



(11) **EP 0 760 197 B1**

(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention of the grant of the patent:
28.01.2009 Bulletin 2009/05

(51) Int Cl.:
H04S 3/00 (2006.01) H04S 1/00 (2006.01)

(21) Application number: **95918832.7**

(86) International application number:
PCT/US1995/004839

(22) Date of filing: **03.05.1995**

(87) International publication number:
WO 1995/031881 (23.11.1995 Gazette 1995/50)

(54) **THREE-DIMENSIONAL VIRTUAL AUDIO DISPLAY EMPLOYING REDUCED COMPLEXITY IMAGING FILTERS**

DREIDIMENSIONALE VIRTUELLE AUDIOANZEIGE UNTER VERWENDUNG VON
ABBILDUNGSFILTERN MIT VERRINGERTER KOMPLEXITÄT

AFFICHAGE AUDIO VIRTUEL TRIDIMENSIONNEL UTILISANT DES FILTRES DE FORMATION
D'IMAGES A COMPLEXITE REDUITE

(84) Designated Contracting States:
**AT BE CH DE DK ES FR GB GR IE IT LI LU MC NL
PT SE**

(72) Inventor: **ABEL, Jonathan, S.**
Palo Alto, CA 94301 (US)

(30) Priority: **11.05.1994 US 241867**
09.09.1994 US 303705

(74) Representative: **Alton, Andrew**
Urquhart-Dykes & Lord LLP
Tower North Central
Merrion Way
Leeds LS2 8PA (GB)

(43) Date of publication of application:
05.03.1997 Bulletin 1997/10

(73) Proprietor: **Aureal Semiconductor Inc.**
Fremont, CA 94538 (US)

(56) References cited:
US-A- 5 105 462 US-A- 5 404 406
US-A- 5 438 623 US-A- 5 440 639

EP 0 760 197 B1

Note: Within nine months of the publication of the mention of the grant of the European patent in the European Patent Bulletin, any person may give notice to the European Patent Office of opposition to that patent, in accordance with the Implementing Regulations. Notice of opposition shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

Description**Technical Field**

5 **[0001]** This invention relates generally to three-dimensional or "virtual" audio. More particularly, this invention relates to a method and apparatus for reducing the complexity of imaging filters employed in virtual audio displays. In accordance with the teachings of the invention, such reduction in complexity may be achieved without substantially affecting the psychoacoustic localization characteristics of the resulting three-dimensional audio presentation.

Background Art

10 **[0002]** Sounds arriving at a listener's ears exhibit propagation effects which depend on the relative positions of the sound source and listener. Listening environment effects may also be present. These effects, including differences in signal intensity and time of arrival, impart to the listener a sense of the sound source location. If included, environmental effects, such as early and late sound reflections, may also impart to the listener a sense of an acoustical environment. By processing a sound so as to simulate the appropriate propagation effects, a listener will perceive the sound to originate from a specified point in three-dimensional space - that is a "virtual" position. See, for example, "Headphone simulation of free-field listening" by Wightman and Kistler, J. Acoust. Soc. Am., Vol. 85, No. 2, 1989.

15 **[0003]** Current three-dimensional or virtual audio displays are implemented by time-domain filtering an audio input signal with selected head-related transfer functions (HRTFs). Each HRTF is designed to reproduce the propagation effects and acoustic cues responsible for psychoacoustic localization at a particular position or region in three-dimensional space or a direction in three-dimensional space. See, for example, "Localization in Virtual Acoustic Displays" by Elizabeth M. Wenzel, Presence, Vol. 1, No. 1, Summer 1992. For simplicity, the present document will refer only to a single HRTF operating on a single audio channel. In practice, pairs of HRTFs are employed in order to provide the proper signals to the ears of the listener.

20 **[0004]** At the present time, most HRTFs are indexed by spatial direction only, the range component being taken into account independently. Some HRTFs define spatial position by including both range and direction and are indexed by position. Although particular examples herein may refer to HRTFs defining direction, the present invention applies to HRTFs representing either direction or position.

