Audio Engineering Society

Convention Paper 8244

Presented at the 129th Convention
2010 November 4–7 San Francisco, CA, USA

Discrete Driving Functions for Horizontal Reproduction using Wave Field Synthesis and Higher Order Ambisonics

César D. Salvador[1]

[1] Universidad de San Martín de Porres, ISONAR, Surquillo, Lima 34, Perú
csalvador@comunicaciones.usmp.edu.pe

ABSTRACT

Practical implementations of physics-based spatial sound reproduction techniques, such as Wave Field Synthesis (WFS) and Higher Order Ambisonics (HOA), require real-time filtering, scaling and delaying operations on the audio signal to be spatialized. These operations form the so-called loudspeaker's driving function. This paper describes a discretization method to obtain rational representations, in the unit circle of the *z*-plane, of the continuous 2.5D-WFS and 2D-HOA driving functions for horizontal reproduction. Visual and numerical comparisons between continuous and discrete driving functions, and between continuous and discrete sound pressure fields, synthesized with circular loudspeaker arrays, are shown. Percentage discretization errors, in the reproducible frequency range and in the whole listening area, are on the order of 1%. A methodology for immersive soundscapes reconstruction composed with nature sounds is also reported as a practical application.

1. INTRODUCTION

Spatial sound reproduction techniques can be classified into those mainly based on the listener's perceptual qualifications (psychoacoustics) and those mainly based on the physical modeling of the propagating sound pressure field (acoustics). Among perceptual methods, there are those that control the gains of each audio channel, such as VBAP; and also those that introduce further delays between audio channels, such as binaural spatialization. Alternatively, most current methods appeal to correct physical modeling, among which there are Wave Field Synthesis (WFS) and Higher Order Ambisonics (HOA). WFS and HOA are actually emerging as optimal formats for spatialization of virtual auditory scenes that look for immerse a listener in an almost real acoustic environment, synthesizing wave fronts with physical methods and rendering them through loudspeaker arrays.

The WFS theoretical framework based on the Huygens principle was initially formulated by Berkhout et al. [1], [2]. The HOA theoretical framework based on spatial harmonics expansions have been widely studied by Daniel [3]. Both techniques are based on the Kirchhoff-Helmholtz integral, which models the propagation of

sound as a wave inside a large listening area surrounded with loudspeaker arrays. WFS and HOA allows to synthesize virtual acoustical environments by rendering room impulse responses with plane wave fronts, as well as to synthesize virtual sources that appear to emanate from a defined position by rendering them with spherical wave fronts. Thus, it provides the listener with consistent spatial localization cues over large listening areas, using a high number of loudspeakers, which in this context are called secondary sources.

Although the most widespread application of WFS an HOA are the mentioned above, the motivation of this work arose in the immersive soundscapes composition field. The concept of soundscapes, introduced by Schafer [4], refers to a combination of sounds that forms or arises from natural or artificial acoustic environments. Most recently, for modeling purposes, soundscapes have been defined by Valle et al. [5] as "a temporal and typological organization of sound objects, related to a certain geo-cultural context, in relation to which a listener can apply a spatial and semiotic transformation." Thus, when listening soundscapes, it can be distinguished: soundmarks (a sound that identifies a community), atmospheres (an overall layer of sound which cannot be decomposed), and events (an isolated single sound of well-defined boundaries). Since atmospheres are perceived as coming from non-localized sources, and soundmarks and events as coming from localized sources, their reproduction model can be done respectively with plane and spherical wave models. Within this framework, it will be reported at the end of this paper a practical example of immersive soundscapes composed with nature sounds, which are rendered using Pure Data and multichannel audio equipment.

The common theoretical base for WFS and HOA has been under consideration by Daniel et al. [6] and Ahrens et al. [7] during the last years. Practical implementations of both techniques require pre-equalization, filtering, scaling and delaying operations on the audio signal to be spatialized. Except for prefiltering, these operations usually need to be computed in real-time, according to the position of each secondary source and the position (spherical wave) or direction of propagation (plane wave) of the virtual source. These operations form the so-called loudspeaker's driving function.

It has been shown that adaptive filters are required for improving WFS and HOA in real scenarios for active listening room compensation [8]. It is also known that most digital signal processing environments include FIR and IIR filters in their libraries. Hence, in order to provide a suitable representation to include H^{∞} optimization [9] and to improve real-time implementations, this paper describes how to obtain rational representations of driving functions, in the unit circle of the z-plane, starting from the continuous 2.5D-WFS and 2D-HOA driving functions, previously derived by Spors et al. [10] and Ahrens et al. [7], respectively.

This paper is organized as follows: Section 2 briefly introduces the continuous WFS and HOA driving functions. Section 3 describes the discretization of these continuous driving functions with special attention on prefiltering and delaying for WFS, and on spherical Hankel functions for HOA. Section 4 presents the simulations of the sound pressure fields synthesized with both continuous and discrete driving functions, as well as corresponding discretization errors. Section 5 presents their application to immersive soundscapes composition, which was the original motivation for the present work. Section 6 resumes the goals of the present work, giving conclusions and ongoing work as well.

2. CONTINUOUS DRIVING FUNCTIONS

2.1. The Kirchoff-Helmholtz integral

The sound pressure field $P(\mathbf{x}, \omega)$ inside a listening area enclosed by a contour ∂V, with a uniform distribution of secondary sources at positions $\mathbf{x}_0$ along this contour, can be expressed, in terms of the secondary source driving functions $D(\mathbf{x}_0, \omega)$ and the Green's functions of the monopoles located at $\mathbf{x}_0$, as the Kirchoff-Helmholtz integral:

$$P(\mathbf{x}, \omega) = -\frac{1}{4\pi} \oint_{\partial V} D(\mathbf{x}_0, \omega) \frac{e^{-j\frac{\omega}{c}|\mathbf{x}-\mathbf{x_0}|}}{|\mathbf{x}-\mathbf{x_0}|} dS_0. \qquad (1)$$

This integral is the common theoretical approach used in both WFS and HOA for deriving the driving functions using spectral division methods and also for calculating sound pressure fields.

2.2. Continuous 2.5D-WFS driving functions

The geometry used in the WFS theory revisited by Spors et al. in [10] is illustrated in figure 1. According to it, the wave field emanating from the virtual source at $\mathbf{x}_s = [\mathrm{x}_s \ \mathrm{y}_s]^{\mathrm{T}}$ can be synthesized in the listening area enclosed by an arbitrarily shaped contour ∂V using

loudspeakers at $\mathbf{x}_0 = [\mathrm{x}_0\ \mathrm{y}_0]^\mathrm{T}$ along this contour as secondary monopole sources. The normal vector to ∂V at $\mathbf{x}_0$ is the column $\mathbf{n}(\mathrm{x}_0)$ and the reference position is $\mathbf{x}_\mathrm{ref} = [\mathrm{x}_\mathrm{ref}\ \mathrm{y}_\mathrm{ref}]^\mathrm{T}$.

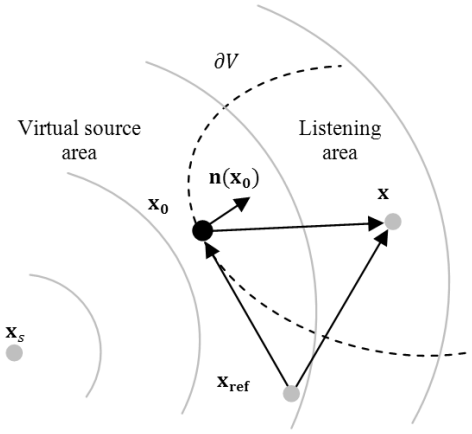

Figure 1: The geometry used in wave field synthesis.

