



Audio Engineering Society Convention Paper 8119

Presented at the 128th Convention
2010 May 22–25 London, UK

The papers at this Convention have been selected on the basis of a submitted abstract and extended precis that have been peer reviewed by at least two qualified anonymous reviewers. This convention paper has been reproduced from the author's advance manuscript, without editing, corrections, or consideration by the Review Board. The AES takes no responsibility for the contents. Additional papers may be obtained by sending request and remittance to Audio Engineering Society, 60 East 42nd Street, New York, New York 10165-2520, USA; also see www.aes.org. All rights reserved. Reproduction of this paper, or any portion thereof, is not permitted without direct permission from the Journal of the Audio Engineering Society.

Surround Sound Panning Technique based on a Virtual Microphone Array

Filippo M. Fazi¹, Toshiro Yamada², Suketu Kamdar², Philip A. Nelson¹, Peter Otto²

¹*Institute of Sound and Vibration Research, University of Southampton, Southampton, SO17 1BJ, U.K.*

²*California Institute for Telecommunications and Information Technology, University of California, San Diego, La Jolla, CA, 92093, U.S.A.*

Correspondence should be addressed to Filippo M. Fazi (ff10isvr.soton.ac.uk)

ABSTRACT

A multi-channel panning technique is presented, which aims at reproducing a plane wave with an array of loudspeakers. The loudspeaker gains are computed by solving an acoustical inverse problem. The latter involves the inversion of a matrix of transfer functions between the loudspeakers and the elements of a virtual microphone array, the center of which corresponds to the location of the listener. The radius of the virtual microphone array is varied consistently with the frequency, in such a way that the transfer function matrix is independent of the frequency. As a consequence, the inverse problem is solved for one frequency only, and the loudspeaker coefficients obtained can be implemented using simple gains.

1. INTRODUCTION

The generation of a desired sound field given an array of loudspeakers is an area of intense ongoing research. Multi-channel reproduction techniques such as Wave Field Synthesis [1], High Order Ambisonics [2] and Vector Based Amplitude Panning [3], among others, have been developed and are currently applied to real systems. These techniques can be used to design algorithms that compute signals for a loudspeaker array allowing for reproduction of a target

sound field. Other methods have been proposed which perform the calculation above by solving an acoustical inverse problem [4], [5], [6], [7]. The goal of these methods is to derive the loudspeaker signals from values of a number of control points in the target field (or on a boundary of the control region). One of the disadvantages of these techniques is the computational cost of their implementation, since the solution of the inverse problem under consideration is frequency dependent. Therefore, the practi-

cal implementations of such techniques usually use a large number of digital filters. In many practical situations these filters should be computed in real time. This is typically the case when the field reproduction is optimized in a limited region surrounding a moving listener.

In this paper a novel method is presented, which allows for the reproduction of a desired sound field (starting with a single plane wave) using a simple network of gains and delays. We assume that the sound system under consideration is a multi-channel surround system, meaning that the system includes several (at least five) loudspeakers, which are arranged in a two-dimensional or three-dimensional displacement around a given array center, coinciding with the origin of the coordinate system. No strict assumptions are made regarding the uniformity of that arrangement.

A simple radiation model is used for the loudspeakers. Each of them generates a sound field which can be approximated, in the vicinity of the listener position, to a single plane wave traveling from the loudspeaker location to the listener. A delay and a gain are applied to each loudspeaker signal, in order to compensate for the distance between the transducer and the listener. In this way, all the plane waves generated by the loudspeakers have the same phase and amplitude at the listener location.

We assume then that a virtual array of microphones is arranged around the listener and translates together with the latter. The virtual microphones are the control points of the sound field. The radius of the array is assumed to be frequency dependent. More specifically, the array radius r varies proportionally with the operating wave number k , so that the product kr is constant. In this way that the values of the target plane wave sampled by the virtual microphone array are independent from the frequency.

A set of gains, one for each loudspeaker, is computed by solving a least squares minimization problem, with the use of two different kind of regularization. The application of similar acoustical inverse problems has been proposed in previous publications by Kirkeby, Poletti and some of the authors [4], [6], [7].

The resulting system is straightforward to implement and can be very effective in single listener systems incorporating a tracking device.

2. NOTATION

Lowercase bold lettering represents vectors. Each vector $\mathbf{x}$ is characterized by a magnitude $|\mathbf{x}|$ and by a unitary vector $\hat{\mathbf{x}}$, identifying the direction of $\mathbf{x}$. It holds that

$$\mathbf{x} = |\mathbf{x}| \hat{\mathbf{x}} \quad (1)$$

Given a relative reference coordinate system with center $\mathbf{x}_s$, the coordinate of a location $\mathbf{x}$ can be expressed with respect to the relative system by

$$\mathbf{x}' = \mathbf{x} - \mathbf{x}_s \quad (2)$$

Given two vectors $\mathbf{x}$ and $\mathbf{y}$, the notation $\mathbf{x} \cdot \mathbf{y}$ indicates the scalar product between the two vectors. Uppercase bold lettering represents matrices.