25 **[0005]** HRTFs are typically derived by experimental measurements or by modifying experimentally derived HRTFs. In practical virtual audio display arrangements, a table of HRTF parameter sets are stored, each HRTF parameter set being associated with a particular point or region in three-dimensional space. In order to reduce the table storage requirements, HRTF parameters for only a few spatial positions are stored. HRTF parameters for other spatial positions are generated by interpolating among appropriate sets of HRTF positions which are stored in the table.

30 **[0006]** As noted above, the acoustic environment may also be taken into account. In practice, this may be accomplished by modifying the HRTF or by subjecting the audio signal to additional filtering simulating the desired acoustic environment. For simplicity in presentation, the embodiments disclosed refer to the HRTFs, however, the invention applies more generally to all transfer functions for use in virtual audio displays, including HRTFs, transfer functions representing acoustic environmental effects and transfer functions representing both head-related transforms and acoustic environmental effects.

35 **[0007]** A typical prior art arrangement is shown in Figure 1. A three-dimensional spatial location or position signal 10 is applied to an HRTF parameter table and interpolation function 11, resulting in a set of interpolated HRTF parameters 12 responsive to the three-dimensional position identified by signal 10. An input audio signal 14 is applied to an imaging filter 15 whose transfer function is determined by the applied interpolated HRTF parameters. The filter 15 provides a "spatialized" audio output suitable for application to one channel of a headphone 17.

40 **[0008]** Although the various Figures show headphones for reproduction, appropriate HRTFs may create psychoacoustically localized audio with other types of audio transducers, including loudspeakers. The invention is not limited to use with any particular type of audio transducer.

45 **[0009]** When the imaging filter is implemented as a finite-impulse-response (FIR) filter, the HRTF parameters define the FIR filter taps which comprise the impulse response associated with the HRTF. As discussed below, the invention is not limited to use with FIR filters.

50 **[0010]** The main drawback to the prior art approach shown in Figure 1 is the computational cost of relatively long or complex HRTFs. The prior art employs several techniques to reduce the length or complexity of HRTFs. An HRTF, as shown in Figure 2a, comprises a time delay D component and an impulse response $g(t)$ component. Thus, imaging filters may be implemented as a time delay function z^{-D} and an impulse response function $g(t)$, as shown in Figure 2b. By first removing the time delay, thereby time aligning the HRTFs, the computational complexity of the impulse response function of the imaging filter is reduced.

55 **[0011]** Figure 3a shows a prior art arrangement in which pairs of unprocessed or "raw" HRTF parameters 100 are

applied to a time-alignment processor 101, providing at its outputs time-aligned HRTFs 102 and time-delay values 103 for later use (not shown). Processor 101 cross-correlates pairs of raw HRTFs to determine their time difference of arrival; these time differences are the delay values 103. Because the time delay value values 103 and the filter terms are retained for later use, there is no psychoacoustic localization loss — the perceptual impact is preserved. Each time-aligned HRTF 102 is then processed by a minimum-phase converter 104 to remove residual time delay and to further shorten the time-aligned HRTFs.

[0012] Figure 3b shows two left-right pairs (R1/L1 and R2/L2) of exemplary raw HRTFs resulting from raw HRTF parameters 100. Figure 3c shows corresponding time-aligned HRTFs 102. Figure 3d shows the corresponding output minimum-phase HRTFs 105. The impulse response lengths of the time-aligned HRTFs 102 are shortened with respect to the raw HRTFs 100 and the minimum-phase HRTFs 105 are shortened with respect to the time-aligned HRTFs 102. Thus, by extracting the delay so as to time align the HRTFs and by applying minimum phase conversion, the filter complexity (its length, in the case of an FIR filter) is reduced.

[0013] Despite the use of the techniques of Figures 2b and 3a, at an audio sampling rate of 48 kHz, minimum phase responses as long as 256 points for an FIR filter are commonly used, requiring processors executing on the order of 25 mips per audio source rendered.

[0014] When computational resources are limited, two additional approaches are used in the prior art, either singly or in combination, to further reduce the length or complexity of HRTFs. One technique is to reduce the sampling rate by down sampling the HRTF as shown in Figure 4a. Since many localization cues, particularly those important to elevation, involve high-frequency components, reducing the sampling rate may unacceptably degrade the performance of the audio display.

