Virtual Sound Source Positioning Using Vector Base Amplitude Panning*

VILLE PULKKI, *AES Member*

Laboratory of Acoustics and Audio Signal Processing, Helsinki University of Technology, FIN-02015 HUT, Finland

A vector-based reformulation of amplitude panning is derived, which leads to simple and computationally efficient equations for virtual sound source positioning. Using the method, vector base amplitude panning (VBAP), it is possible to create two- or three-dimensional sound fields where any number of loudspeakers can be placed arbitrarily. The method produces virtual sound sources that are as sharp as is possible with current loudspeaker configuration and amplitude panning methods. A digital tool that implements two- and three-dimensional VBAP with eight inputs and outputs has been realized.

0 INTRODUCTION

The acoustical sound field around us is very complex. Direct sounds, reflections, and refractions arrive at the listener's ears, who then analyzes incoming sounds and connects them mentally to sound sources. Spatial hearing is an important part of the cognition of the surrounding world.

The perception of the direction of the sound source relies heavily on the two main localization cues: interaural level difference (ILD) and interaural time difference (ITD) [1]. These frequency-dependent differences occur when the sound arrives at the listener's ears after having traveled paths of different lengths or being shadowed differently by the listener's head. In addition to ILD and ITD some other cues, such as spectral coloring, are used by humans in sound source localization.

Bringing a virtual three-dimensional sound field to a listening situation is one goal of the research in the audio reproduction field. The first recordings were monophonic; they created pointlike sound fields. A big step was two-channel stereophonic reproduction, with which the sound field was enlarged to a line between two loudspeakers. Two-channel stereophony is still the most used reproduction method in domestic and professional equipment.

Various attempts to enlarge the sound field have been proposed. Horizontal-only (pantophonic) sound fields have been created with various numbers of loudspeakers and with various systems of encoding and decoding and matrixing. A review of such systems is presented in [2]. In most systems the loudspeakers are situated in a two-dimensional (horizontal) plane. Some attempts to produce periphonic (full-sphere) sound fields with three-dimensional loudspeaker placement exist, such as holophony [3] or three-dimensional Ambisonics [4].

Periphonic sound fields can be produced in two-channel loudspeaker or headphone listening by filtering the sound material with digital models of the free-field transfer functions between the listener's ear canal and the desired place of the sound source [head-related transfer functions (HRTFs)] [5]. The spectral information of the direction of the sound source is thus added to the signal emanating from the loudspeakers. The system, however, has quite strict boundary conditions, which limits its use.

In most systems the positions of the loudspeakers are fixed. In the Ambisonics systems, the number and placement of the loudspeakers may be variable. However, the best possible localization accuracy is achieved with orthogonal loudspeaker placement. If the number of loudspeakers is greater, the accuracy is not improved appreciably.

A natural improvement would be a virtual sound source positioning system that would be independent of the loudspeaker arrangement and could produce virtual sound sources with maximum accuracy using the current loudspeaker configuration.

The vector base amplitude panning (VBAP) described in this paper, first introduced in [6], is a new approach to the problem. The approach enables the use of an unlimited number of loudspeakers in an arbitrary two- or three-dimensional placement around the listener. The loudspeakers are required to be nearly equidistant from

* Manuscript received 1996 November 14; revised 1997 April 5.

the listener, and the listening room is assumed to be not very reverberant. Multiple moving or stationary sounds can be positioned in any direction in the sound field spanned by the loudspeakers.

In VBAP the amplitude panning method is reformulated with vectors and vector bases. The reformulation leads to simple equations for amplitude panning, and the use of vectors makes the panning methods computationally efficient. Two- and three-dimensional VBAP methods are presented. Applications of the methods are discussed and some comparisons are made between the three-dimensional methods proposed in this paper and those presented elsewhere. A panning tool that is capable of producing sound fields with multiple moving virtual sources in two or three dimensions is also discussed.

1 TWO-DIMENSIONAL AMPLITUDE PANNING

Two-dimensional amplitude panning, also known as intensity panning, is the most popular panning method. The applications range from small domestic stereophonic amplifiers to professional mixers. The method is, however, an approximation of real-source localization.

In the simple amplitude panning method two loudspeakers radiate coherent signals, which may have different amplitudes. The listener perceives an illusion of a single auditory event (virtual sound source, phantom sound source), which can be placed on a two-dimensional sector defined by locations of the loudspeakers and the listener by controlling the signal amplitudes of the loudspeakers. A typical loudspeaker configuration is illustrated in Fig. 1. Two loudspeakers are positioned symmetrically with respect to the median plane. Amplitudes of the signals are controlled with gain factors g_1 and g_2, respectively. The loudspeakers are typically positioned at $\varphi_0 = 30°$ angles.

The direction of the virtual source is dependent on the relation of the amplitudes of the emanating signals. If the virtual source is moving and its loudness should be constant, the gain factors that control the channel levels have to be normalized. The sound power can be set to a constant value C, whereby the following approximation can be stated:

$$g_1^2 + g_2^2 = C\,. \tag{1}$$

Some other ways to normalize the gain factors are presented in [7]. The parameter $C > 0$ can be considered a volume control of the virtual source. The perception of the distance of the virtual source depends within some limits on C—the louder the sound, the closer it is located. To control the distance accurately, some psychoacoustical phenomena should be taken into account, and some other sound elements should be added, such as reflections and reverberation [1].

When the distance of the virtual source is left unattended, the virtual source can be placed on an arc between the loudspeakers, the radius of which is defined by the distance between the listener and the loudspeakers. The arc is called the active arc, as seen in Fig. 1.

In the ideal panning process only the direction where the virtual source should appear is defined and the panning tool performs the gain factor calculation. In the next two subsections some different ways of calculating the factors will be presented.

1.1 Trigonometric Formulation

The directional perception of a virtual sound source produced by amplitude panning follows approximately the stereophonic law of sines originally proposed by Blumlein [8] and reformulated in phasor form by Bauer [9],

$$\frac{\sin\varphi}{\sin\varphi_0} = \frac{g_1 - g_2}{g_1 + g_2} \tag{2}$$

where $0° < \varphi_0 < 90°$, $-\varphi_0 \leq \varphi \leq \varphi_0$, and $g_1, g_2 \in [0, 1]$. In Eq. (2) φ represents the angle between the x axis and the direction of the virtual source; $\pm\varphi_0$ is the angle between the x axis and the loudspeakers. This equation is valid if the listener's head is pointing directly forward. If the listener turns his or her head following the virtual source, the tangent law is more correct [10],

$$\frac{\tan\varphi}{\tan\varphi_0} = \frac{g_1 - g_2}{g_1 + g_2} \tag{3}$$

where $0° < \varphi_0 < 90°$, $-\varphi_0 \leq \varphi \leq \varphi_0$, and $g_1, g_2 \in [0, 1]$. Eqs. (2) and (3) have been calculated with the assumption that the incoming sound is different only in magnitude, which is valid for frequencies below 500–600 Hz. When keeping the sound power level constant, the gain factors can be solved using Eqs. (2) and (1) or using Eqs. (3) and (1). The slight difference between Eqs. (2) and (3) means that the rotation of the head causes small movements of the virtual sources. However, in subjective tests [11] it was shown that this effect is negligible.