The 2.5D driving function of a loudspeaker at $\mathbf{x_0}$, for the synthesis of a plane wave with spectrum $S_p(\omega)$, $S_P(\mathbf{x},\omega) = S_p(\omega)\exp(-j\omega\mathbf{n}_p^T\mathbf{x}/c)$, propagating in the $\mathbf{n}_p$ direction, is

$$D_p(\mathbf{x_0},\omega) = -2a_p(\mathbf{x_0})\sqrt{2\pi|\mathbf{x}_\mathrm{ref}-\mathbf{x}_0|}\,\mathbf{n}_p^\mathrm{T}\mathbf{n}(\mathrm{x}_0) \times \tfrac{1}{\sqrt{c}}(j\omega)^{1/2}e^{-j\omega\frac{\mathbf{n}_p^T\mathbf{x}_0}{c}}S_p(\omega), \quad (2)$$

where, $|\cdot|$ denotes the Euclidian norm, c the velocity of sound in air, and the selection loudspeaker functions is

$$a_p(\mathbf{x_0}) = \begin{cases} 1 & \text{, if } \mathbf{n}_p^T\mathbf{n}(\mathbf{x_0}) > 0, \\ 0 & \text{, otherwise.} \end{cases} \quad (3)$$

The 2.5D driving function of a loudspeaker at $\mathbf{x_0}$, for the synthesis of the non focused spherical wave $S_s(\mathbf{x},\omega) = S_s(\omega)\exp(-j\omega|\mathbf{x}-\mathbf{x}_s|/c)/|\mathbf{x}-\mathbf{x}_s|$, with spectrum $S_s(\omega)$ emanating from $\mathbf{x}_s$, is

$$D_s(\mathbf{x_0},\omega) = -2a_s(\mathbf{x_0})\sqrt{2\pi|\mathbf{x_{ref}}-\mathbf{x_0}|}\,\frac{(\mathbf{x_0}-\mathbf{x}_s)^T\mathbf{n}(\mathbf{x_0})}{|\mathbf{x_0}-\mathbf{x}_s|^2} \times \tfrac{1}{\sqrt{c}}\left(j\omega+\tfrac{c}{|\mathbf{x_0}-\mathbf{x}_s|}\right)(j\omega)^{-1/2}e^{-j\omega\frac{|\mathbf{x_0}-\mathbf{x}_s|}{c}}S_s(\omega), \quad (4)$$

where the loudspeaker selection function is:

$$a_s(\mathbf{x_0}) = \begin{cases} 1 & \text{, if } (\mathbf{x_0}-\mathbf{x}_s)^T\mathbf{n}(\mathbf{x_0}) > 0, \\ 0 & \text{, otherwise.} \end{cases} \quad (5)$$

The rendering of sources positioned in between the loudspeakers and the listener, called focused sources, can be derived using a complex conjugate version of equation (1) [11]. This is possible thanks to the time-reversal invariance of the wave equation: "for each burst of sound diverging from a source, there exists a set of waves that retraces its paths and converges simultaneously at the original source site as if time were running backwards" [12], [13]. Here, it is used a modified version of the driving functions proposed in [14] and [15]:

$$D_{fs}(\mathbf{x_0},\omega) = \sqrt{\frac{|\mathbf{x_{ref}}-\mathbf{x_0}|}{|\mathbf{x_{ref}}-\mathbf{x_0}|+|\mathbf{x_0}-\mathbf{x}_s|}}\,\frac{(\mathbf{x_0}-\mathbf{x}_s)^T\mathbf{n}(\mathbf{x_0})}{|\mathbf{x_0}-\mathbf{x}_s|^{3/2}} \times \tfrac{-\mathrm{j}}{\sqrt{2\pi c}}(j\omega)^{1/2}e^{j\omega\frac{|\mathbf{x_0}-\mathbf{x}_s|}{c}}\hat{S}_s(\omega). \quad (6)$$

2.3. Continuous 2D-HOA driving functions

The spatial coordinate system used by Ahrens et al. in [7] for deriving the continuous HOA driving functions for plane wave reproduction is illustrated in figure 2. In the wave number domain, $k = \omega/c$ corresponds to r, the angle θ corresponds to α, and ϕ corresponds to β.

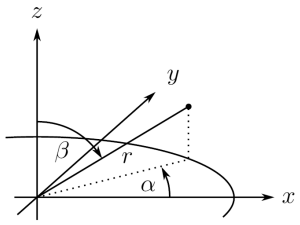

Figure 2: The spatial coordinate system used in HOA.

Ambisonics systems are frequently restricted to reproduction in the horizontal plane ($\beta = 0$). The secondary sources are arranged on a circle of radius r_0 at discrete angles α_0. The 2D driving function for horizontal reproduction of a plane wave with spectrum $S_p(\omega)$, traveling in the direction of the angle θ_PW, using a circular array of loudspeakers uniformly distributed at positions (α_0, r_0), is:

$$D_p^{HOA}(\alpha_0, r_0, \omega) = S_p(\omega)\frac{2j}{kr_0}\sum_{m=-\frac{L-1}{2}}^{\frac{L-1}{2}}\frac{(-j)^{|m|}}{h_{|m|}^{(2)}(kr_0)}e^{jm(\alpha_0-\theta_\mathrm{PW})}, \quad (7)$$

with $h_n^{(2)}(\cdot)$ being the $n-$th order spherical Hankel function of second kind and L being and odd number of loudspeakers.

The next section deals with discretization of the WFS and HOA reproduction models. Space discretization is mainly done on equation (1) and time discretization on equations (2), (4), (6) and (7).

3. DISCRETIZATION

3.1. Space discretization

Although the theory presented assumes a spatial continuous distribution of secondary sources, practical implementations of WFS and HOA will consist of secondary sources that are placed at spatially discrete positions. Let suppose that discrete-time driving functions $D(\mathbf{x}_0, z)$ have already been obtained at a sample rate f_s. When a finite set of loudspeakers are uniformly distributed along a circumference, the discrete space formulation of equation (1), for continuous and discrete time, are stated as follows:

$$P_{cont-time}(\mathbf{x}, \omega) = -\frac{\Delta\mathbf{x_0}}{4\pi}\sum_{l=0}^{L} D(\mathbf{x}_{0,l}, \omega)\frac{e^{-j\omega\frac{|\mathbf{x}-\mathbf{x}_{0,l}|}{c}}}{|\mathbf{x}-\mathbf{x}_{0,l}|}, \quad (8)$$

$$P_{disc-time}(\mathbf{x}, z) = -\frac{\Delta\mathbf{x_0}}{4\pi}\sum_{l=0}^{L} D(\mathbf{x}_{0,l}, z)\frac{z^{-f_s\frac{|\mathbf{x}-\mathbf{x}_{0,l}|}{c}}}{|\mathbf{x}-\mathbf{x}_{0,l}|}, \quad (9)$$

where $\mathbf{x}_{0,l}$ is the position of the lth loudspeaker and $\Delta\mathbf{x_0}$ is the distance between two adjacent loudspeakers that defines the spatial aliasing frequency $f_a = c/\Delta x_0$. Incoming simulations of sound pressure fields inside the listening area in section 4, with contours chosen to be line, square and circular arrays, are calculated with these equations. They will allow for visual comparisons on the following discrete time version of continuous driving functions. It is important to mention that for HOA, the Green functions of the type $\exp(-j\omega d/c)/d$ are expressed in spherical coordinates before applying equations (8) and (9). All percentage errors between continuous and discrete-time wave field synthesis are computed using:

$$e_r(P_1, P_2) = \frac{|P_1 - P_2|_\infty}{|P_1|_\infty} \times 100\%, \quad (10)$$

where $|\cdot|_\infty$ denotes the L^∞ norm.

