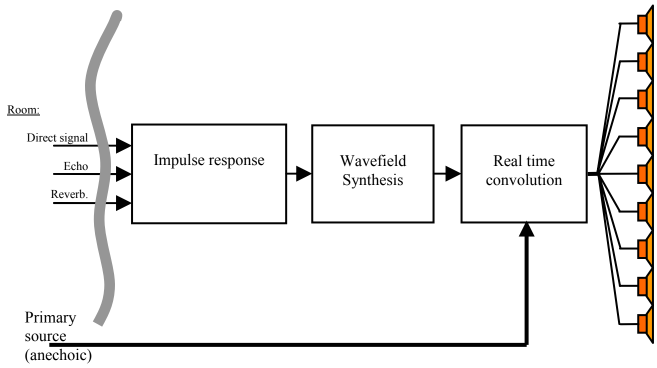


Figure 3: *Typical configuration of a 3D sound WFS system*

3. SOFTWARE TOOLS

A series of software tools to simulate, analyze and implement WFS systems have been developed. This software helps us in the development and improvement of WFS systems. The process for building a successful prototype for WFS is quite complex and this software carries out all the stages. It comprises three main modules:

- Array design and driving signals derivation
- Simulation of the sound field
- Real-time playing tool

The first module helps the user in the design of the array of loudspeakers to be employed in the reproduction by computing the equations of signal excitation for each loudspeaker. The software takes into account non ideal parameters of the loudspeakers employed for building the array, as the variation of its sensitivity with frequency and the directive characteristic [7]. Different array shapes can be chosen: linear, in U, in open U, etc. The second module simulates the wave field generated by the arrays and analyse both performance and quality of the acoustic field. Figure 4 shows a block diagram of these two stages. These two modules have been programmed in Matlab. Figure 5 shows a GUI for access to the programmed Matlab functions easily.

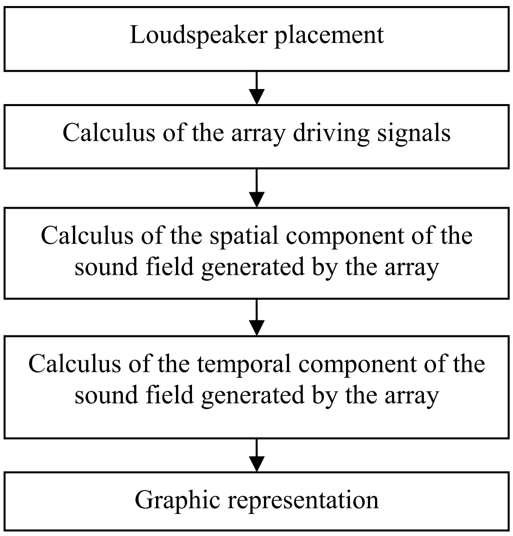


Figure 4: *Block diagram of the algorithm used in the simulation software for WFS.*

In the other hand, the real-time playing tool consists of a user friendly application with a window graphic user interface (GUI). It helps the user to place different instruments or sound sources in the stage producing the excitation signals for the loudspeakers. This application includes a real-time multichannel convolver, which is capable of producing the excitation signal for arrays of up to 96 loudspeakers and generating the output signals to the sound cards attached to the computer. The impulse responses for each loudspeaker are provided by the previous module written in Matlab.

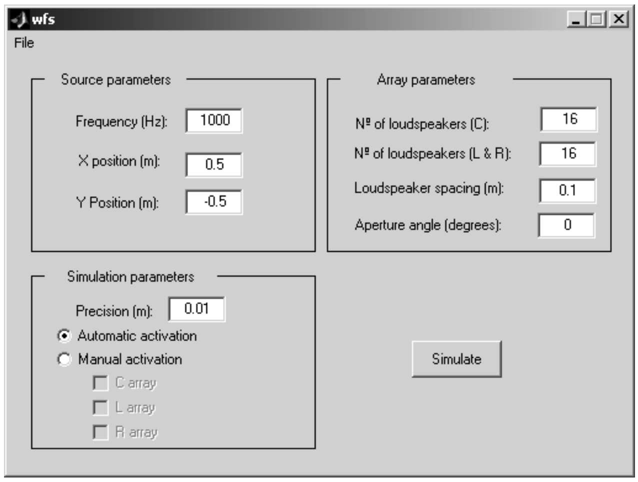


Figure 5: *Graphic user interface for easy access to the simulation algorithm written in MATLAB.*