[0015] Another technique, shown in Figure 4b, is to apply a windowing function to the HRTF by multiplying the HRTF by a windowing function in the time domain or by convolving the HRTF with a corresponding weighting function in the frequency domain. This process is most easily understood by considering the multiplication of the HRTF by a window in the time domain — the window width is selected to be narrower than the HRTF, resulting in a shortened HRTF. Such windowing results in a frequency-domain smoothing with a fixed weighting function. This known windowing technique degrades psychoacoustic localization characteristics, particularly with respect to spatial positions or directions having complex or long impulse responses. Thus, there is a need for a way to reduce the complexity or length of HRTFs while maintaining the perceptual impact and psychoacoustic localization characteristics of the original HRTFs.

[0016] US 5,105,462 describes methods and apparatus for creating the illusion of distinct sound sources distributed throughout a three-dimensional space containing the listener. Each channel of a left/right stereo signal is separately processed and then combined for playback. The sound processing involves dividing each monaural or single channel signal into two signals and then adjusting the differential phase and amplitude of the two channel signals on a frequency dependent basis in accordance with an empirically derived transfer function that has a specific phase and amplitude adjustment for each predetermined frequency interval over the audio spectrum. Each transfer function is empirically derived to relate to a different sound source location and by providing a number of different transfer functions and selecting them accordingly the sound source can be made to appear to move.

[0017] In accordance with the present invention, there is provided a three-dimensional virtual audio display method comprising: generating a set of head-related transfer function parameters in response to a spatial location or direction signal, wherein said set of head-related transfer function parameters are selected from, or interpolated among, head-related transfer function parameters derived by smoothing frequency components of a known head-related transfer function over a bandwidth which is a non-constant function of frequency, and noting the head-related transfer function parameters of the transfer function of a resulting compressed transfer function; and filtering an audio signal in response to said set of head-related transfer function parameters.

[0018] The smoothing according to the present invention is best explained by considering its action in the frequency domain: the frequency components of known transfer functions are smoothed over bandwidths which are a non-constant function of frequency. The parameters of the resulting transfer functions, referred to herein as "compressed" transfer functions, are used to filter the audio signal for the virtual audio display. The compressed head-related transfer function parameters may be prederived or may be derived in real time. Preferably, the smoothing bandwidth is a function of the width of the ear's critical bands (i.e., a function of "critical bandwidth"). The function may be such that the smoothing bandwidth is proportional to critical bandwidth. As is well known, the ear's critical bands increase in width with increasing frequency, thus the smoothing bandwidth also increases with frequency.

[0019] The wider the smoothing bandwidth relative to the critical bandwidth, the less complex the resulting HRTF. In the case of an HRTF implemented as an FIR filter, the length of the filter (the number of filter taps) is inversely related to the smoothing bandwidth expressed as a multiple of critical bandwidth.

[0020] By applying the teachings of the present invention which take critical bandwidth into account, for the same reduction in complexity or length, the resulting less complex or shortened HRTFs have less degradation of perceptual impact and psychoacoustic localization than HRTFs made less complex or shortened by prior art windowing techniques such as described above.

[0021] An example HRTF ("raw HRTF") and shortened versions produced by a prior art windowing method ("prior art HRTF") and by the method according to the present invention ("compressed HRTF") are shown in Figures 5a (time domain) and 5b (frequency domain). The raw HRTF is an example of a known HRTF that has not been processed to reduce its complexity or length. In Figure 5a, the HRTF time-domain impulse response amplitudes are plotted along a time axis of 0 to 3 milliseconds. In Figure 5b the frequency-domain transfer function power of each HRTF is plotted along a log frequency axis extending from 1 kHz to 20 kHz. In the time domain, Figure 5a, the prior art HRTF exhibits some shortening, but the compressed HRTF exhibits even more shortening. In the frequency domain, Figure 5b, the effect of uniform smoothing bandwidth on the prior art HRTF is apparent, whereas the compressed HRTF shows the effect of an increasing smoothing bandwidth as frequency increases. Because of the log frequency scale of Figure 5b, the compressed HRTF displays a constant smoothing with respect to the raw HRTF. Despite their differences in time-domain length and frequency-domain frequency response, the raw HRTF, the prior art HRTF, and the compressed HRTF provide comparable psychoacoustic performance.