3.2. Time discretization

A discretization in time of 2.5D WFS driving functions for the rendering of plane, focused spherical and non-focused spherical waves, using monopole secondary sources, requires special attention to the prefiltering and delaying stages. The pre-filter stage $(j\omega)^\alpha$ involves half-order systems as a consequence of using the asymptotic expansion of the Hankel function for large arguments in the 2D Green function [16], which allows to use a scaled version of the 3D Green function. The time-delay stage involves fractional sample delays due to rational multiples of the sample time.

A discretization in time of 2D HOA driving functions for the rendering of plane waves, using monopole secondary sources, requires special attention to spherical Hankel functions of second kind. A series expansion representation of such function will allow to equation (7) be represented by rational funtions of $j\omega$.

In general, two continuous to discrete mappings will be used, both at a sampling rate f_s. For filters and prefilters in 2.5D-WFS, and Hankel functions in 2D-HOA, all of them represented by rational functions, the following Al-alaoui approximation is used

$$j\omega \approx \frac{8f_s}{7}\frac{1 - z^{-1}}{1 + \frac{1}{7}z^{-1}}. \quad (11)$$

This continuous to discrete map corresponds to an interpolation between the Euler and Tustin maps [17], [18]. This map has been chosen due to its constant magnitude discretization error (see figure 3).

As continuous delays in WFS appears as exponential functions, the following standard mapping is used on them

$$z = \exp(j\omega/f_s). \quad (12)$$

In all cases, the objective is to find rational functions of z which can be programmed as IIR or FIR filters. Next section starts with the discretization of the WFS one-zero filter in equation (4), follows with the discretization of WFS fractional prefilters and delays, and finishes with the discretization of HOA spherical Hankel functions of the second kind.

3.2.1. WFS filter for non-focused spherical wave

The Al-alaoui approximation in (11) of the variable zero $(j\omega + c/|\mathbf{x_0} - \mathbf{x}_s|)$ in the non-focused spherical wave driving function of equation (4) leads to:

$$F(z) = \left(\frac{8f_s}{7} + \frac{c}{|\mathbf{x_0}-\mathbf{x}_s|}\right)\frac{1-\frac{|\mathbf{x_0}-\mathbf{x}_s|-\frac{c}{8f_s}}{7c}z^{-1}}{|\mathbf{x_0}-\mathbf{x}_s|+\frac{7c}{8f_s}}{1+\frac{1}{7}z^{-1}}. \quad (13)$$

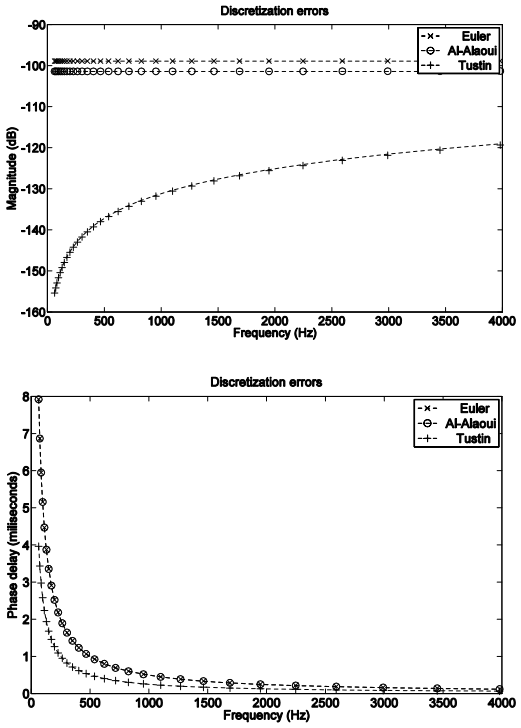

Figure 3: Continuous to discrete maps.

3.2.2. WFS fractional order prefilters

The factor $(j\omega)^{0.5}$ on equations (2) and (6) shows that the rendering of plane and focused spherical waves requires half order differentiation. The factor $(j\omega)^{-0.5}$ on equation (4) shows that the rendering of non-focused spherical waves requires half order integration. The continuous to discrete Al-alaoui approximation of $(j\omega)^{\alpha}$, using equation (11) at a sampling rate f_s, can be expanded into a power series as follows

$$(j\omega)^{\alpha} \approx \left(\frac{8f_s}{7}\frac{1-z^{-1}}{1+\frac{1}{7}z^{-1}}\right)^{\alpha} \approx \sum_{k=0}^{K} h_k z^{-k}. \qquad (14)$$

Following the Grünwald-Letnikov approach for fractional integration and differentiation [19], we can compute the Taylor series of the middle term. The coefficients of the $K+1$ order polynomial in z now is

$$h_k = \left(\frac{8f_s}{7}\right)^{\alpha} \sum_{j=0}^{k} (-1)^j \left(\frac{1}{7}\right)^{k-j} \binom{\alpha}{j}\binom{-\alpha}{k-j} \qquad (15)$$

where $\binom{\alpha}{j} = \frac{\Gamma(\alpha+1)}{\Gamma(j+1)\Gamma(\alpha-j+1)}$ is the generalization of the binomial function for rational numbers, and $\Gamma(\alpha) = \int_0^{\infty} t^{\alpha-1}e^{-t}dt$ is the gamma function: an interpolation that extents the factorial function to rational numbers. Since the later polynomial representation corresponds to a very large FIR filter, it is preferable to transform it into a short IIR filter as follows:

$$(j\omega)^{\alpha} \approx \sum_{k=0}^{K} h_k z^{-k} = \frac{\sum_{k=0}^{m} b_k z^{-k}}{1+\sum_{k=1}^{n} a_k z^{-k}} = g\frac{(1-\zeta_1 z^{-1})(1-\zeta_2 z^{-1})\dots(1-\zeta_m z^{-1})}{(1-\pi_1 z^{-1})(1-\pi_2 z^{-1})\dots(1-\pi_n z^{-1})}. \qquad (16)$$

To compute the numerator and denominator from h_k the Shank's method for least squares approximation is used according to [20] and [21]. The idea is to interpret b_k as the convolution product $(h * a)_k$. Once its associated Toeplitz matrix is identified, its lower half matrix is taken. Then $\mathbf{a} = [a_1 \quad a_2 \quad \dots \quad a_n]^{\mathbf{T}}$ is first computed from $\mathbf{h_2} = [h_{m+1} \quad h_{m+2} \quad \dots \quad h_{N-1}]^{\mathbf{T}}$, with the following pseudo inverse matrix:

$$\mathbf{a} = -\left(\mathbf{H}_2^T\mathbf{H}_2\right)^{-1}\mathbf{H}_2^T\mathbf{h}_2, \qquad (17)$$

where

$$\mathbf{H_2} = \begin{bmatrix} h_m & h_{m-1} & \cdots & h_{m-n+1} \\ h_{m+1} & h_m & \dots & h_{m-n+2} \\ \vdots & \vdots & \ddots & \vdots \\ h_{N-2} & h_{N-3} & \dots & h_{N-n-1} \end{bmatrix}.$$