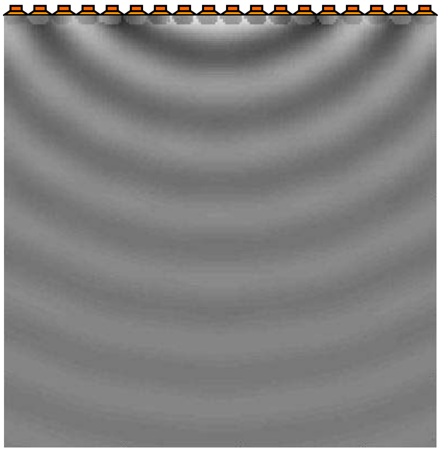
The software accepts the sound of primary sources from the inputs of the sound cards or from sound files stored on the HD or CD. This convolver supports impulses responses long enough to simulate room impulse responses. The experience gained on previous works [8] on real-time convolvers has been useful in the development of this powerful tool. It has been programmed in C++Builder / Delphi for MS-Windows OS and a port to Linux is in progress.

The hardware infrastructure used to achieve real-time rendering consists of a single PC (Pentium 4 2.4 GHz, 1GB RAM), a PCI-424-based MOTU Audio System and up to four MOTU 24 I/O sound cards, depending of the number of channels needed. All the sound card programming is done via the Steinberg Audio Stream I/O API (ASIO).

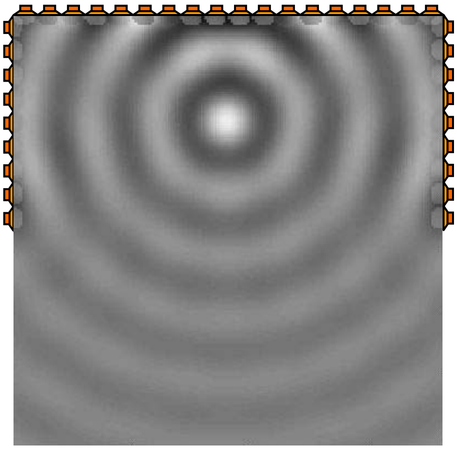
4. SIMULATIONS

Using the developed software, it is possible to simulate the acoustic field produced in the listening area. In order to show an example of the possibilities of this software, the field produced by a sinusoidal virtual source is shown on figure 6. The figures show a snapshot of the acoustic field in an extended area in a time instant, as similar made in previous works [9].

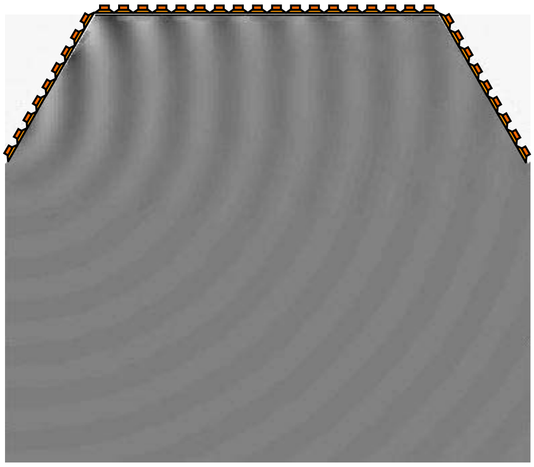
The purpose of the simulation presented in figure 6 is to compare different array shapes. In the three cases the virtual source has a frequency of 650 Hertz and the spacing between each loudspeaker is 20 cm. In a) a simple linear array is used and the virtual source in located just behind the array. In b) a configuration of three arrays in U is used and the source is in front of the central array. In c) a configuration in open U with 45º aperture is used and the source is located behind the left array.



a)