[0022] When the amount of prior art windowing and compression according to the present invention are chosen so as to provide substantially similar psychoacoustic performance with respect to raw HRTFs, preliminary double-blind listening tests indicate a preference for compressed HRTFs over prior art windowed HRTFs. Somewhat surprisingly, compressed HRTFs were also preferred over raw HRTFs. This is believed to be because the HRTF fine structure eliminated by the smoothing process is uncorrelated from HRTF position to HRTF position and may be perceived as a form of noise.

[0023] The present invention may be implemented in at least two ways. In a first way, an HRTF is smoothed by convolving the HRTF with a frequency dependent weighting function in the frequency domain. This weighting function differs from the frequency domain dual of the prior art time-domain windowing function in that the weighting function varies as a function of frequency instead of being invariant. Alternatively, a time-domain dual of the frequency dependent weighting function may be applied to the HRTF impulse response in the time domain. In a second way, the HRTF's frequency axis is warped or mapped into a non-linear frequency domain and the frequency-warped HRTF is either multiplied by a conventional window function in the time domain (after transformation to the time domain) or convolved with the non-varying frequency response of the conventional window function in the frequency domain. Inverse frequency warping is subsequently applied to the windowed signal.

[0024] The present invention may be implemented using any type of imaging filter, including, but not limited to, analog filters, hybrid analog/digital filters, and digital filters. Such filters maybe implemented in hardware, software or hybrid hardware/software arrangements, including, for example, digital signal processing. When implemented digitally or partially digitally, FIR, IIR (infinite-impulse-response) and hybrid FIR/IIR filters may be employed. The present invention may also be implemented by a principal component filter architecture. Other aspects of the virtual audio display may be implemented using any combination of analog, digital, hybrid analog/digital, hardware, software, and hybrid hardware/software techniques, including, for example, digital signal processing.

[0025] In the case of an FIR filter implementation, the HRTF parameters are the filter taps defining the FIR filter. In the case of an IIR filter, the HRTF parameters are the poles and zeroes or other characteristics defining the IIR filter. In the case of a principal component filter, the HRTF parameters are the position-dependent weights.

[0026] In another aspect of the invention, there is provided three-dimensional virtual audio display apparatus comprising: means for smoothing frequency components of a known head related transfer function over a bandwidth which is a non-constant function of frequency; means for noting the parameters of the transfer function of a resulting compressed transfer function; means for generating a set of head-related transfer function parameters in response to a spatial location or direction signal, said set of head-related transfer function parameters being selected from, or interpolated among, said parameters of the transfer function of the resulting compressed transfer function; and means for filtering an audio signal in response to said set of head-related transfer function parameters.

Figure 1 is a functional block diagram of a prior art virtual audio display arrangement.

Figure 2a is an example of the impulse response of a head-related transfer function (HRTF).

Figure 2b is a functional block diagram illustrating the manner in which an imaging filter may represent the time-delay and impulse response portions of an HRTF.

Figure 3a is a functional block diagram of one prior art technique for reducing the complexity or length of an HRTF.

Figure 3b is a set of example left and right "raw" HRTF pairs.

Figure 3c is the set of HRTF pairs as in Figure 3b which are now time aligned to reduce their length.

Figure 3d is the set of HRTF, pairs as in Figure 3c which are now minimum phase converted to further reduce their length.

Figure 4a is a functional block diagram showing a prior art technique for shortening an HRTF impulse response by reducing the sampling rate.

Figure 4b is a functional block diagram showing a prior art technique for shortening an HRTF impulse response by multiplying it by a window in the time domain.

Figure 5a is a set of three waveforms in the time domain, illustrating an example of a "raw" HRTF, the HRTF shortened by prior art techniques and the HRTF compressed according to the teachings of the present invention.

Figure 5b is a frequency domain representation of the set of HRTF waveforms of Figure 5a.