The next step is to compute the impulse response of the filter $1/(1+\sum_{k=1}^{n} a_k z^{-k})$, denoted $g_k = \delta_k - \sum_{l=1}^{n} a_l g_{k-l}$, $k = 0,1,\dots,N-1$. With this impulse response, a new pseudo inverse matrix appears to compute $\mathbf{b} = [b_0 \quad b_1 \quad \dots \quad b_m]^{\mathbf{T}}$ from $\mathbf{h} = [h_0 \quad h_1 \quad \dots \quad h_{N-1}]^{\mathbf{T}}$ as follows:

$$\mathbf{b} = (\mathbf{G}^{\mathbf{T}}\mathbf{G})^{-1}\mathbf{G}^{\mathbf{T}}\mathbf{h}, \qquad (18)$$

where

$$\mathbf{G} = \begin{bmatrix} g_0 & 0 & \cdots & 0 \\ g_1 & g_0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ g_{N-1} & g_{N-2} & \dots & g_{N-m-1} \end{bmatrix}.$$

Equations (14), (16) and (17) allows to compute the half order prefilters using $\alpha = 0.5$ for plane and focused spherical waves and $\alpha = -0.5$ for non focused spherical waves. Thus, prefilters can be expressed as

$$P_\alpha(z) = g\frac{(1-\zeta_1 z^{-1})(1-\zeta_2 z^{-1})\dots(1-\zeta_m z^{-1})}{(1-\pi_1 z^{-1})(1-\pi_2 z^{-1})\dots(1-\pi_n z^{-1})} \quad (19)$$

Next section deals with fractional delays for WFS.

3.2.3. WFS Fractional delays

The ideal impulse response of a fractional delay $z^{-\tau}$, where τ is a real number is $b_k = \sin(\pi[k-\tau])/\pi[k-\tau]$, for all k. When the desired delay τ assumes an integer value, the impulse response reduces to $\delta_{k-\tau}$. Several FIR and IIR filters for approximating this ideal impulse response have been reviewed by Laakso et al. in [22], [23], from where it has been chosen two methods: FIR Lagrange interpolation and IIR Tiran all pass filtering, due to their good behavior at low frequencies in terms of magnitude and phase delay accuracy.

In both approaches, the total delay τ is separated into an integer delay z^{-M} and a non-integer delay $T_{\tau-M}(z)$, such that it can be implemented with a shift operator followed by a FIR filter or an IIR filter. Let the order of the filter be equal to q, then the optimum choice for the integer delay M is given by

$$M = \begin{cases} round(\tau) - \dfrac{q}{2} & for\ even\ q \\ floor(\tau) - \dfrac{q-1}{2} & for\ ood\ q \end{cases} \quad (20)$$

which is appropriate for arbitrarily long delays as it happens in WFS. The objective, in general, is to express the total delay as

$$z^{-M}T_{\tau-M}(z) = z^{-M}\frac{b_0 + b_1 z^{-1} + \dots + b_q z^{-q}}{a_0 + a_1 z^{-1} + \dots + a_q z^{-q}} \quad (21)$$

The coefficients of the FIR Lagrange interpolating filter that implements the non-integer delay $T_{\tau-M}(z)$ are

$$b_{k-M} = \prod_{\substack{i=M \\ i\neq k}}^{M+q} \frac{\tau - i}{k - i} \quad (22)$$

where $k = M,\ M+1,\ \dots,\ M+q$, τ can also be integer, and $a_k = \delta_k$.

The coefficients of the IIR Tiran all-pass filter that implements the non-integer delay $T_{\tau-M}(z)$ are

$$a_k = \begin{cases} (-1)^k \binom{q}{k} \prod_{i=0}^{q} \dfrac{\tau - M - q + i}{\tau - M - q + k + i}, & \tau \notin \mathbb{Z} \\ \delta_k & \tau \in \mathbb{Z} \end{cases} \quad (23)$$

and

$$b_k = \begin{cases} a_{q-k}, & \tau \notin \mathbb{Z} \\ \delta_{k-(\tau-M)} & \tau \in \mathbb{Z} \end{cases} \quad (24)$$

where $k = 0,1,\dots,q$.

Up to here, all the frequency dependent stages in the 2.5D WFS driving functions have been discretisized. Next section resumes these results.

3.2.4. HOA spherical Hankel functions

The following series expansion of the spherical Hankel functions [24] will allow to the driving function in (7) be represented by rational functions of $j\omega$:

$$h_m^{(2)}(x) = x^m \sum_{k=0}^{\infty} \frac{1}{k!(2k+2m+1)!!}\left(\frac{-x^2}{2}\right)^{-k} + j(-x)^{-m-1}\sum_{k=0}^{\infty}\frac{1}{k!(2k-2m+1)!!}\left(\frac{-x^2}{2}\right)^{-k}. \quad (25)$$

3.3. Discrete driving functions

3.3.1. Discrete 2.5D-WFS driving functions

In practice, filtering F, delaying $z^{-M}T_{\tau-M}$ and scaling A operations need to be applied in real-time to audio samples in order to produce each loudspeaker's driving function, according to the distance $|\mathbf{x_0} - \mathbf{x}_s|$ between the virtual source and the loudspeaker in case of spherical model, or according to the direction of propagation $\mathbf{n}_p^T\mathbf{x_0}$ in case of plane waves. An additional prefiltering operation P_α independent of position is also required and it can be computed only once for all the loudspeakers. After an appropriate ordering of factors, a discrete version of the 2.5D driving functions for plane, non-focused spherical and focused spherical waves appears respectively in equations (26), (28) and (31), which is preferable to state in gain-zero-pole form, due to that most of real time DSP environments include zero-pole filters.

The discrete driving function of a loudspeaker at $\mathbf{x_0}$ for the synthesis of an audio source with spectrum S using a plane model propagating in the $\mathbf{n}_p$ direction is

$$D_p = z^{-M} T_{\tau-M} A P_\alpha S, \tag{26}$$

where P_α is computed with (15), (17) and (18) for $\alpha = 0.5$,

$$A = -2a_p(\mathbf{x_0}) \sqrt{\frac{2\pi|\mathbf{x_{ref}}-\mathbf{x_0}|}{c}} \mathbf{n}_p^T \mathbf{n}(\mathbf{x_0}), \tag{27}$$

M is computed with (20), and $T_{\tau-M}$ is computed with (22), or (23) and (24), for $\tau = f_s \ \mathbf{n}_p^T \mathbf{x_0}/c$.

The discrete driving function of a loudspeaker at $\mathbf{x_0}$ for the synthesis of an audio source with spectrum S using a non focused spherical model emanating from $\mathbf{x}_s$ is

$$D_S = z^{-M} T_{\tau-M} A F P_\alpha S, \tag{28}$$

where P_α is computed with (15), (17) and (18) for $\alpha = -0.5$,

$$F = \left(\frac{8f_s}{7} + \frac{c}{|\mathbf{x_0}-\mathbf{x}_s|}\right) \frac{1 - \frac{|\mathbf{x_0}-\mathbf{x}_s| - \frac{c}{8f_s}}{|\mathbf{x_0}-\mathbf{x}_s| + \frac{7c}{8f_s}} z^{-1}}{1 + \frac{1}{7} z^{-1}}, \tag{29}$$

$$A = -2a_s(\mathbf{x_0}) \sqrt{\frac{2\pi|\mathbf{x_{ref}}-\mathbf{x_0}|}{c}} \frac{(\mathbf{x_0}-\mathbf{x}_S)^T \mathbf{n}(\mathbf{x_0})}{|\mathbf{x_0}-\mathbf{x}_S|^2}, \tag{30}$$

M is computed with (20) and $T_{\tau-M}$ is computed with (22), or (23) and (24), for $\tau = f_s \ |\mathbf{x_0} - \mathbf{x}_s|/c$.