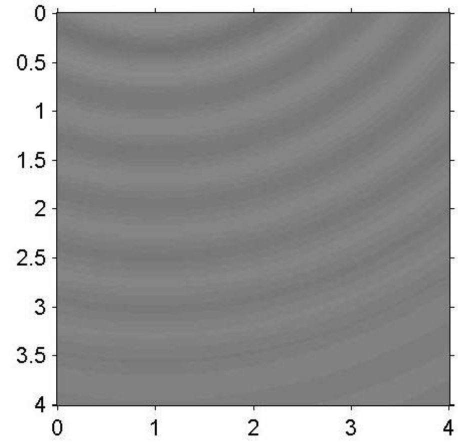
b)



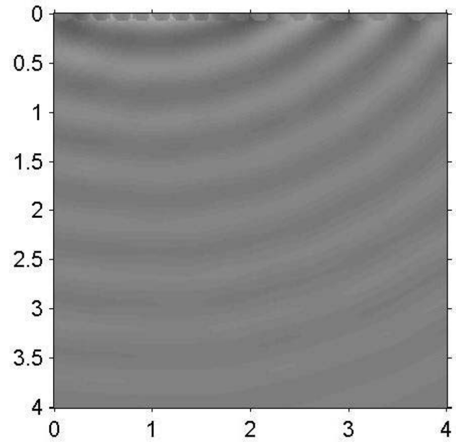
c)

Figure 6: *Simulation of the acoustic field produced by three diferent array configuration and with different virtual source positions.*

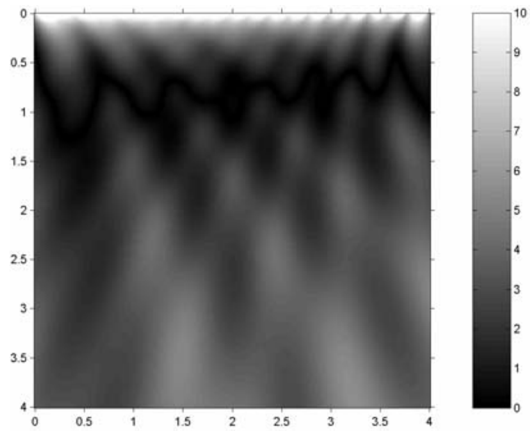
In figure 7 another type of simulation is presented. The pressure error between the real field that would produce a real source a) is compared with the one produced by the array of loudspeakers. In c) the pressure error in dB is shown. Looking at figure 7c is easy to realize that in the areas very close to the loudspeakers, the error is very high. This is due to the near field effect of loudspeakers. Also it can be observed that in moving far away from the array the error also increases.



a)



b)



c)

Figure 7: *Comparison between the field produced by a real source a) and the one produced by a WFS system b). Error in dB between them c).*

5. PRACTICAL CONSIDERATIONS

The use of more than one line of loudspeakers array, particularly the use of side arrays in WFS systems has been discussed in different papers, for example in [6]. Next, some practical considerations about these multi array set-ups are going to show using our simulation software and some examples and results.

When performing WFS reconstructions of the pressure field, the simultaneous activation of all the line arrays available for each virtual source in not always the best solution for obtaining the best reliable field synthesis. As a general rule, only de arrays with the virtual source behind them must be activated. In order to show clearly this question, the figures 8 and 9 show two typical set-ups of WFS arrays, consisting on three line arrays each one. Figure 8 shows a set-up in U with a frontal array and two side arrays. Figure 9 shows a set-up in open-U that is useful for creating wider areas. The area surrounding the loudspeakers strips has been divided in 6 zones regarding the position of the virtual source to be synthesized. The purpose is to connect the zone when the virtual source is with the loudspeaker strips to be activated.

Some simulations have been carried on in order to verify the best excitation. Figure 10 compares the error obtained when a source is placed in zone 2. Figure 10 a) shows the error when activating only the C array and b) shows the error when activating the three arrays. It is clear that the side arrays make the synthesis worse, so they must be turned off. In the figure 11 the simulation is repeated but with a source in the zone 1. In 11 a) only L and C arrays are activated compared with b) when all three are. Likewise the results are clear.