Figure 6a is a functional block diagram showing an embodiment for deriving compressed HRTFs according to the present invention.

Figure 6b shows the frequency response of an exemplary input HRTF.

Figure 6c shows the impulse response of the exemplary input HRTF impulse response.

Figure 6d shows the frequency response of the compressed output HRTF.

Figure 6e shows the impulse response of the compressed output HRTF.

Figure 7a shows an alternative embodiment for deriving compressed HRTFs according to the present invention.

Figure 7b shows the impulse response of an exemplary input HRTF impulse response.

Figure 7c shows the frequency response of the exemplary input HRTF.

Figure 7d shows the frequency response of the input HRTF after frequency warping.

Figure 7e shows the frequency response of the compressed output HRTF.

Figure 7f shows the frequency response of the compressed output HRTF after inverse frequency warping.

Figure 7g shows the impulse response of the compressed output HRTF after inverse frequency warping.

Figure 8 shows three of a family of windows useful in understanding the operation of the embodiments of Figures 6a and 7a.

Figure 9 is a functional block diagram in which the imaging filter is embodied as a principal component filter.

Figure 10 is a functional block diagram showing another aspect of the present invention.

Figure 6a shows an embodiment for deriving compressed HRTFs according to the present invention. According to this embodiment, an input HRTF is smoothed by convolving the frequency response of the input HRTF with a frequency dependent weighting function in the frequency domain. Alternatively, a time-domain dual of the frequency dependent weighting function may be applied to the HRTF impulse response in the time domain.

Figure 7a shows an alternative embodiment for deriving compressed HRTFs according to the present invention. According to this embodiment, the frequency axis of the input HRTF is warped or mapped into a non-linear frequency domain and the frequency-warped HRTF is convolved with the frequency response of a non-varying weighting function in the frequency domain (a weighting function which is the dual of a conventional time-domain windowing function). Inverse frequency warping is then applied to the smoothed signal. Alternatively, the frequency-warped HRTF may be transformed into the time domain and multiplied by a conventional window function.

[0027] Referring to Figure 6a, an optional nonlinear scaling function 51 is applied to an input HRTF 50. A smoothing function 54 is then applied to the HRTF 52. If nonlinear scaling is applied to the input HRTF, an inverse scaling function 56 is then applied to the smoothed HRTF 54. A compressed HRTF 57 is provided at the output. As explained further below, the nonlinear scaling 51 and inverse scaling 56 can control whether the smoothing mean function is with respect to signal amplitude or power and whether it is an arithmetic averaging, a geometric averaging or another mean function.

[0028] The smoothing processor 54 convolves the HRTF with a frequency-dependent weighting function. The smoothing processor may be implemented as a running weighted arithmetic mean,

$$S(f) = \frac{1}{2b_f + 1} \sum_{n=-b_f}^{b_f} W_f(n) \cdot H(f - n),$$

Equation 1.

where at least the smoothing bandwidth b_f and, optionally, the window shape W_f are a function of frequency. The width of the weighting function increases with frequency; preferably, the weighting function length is a multiple of critical bandwidth: the shorter the required HRTF impulse response length, the greater the multiple.

[0029] HRTFs typically lack low-frequency content (below about 300 Hz) and high-frequency content (above about 16 kHz). In order to provide the shortest possible (and, hence, least complex) HRTFs, it is desirable to extend HRTF frequency response to or even beyond the normal lower and upper extremes of human hearing. However, if this is done, the width of the weighting function in the extended low-frequency and high-frequency audio-band regions should be wider relative to the ear's critical bands than the multiple of critical bandwidth used through the main, unextended portion of the audio band in which HRTFs typically have content.