The discrete driving function of a loudspeaker at $\mathbf{x_0}$ for the synthesis of an audio source with spectrum S using a focused spherical model emanating from $\mathbf{x}_s$ is

$$D_{fS} = z^{M} T^*_{\tau-M} A P_\alpha S, \tag{31}$$

where P_α is computed with (15), (17) and (18) for $\alpha = 0.5$,

$$A = \frac{-\mathrm{j}}{\sqrt{2\pi c}} \sqrt{\frac{|\mathbf{x_{ref}}-\mathbf{x_0}|}{|\mathbf{x_{ref}}-\mathbf{x_0}|+|\mathbf{x}_S-\mathbf{x_0}|}} \frac{(\mathbf{x_0}-\mathbf{x}_S)^T \mathbf{n}(\mathbf{x_0})}{|\mathbf{x_0}-\mathbf{x}_S|^{3/2}} \tag{32}$$

M is computed with (20), and $T^*_{\tau-M}$ is computed with (22), or (23) and (24), for $\tau = f_s \ |\mathbf{x_0} - \mathbf{x}_s|/c$. Here, $*$ denotes the complex conjugate operator.

3.3.2. Discrete 2D-HOA driving functions

The Al-alaoui map (11) of the 2D-HOA driving function (7), using the series expansion of the spherical Hankel function of second kind (25), leads to:

$$D_p^{HOA} = \gamma \mathrm{p}(z) \times \sum_{m=-\frac{L-1}{2}}^{\frac{L-1}{2}} \frac{\gamma^m e^{jm(\alpha_0 - \theta_{\mathrm{PW}})}}{\sum_k c_{1,k}[\mathrm{p}(z)]^{m+2k} - j\sum_k c_{1,k}[\mathrm{p}(z)]^{-m-1+2k}}, \tag{33}$$

where

$$\gamma = \frac{7c}{4f_s r_0}, \tag{34}$$

$$\mathrm{p}(z) = \frac{1 - z^{-1}}{1 + \frac{1}{7} z^{-1}}, \tag{35}$$

$$c_{1,k} = \frac{(-\gamma^{-2})^k}{k!(2k+2m+1)!!}, \tag{36}$$

$$c_{2,k} = \frac{(-1)^m (-\gamma^{-2})^k}{k!(2k-2m+1)!!}. \tag{37}$$

In the next section, the sound pressure fields in the whole listening area are synthesized using the continuous and discrete driving functions.

4. SIMULATION RESULTS

Figures 9 to 11 allow for visual comparisons of sound pressure fields synthesized from continuous and discrete driving functions. Percentage discretization errors of driving functions are also shown for different frequencies. Sound pressure fields are computed with real part of equations (8) and (9), and percentage errors with equation (10).

In order to select suitable parameters for WFS simulations, it has previously been evaluated the discretization error for different orders of both prefiltering and delaying stages. These results are shown in figures 4 and 5, from where it can be seen that a 5th order prefilter provides the best result, and integer delays are enough for delaying. If necessary, fractional delays with 1st order filters on either a FIR or IIR approach are enough. On incoming simulations, the following discretization parameters for prefilters in equations (15), (17) and (18) have been chosen: $K = 150$, $m = n = 6$, and $N = 25$. They return the half differentiator (Table 1) and the half integrator (Table 2) used for prefiltering on equations (26), (28) and (31).

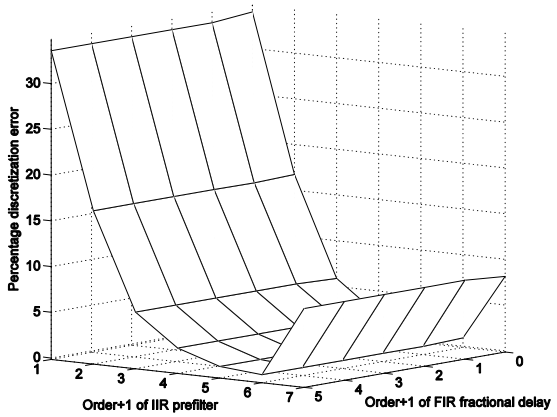

Figure4. Discretization errors for different orders of IIR prefilters and FIR fractional delays used in 2.5D-WFS.

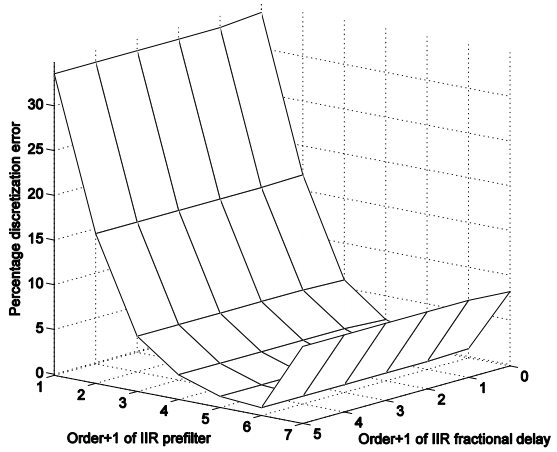

Figure5. Discretization errors for different orders of IIR prefilters and IIR fractional delays used in 2.5D-WFS.

Gain g	224.4994					
Zeros ζ_k	0.9887	0.8972	0.7112	0.4478	0.1628	-0.0590
Poles π_k	0.9547	0.8158	0.5867	0.3029	-0.1214	0.0386

Table1. 2.5D-WFS prefilter P_α for the synthesis of plane and focused spherical waves ($\alpha = 0.5$, $K = 150$, $m = n = 6$, and $N = 25$).

Gain g	0.0045					
Zeros ζ_k	0.9695	0.8565	0.6560	0.3432	-0.1251	0.0819
Poles π_k	0.9928	0.9251	0.7674	0.5093	0.2045	-0.0458

Table2. 2.5D-WFS prefilter P_α for the synthesis of non focused spherical waves ($\alpha = -0.5$, $K = 150$, $m = n = 6$, and $N = 50$).

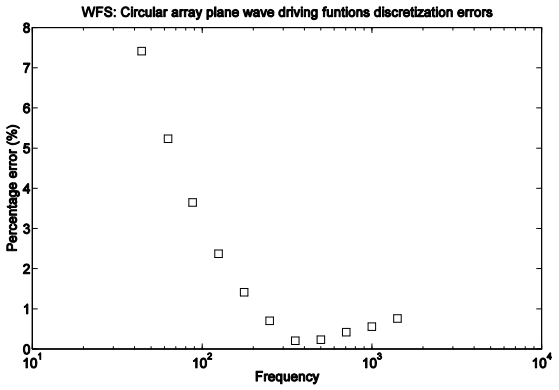

Figure6. Discretization errors of WFS plane wave driving functions, in the reproducible frequency range.

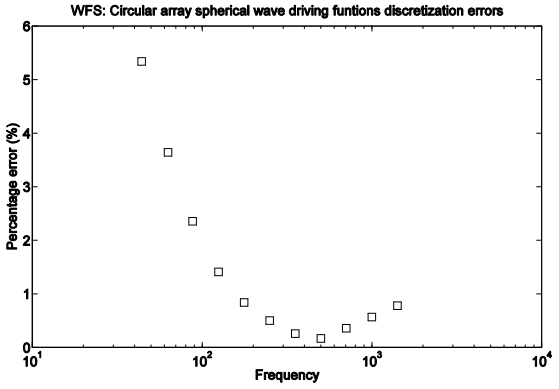

Figure7. Discretization errors of WFS focused wave driving functions, in the reproducible frequency range.