There are three main configurations, according to physical symmetry:

- Zones 2, 4 and 6: Virtual sources are located behind a single array.
- Zones 1 and 3: Virtual sources are located behind two arrays simultaneously.
- Zone 5: Virtual sources are located in front of the three arrays.

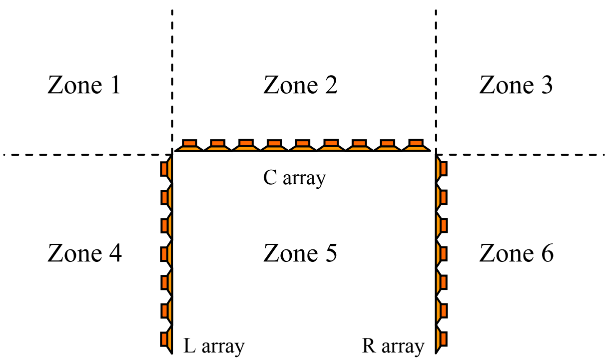


Figure 8: Closed U virtual source zones.

As expected, any array located in front of the source must be activated. On the contrary, arrays behind the source must be deactivated. The exception to this rule occurs in the zone 5, where the source is in front of all arrays. In this situation, at least one array must be activated or no wave field will be produced, the first thought is to activate all three arrays because the virtual source is in the same relative position for all. In Table 1 the combination of zones and arrays is presented.

Zone	1	2	3	4	5	6
Arrays	L, C	C	C, R	L	L, R, C *	R

Table 1: *Array activation by source location.*

Placing virtual sources in zone 5 is particularly complicated, but is possible to simulate sources located between the listener and the frontal array. Despite the first thought would be to activate the three arrays in order to simulate sources in this zone, our simulations show that this is not true. In order to achieve better results the three arrays must not be activated simultaneously. The zone 5 must be subdivided in two new zones, the left and the right parts, 5a and 5b respectively. Only L and C arrays must be activated in zone 5a, and only R and C in zone 5b. An example that demonstrates this behaviour is presented in figure 12. A look to the figure is enough to check the proposed rule. Same considerations must be applied to an open U array configuration (figure 9).

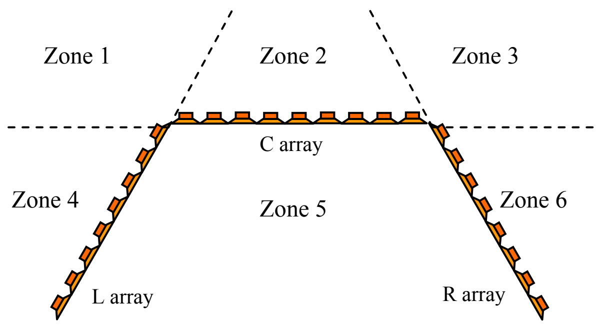
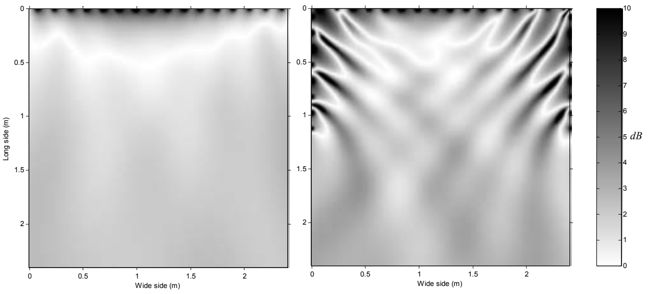


Figure 9: *Open U virtual source zones.*

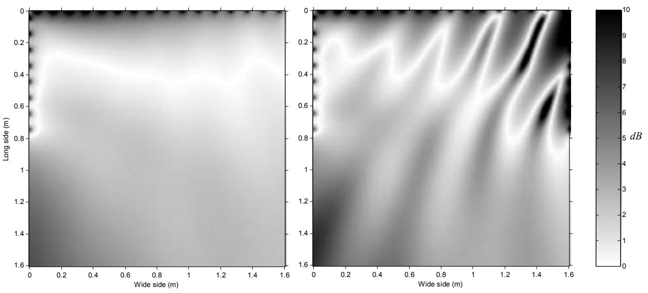
It must take into account that the arrays can not be activated or deactivated instantly when the virtual source moves from one zone to another, since the ear could detect the wave field origin change. Thus the array activation/deactivation must be gradual with the virtual source displacement.