3. THEORY

The surround sound system under consideration is assumed to include L loudspeakers. Each loudspeaker is arranged at a given location $\mathbf{y}_\ell$, as shown in Figure 1. It is also assumed that the listener occupies a location in the interior of the loudspeaker array, identified by the vector $\mathbf{x}_s$ and hereafter also referred to as the *sweet spot*. It is possible to choose the listener position as the origin of a relative reference coordinate system. We will assume that this relative system can translate with respect to the absolute reference system, but no relative rotation is allowed. Each absolute position $\mathbf{x}$ can be expressed with respect to the listener position, and the relative position $\mathbf{x}'$ is given by equation (2).

The loudspeaker locations can therefore be expressed with respect to the listener position, yielding a set of relative coordinates $\mathbf{y}'_\ell$.

Assuming that the sweet spot is not in the near-field of the given loudspeaker, it is possible to approximate the sound field generated by one loudspeaker with a plane wave arriving at the listener position from the direction of the loudspeaker as seen from the sweet spot. It should be noted that this relative direction $\hat{\mathbf{y}}'_\ell$ is in general different from the loudspeaker direction $\hat{\mathbf{y}}_\ell$, seen from the origin of the absolute coordinate system.

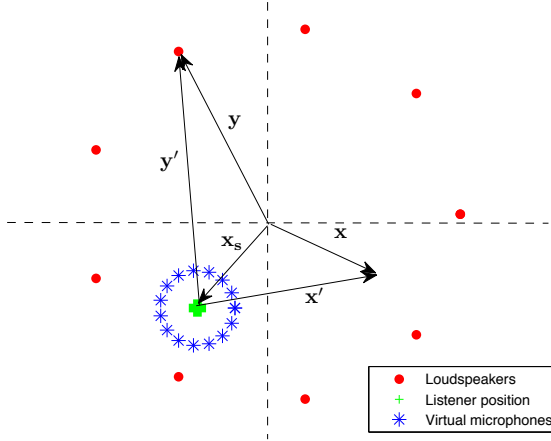


Fig. 1: Reference geometry.

Therefore, the field $p_\ell(\hat{\mathbf{x}}, \omega)$ generated by the ℓ -th loudspeaker can be expressed, in the vicinity of the listener and for a given frequency ω , as a plane wave of the form

$$p_\ell(\mathbf{x}', \omega) = \frac{1}{|\mathbf{y}'_\ell|} e^{-jk(|\mathbf{y}'_\ell| - \hat{\mathbf{y}}'_\ell \cdot \mathbf{x}')} \quad (3)$$

where $k = \omega/c$ is the wave number (the speed of sound c is uniform, since the medium of propagation is assumed to be homogeneous and non dispersive) and the time convention $e^{j\omega t}$ has been chosen, so that the expression above represents a wave propagating in the direction $-\hat{\mathbf{y}}'_\ell$ (or, equivalently, arriving from the direction $\hat{\mathbf{y}}'_\ell$). The factor $1/|\mathbf{y}'_\ell|$ has been introduced in order to take into account the attenuation that the wave undergoes while propagating (it is assumed that $|\mathbf{x}'| \ll |\mathbf{y}'_\ell|$ and therefore $|\mathbf{y}'_\ell - \mathbf{x}'| \approx |\mathbf{y}'_\ell|$ for the energy decay).

In view of this, the sound field $p(\mathbf{x}', \omega)$ reproduced by the entire loudspeaker array in the vicinity of the listener position and for a single frequency ω is given by

$$p(\mathbf{x}', \omega) = \sum_{\ell=1}^L s_\ell(\omega) \frac{1}{|\mathbf{y}'_\ell|} e^{jk \hat{\mathbf{y}}'_\ell \cdot \mathbf{x}'} e^{-jk|\mathbf{y}'_\ell|} \quad (4)$$

where $s_\ell(\omega)$ is the signal feeding the ℓ -th loudspeaker, expressed as a function of the frequency.

3.1. Target sound field

We assume that the target sound field $\tilde{p}_T(\mathbf{x}', \omega)$ can

be expressed (with respect to the listener) by a linear superposition of plane waves, and is therefore given by

$$\tilde{p}_T(\mathbf{x}', \omega) = \int_{\Omega} e^{jk \hat{\mathbf{k}}' \cdot \mathbf{x}'} \varphi(\hat{\mathbf{k}}', \omega) d\Omega(\hat{\mathbf{k}}') \quad (5)$$

where Ω is the unitary sphere for a three-dimensional problem or a circle for the two-dimensional case. The integral above can be discretized by uniformly (or almost uniformly) sampling Ω at N angular locations $\hat{\mathbf{k}}'_n$, and is approximated by a summation of the form

$$\tilde{p}_T(\mathbf{x}', \omega) = \sum_{n=1}^N e^{jk \hat{\mathbf{k}}'_n \cdot \mathbf{x}'} \varphi(\hat{\mathbf{k}}'_n, \omega) \Delta\Omega_n \quad (6)$$

The factors $\Delta\Omega_n$ depend on the sampling scheme adopted. In the case of Ω being a unitary circle uniformly sampled at N positions, we have that $\Delta\Omega_n = 2\pi/N$.