[0030] Below about 500 Hz, HRTFs are approximately flat spectrally because audio wavelengths are large compared to head size. Thus, a smoothing bandwidth wider than the above-mentioned multiple of critical bandwidth preferably is used. At high frequencies, above about 16 kHz, a smoothing bandwidth wider than the above-mentioned multiple of

critical bandwidth preferably is also used because human hearing is poor at such high frequencies and most localization cues are concentrated below such high frequencies. Thus, the weighting bandwidth at the low-frequency and high-frequency extremes of the audio band preferably may be widened beyond the bandwidths predicted by the equations set forth herein. For example, in one practical embodiment of the invention, a constant smoothing bandwidth of about 250 Hz is used for frequencies below 1 kHz, and a third-octave bandwidth is used above 1 kHz. One-third octave bandwidth approximates critical bandwidth; at 1 kHz the one-third octave bandwidth is about 250 Hz. Thus, below 1 kHz the smoothing bandwidth is wider than the critical bandwidth. In some cases, power noted at low frequencies (say, in the range 300 to 500 Hz) is extrapolated to DC to fill in data not accurately determined using conventional HRTF measurement techniques.

[0031] Although a weighting function having the same multiple of critical bandwidth may be used in processing all of the HRTFs in a group, weighting functions having different critical bandwidth multiples may be applied to respective HRTFs so that not all HRTFs are compressed to the same extent — this may be necessary in order to assure that the resulting compressed HRTFs are generally of the same complexity or length (certain ones of the raw HRTFs will be of greater complexity or length depending on the spatial location which they represent and may therefore require greater or lesser compression). Alternatively, HRTFs representing certain directions or spatial positions may be compressed less than others in order to maintain the perception of better overall spatial localization while still obtaining some overall lessening in computational complexity. The amount of HRTF compression may be varied as a function of the relative psychoacoustic importance of the HRTF. For example, early reflections, which are rendered using separate HRTFs because they arrive from different directions, are not as important to spatialize as accurately as is the direct sound path. Thus, early reflections could be rendered using "over shortened" HRTFs without perceptual impact. Another way to view the smoothing 54 of Figure 6a is that for each frequency f ,

$$S_{\theta}(f) = \sum_{n=0}^N W_{f,\theta}(n) \cdot H_{\theta}(n), \quad \text{Equation 2.}$$

$$\text{where } \sum_{n=0}^{n=N} W_{f,\theta}(n) = 1, \quad \text{Equation 3.}$$

$$W_{f,\theta}(n) \geq 0, \text{ for all } n, \quad \text{Equation 4.}$$

$H_{\theta}(n)$ is the input HRTF 52 at position θ , $S_{\theta}(f)$ is the compressed HRTF 54, n is frequency, and N is one half the Nyquist frequency. Thus, there are a family of weighting functions $W_{f,\theta}(n)$, each defined on an interval 0 to N , which have a width which is a function of their center frequency f and, optionally, also a function of the HRTF position θ . The summation of each weighting function is 1 (Equation 3). Figure 8 shows three members of a family of Gaussian-shaped weighting functions with their amplitude response plotted against frequency. Only three of the family of weighting functions are shown for simplicity. The center window is centered at frequency n_0 and has a bandwidth $b_{f=n}$. The weighting functions need not have a Gaussian shape. Other shaped weighting functions, including rectangular, for simplicity, may be employed. Also, the weighting functions need not be symmetrical about their center frequency.

[0032] Taking into account the nonlinear scaling function 51 and the inverse scaling function 56, Figure 6a may be more generally characterized as

$$S_{\theta}(f) = G^{-1} \left\{ \sum_{n=0}^N W_{f,\theta}(n) \cdot G \{ H_{\theta}(n) \} \right\} \quad \text{Equation 5.}$$

where G is the scaling 51 and G^{-1} is the inverse scaling.

[0033] While the smoothing 54 thus far described provides an arithmetic mean function, depending on the statistics of the input HRTF transfer function, a trimmed mean or median might be favored over the arithmetic mean.