Some kind of amplitude panning method is used in the Ambisonics encoding system [12]. In pantophonic Ambisonics the entire sound field is decoded to three channels using a modified amplitude panning method. Two of the channels, X and Y, contain the components of the sound on the x axis and the y axis, respectively. The third, W, contains a monophonic mix of the sound material. The signal to be stored on the channels is calcu-

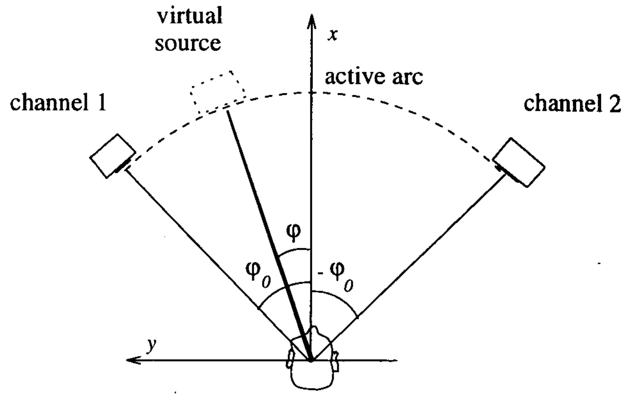

Fig. 1. Two-channel stereophonic configuration.

lated by multiplying the input signal samples by the channel-specific gain factor. The gain factors g_x, g_y, and g_w are formulated as

$$g_x = \cos\theta \tag{4}$$

$$g_y = \sin\theta \tag{5}$$

$$g_w = 0.707 \tag{6}$$

where θ is the azimuth angle of the virtual sound source, as illustrated in Fig. 2.

This method differs from the standard amplitude panning method in that the gain factors g_x and g_y may have negative values. The negative values imply that the signal is stored on the recorder in antiphase when compared with the monophonic mix in the W channel. When the decoded sound field is encoded, the antiphase signals on a channel are applied to the loudspeakers in a negative direction of the respective axis. The decoding stage is performed with matrixing equations [12], which are not discussed in this paper. In the equations some additions or subtractions are performed between the signal samples on the W channel and on the X and Y channels. Equations for various loudspeaker configurations can be formulated.

The absolute values of the gain factors used in two-dimensional Ambisonics satisfy the tangent law [Eq. (3)], which the reader may verify, for example, for values of $0° < \theta < 90°$, by setting $\theta = \varphi_0 + \varphi$, $\varphi_0 = 45°$, $g_2 = g_x$, and $g_1 = g_y$, and by substituting Eqs. (4) and (5) into the relation $(g_y - g_x)/(g_y + g_x)$.

1.2 Vector Base Formulation

In the two-dimensional VBAP method presented in this section, the two-channel stereophonic loudspeaker configuration is reformulated as a two-dimensional vector base. The base is defined by unit-length vectors $\boldsymbol{l}_1 = [l_{11} \quad l_{12}]^{\mathrm{T}}$ and $\boldsymbol{l}_2 = [l_{21} \quad l_{22}]^{\mathrm{T}}$, which are pointing toward loudspeakers 1 and 2, respectively, as seen in Fig. 3. The superscript T denotes the matrix transposition. The unit-length vector $\boldsymbol{p} = [p_1 \quad p_2]^{\mathrm{T}}$, which points toward the virtual source, can be treated as a linear combination of loudspeaker vectors,

$$\boldsymbol{p} = g_1\boldsymbol{l}_1 + g_2\boldsymbol{l}_2 . \tag{7}$$

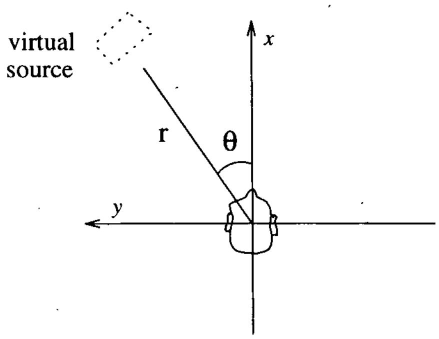

Fig. 2. Coordinate system of two-dimensional Ambisonics system. θ—azimuth angle; r—distance of virtual source.

In Eq. (7) g_1 and g_2 are gain factors, which can be treated as nonnegative scalar variables. We may write the equation in matrix form,

$$\boldsymbol{p}^{\mathrm{T}} = \boldsymbol{g}\boldsymbol{L}_{12} \tag{8}$$

where $\boldsymbol{g} = [g_1 \quad g_2]$ and $\boldsymbol{L}_{12} = [\boldsymbol{l}_1 \quad \boldsymbol{l}_2]^{\mathrm{T}}$. This equation can be solved if $\boldsymbol{L}_{12}^{-1}$ exists,

$$\boldsymbol{g} = \boldsymbol{p}^{\mathrm{T}}\boldsymbol{L}_{12}^{-1} = [p_1 \quad p_2]\begin{bmatrix} l_{11} & l_{12} \\ l_{21} & l_{22} \end{bmatrix}^{-1} . \tag{9}$$

The inverse matrix $\boldsymbol{L}_{12}^{-1}$ satisfies $\boldsymbol{L}_{12}\boldsymbol{L}_{12}^{-1} = \boldsymbol{I}$, where $\boldsymbol{I}$ is the identity matrix. $\boldsymbol{L}_{12}^{-1}$ exists when $\varphi_0 \neq 0°$ and $\varphi_0 \neq 90°$, both problem cases corresponding to quite uninteresting stereophonic loudspeaker placements. For such cases the one-dimensional VBAP can be formulated, which is not discussed here because of its triviality.

Gain factors g_1 and g_2 calculated using Eq. (9) satisfy the tangent law of Eq. (3), which is proved in the Appendix. When the loudspeaker base is orthogonal, $\varphi_0 = 45°$, the gain factors are also equivalent to those calculated for the Ambisonics encoding system, with the exception that the gain factors in Ambisonics may have negative values. In such cases, however, the absolute values of the factors are equal.

When $\varphi_0 \neq 45°$, the gain factors have to be normalized using the equation

$$\boldsymbol{g}^{\mathrm{scaled}} = \frac{\sqrt{C}\boldsymbol{g}}{\sqrt{g_1^2 + g_2^2}} . \tag{10}$$

Now gain factors $\boldsymbol{g}^{\mathrm{scaled}}$ satisfy Eq. (1).