We will concentrate on the case of a single pulse-like plane wave arriving from the direction $\hat{\mathbf{k}}'_0$, with respect to the listener, that is when the density function $\varphi(\hat{\mathbf{k}}')$ in the integral above is a Dirac delta function $\delta(\hat{\mathbf{k}}' - \hat{\mathbf{k}}'_0)$. In this case we have that

$$p_T(\mathbf{x}', \omega) = e^{jk \hat{\mathbf{k}}'_0 \cdot \mathbf{x}'} \quad (7)$$

3.2. Distance compensation

It is possible to apply to each loudspeaker a negative delay $e^{jk|\mathbf{y}'_\ell|}$ (anticipation) and a gain $|\mathbf{y}'_\ell|$ in order to compensate for the distance of the transducer from the listener. The loudspeaker signals for reproducing the target field given by equation (7) can therefore be expressed as follows

$$s_\ell(\omega) = g_\ell(\omega) |\mathbf{y}'_\ell| e^{jk|\mathbf{y}'_\ell|} \quad (8)$$

where the frequency dependent coefficient $g_\ell(\omega)$ is to be determined. The reproduced field is

$$p(\mathbf{x}, \omega) = \sum_{\ell=1}^L g_\ell(\omega) e^{jk \hat{\mathbf{y}}'_\ell \cdot \mathbf{x}'} \quad (9)$$

When, for example, all coefficients $g_\ell(\omega)$ are unity for all frequency, each speakers generate a single pulse and the L pulses arrive simultaneously at the listener position and interfere constructively.

3.3. The virtual microphone array

The ideal objective at this stage is to compute a set of gains $g_\ell(\omega)$ such that the reproduced field $p(\mathbf{x}, \omega)$ given by equation (9) is *as close as possible*, and ideally identical, to the target field $p_T(\mathbf{x}, \omega)$ given by equation (7). In order to define what is meant by the concept of distance introduced above, the following criterium is introduced.

We assume that a *virtual array of microphones* is arranged around the listener position and moves together with it. The virtual microphones are arranged at the same distance r from the listener position and their relative positions $\mathbf{x}'_m = r\hat{\mathbf{x}}'_m$ do not vary with listener movements, while the opposite is true for their absolute positions $\mathbf{x}_m = \mathbf{x}'_m + \mathbf{x}_s$. Although no *a priori* constrain exists on the angular arrangement of the virtual microphones, in this work we will assume that, in the case of a two-dimensional loudspeaker array (all speakers laying on the same plane), they are uniformly arranged on a circle of radius r and centered at the listener position, and in the case of a three dimensional loudspeaker array, they are almost uniformly arranged on the surface of a sphere with same center (example of almost uniform sampling of a sphere can be found in [8]). In view of the uniqueness of the interior Dirichlet problem¹ (see, for example, [7]) if the target sound field is reproduced accurately on the surface of the control boundary identified by the microphone array, then the target field is reproduced accurately also in the interior of that region and in part of its exterior. In the case under consideration, in which the acoustic field is sampled on the surface identified by virtual microphone array, the arguments above are valid if the average distance between neighboring virtual microphones is smaller than half of the operating wave length $\Lambda = 2\pi/k$.

The values of the target sound field and of the reproduced sound field measured by the virtual microphones can be expressed by the vectors $\mathbf{p}_T$ and $\mathbf{p}$,

respectively. These are given by

$$\mathbf{p}_T := \begin{bmatrix} e^{jkr} \hat{\mathbf{k}}'_0 \cdot \hat{\mathbf{x}}'_1 \\ e^{jkr} \hat{\mathbf{k}}'_0 \cdot \hat{\mathbf{x}}'_2 \\ \vdots \\ e^{jkr} \hat{\mathbf{k}}'_0 \cdot \hat{\mathbf{x}}'_M \end{bmatrix} \quad \mathbf{p} := \begin{bmatrix} p(\mathbf{x}_1, \omega) \\ p(\mathbf{x}_2, \omega) \\ \vdots \\ p(\mathbf{x}_M, \omega) \end{bmatrix} \quad (10)$$

In view of equation (9), vector $\mathbf{p}$ can be expressed by

$$\mathbf{p} = \mathbf{H} \mathbf{g} \quad (11)$$

where matrix $\mathbf{H}$ and vector $\mathbf{g}$ are respectively given by

$$\mathbf{H} := \begin{bmatrix} e^{jkr} \hat{\mathbf{y}}'_1 \cdot \hat{\mathbf{x}}'_1 & \dots & e^{jkr} \hat{\mathbf{y}}'_L \cdot \hat{\mathbf{x}}'_1 \\ \vdots & \ddots & \vdots \\ e^{jkr} \hat{\mathbf{y}}'_1 \cdot \hat{\mathbf{x}}'_M & \dots & e^{jkr} \hat{\mathbf{y}}'_L \cdot \hat{\mathbf{x}}'_M \end{bmatrix} \quad (12)$$

$$\mathbf{g} := \begin{bmatrix} g_1(\omega) \\ g_2(\omega) \\ \vdots \\ g_L(\omega) \end{bmatrix} \quad (13)$$

It should be noted that the scalar product $\hat{\mathbf{y}}'_\ell \cdot \hat{\mathbf{x}}'_m$ corresponds to the cosine of the angle between the directions identified by the ℓ -th loudspeaker and the m -th virtual microphone with respect to the sweet spot.