[0034] Because the human ear appears to be sensitive to the total filter power in a critical band, it is preferred to implement the nonlinear scaling 51 of Figure 6a as a magnitude squared operation and the output inverse scaler 56 as

a square root. It may be desirable to apply certain pre-processing or post-processing such as minimum phase conversion. Alternatively, or in addition to the magnitude squared scaling and square root inverse scaling, the arithmetic mean of the smoothing 54 becomes a geometric mean when the nonlinear scaling 51 provides a logarithm function and the inverse scaling 56 an exponentiation function. Such a mean is useful in preserving spectral nulls thought to be important

5 [0035] Figures 6b and 6c show an exemplary input HRTF frequency spectrum and input impulse response, respectively, in the frequency domain and the time domain. Figures 6d and 6e show the compressed output HRTF 57 in the respective domains. The degree to which the HRTF spectrum is smoothed and its impulse response is shortened will depend on the multiple of critical bandwidth chosen for the smoothing 54. The compressed HRTF characteristics will also depend

10 [0036] Refer now to Figure 7a. In this embodiment the frequency axis of the input HRTF is altered by a frequency warping function 121 so that a constant-bandwidth smoothing 125 acting on the warped frequency spectrum implements the equivalent of smoothing 54 of Figure 6a. The smoothed HRTF is processed by an inverse warping 129 to provide the output compressed HRTF. In the same manner as in Figure 6a, nonlinear scaling 51 and inverse scaling 56 optionally

15 [0037] The frequency warping function 121 in conjunction with constant bandwidth smoothing serves the purpose of the frequency-varying smoothing bandwidth of the Figure 6a embodiment. For example, a warping function mapping frequency to Bark may be used to implement critical-band smoothing. Smoothing 125 may be implemented as a time-domain window function multiplication or as a frequency-domain weighting function convolution similar to the embodiment

20 of Figure 6a except that the weighting function width is constant with frequency. As with respect to Figure 6a, it may be desirable to apply certain pre-processing or post-processing such as minimum phase conversion. [0038] The order in which the frequency warping function 121 and the scaling function 51 are applied may be reversed. Although these functions are not linear, they do commute because the frequency warping 121 affects the frequency

25 domain while the scaling 51 affects only the value of the frequency bins. Consequently, the inverse scaling function 56 and the inverse warping function 129 may also be reversed. [0039] As a further alternative, the output HRTF may be taken after block 125, in which case inverse scaling and inverse warping may be provided in the apparatus or functions which receive the compressed HRTF parameters.

30 [0040] Figures 7b and 7c show an exemplary input HRTF input response and frequency spectrum, respectively. Figure 7d shows the frequency spectrum of the HRTF mapped into Bark. Figure 7e shows the spectrum of the HRTF after smoothing 125. After undergoing inverse frequency warping, the resulting compressed HRTF has a spectrum as shown in Figure 7f and an impulse response as shown in Figure 7g. It will be noted that the resulting HRTF characteristics are the same as those of the embodiment of Figure 6a.

35 [0041] The imaging filter may also be embodied as a principal component filter in the manner of Figure 9. A position signal 30 is applied to a weight table and interpolation function 31 which is functionally similar to block 11 of Figure 1. The parameters provided by block 31, the interpolated weights, the directional matrix and the principal component filters are functionally equivalent to HRTF parameters controlling an imaging filter. The imaging filter 15' of this embodiment filters the input signal 33 in a set of parallel fixed filters 34, principal component filters, PC_0 through PC_N , whose outputs are mixed via a position-dependent weighting to form an approximation to the desired imaging filter. The accuracy of the approximations increase with the number of principal component filters used. More computational resources, in the

40 form of additional principal component filters, are needed to achieve a given degree of approximation to a set of raw HRTFs than to versions compressed in accordance with this embodiment of the present invention. [0042] Another aspect of the invention is shown in the embodiment of Figure 10. A three-dimensional spatial location or position signal 70 is applied to an equalized HRTF parameter table and interpolation function 71, resulting in a set of interpolated equalized HRTF parameters 72 responsive to the three-dimensional position identified by signal 70. An

45 input audio signal 73 is applied to an equalizing filter 74 and an imaging filter 75 whose transfer function is determined by the applied interpolated equalized HRTF parameters. Alternatively, the equalizing filter 74 may be located after the imaging filter 75. The filter 75 provides a spatialized audio output suitable for application to one channel of a headphone 77. [0043] The sets of equalized head-related transfer function parameters in the table 71 are prederived by splitting a group of known head-related transfer functions into a fixed head-related transfer function common to all head-related

50 transfer functions in the group and a variable, position-dependent head-related transfer function associated with each of the known head-related transfer functions, the combination of the fixed and each variable head-related transfer function being substantially equal to the respective original known head-related transfer function. The equalizing filter 74 thus represents the fixed head-related transfer function common to all head-related transfer functions in the table. In this manner the HRTFs and imaging filter are reduced in complexity.