1.3 Two-Dimensional VBAP for More Than Two Loudspeakers

In many existing audio systems there are more than two loudspeakers in the horizontal plane, such as in Dolby Surround systems. Such systems can also be reformulated with vector bases. A set of loudspeaker pairs is selected from the system, and the signal is applied at

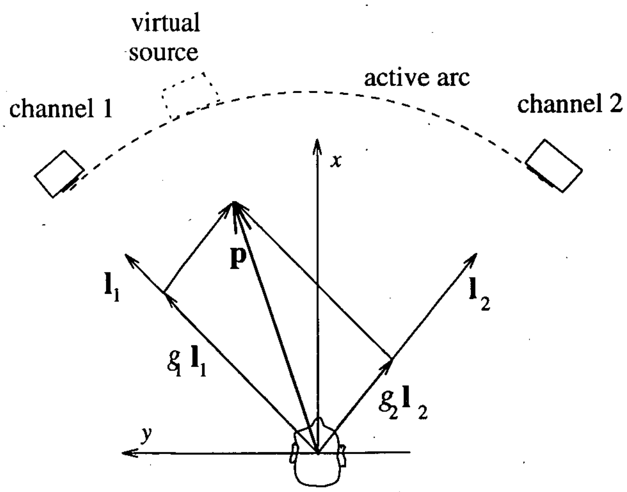

Fig. 3. Stereophonic configuration formulated with vectors.

any one time to only one pair. Thus the loudspeaker system consists of many vector bases competing among themselves. Each loudspeaker may belong to two pairs. In Fig. 4 a loudspeaker system in which the two-dimensional VBAP can be applied is illustrated. A system for virtual source positioning, which similarly uses only two loudspeakers at any one time, has been implemented in an existing theater [13].

The virtual source can be produced by the loudspeaker base on the active arc of which the virtual source is located. Thus the sound field that can be produced with VBAP is a union of the active arcs of the available loudspeaker bases. In two-dimensional cases the best way of choose the loudspeaker bases is to let the adjacent loudspeakers form them. In the loudspeaker system illustrated in Fig. 4 the selected bases would be $\boldsymbol{L}_{12}$, $\boldsymbol{L}_{23}$, $\boldsymbol{L}_{34}$, $\boldsymbol{L}_{45}$, and $\boldsymbol{L}_{51}$. The active arcs of the bases are thus nonoverlapping.

The use of nonoverlapping active arcs provides continuously changing gain factors when moving virtual sources are applied. When the sound moves from one pair to another, the gain factor of the loudspeaker, which is not used after the change, becomes gradually zero before the change-over point. See Section 6.1 for a demonstration of this effect in a three-dimensional case.

The fact that all other loudspeakers except the selected pair are idle may seem a waste of resources. In this way, however, good localization accuracies can be achieved for the principal sound, whereas the other loudspeakers may produce reflections and reverberation as well as other sound elements.

1.4 Implementing Two-Dimensional VBAP for More Than Two Loudspeakers

A digital panning tool that performs the panning process is now considered. Sufficient hardware consists of a signal processor that can perform input and output with multiple analog-to-digital (A/D) and digital-to-analog (D/A) converters and has enough processing power for the computation needed. The tool has to include also a user interface.

When the tool is initialized, the directions of the loudspeakers are measured relative to the best listening position and loudspeaker pairs are formed from adjacent loudspeakers. $\boldsymbol{L}_{nm}^{-1}$ matrices are calculated for each pair and stored in the memory of the panning system.

During run time the system performs the following steps in an infinite loop:

- New direction vectors $\boldsymbol{p}_{(1,\ldots,n)}$ are defined.
- The right pairs are selected.
- The new gain factors are calculated.
- The old gain factors are cross faded to new ones and the loudspeaker bases are changed if necessary.

The pair can be selected by calculating unscaled gain factors with Eq. (9) using all selected vector bases, and by selecting the base that does not produce any negative factors. In practice it is recommended to choose the pair with the highest smallest factor, because a lack of numerical accuracy during calculation may produce slightly negative gain factors in some cases. The negative factor must be set to zero before normalization.

A digital panning tool that can be used to position sounds on the horizontal plane with a variable number of loudspeakers has been constructed and will be discussed in Section 4.

2 THREE-DIMENSIONAL AMPLITUDE PANNING

The typical two-channel stereophonic listening configuration is extended with a third loudspeaker placed in an arbitrary position at the same distance from the listener as the other loudspeakers. However, the loudspeaker should not be placed on the two-dimensional plane defined by the listener and the two other loudspeakers. The virtual source can now appear within a triangle formed by the loudspeakers when viewed from the listener's position, as illustrated in Fig. 5. The term three-dimensional amplitude panning denotes a method for positioning a virtual sound source into a triangle formed by three sound sources, which are driven by coherent electrical signals with different amplitudes.

Now the relation of three gain factors defines the virtual source direction perceived by the listener. Eq. (1) can be generalized into a three-dimensional form as

$$g_1^2 + g_2^2 + g_3^2 = C\,. \tag{11}$$

The virtual source can thus be placed on the surface of the three-dimensional sphere, the radius of which is defined by the distance between the listener and the loudspeakers. The region on the surface of the sphere onto which the virtual source can be positioned is called the active triangle.

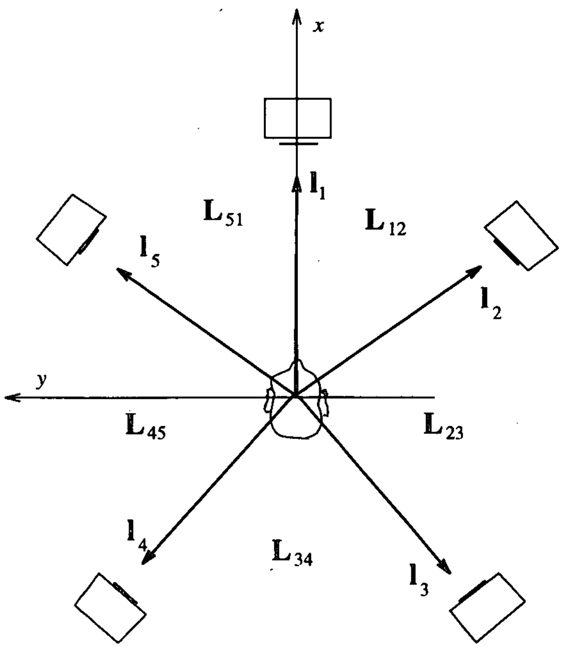

Fig. 4. Two-dimensional VBAP with five loudspeakers. Vectors $\boldsymbol{l}_n$ point to loudspeakers; loudspeaker vector bases $\boldsymbol{L}_{nm}$ are formed from adjacent loudspeakers.

2.1 On the Trigonometric Formulation

The general trigonometric formulation of three-dimensional amplitude panning for arbitrary loudspeaker placement remains unexplored. Spherical trigonometry is complicated and computationally inefficient.

The periphonic Ambisonics system [4], [12] decodes the three-dimensional sound field to four channels. Channels X, Y, and Z correspond to the axes of the Cartesian coordinate system, and channel W contains a monophonic mix of the input material.