The distance $J(\omega)$ between the target and reproduced field (or error), which we seek to minimize, is given by

$$J(\omega) := \|\mathbf{p}_T - \mathbf{H} \mathbf{g}\|^2 \quad (14)$$

This is a classical least squares problem, in which the vector $\mathbf{g}$ is to be determined. It is clear that the solution of the above depends strongly on the number and arrangement of the virtual microphones, and that the solution will be different for different frequencies. Since the coefficients $g_\ell(\omega)$ are frequency dependent, they need to be implemented as digital filters, and their real time calculation can be computationally expensive.

3.4. Frequency dependent array radius

We will now apply a strategy, which allows to make the above calculation independent of the frequency, in such a way that the coefficients g_ℓ can be implemented as simple gains.

The distance r of the virtual microphones from the listening position is not fixed, but is now a function

¹For the solution of the Dirichlet problem to be unique, the operating wave number k should not be one of the Dirichlet eigenvalues for the region identified by the virtual microphone array [7]. However, since later in this work the factor kr is assumed to be fixed, the nonuniqueness problem above does not arise

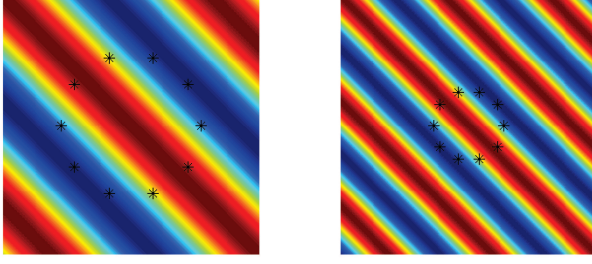


Fig. 2: Plane waves at different frequencies. The * represent virtual microphones. The frequency dependent radius allows the virtual microphones to measure the same values of the target plane wave, independently of the frequency.

of the frequency. More specifically, if for a given frequency ω_0 a given distance $r(\omega_0)$ is chosen, then the virtual microphone distance $r(\omega)$ for all other frequencies is defined by

$$r(\omega) = \frac{\omega_0}{\omega} r(\omega_0) \quad (15)$$

It is clear that with this choice the factor kr included in the definition (10) and (12) of the vector $\mathbf{p}_T$ and of the matrix $\mathbf{H}$, respectively, becomes a frequency independent parameter γ . It can be indeed noted that

$$kr(\omega) = \frac{\omega}{c} \frac{\omega_0}{\omega} r(\omega_0) = k_0 r(\omega_0) := \gamma \quad (16)$$

The physical explanation of this fact is illustrated by Figure 2. This shows that, given a reference frequency ω_0 and a reference array radius $r(\omega_0)$, when the frequency ω is half of ω_0 , and the corresponding wave length Λ is therefore twice the reference wavelength Λ_0 , the radius $r(\omega)$ of the array is twice the reference radius ω_0 . In this way, the values of the target plane wave, sampled by the virtual microphone array, are independent from the frequency.

The ideal choice of the optimal constant γ is not trivial, and depends in general on the configuration of both the virtual microphone array and the loudspeaker array. It seems sensible to choose γ as a function of the number M of virtual microphones. In order to avoid spatial aliasing, the average distance between microphones should be less than half

of the reference wave length Λ_0 ; hence a smaller radius seems an optimal choice. For a circular array, for example, the distance between neighboring virtual microphones is $2\pi r(\omega)/M$, therefore the non-aliasing condition is satisfied if $\gamma < M/2$. On the other hand, the parameter γ should not be too small, since an increase of the array radius has the effect of reducing the condition number of the matrix $\mathbf{H}$ and make the system more stable (the reader can refer to [9] for an explanation of this phenomenon). Numerical simulations indicate that the optimal choice of the parameter γ is affected also by the number and arrangement of the loudspeakers. In fact, if the loudspeaker array is non-uniform, a choice of a small parameter γ leads to better results. It should be noted that, since matrix $\mathbf{H}$ is computed using the *relative* coordinates of the loudspeaker array, the uniformity of the array is to be considered as seen from the sweet spot rather than on the basis of the absolute loudspeaker arrangement. For example, a perfectly uniform circular array can be regarded as strongly irregular if the listener is very close to one of the loudspeakers. As a general guidance, the consequences of choosing a too large γ are much worse than those deriving from the choice of a too small γ , since the effects of ill-conditioning are limited by the absence of errors.