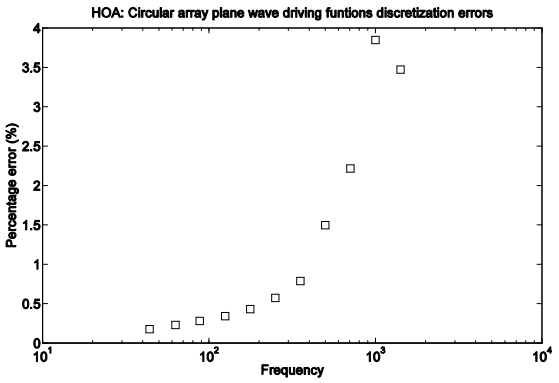

Figure8. Discretization errors of HOA plane wave driving functions, in the reproducible frequency range.

In the top of figures 9 to 11, the sound pressure field of a pure tone of 500Hz, synthesized with continuous (left) and discrete (left) driving functions through a circular array loudspeaker, are shown. The top of figures 9 and 10 shows a plane wave and focused spherical wave using WFS. The top of figure 11 shows a plane wave.

5. AN APPLICATION TO IMMERSIVE SOUNDSCAPES

This section returns to the concept of soundscape. Although the reproduction of sound using spherical and plane wave models is widely used in room acoustics (spherical waves for audio tracks and plane waves for room impulse responses), it is proposed here an application for the composition and recreation of soundscapes. Since atmospheres are perceived as coming from non-localized sources, and soundmarks and events from localized sources, their reproduction model can be done respectively with plane and spherical waves. Within this framework this paper reports a practical auralization of an immersive soundscape composed with nature sounds recorded in the highlands and beaches of Lima, Peru. Next section explains the considerations for data registration.

5.1. Data registration

Monaural samples, recorded with omnidirectional and wind-shielded microphones, should be used with this reproduction scheme. It is preferable the use of a digital recorder with 24 quantization and capable of capturing at least 44100 samples per second. It is also strongly recommended to normalize the samples at -3dB and then to eliminate the DC offset. For the following application, audio samples have been recorded using digital recorders such that the Sound Device 722 and the Sony PCMD50, both with an external, omnidirectional and wind-screened microphone.

Once the collection of samples has been recorded, a spectral representation of the sound objects should be done. This representation is very helpful to classify them in background and foreground sounds. For example, the spectrograms of a background sound such that the wind blowing through the trees have a smooth covering of all almost the whole time-frequency domain. This is not the case for example for the foreground sound of the song of a bird, whose spectrograms are composed of several increasing and decreasing curved lines referring to the fundamental and harmonic frequencies.

5.2. Auralization of nature sounds

In the previous context, the multichannel auralization has been implemented using Pure Data [25] with a graphical interface based on reacTIVision [26], which allows the composer to render up to five virtual sources over a 10mx10m virtual listening area (although the physical space is aprox. 4mx4m) through 24 audio channels, giving the spectator the sensation of being immerse in a real scene due to the spatial component added to the environ-ment. The equipment used for this purpose was: one Power Mac G5, one M-audio Profire Lightbridge audio interface, three Behringer ADA8000 digital to analog converters, three QSC168X eight channel amplifiers, and 24 Behringer 1CBK loudspeakers.

Figure 12 shows a practical implementation of an immersive soundscape composed with nature sounds recorded in the highlands and beaches of Lima, Peru. A real system was developed in ISONAR Sound Research Workshop, at University of San Martin de Porres. The study seeks to standardize criteria for listening, through education and recreational participation, deepening about the symbolic and significant features that sounds have, enhancing the ability for each student to appreciate, in relation to time, frequency and space. We expect that this immersive space will contribute, in a practical way, on the training of university students, in the career of communication science, allowing them to improve the quality of their audiovisual creations such as documentaries, films, radio and television, with proper manipulation of spatial hearing and sound [27].

Figure12. The ISONAR space for horizontal auralization of immersive soundscapes.

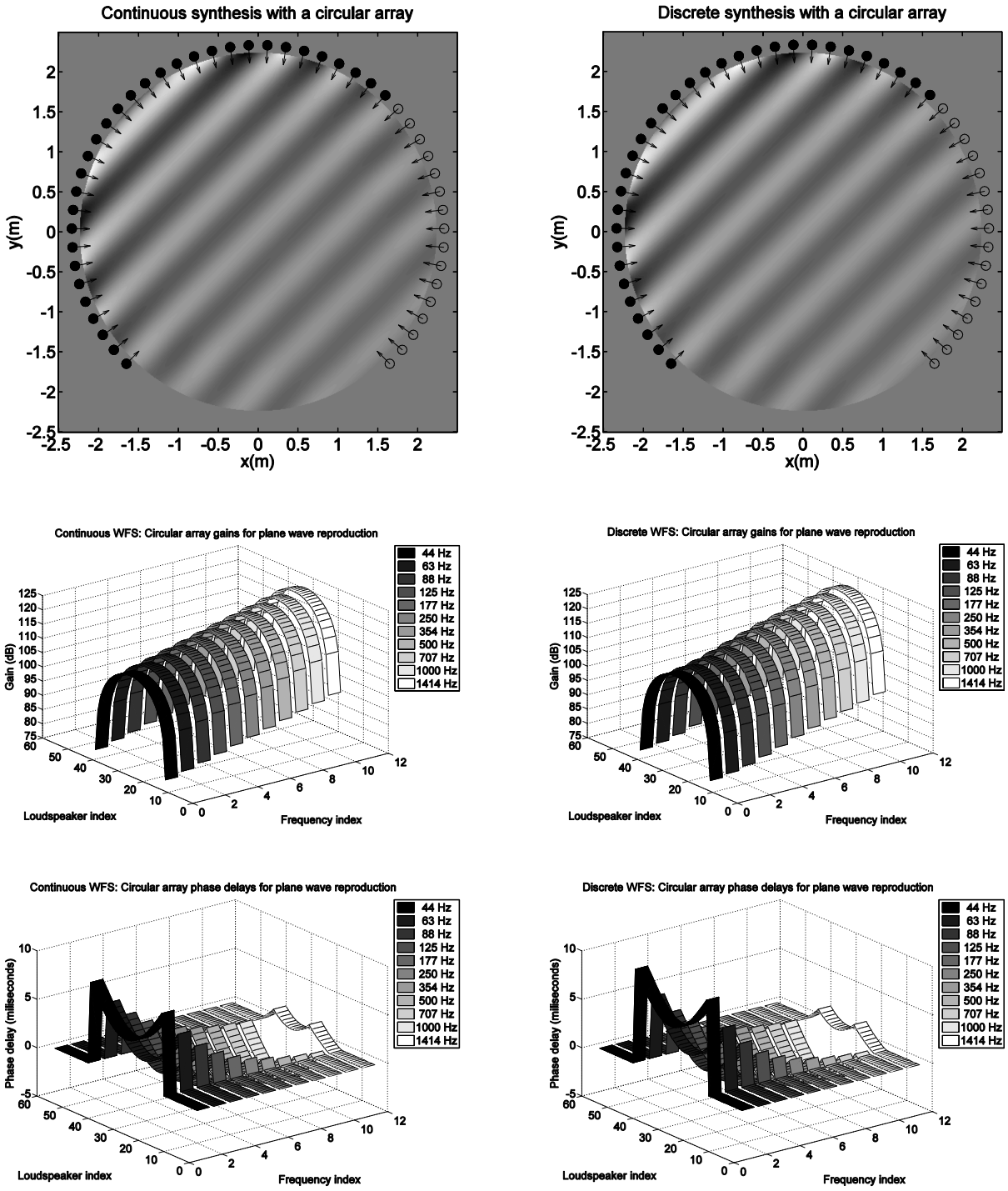

Figure9. 2.5D-WFS plane wave reproduction with a circular array of 48 loudspeakers. At the top: continuous (left) and discrete (right) sound pressure fields for a tone of 500 Hz. At the middle and bottom: Continuous and discrete driving functions for different frequencies.