The formulation is analogous to the two-dimensional formulation mentioned in Section 1. The channel-specific gain factors g_x, g_y, g_z, and g_w are calculated as

$$g_x = \cos\theta\cos\gamma \tag{12}$$

$$g_y = \sin\theta\cos\gamma \tag{13}$$

$$g_z = \sin\gamma \tag{14}$$

$$g_w = 0.707 \tag{15}$$

where γ is the elevation angle and θ the azimuth angle, as presented in Fig. 6. The gain factor calculation is similar to a coordinate transformation between spherical coordinates and Cartesian coordinates. The gain factors are thus Cartesian coordinates of the unit-length vector pointing in the direction (θ, γ) in the defined spherical coordinate system. The decoding stage is not discussed in this paper.

2.2 Vector Base Formulation

The two-dimensional VBAP method may now be generalized to the three-dimensional VBAP method. Let the loudspeakers be positioned on the surface of a three-dimensional unit sphere, equidistant from the listener. The three-dimensional unit vector $\boldsymbol{l}_1 = [l_{11} \quad l_{12} \quad l_{13}]^{\mathrm{T}}$, the origin of which is the center of the sphere, points to the direction of loudspeaker 1. The unit vectors $\boldsymbol{l}_1$, $\boldsymbol{l}_2$, and $\boldsymbol{l}_3$ then define the directions of loudspeakers 1, 2, and 3, respectively. The direction of the virtual sound source is defined as a three-dimensional unit vector $\boldsymbol{p} = [p_1 \quad p_2 \quad p_3]^{\mathrm{T}}$. A sample configuration is presented in Fig. 5.

We express the virtual source vector $\boldsymbol{p}$ as a linear combination of three loudspeaker vectors $\boldsymbol{l}_1$, $\boldsymbol{l}_2$, and $\boldsymbol{l}_3$, analogically to the two-dimensional case, and express it in matrix form,

$$\boldsymbol{p} = g_1\boldsymbol{l}_1 + g_2\boldsymbol{l}_2 + g_3\boldsymbol{l}_3 \tag{16}$$

$$\boldsymbol{p}^{\mathrm{T}} = \boldsymbol{g}\boldsymbol{L}_{123}\,. \tag{17}$$

Here g_1, g_2, and g_3 are gain factors, $\boldsymbol{g} = [g_1 \quad g_2 \quad g_3]$, and $\boldsymbol{L}_{123} = [\boldsymbol{l}_1 \quad \boldsymbol{l}_2 \quad \boldsymbol{l}_3]^{\mathrm{T}}$. Vector $\boldsymbol{g}$ can be solved,

$$\boldsymbol{g} = \boldsymbol{p}^{\mathrm{T}}\boldsymbol{L}_{123}^{-1} = [p_1 \quad p_2 \quad p_3] \begin{bmatrix} l_{11} & l_{12} & l_{13} \\ l_{21} & l_{22} & l_{23} \\ l_{31} & l_{32} & l_{33} \end{bmatrix}^{-1} \tag{18}$$

if $\boldsymbol{L}_{123}^{-1}$ exists, which is true if the vector base defined by $\boldsymbol{L}_{123}$ spans a three-dimensional space. Eq. (18) makes a projection of vector $\boldsymbol{p}$ to a vector base defined by $\boldsymbol{L}_{123}$ in a similar way as in the two-dimensional case. The components of vector $\boldsymbol{g}$ can be used as gain factors after scaling, which is given by

$$\boldsymbol{g}^{\text{scaled}} = \frac{\sqrt{C}\boldsymbol{g}}{\sqrt{g_1^2 + g_2^2 + g_3^2}}\,. \tag{19}$$

When the three loudspeakers are placed in an orthogonal grid, the gain factors calculated with the three-dimensional VBAP are equivalent to the absolute values of gain factors calculated in the three-dimensional Ambisonics system. This is easily proved. From the orthogonality of the loudspeaker vector base we see that $\boldsymbol{L}_{123}$ =

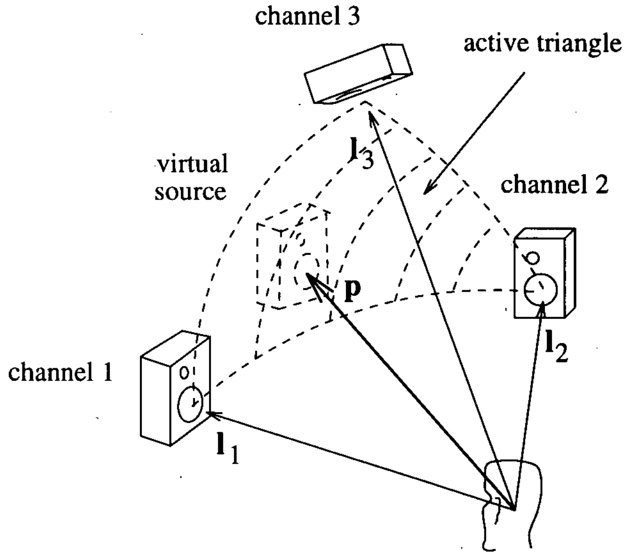

Fig. 5. Sample configuration for three-dimensional amplitude panning. Loudspeakers form a triangle into which the virtual source can be placed.

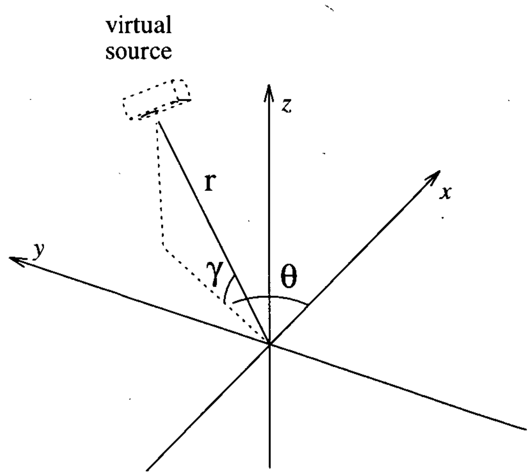

Fig. 6. Coordinate system used in three-dimensional Ambisonics encoding and decoding. θ—azimuth angle; γ—elevation angle; r—distance of virtual source.

$\boldsymbol{l} = \boldsymbol{L}_{123}^{-1}$. Using Eq. (18) we see directly that $\boldsymbol{g} = \boldsymbol{p}^{\mathrm{T}}$. The gain factors are thus the Cartesian coordinates of the head of the virtual source direction vector $\boldsymbol{p}$, similarly as in the three-dimensional Ambisonics system.