3.5. Solution of the inverse problem

For a given choice of the parameter γ , the solution of the least square problem for the minimization of the frequency independent cost function J expressed by equation (14) can be calculated by [4]

$$\mathbf{g} = \mathbf{H}^\dagger \mathbf{p}_T \quad (17)$$

where the Moore-Penrose pseudo-inverse matrix $\mathbf{H}^\dagger$ is given by

$$\mathbf{H}^\dagger := (\mathbf{H}^H \mathbf{H})^{-1} \mathbf{H}^H \quad (18)$$

if $L < M$ (more virtual microphones than loudspeakers- overdetermined system), or by

$$\mathbf{H}^\dagger := \mathbf{H}^H (\mathbf{H} \mathbf{H}^H)^{-1} \quad (19)$$

if $L > M$ (underdetermined system). The superscript $\mathbf{H}^H$ indicates the Hermitian transpose of the matrix $\mathbf{H}$ (namely, the complex conjugate of $\mathbf{H}$ transpose). If $L = M$, the two formulae above are equivalent and $\mathbf{H}^\dagger = \mathbf{H}^{-1}$. It is assumed that the

matrix $\mathbf{H}$ is full rank. This condition is satisfied if each couple of two different loudspeakers ℓ and i identifies two different relative directions $\hat{\mathbf{y}}'_\ell$ and $\hat{\mathbf{y}}'_i$, and if different virtual microphones are arranged at different positions. In this case, either all rows or all columns of $\mathbf{H}$ are linearly independent.

The solution method above leads to some undesired effects. Firstly, loudspeakers may be active, whose relative directions $\hat{\mathbf{y}}'_\ell$ are much distant (in terms of angle between the vectors) from the relative direction of the incoming plane wave $\hat{\mathbf{k}}'_0$. This phenomenon is very pronounced when the direction of arrival of the target plane wave is between two loudspeakers and the listener is off-centered toward the direction of the plane wave, or equivalently when in an non-uniform array the target plane wave impinges from a direction that is distant from all loudspeakers.

Secondly, in the case of $L \leq M$, when the relative direction of the incoming plane wave $\hat{\mathbf{k}}'_0$ coincides with the relative direction $\hat{\mathbf{y}}'_\ell$ identified by one of the loudspeakers, only that loudspeaker generates sound. As soon as $\hat{\mathbf{k}}'_0$ varies, several other loudspeakers become active. Although the above corresponds mathematically to an exact solution, it has the psychoacoustical effect that the perceived sound scene *collapses* to the single active loudspeaker.

In order to overcome these issues, two different kinds of regularization are applied. Firstly, a classical Tikhonov (or Tihonov) regularization is applied [10]. This technique aims at minimizing the norm of the solution, corresponding to the magnitude of the vector $\mathbf{g}$. This regularization also avoids having only one loudspeaker active.

Secondly, a regularization technique proposed by Poletti [6] is applied. This technique minimizes the magnitude of the gains g_ℓ corresponding to relative directions $\hat{\mathbf{y}}'_\ell$, which are distant (in term of angle between the two vectors) from the direction of arrival of the target plane wave, $\hat{\mathbf{k}}'_0$.

The function J to be minimized becomes

$$J := |\mathbf{p}_T - \mathbf{H} \mathbf{g}|^2 + \beta_T |\mathbf{g}|^2 + \beta_P |\mathbf{b}^T \mathbf{g}|^2 \quad (20)$$

The vector $\mathbf{b}$ is given by

$$\mathbf{b} := \begin{bmatrix} b(\hat{\mathbf{y}}'_1 \cdot \hat{\mathbf{k}}'_0) \\ b(\hat{\mathbf{y}}'_2 \cdot \hat{\mathbf{k}}'_0) \\ \vdots \\ b(\hat{\mathbf{y}}'_L \cdot \hat{\mathbf{k}}'_0) \end{bmatrix} \quad (21)$$

where $b(\cdot)$ is a monotonically decreasing function, whose argument is the cosine of the angle between the unitary vectors $\hat{\mathbf{y}}'_\ell$ and $\hat{\mathbf{k}}'_0$. This function should be zero, or close to zero, when the two vectors are aligned, and should increase together with the angle between the two vectors. An example, as proposed by Poletti, is

$$b(\hat{\mathbf{y}}'_\ell \cdot \hat{\mathbf{k}}'_0) = \frac{1}{2}(1 - \hat{\mathbf{y}}'_\ell \cdot \hat{\mathbf{k}}'_0) \quad (22)$$

where β_T and β_P are regularization parameter for the Tikhonov and Poletti regularization schemes, respectively. The regularized pseudo-inverse matrix $\mathbf{H}_R^\dagger$ can be computed with the following formula

$$\mathbf{H}_R^\dagger := (\mathbf{H}^H \mathbf{H} + \beta_T \mathbf{I} + \beta_P \mathbf{B}^H \mathbf{B})^{-1} \mathbf{H}^H \quad (23)$$

where $\mathbf{B}$ is the $L \times L$ diagonal matrix such that $\mathbf{B} := \text{diag}(\mathbf{b})$. The gain vector $\mathbf{g}$ is again given by equation (17), replacing the matrix $\mathbf{H}^\dagger$ with $\mathbf{H}_R^\dagger$.

As shown in the examples below, this technique allows for an accurate reproduction of the target field in the vicinity of the listener.

4. DSP IMPLEMENTATION

A block diagram of the Digital Signal Processing required for the implementation of the panning technique presented above is shown in Figure 3. It is assumed that the absolute positions $\mathbf{y}_\ell$ of the loudspeakers are given, and also that the location of the listener is known. The latter can be either static, or given in real time as an input to the system if a listener tracking device is available. The relative virtual microphone locations are also given as a data.