This consideration is implicit in the loudspeaker driving signal, since the driving signal uses the angle formed between the virtual source and the main loudspeaker axis. This angle is 90º when the virtual source changes of zone and a new array must be activated/deactivated, then the array has no influence in the generated wave field until the virtual source has went deeper in the new zone.



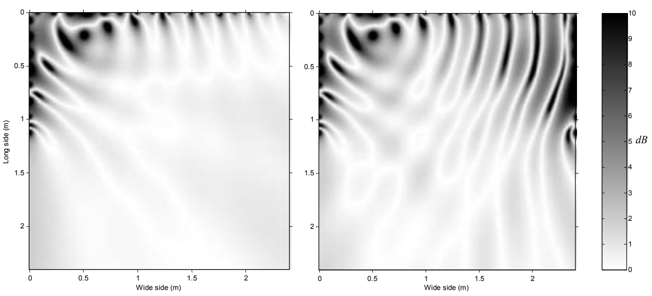
a) b) .

Figure 10: Error simulating a virtual source in zone 2 with C array (left) and with the three arrays (right).



a) b) .

Figure 11: Error simulating a virtual source in zone 1 with L and C arrays a) and with the three arrays b).



a) b) .

Figure 12:- Error simulating a virtual source in zone 5a with L and C arrays a) and with the three arrays b).

6. CONCLUSIONS

A software tool to carry out design, simulation and real-time implementation stages on WFS has been presented. This tool provides to us the platform for continuing researching and testing the WFS method for immersive sound applications and brings to the audio community a feasible an easy way to start working with wave field synthesis systems.

Also, different experiments have been carried out with this software as examples of its functionality. Finally some practical considerations regarding the correct activation of the different arrays strips of a WFS set-up have been discussed.

The developed simulation software in Matlab and the real-time convolver is available for download at the next internet address: http://www.gaips.upv.es/wfs.htm.

7. REFERENCES

[1] J.J. Lopez, A. González, "Measurement of cross-talk cancellation and equalization zones in 3-D sound reproduction under real listening conditions", proc. of the AES 16th Int. Conf. on Spatial Sound Reproduction, Rovaniemi, Finland., 1999

[2] J. Daniel et al., "Ambisonics Encoding of Other Audio Formats for Multiple Listening Conditions", Proc. 105th AES Conference, San Francisco 1998, preprint 4795

[3] U. Horbach and M. Boone, "Future Transmission and Rendering Formats for Multichannel Sound", Proc. 16th AES Conference on Spatial Sound Reproduction, Finland, 1999

[4] A. J. Berkhout, "A Holographic Approach to Acoustic Control", J. Audio Eng. Soc., 36: 977-995, 1988

[5] A. J. Berkhout, D. de Vries, P Vogel, "Acoustic Control by Wave field Synthesis", J. Acoust. Soc. Am., vol 93, pp. 2765-2778, 1993

[6] M.M. Boone, E.N.G. Verheijen, P.F. van Tol, P.F. "Spatial sound field reproduction by wave field synthesis", J.Audio Eng.Soc. vol 43, pp 1003-1012, 1995

[7] D. de Vries, "Sound Reinforcement by Wave field Synthesis: Adaptation of the Synthesis Operator to the Loudpeaker Directivity Characteristics", J. Audio Eng. Soc., vol 44, pp. 1120-1131, 1996

[8] J.J. Lopez, A. González, "PC Based Real-Time Multichannel convolver for Ambiophonic Reproduction", proc. of the AES 19th Int. Conf., Schloss Elmau, Germany, 2001

[9] U. Horbach et al., "Numerical Simulation of Wave Fields created by Loudspeaker Arrays", 107th AES Convention, New York 1999, preprint 5021