2.3 Three-Dimensional VBAP for More Than Three Loudspeakers

The three-dimensional VBAP can be applied to systems that consist of more than three loudspeakers in an arbitrary three-dimensional placement. The formulation of such a system is very much the same as in the two-dimensional case presented in Section 1.3. Some differences exist, however. The number of loudspeakers in a base is obviously three, and each loudspeaker can belong to several bases. The active triangles of bases should not be intersecting, and they should be selected so that maximum localization accuracy in each direction is provided. A sample configuration with five loudspeakers is illustrated in Fig. 7. In this case the selected loudspeaker bases are $\boldsymbol{L}_{145}$, $\boldsymbol{L}_{345}$, and $\boldsymbol{L}_{235}$.

A digital panning tool that is able to select the loudspeaker triplet and to calculate the gain factors can be constructed as in the two-dimensional case. The tool demands a little more computing power than the two-dimensional panning tool, but it can still be implemented easily with a modern floating-point signal processor.

The tool is initialized in a similar way and runs similarly to the two-dimensional case. Selection of the triplet is performed as in the two-dimensional case. See Section 1.4 for details. A digital panning tool with eight input and output channels has been constructed [6] and will be discussed in Section 4.

3 SOME FEATURES OF VBAP

In VBAP, as in all amplitude panning methods, the virtual source can not be positioned outside the active arc or region. This holds even if the listener is in an arbitrary position. Thus the maximum error in the virtual source localization is proportional to the dimensions of the active region. Therefore when good localization accuracies on a large listening area are desired, the dimensions of the active regions must be decreased. This is done by applying more loudspeakers on the desired region of the sound field, such as around and behind the screen in movie theaters.

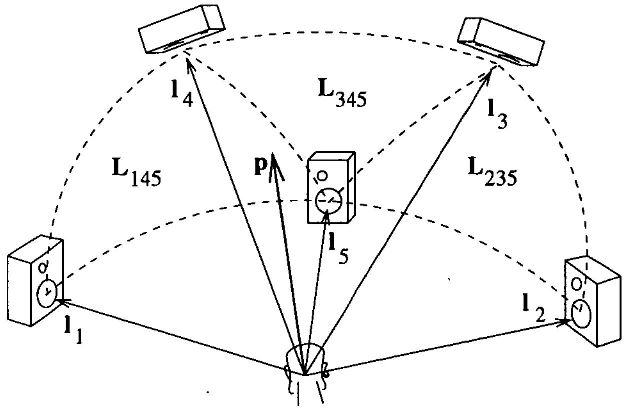

Fig. 7. Three-dimensional VBAP with five loudspeakers. Vectors $\boldsymbol{l}_n$ point to loudspeakers; selected loudspeaker bases $\boldsymbol{L}_{ijk}$ are shown by dashed lines.

The decreasing sizes of the triangles permit also differing distances between a loudspeaker and the listener (loudspeaker distances). Since the coherent signal is applied to only three loudspeakers at one time, only the difference between the loudspeaker distances of the three particular loudspeakers adds error to the perceived direction of the virtual sound source. The loudspeaker distances can be much greater in one end of the listening room if there are enough loudspeakers in between to ensure smooth enough changes in distances. The different loudspeaker distances can be compensated by time shifting and gaining the loudspeaker signals, which enables even freer loudspeaker placement.

VBAP has three important properties:

1) If the virtual source is located in the same direction as any of the loudspeakers, the signal emanates only from that particular loudspeaker, which provides maximum sharpness of the virtual source.

2) If the virtual source is located on a line connecting two loudspeakers, the sound is applied only to that pair, following the tangent law. The gain factor of the third loudspeaker is zero.

3) If the virtual source is located at the center of the active triangle, the gain factors of the loudspeakers are equal.

These properties imply that VBAP produces virtual sound sources that are as sharp as it is possible with present loudspeaker configurations.

4 DSP TOOL FOR TWO- AND THREE-DIMENSIONAL VBAP

4.1 System Overview

A tool for two- and three-dimensional VBAP has been constructed [6], [14]. The tool consists of eight A/D and eight D/A converters and two Loughborough Sound Images MDC40S modules which adhere to Texas Instruments TIM-40 specification [15], each module having a TMS320C40 (C40) processor. The system has as a host a Macintosh computer. A single C40 is used in the VBAP implementation, the other one is reserved for future extensions. An overview of the system is presented in Fig. 8.

The sample rate of the digital signal processing (DSP) tool is currently 32 kHz, but higher rates (44.1 and 48 kHz) are also supported. The number of loudspeakers can range from two or three to eight, and they can be located anywhere on the edge of a two-dimensional circle or on the surface of a three-dimensional sphere. The maximum number of input channels is eight. Each of them has a virtual source direction vector of its own.

The software implementation was carried out using the QuickSig and QuickC30 DSP programming environments [16]. QuickSig and QuickC30 are based on Common Lisp and its object-oriented extension CLOS, and they support low-level TMS320C3x assembly language with high-level object-based programming.

4.2 VBAP Implementation

Each input channel is panned into two or three output channels. Panning is performed additively: when multiple input channels share the same output channels, the signal values are added together.

The tool has two levels of interpolation for virtual source direction movement. The user may update the direction vectors $\boldsymbol{p}_{1,\ldots,8}$ approximately once per second. The panning tool calculates for example 50 vectors $\boldsymbol{p}_{1,\ldots,8}(1, \ldots, 50)$ between new and old direction vectors. With each interpolating direction vector set $\boldsymbol{p}_{1,\ldots,8}(n)$ new loudspeaker triplets are selected and new gain factors are calculated using the VBAP method. The gain factor calculation is carried out at all eight input channels during approximately 32 sample intervals.

The previous gain factors $\boldsymbol{p}_{1,\ldots,8}(n-1)$ are cross faded to calculate factors $\boldsymbol{p}_{1,\ldots,8}(n)$ linearly. One interpolation is completed with equal steps during approximately 100 sample intervals. All eight gain factor triplets are cross faded simultaneously. The gain factors do not exactly satisfy Eq. (11) during fading. However, when the angle between starting-point and end-point direction vectors is small ($\approx 1°$), no disturbing effects can be heard.

If the movements of virtual sound sources are designed beforehand, the virtual source direction vectors $\boldsymbol{p}_{1,\ldots,8}$ are then calculated for each desired arrangement of virtual source directions and stored in the signal processor's memory. Movements can also be controlled in real time from the host computer. In such cases the direction vectors $\boldsymbol{p}_{1,\ldots,8}$ are written in the signal processor's memory during run time. This enables three-dimensional live sound positioning. Fig. 9 illustrates a possible system configuration to be used with the three-dimensional panning tool.

5 USING VBAP

5.1 Off-Line Panning

An off-line panned sound field is produced for a fixed loudspeaker configuration and stored in a multichannel audio recording device. Off-line panning can be used if the loudspeaker placement is fixed, such as any surround system or a hypothetical three-dimensional loudspeaker placement standard for cinemas and theaters.

In off-line panning each channel contains the signal to be radiated from the respective loudspeaker. No special hardware is needed in the listening phase. The three-dimensional sound field thus produced could have a large number of virtual sources, which could all be moving or stationary.