The first signal processing step is to determine of the relative coordinate $\mathbf{y}'_\ell$ of the loudspeakers with respect to the listener. This operation involves a simple vector translation (it is recalled that there is no relative rotation between the absolute and relative coordinate systems), given by formula 4. In view of equation 8, it is possible to calculate the gains and

Loudsp. #	ϕ [deg]	r [m]
1	0	4.11
2	33	4.41
3	84	4
4	114	4.7
5	179	4.4
6	214	5.2
7	263	4.7
8	292	4.6

Table 1: Loudspeaker positions (polar coordinates)

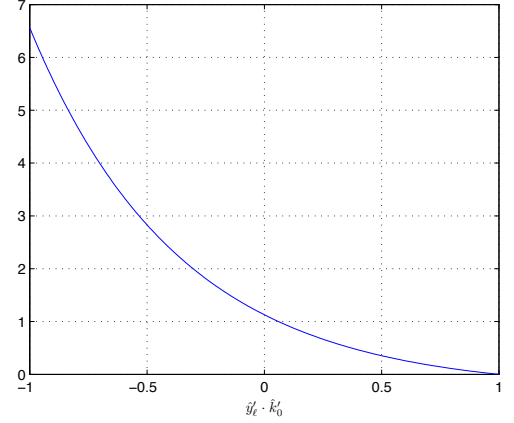
reproduced sound field. Therefore, in the case of a target sound field $\tilde{p}$ as given by equation (6) that includes multiple plane waves, the terms above can be calculated once for all plane waves, and only the computation of the vectors $\mathbf{p}_{\mathbf{T}_n}$ and of the product $\mathbf{H}_R^\dagger \mathbf{p}_{\mathbf{T}_n}$ should be calculated N times.

To validate our proposed panning technique, we have written a software program that runs inside Cycling 74's Max/MSP software environment (Figure 4). To calculate $\mathbf{H}_R^\dagger$, we used an implementation of LAPACK library. While the implementation was not fully optimized, we were able to run more than 110 instances of the panner on a laptop with 2.5GHz Intel Core 2 Duo, while the source angle was updated at 10 milliseconds interval.

We tested the performance of the algorithms in the Spatialization Laboratory at the Center for Research in Computing and the Arts (CRCA). A horizontal, non-uniform array of eight loudspeakers has been used for this purpose. The position of the transducers, in polar coordinates, are reported in Table 1. The performance of the algorithm was evaluated subjectively and provided convincing results.

5. NUMERICAL SIMULATIONS

The performance of the algorithms are demonstrated also by numerical simulations. The following simulations also use the same loudspeaker arrangement in Table 1. Ten virtual microphones are uniformly arranged on a circle around the listener. The choice of $\gamma = 1$ has been made, corresponding to an virtual microphone array radius of 110 cm for a frequency of 500 Hz (55 cm at 1000 Hz, 220 cm at 250 Hz, etc.). The regularization parameters chosen are $\beta_T = 1$

Fig. 5: Function $b(\phi_{\mathbf{y}'_\ell \cdot \mathbf{k}'_0})$ used for the Poletti regularization.

and $\beta_P = 10$, while the function $b(\hat{\mathbf{y}}'_\ell \cdot \hat{\mathbf{k}}'_0)$ used for the Poletti regularization is given by

$$b(\hat{\mathbf{y}}'_\ell \cdot \hat{\mathbf{k}}'_0) = \frac{1}{2} e^{-\frac{\pi}{6}} (e^{\frac{\pi}{2}(1-\hat{\mathbf{y}}'_\ell \cdot \hat{\mathbf{k}}'_0)} - 1.0) \quad (24)$$

and is reported in Figure 5.

Figure 6 reports an example of target and reproduced fields for two different listener locations. The color scale represent the real part of a complex valued field, with a range from -1 (blue) to 1 (red). The loudspeaker positions are indicated by black dots, while the virtual microphone are represented by $*$ symbols. A black cross indicates the listener position. The black arrows represent the gains of the loudspeakers. They are directed from the corresponding loudspeaker towards the listener and their length is proportional to the magnitude (dB) of the corresponding loudspeaker gain. The target sound field is a plane of 210 Hz wave arriving from 45° , as shown in Figure 6a.

Figure 7 is analogous to Figure 6, but with the operating frequency corresponding to 808 Hz.

The normalized reproduction error is defined by

$$\epsilon_N(\mathbf{x}, \omega) := \frac{|p_T(\mathbf{x}, \omega) - p(\mathbf{x}, \omega)|^2}{|p_T(\mathbf{x}, \omega)|^2} \quad (25)$$

where $p_T(\mathbf{x}, \omega)$ is the target sound field and $p(\mathbf{x}, \omega)$ is the field reproduced by the loudspeaker array. Since