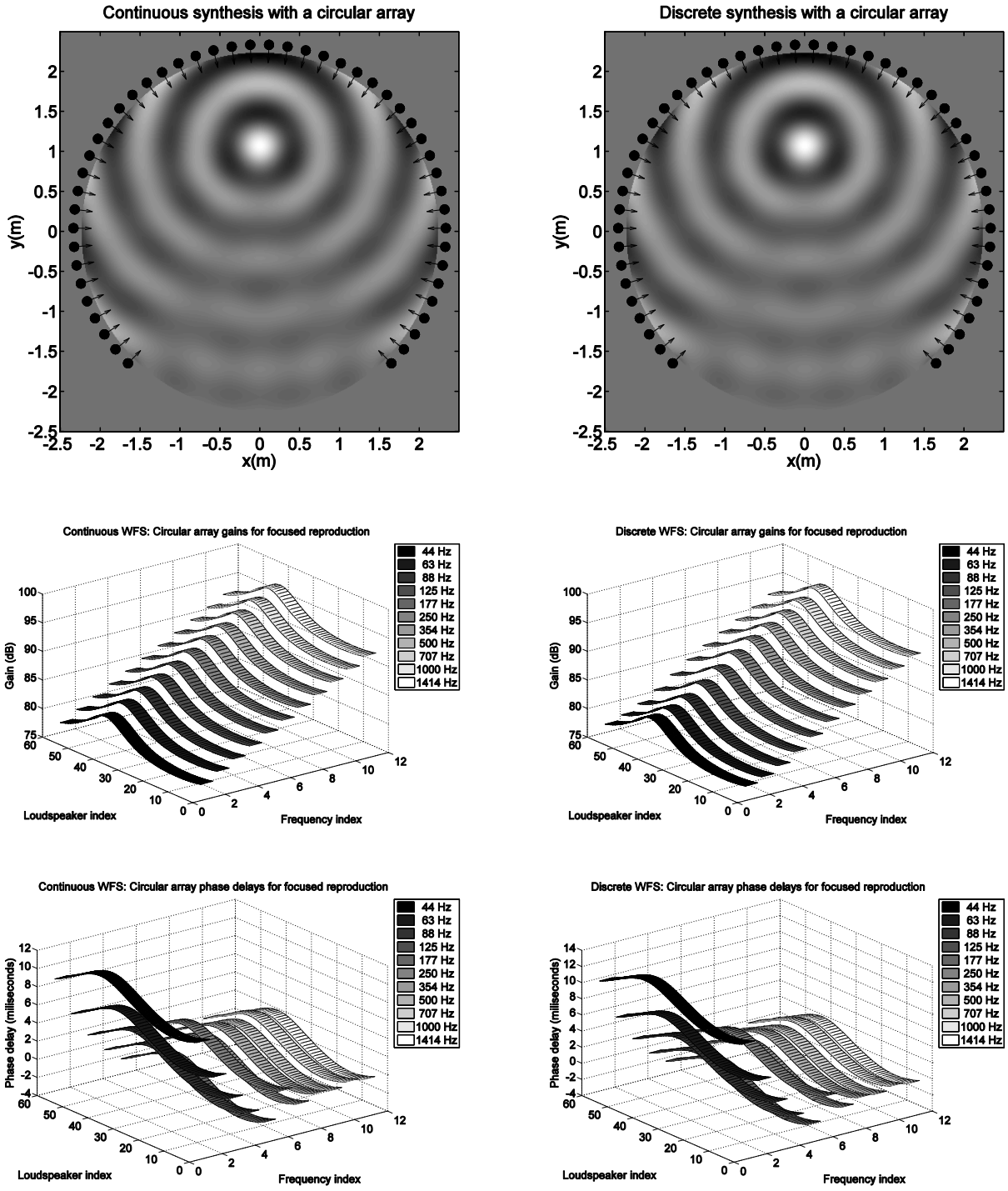

Figure10. 2.5D-WFS focused spherical wave field reproduction with a circular array of 48 loudspeakers. At the top: continuous (left) and discrete (right) sound pressure fields for a tone of 500 Hz. At the middle and bottom: Continuous and discrete driving functions for different frequencies.