Off-line panning can also be performed asynchronously. The output channels can be stored in a multichannel audio file in a computer. The file can be stored afterward in a multichannel recorder. The VBAP method could be a part of existing software, such as CSound [17] or ProTools [18].

5.2 On-Line Panning

If the loudspeaker configuration is known to be variable, or if the sound signal is to be placed in the sound field during the reproduction phase, the material can be stored in a multichannel recorder without panning. During the reproduction phase a digital panning tool initialized with the current loudspeaker placement should perform the panning process following the user's commands or the information stored in the recorder.

Hybrid systems can be constructed as well. Some part of the material can be panned off line to a few channels forming a sound field, which can be panned on line to the position desired in the sound field. In movies, for example, the sound track could be panned off line to two or three channels to be placed in a stationary position in the sound field, whereas lines and effects could be in separate tracks, which could be positioned on line anywhere in the sound field.

In theaters the on-line panning with VBAP would be very useful. The system would position the virtual sound source at the spot where the instrument or the singer is located on the stage. The places of the microphones

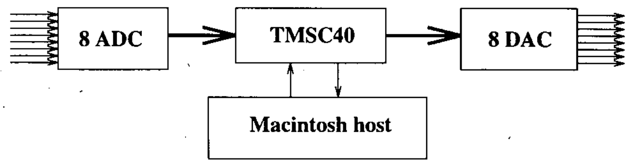

Fig. 8. Overview of hardware running two- or three-dimensional VBAP.

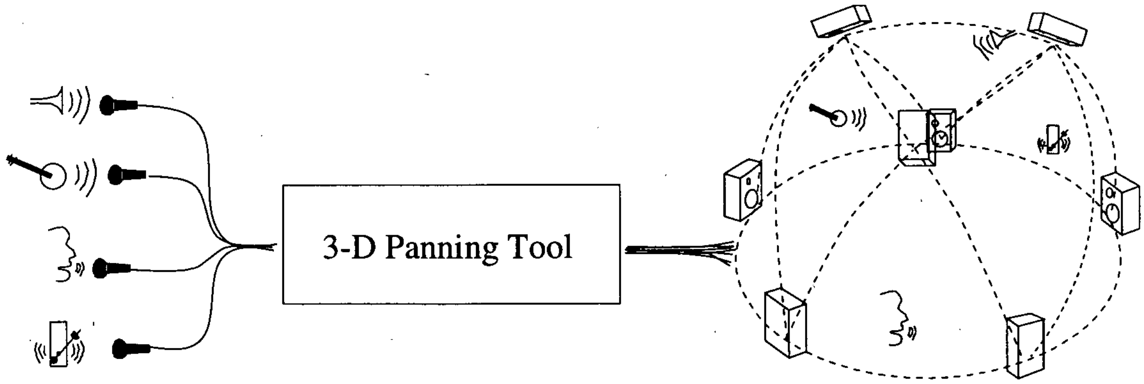

Fig. 9. Possible use of three-dimensional VBAP panning tool. Number of sound sources can vary up to eight; loudspeaker placement is arbitrary; virtual sources may be moving or stationary.

could be tracked automatically, which could enable the positioning of multiple virtual sound sources at the same time.

The VBAP could also be applied in virtual reality, multimedia, computer games, on-stage monitoring, artificial reverberation, to name a few.

5.3 Three-Dimensionally Panned Piece of Computer Music

A piece of computer music panned with the three-dimensional panning tool has been performed in a public concert in an existing concert hall, the Chamber Music Hall at the Sibelius Academy in Helsinki, Finland. In the hall there are 96 loudspeakers, grouped as 32 channels [19]. The 32 channels were further grouped into eight channels. A virtual source was created with each of the six channels of the piece. The virtual sources were moving in the sound field as the composer desired. The panning was completed with the three-dimensional panning tool discussed in Section 4. The processed material was stored on an eight-channel Alesis ADAT digital audio tape.

6 SOME EXPERIMENTS WITH VBAP

6.1 Gain Factor Monitoring of a Moving Virtual Source

A test run for monitoring the values of gain factors during the virtual source movement was completed with the three-dimensional panning tool described in Section 4. The loudspeaker configuration is illustrated in Fig. 10. Fig. 11 shows the changes in gain factors during the virtual source movement. The virtual source was moved along the solid line from loudspeaker 3 through points A, B, and C to loudspeaker 4. At each point marked with a letter the gain factors were monitored twice.

In the beginning of the test, the user commanded the computer to position the virtual source at location (0, 50) (azimuth and elevation angles), which was the point where loudspeaker 3 (L3) was situated. The gain factor g_3 was adjusted to the value 1.0 by the tool, while others shared 0.0, as expected. Next the user prompted 10 points on the arc between loudspeakers 2 and 3 with equal steps, ending at point B. From Fig. 11 we can see that g_3 is decreased while g_2 is increased, and g_1 remains zero.

From point A the source was moved similarly to point B, which was situated in the center of the triangle. The tool changes gain factor values smoothly to nearly equal values, $g_1 = g_2 = g_3 = 0.578$, which is as stated by the theory. From point B the virtual source was moved to point C, which was in the middle of the arc between loudspeakers 2 and 3. Gain factors obtained values of $g_2 = g_3 = 0.707$ and $g_1 = 0.0$. From point C the virtual source was moved to the point defined by loudspeaker 4 (60, 40) in 20 equal-sized steps. Immediately after point C the tool changed the loudspeaker triplet to which the signal was panned from the triplet 123 to triplet 234. In Fig. 11 it can be seen that g_4 increases smoothly to the value 1.0, whereas g_3 and g_2 decrease to the value 0.0. The test thus showed that the properties of VBAP stated in Section 3 can be achieved with a digital tool.

6.2 Tests with Extreme Loudspeaker Triangles

Some tests were run with different sized loudspeaker triangles using the three-dimensional panning tool described in Section 4. In the tests the virtual source was positioned at a few locations on the triangle, the geometry of which was varied. The gain factors were monitored. The first tests were conducted to see how small triangles could be used. A tested triangle had 5° sides;

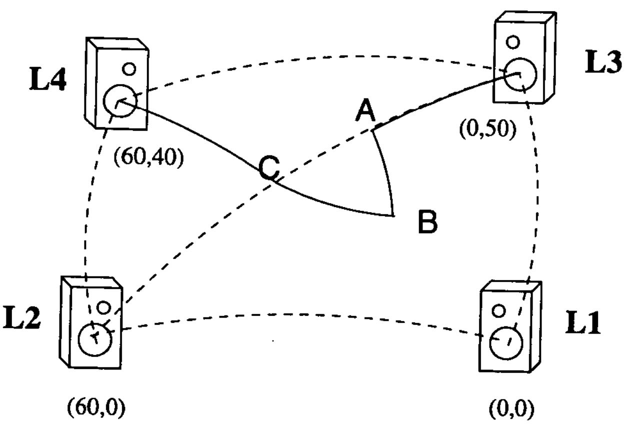

Fig. 10. Loudspeaker placement and virtual source movement in gain factor monitoring test. Dashed lines are boundaries of active triangles of loudspeaker bases. Virtual source movement is marked by solid line.