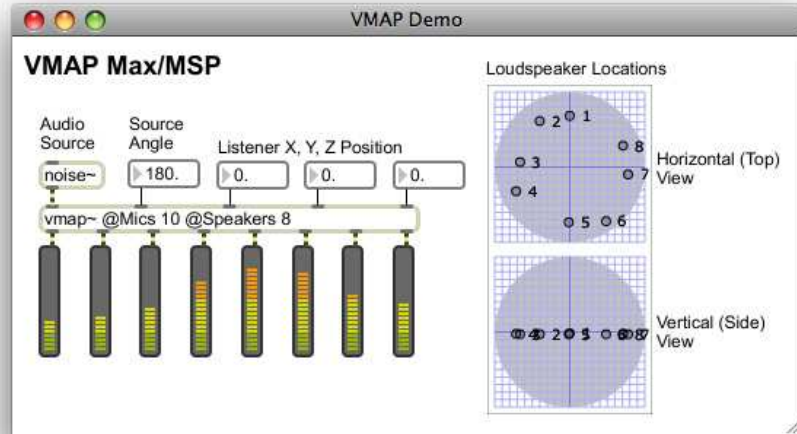


Fig. 4: vmap Max/MSP external running on Mac OS X

the target field is a plane wave, the denominator of the fraction above equals unity. In each figure, the contour line defines the area where the normalized reproduction error is less than 20%.

It can be seen that at 210 Hz an accurate reproduction of the target sound field is achieved, especially when the listener *sees* many loudspeaker in the direction of arrival of the target plane wave. This is the case of Figure 6b, but when the target plane wave arrives from a direction between two loudspeakers near the listener, as in the case of Figure 6c, the region of accurate reproduction is smaller. Yet the magnitude of the loudspeaker gains is proportional to the angle identified by the scalar product $\hat{\mathbf{y}}'_\ell \cdot \hat{\mathbf{k}}'_0$. In other words, the loudspeaker arranged in the opposite direction from the direction of arrival of the target plane wave are not contributing to the reproduced sound field.

Figure 8 gives an example of the panning algorithm for a case analogous to Figure 6c, but without the application of any regularization scheme in the computation of the matrix $\mathbf{H}^\dagger$. It can be clearly seen that the region of accurate reproduction is similar in size (if not larger) to that in Figure 6c, but all loudspeakers contribute to the reproduced field, regardless of their relative direction with respect to the direction of the target plane wave. Also, it can

be noted that the magnitude of the reproduced field in regions that are distant from the listener, and in general the acoustic energy generated by the loudspeaker array, are much larger in Figure 6c.

At high frequency the size of the reproduction region is very small, but the loudspeakers that generate most of the energy are still close to the direction of arrival of the target plane wave. It can be observed that the arrows representing the loudspeaker gains are identical for Figure 6 and Figure 7, since the gains are not a function of the frequency.

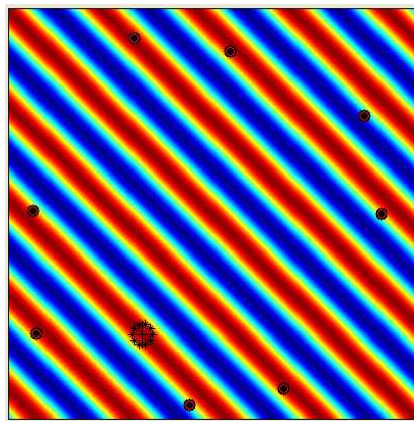
It can also be seen that the phase of the sound field is always 0° , as a consequence of the delays applied for distance compensation.

6. CONCLUSION AND FURTHER WORK

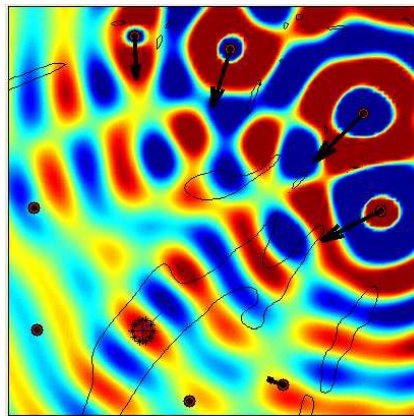
An novel panning method for surround sound systems has been presented. The Digital Signal Processing implementation involves a simple network of gains and delays, which are computed depending on the target field and on the location of the listener.

For its simplicity, the method presented is very well suited for applications where a small computational load is required and/or when the position of the listener is not static.

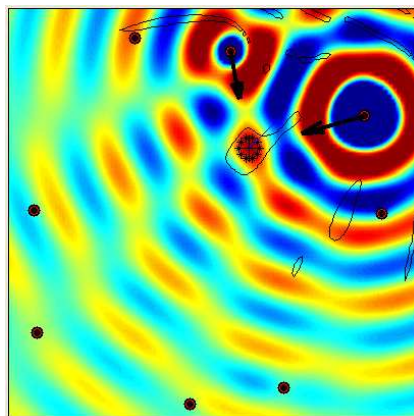
The design of the algorithm presented involves the choice of a series of important parameters, such as



(a)

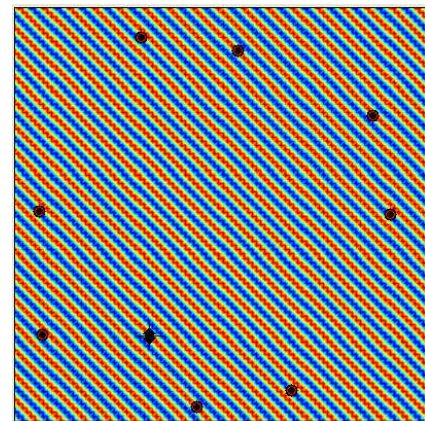


(b)