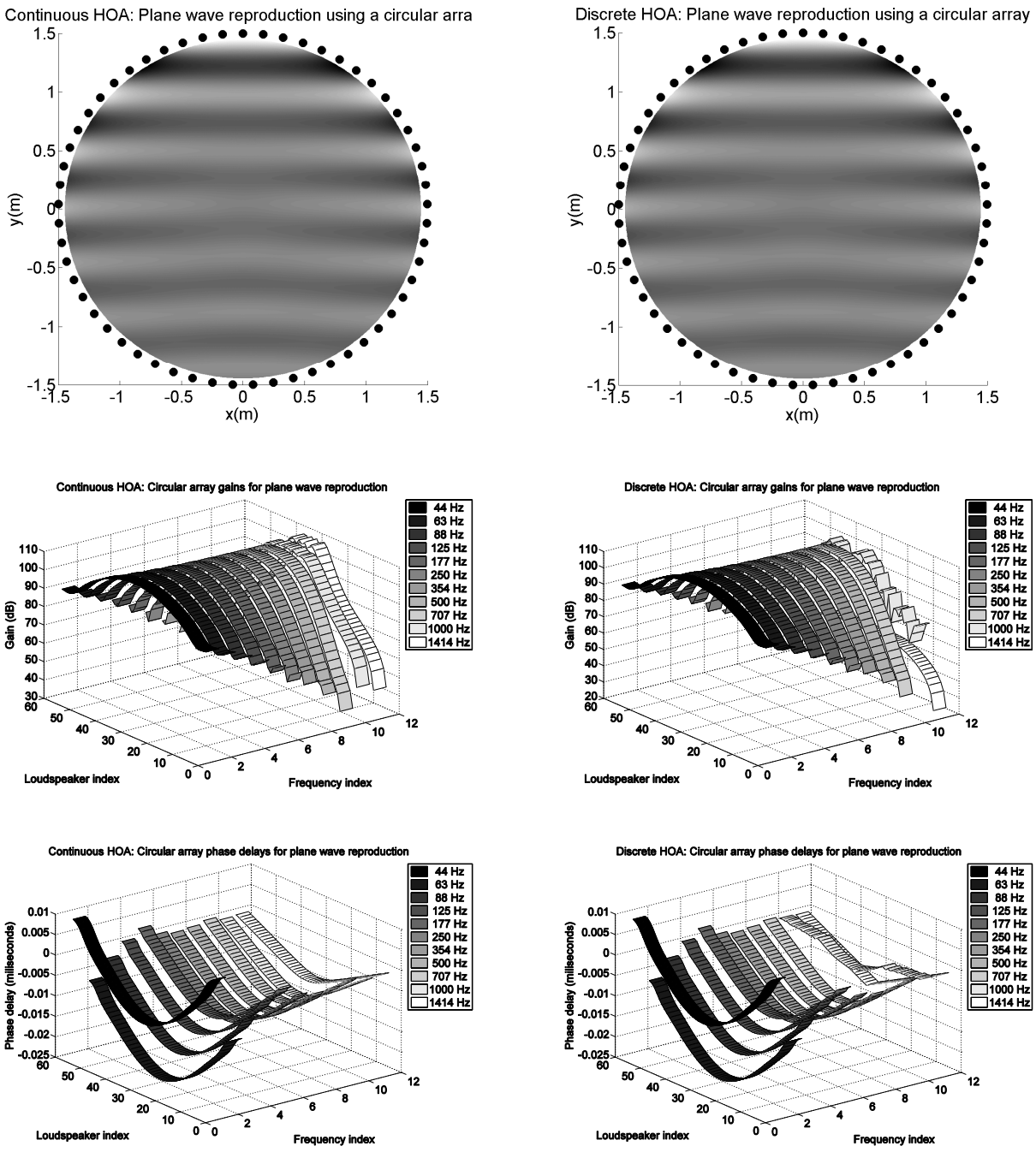

Figure11. 2D-HOA plane wave reproduction with a circular array of 57 loudspeakers. At the top: continuous (left) and discrete (right) sound pressure fields for a tone of 500 Hz. At the middle and bottom: Continuous and discrete driving functions for different frequencies.

6. CONCLUSION

A discretization of the driving functions for 2.5D Wave Field Synthesis (WFS) and 2D Higher Order Ambisonics (HOA) has been described, with special attention on prefiltering, delaying and spherical Hankel functions. On WFS prefiltering, the following methods has been used: Al-alaoui discretization that corresponds to an interpolation between the Euler and Tustin approximations, the Grünwald-Letnikov approach for fractional integration and differentiation that allows a binomial expansion of the Al-alaoui map whose result is a large FIR filter, and the Shank's least squares approximation method to transform the previous FIR filter into a short IIR one that generates the corresponding half order system. On WFS delaying, the Lagrange interpolation and the Tiran all-pass filtering approaches has been employed to compute respectively the FIR and IIR filters that generate the fractional sample delays.

Simulations for WFS showed that the discretization of prefiltering directly affects to discretization errors, where the best choice was a 5^{th} order filter (length = 6). On the other hand, the fractional delays does not affect to discretization, where integer delays or 1^{st} order FIR or IIR filters are enough. Simulations for HOA shows that a good behaviour in the reproducible frequency range up to about 1000 Hz. All simulations in figures 6, 7 and 8, show a deviation from the continuous formulation on the order of 1%, up to 1000Hz. As a practical example, it has been reported the implementation of an immersive soundscape composed of nature sounds.

7. ACKNOWLEDGEMENTS

This work was supported by the ISONAR Sound Research Workshop and the Research Institute, both at Universidad de San Martin de Porres, Lima, Peru.

8. REFERENCES

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

[2] A. Berkhout, D. de Vries and P. Vogel, "Acoustic control by wave field synthesis," in Journal of the Acoustic Society of America, vol. 93, no. 5, pp. 2664-2778, May 1993.

[3] J. Daniel, "Représentation de champs acoustiques, application à la transmission et à la reproduction de scenes sonores complexes dans un context multimedia," PhD thesis Université Paris 6, 2001.

[4] R. Schafer, The tuning of the world, Knopf, New York, 1977.

[5] A. Valle, V. Lombardo, and M. Schirosa, "A graph-based system for the dynamic generation of soundscapes," Proc. of the 15th Int. Conf. on Auditory Display, Copenhagen, 2009.

[6] J. Daniel, S. Moreau and R. Nicol, "Further Investigations of High-Order Ambisonics and Wavefield Synthesis for Holophonic Sound Imaging," in 114th Audio Engineering Society Convention, Amsterdam, The Netherlands, March 2003.

[7] J. Ahrens and S. Spors, "Analytical driving functions for higher order ambisonics," in IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), April 2008.

[8] S. Spors, H. Buchner, R. Rabenstein and W. Herbordt, "Active listening room compensation for massive multichannel sound reproduction systems using wave-domain adaptive filtering," in Journal of the Acoustic Society of America, Vol. 122, No. 1, pp. 354-369, July 2007.

[9] C. Chui and G. Chen, Discrete H^{∞} optimization: With applications in signal processing and control systems, Springer, Berlin, 1997.

[10] S. Spors, J. Ahrens and R. Rabenstein, "The theory of wave field synthesis revisited," in 124th Audio Engineering Society Convention, Amsterdam, The Netherlands, May 2008.

[11] D. de Vries, Wave Field Synthesis, Audio Engineering Society Monograph, April 2009.

[12] S. Yon, M. Tanter and M. Fink, "Sound focusing in rooms: The time-reversal approach," in Journal of the Acoustic Society of America, vol. 113, no. 3, pp. 1533-1543, March 2003.

[13] S. Yon, M. Tanter and M. Fink, “Sound focusing in rooms. II. The spatio-temporal inverse filter,” in Journal of the Acoustic Society of America, vol. 114, no. 6, pp. 3044-3052, December 2003.

[14] S. Spors, H. Wierstorf, M. Geier and J. Ahrens, “Physical and perceptual properties of focused sources in wave field synthesis,” in 127th Audio Engineering Society Convention, New York, USA, October 2009.

[15] J. Ahrens and S. Spors, “Notes on rendering of focused directional sound sources in wave field synthesis,” in 34rd German Annual Conference on Acoustics, Dresden, Germany, March 2008.

[16] P. Morse and K. Ingard, Theoretical acoustics, Princeton: Princeton University Press, 1986.

[17] M. Al-alaoui, “Novel digital integrator and differentiator,” in Electronics letters, vol. 29, pp. 376-378, 1993.

[18] M. Al-alaoui, “Al-alaoui operator and the new transformation polynomials for discretization of analogue systems,” in Electronic Engineering, Springer-Verlag, vol. 90, pp. 455-467, 2008.

[19] Oldham and J. Spanier, The fractional calculus, Academic Press, New York, USA, 1974.

[20] R. Barbosa, J. Tenreiro Machado and I. Ferreira, “Pole-zero approximations of digital fractional-order integrators and differentiators using signal modeling techniques,” in Proceedings of the 16th IFAC World Congress, Prague, Czech Republic, July 2005.

[21] R. Barbosa and J. Tenreiro Machado, “Implementation of discrete-time fractional-order controllers based on LS aproximations,” in Acta Polytechnica Hungarica, vol. 3, no. 4, pp. 5-22, 2006.

[22] T. Laakso, V. Välimäki, M. Karjalainen and U. Laine, “Splitting the unit delay: Tools for fractional delay filter design,” IEEE Signal Processing Magazine, vol. 13, no. 1, pp. 30-60, January 1996.

[23] V. Välimäki and T. Laakso, "Principles of fractional delay filters," IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'00), Istanbul, Turkey, 5-9 June 2000.

[24] M. Abramowitz and I. Stegun, Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, 9th printing. New York: Dover, 1972.

[25] Pure Data: http://www.pure-data.org

[26] reacTIVision: http://reactivision.sourceforge.net

[27] Lima Sonora: http://www.limasonora.com