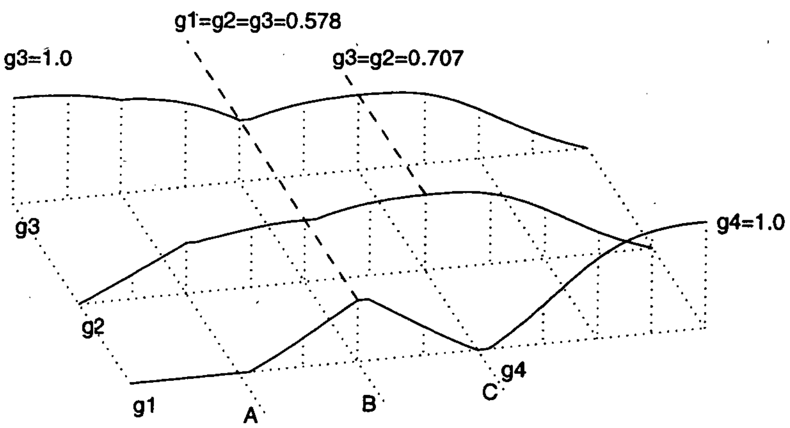

Fig. 11. Gain factor magnitudes during virtual source movement. Gain factors are numbered according to loudspeaker numbering in Fig. 10. Note that g_1 changes to g_4 at point C.

the error of the gain factors calculated was found to be negligible. In the second test a triangle that also hardly spanned a three-dimensional space was configured. The sides of the triangle were 5°, 175°, and 175°. Still the gain factors had the expected values. Thus it can be stated that if the three-dimensional VBAP tool has 32-bit floating-point calculation accuracy, it does not set limitations on the placement of the loudspeakers.

7 COMPARING THREE-DIMENSIONAL VBAP WITH EXISTING THREE-DIMENSIONAL PANNING METHODS

In this section the three-dimensional VBAP is compared with the Ambisonics system, with very large arrays of loudspeakers, and with HRTF-based systems. The audio control equipment and the software of the Level Control Systems (LCS) [20] cannot be compared with the VBAP method because there is not enough information available. However, there seem to exist some similarities between VBAP and LCS.

7.1 Ambisonics

The Ambisonics surround sound system is a large system for storing and reproducing two- and three-dimensional sound fields [4], [12]. The main drawback of the Ambisonics system is that the sound-storing format supports the best loudspeakers placed on the axes of the Cartesian coordinate system, as seen in Fig. 6. The number of loudspeakers can be increased. However, in such cases the number of loudspeakers radiating a coherent signal increases, which does not improve the quality of the virtual sources.

In VBAP each selected loudspeaker triplet is a coordinate system of its own. The gain factor calculation in the VBAP method equals that of the Ambisonics in an orthogonal loudspeaker placement. VBAP thus generalizes the gain factor calculation of Ambisonics to nonorthogonal situations, which provides great flexibility in loudspeaker placement with maximum accuracy. A drawback of the VBAP method is that in many cases it needs more than four channels for sound storage.

7.2 Very Large Arrays of Loudspeakers

In very large arrays of loudspeakers (VLALs) the sound is always emanating from one of many loudspeakers. These types of systems are used mainly for scientific purposes, as in [21]. VBAP can be considered an assistant system to a VLAL, since it does not weaken the localization quality achieved with the array system. If the direction vector of the virtual source is equal to the direction vector of any of the loudspeakers, the sound is emanating only from it, as stated in Section 3. Thus the system gives the same localization accuracy as the VLAL, but the sound can also be placed and moved between loudspeakers.

7.3 HRTF-Based Systems

The use of head-related transfer functions (HRTFs) in producing three-dimensional sound fields for loudspeaker reproduction has been explored [5]. When filtering signals using digital models of HRTFs it is possible to create illusions of directions of virtual sound sources for the listener in stereophonic headphone or loudspeaker listening.

Using HRTF-based systems a three-dimensional sound field can be produced with only two loudspeakers, which permits reproduction in common domestic stereophonic equipment. VBAP is a method for panning virtual sources to any number of loudspeakers in an arbitrary placement. The comparison of the qualities of the virtual sources produced with HRTFs and VBAP is an interesting topic, which can be studied after listening tests on VBAP have been conducted.

The filters that model HRTFs are computationally quite expensive. In some cases the filtering of one sample of the sound material requires approximately 50–150 multiplications and additions per sample of a virtual source. There exist some simplified models [22], and by using them the computing requirement can be decreased to as few as 12 multiplications and additions per a virtual source sample. In three-dimensional VBAP only three multiplications are required per a virtual source sample.

8 CONCLUSIONS

New methods for virtual source positioning in two- and three-dimensional sound fields have been introduced. The two-dimensional amplitude panning has been reformulated with vectors and vector bases to a two-dimensional VBAP method. The two-dimensional VBAP method follows a traditional panning method. It has been generalized in this paper to a three-dimensional VBAP method. The three-dimensional VBAP is a general method for virtual source positioning in a three-dimensional sound field formed by loudspeakers in an arbitrary three-dimensional placement.

VBAP is computationally efficient and accurate. The loudspeakers may be in arbitrary two- or three-dimensional positioning. VBAP gives a maximum localization sharpness that can be achieved with amplitude panning since it uses at any one time the minimum number of loudspeakers needed: one, two, or three. The number of virtual sound sources or loudspeakers is not limited by the method.

By using VBAP it is possible to make recordings for any loudspeaker configuration or create recordings that are independent of loudspeaker placement. Multiple virtual sources can be positioned in two- or three-dimensional sound fields, even with very complex loudspeaker configurations.

VBAP is more flexible than the Ambisonics sound reproduction system because of the free loudspeaker placement. VBAP is a computationally more efficient three-dimensional sound positioning method than are HRTF-based methods.

A digital tool has been constructed which can position eight virtual sources in a two- or three-dimensional sound field formed by eight arbitrarily placed loudspeakers. The tool is also able to move virtual sources in the sound field independently of each other. A three-dimensionally panned piece of computer music has been performed successfully in an existing concert hall using the proposed technique.

9 ACKNOWLEDGMENT

The author would like to thank Professor Matti Karjalainen, Vesa Välimäki, Toomas Altosaar, Jyri Huopaniemi, Samuli Siltanen, and the personnel of Acoustics Laboratory at Helsinki University of Technology for support in the project. Pauli Laine and Andrew Bentley at the Sibelius Academy Computer Music Studio and Juhani Borenius at the Finnish Broadcasting Company Ltd. have also given valuable support. The author would like to thank also Professor Teuvo Kohonen at Helsinki University of Technology. The VBAP theory presented in this paper uses partly the same way of thinking as his theory of the self-organizing map [23]. The author's wife, Sari Löytymäki, deserves also warm thanks. This project has been partially financed by the Sibelius Academy, Helsinki, Finland.