(c)

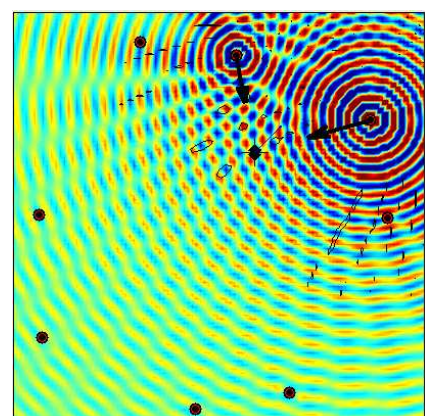
Fig. 6: Target field (a) and reproduced fields for two different listener positions (b,c). The frequency is 210 Hz.



(a)



(b)



(c)

Fig. 7: Target field (a) and reproduced fields for two different listener positions (b,c). The frequency is 808 Hz.

□

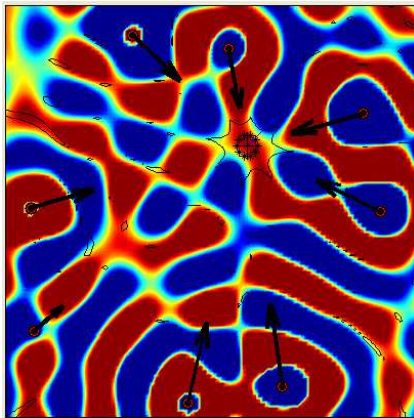


Fig. 8: Reproduced field using a panning algorithm without regularization. The frequency is 210 Hz.

the number M of virtual microphones and their arrangement, the factor $\gamma = kr$ and the regularization parameters β_T , β_P and the function $b(\hat{\mathbf{y}}'_\ell \cdot \hat{\mathbf{k}}'_0)$. No specific guidance has been given for the optimal choice of these parameters. Further work might include evaluating performance for various values of the parameters above, with the aim of automatically computing and optimizing them given a certain loudspeaker arrangement.

It would be possible to undertake a comparative study between the algorithm proposed and other widely used techniques, such as High Order Ambisonics or Vector Based Amplitude panning. Although it is not shown in this paper, the algorithm proposed could be interpreted as a variation of a real time High Order Ambisonics encoding and decoding, optimized for the given loudspeaker array and for the given listener location.

7. ACKNOWLEDGMENTS

Mr. Fazi and Prof. Nelson would like to acknowledge Dr Kyeongok Kang and Dr Jeongil Seo of the Electronic and Telecommunications Research Institute (ETRI - South Korea) for their valuable contribution to the development of some of the theoretical results used in this work.

This work was conducted with equipment and laboratory facilities provided by the Center for Research and Computing in the Arts (CRCA) at the University of California, San Diego.

This work has been partially funded by the Royal Academy of Engineering, by the Engineering and Physical Sciences Research Council and by the UK Government's Office of Science and Innovation (UK US Collaboration Programme).

This work was partially supported by Award No US-2008-107/SA-C0064, made by King Abdullah University of Science and Technology (KAUST) to the University of California, San Diego.

8. REFERENCES

- [1] A. J. Berkhout, "A holographic approach to acoustic control," *Journal of the Audio Engineering Society*, vol. 36, no. 12, pp. 977–995, 1988.
- [2] J. Daniel, R. Nicol, and S. Moureau, "Further investigations of high-order ambisonics and wavefield synthesis for holophonic sound imaging," in *114th International Convention of the Audio Engineering Society*, 2003.
- [3] V. Pulkki, "Virtual sound source positioning using vector base amplitude panning," *Journal of the Audio Engineering Society*, vol. 45, no. 6, pp. 456–466, 1997.
- [4] O. Kirkeby and P. A. Nelson, "Reproduction of plane-wave sound fields," *Journal of the Acoustical Society of America*, vol. 94, no. 5, pp. 2992–3000, 1993.
- [5] M. A. Poletti, "A unified theory of horizontal holographic sound systems," *Journal of the Audio Engineering Society*, vol. 48, pp. 1155–1182, 2000.
- [6] M. A. Poletti, "Robust two-dimensional surround sound reproduction for nonuniform loudspeaker layouts," *Journal of the Audio Engineering Society*, vol. 55, no. 7/8, pp. 598–610, 2007.
- [7] F. M. Fazi, P. A. Nelson, J. E. N. Christensen, and J. Seo, "Surround system based on three-dimensional sound field reconstruction," in *125th International Convention of the Audio Engineering Society*, (San Francisco, USA), 2008.

- [8] J. Fliege, "<http://www.personal.soton.ac.uk/jflw07/nodes/nodes.html>."
- [9] F. M. Fazi and P. A. Nelson, "The ill-conditioning problem in sound field reconstruction," in *123rd International Convention of the Audio Engineering Society*, (New York), 2007.
- [10] Deschamp.Ga and H. S. Cabayan, "Antenna synthesis and solution of inverse problems by regularization methods," *Ieee Transactions on Antennas and Propagation*, vol. Ap20, no. 3, pp. 268–274, 1972.