10 REFERENCES

[1] J. Blauert, *Spatial Hearing* (MIT Press, Cambridge, MA, 1983).

[2] G. Steinke, "Surround Sound—The New Phase: An Overview," presented at the 100th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 44, p. 651 (1996 July/Aug.), preprint 4286.

[3] R. Condamines, "Le relief sonore tètraphonique," *Rev. d'Acoustique*, pp. 277–283 (1972).

[4] M. A. Gerzon, "Periphony: With-Height Sound Reproduction," *J. Audio Eng. Soc.*, vol. 21, pp. 2–10 (1973 Jan./Feb.).

[5] H. Møller, "Fundamentals of Binaural Technology," *Appl. Acoust.*, vol. 36, pp. 171–218 (1992).

[6] V. Pulkki, J. Huopaniemi, and T. Huotilainen, "DSP Tool for 8-Channel Audio Mixing," in *Proc. Nordic Acoustical Meeting 96* (1996 June), pp. 307–314.

[7] F. R. Moore, *Elements of Computer Music* (Prentice Hall, Englewood Cliffs, NJ, 1990).

[8] A. D. Blumlein, U.K. patent 394,325, 1931. Reprinted in *Stereophonic Techniques* (Audio Engineering Society, New York, 1986).

[9] B. B. Bauer, "Phasor Analysis of Some Sterephonic Phenomena," *J. Acoust. Soc. Am.*, vol. 33, pp. 1536–1539 (1961 Nov.).

[10] B. Bernfeld, "Attempts for Better Understanding of the Directional Sterephonic Listening Mechanism," presented at the 44th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 21, p. 308 (1973 May).

[11] D. M. Leakey, "Some Measurements on the Effect of Interchannel Intensity and Time Difference in Two Channel sound Systems," *J. Acoust. Soc. Am.*, vol. 31, pp. 977–986 (1959).

[12] D. G. Malham and A. Myatt, "3-D Sound Spatialization Using Ambisonic Techniques," *Computer Music J.*, vol. 19, no. 4, pp. 58–70 (1995).

[13] J. Borenius, "Moving Sound Image in the Theaters," *J. Audio Eng. Soc.*, vol. 25, pp. 200–203 (1977 Apr.).

[14] V. Pulkki, J. Huopaniemi, T. Huotilainen, and M. Karjalainen, "DSP Approach to Multichannel Audio Mixing," in *Proc. Int. Computer Music Conf. ICMC-96* (1996 Aug.), pp. 93–96.

[15] R. Weir, "TIM-40-TMS320C4X Module Specification, 1992," Texas Instruments Inc. (1992).

[16] M. Karjalainen, "Object-Oriented Programming of DSP Processors: A Case Study of QuickC30," in *Proc. Int. Conf. on Acoustics, Speech, and Signal Processing* (IEEE Signal Processing Society, 1992), pp. 601–604.

[17] URL: http://www.leeds.ac.uk/music/Man/c_front.html.

[18] URL: http://www.digidesign.com.

[19] P. Laine, "Sibelius Academy Computer Music Studio," in *Proc. Int. Computer Music Conf. ICMC-93* (1993, Sept.), pp. 302–305.

[20] URL: http://www.lcsaudio.com/lcs.html.

[21] J. Sandvad and D. Hammershøi, "Binaural Synthesis for Auditory Virtual Environments," in *Proc. Nordic Acoustical Meeting 96* (1996 June), pp. 343–350.

[22] J. Huopaniemi and M. Karjalainen, "Review of Digital Filter Design and Implementation Methods for 3-D Sound," presented at the 102nd Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 45, p. 413 (1997 May), preprint 4461.

[23] Teuvo Kohonen, *Self-Organizing Maps*, vol. 30 of Springer Ser. in Information Sciences (Springer, Berlin, Heidelberg, New York, 1995).

APPENDIX

The statement that gain factors calculated using Eq. (9) will satisfy the tangent law [Eq. (3)] will now be proved. At first, some known expressions that will be useful in the proof are written down:

$$\boldsymbol{L}^{-1} = \begin{bmatrix} l_{11} & l_{12} \\ l_{21} & l_{22} \end{bmatrix}^{-1} = \frac{1}{l_{11}l_{22} - l_{21}l_{12}} \begin{bmatrix} l_{22} & -l_{12} \\ -l_{21} & l_{11} \end{bmatrix} \tag{20}$$

$$l_{11} = l_{21} = \cos\varphi_0 \tag{21}$$

$$l_{12} = -l_{22} = \sin\varphi_0 \tag{22}$$

$$p_1 = \cos\varphi \tag{23}$$

$$p_2 = \sin\varphi\,. \tag{24}$$

Eq. (9) can be written in the form

$$\boldsymbol{g} = \frac{1}{l_{11}l_{22} - l_{21}l_{12}} [p_1 l_{22} - p_2 l_{21} \quad p_2 l_{11} - p_1 l_{12}]\,. \tag{25}$$

Then, using Eqs. (21)–(24), it can be seen that

$$g_1 = \frac{\cos\varphi \sin\varphi_0 + \sin\varphi \cos\varphi_0}{2\cos\varphi_0 \sin\varphi_0} \tag{26}$$

$$g_2 = \frac{\cos\varphi \sin\varphi_0 - \sin\varphi \cos\varphi_0}{2\cos\varphi_0 \sin\varphi_0}. \tag{27}$$

The relation $(g_1 - g_2)/(g_1 + g_2)$ may now be calculated using Eqs. (26) and (27),

$$\frac{g_1 - g_2}{g_1 + g_2} = \frac{2\sin\varphi \cos\varphi_0}{2\cos\varphi \sin\varphi_0} = \frac{\tan\varphi}{\tan\varphi_0}. \tag{28}$$

This is the tangent law [Eq. (3)], which completes the proof.

THE AUTHOR

Ville Pulkki was born in Jyväskylä, Finland, in 1969. He received the M.Sc. degree in engineering from the Helsinki University of Technology (HUT), Espoo, Finland, in 1994. Since 1994 he has been majoring in studies at the Musical Education Department at the Sibelius Academy, Helsinki, Finland, and has been a Ph.D. student at HUT.

Mr. Pulkki was with the Laboratory of Computer and Information Sciences and the Neural Networks Research Centre at HUT from 1991 to 1995. He worked as a research assistant and later as a research scientist. Since 1995 he has been working as a research scientist in the Laboratory of Acoustics and Audio Signal Processing at HUT.

His research interests include three-dimensional sound, virtual acoustics, and music. Mr. Pulkki is a member of the AES and the Polar Bear Club of Finland.