55 [0044] The equalization filter characteristics are chosen to minimize the complexity of the imaging filters. This minimizes the size of the equalized HRTF table, reduces the computational resources for HRTF interpolation and image filtering and reduces memory resources for tabulated HRTFs. In the case of FIR imaging filters, it is desired to minimize filter length. [0045] Various optimization criteria may be used to find the desired equalization filter. The equalization filter may

approximate the average HRTF, as this choice makes the position-dependent portion spectrally flat (and short in time) on average. The equalization filter may represent the diffuse field sound component of the group of known transfer functions. When the equalization filter is formed as a weighted average of HRTFs, the weighting should give more importance to longer or more complex HRTFs.

[0046] Different fixed equalization may be provided for left and right channels (either before or after the position variable HRTFs) or a single equalization may be applied to the monaural source signal (either as a single filter before the monaural signal is split into left and right components or as two filters applied to each of the left and right components). As might be expected from human symmetry, the optimal left-ear and right-ear equalization filters are often nearly identical. Thus, the audio source signal may be filtered using a single equalization filter, with its output passed to both position-dependent HRTF filters.

[0047] Further benefits may be achieved by smoothing either the equalized HRTF parameters, the parameters of the fixed equalizing filter or both the equalized HRTF parameters and equalizing filter parameters in accordance with the teachings of the present invention.

[0048] Also, using different filter structures for the equalization filter and the imaging filter may result in computational savings: for example, one may be implemented as an IIR filter and the other as an FIR filter. Because it is a fixed filter typically with a fairly smooth response, the equalizing filter may best be implemented as a low-order IIR filter. Also, it could readily be implemented as an analog filter.

[0049] Any filtering technique appropriate for use in HRTF filters, including principal component methods, may be used to implement the variable, position-dependent portion equalized HRTF parameters. For example, Figure 10 may be modified to employ as imaging filter 75 a principal component imaging filter 15' of the type described in connection with the embodiment of Figure 9.

Claims

1. A three-dimensional virtual audio display method comprising:

generating a set of head-related transfer function parameters in response to a spatial location or direction signal, wherein said set of head-related transfer function parameters are selected from, or interpolated among, head-related transfer function parameters derived by smoothing (54, 125) frequency components (50) of a known head-related transfer function over a bandwidth which is a non-constant function of frequency, and noting the head-related transfer function parameters of a resulting compressed head-related transfer function (57, 130); and filtering an audio signal in response to said set of head-related transfer function parameters.

2. An audio display method according to claim 1 wherein the bandwidth is a function of the width of the ear's critical band.

3. An audio display method according to claim 2 wherein the smoothing (54) comprises, for each frequency component in at least part of an audio band of the display, applying a mean function to the frequency components within the bandwidth containing the frequency component.

4. An audio display method according to claim 3 wherein the mean function is a function of the amplitude of the frequency components.

5. An audio display method according to claim 3 wherein the mean function is a function of the power of the frequency components.

6. An audio display method according to claim 4 or claim 5 wherein said mean function determines the median.

7. An audio display method according to claim 4 or claim 5 wherein said mean function determines the weighted arithmetic mean.

8. An audio display method according to claim 4 or claim 5 wherein said mean function determines the weighted geometric mean.

9. An audio display method according to claim 4 or claim 5 wherein said mean function determines a trimmed mean.

10. An audio display method according to claim 2 wherein the smoothing comprises convolving the head-related transfer function with a frequency dependent weighting function said weighting function having a rectangular shape.

11. An audio display method according to claim 1 wherein the bandwidth is proportional to the width of the ear's critical band.
- 5 12. An audio display method according to claim 11 wherein said head-related transfer function parameters are extended at low and high frequencies and wherein said bandwidth is wider than a bandwidth proportional to the width of the ear's critical band in said low- and high-frequency regions.
- 10 13. An audio display method according to claim 1 wherein the smoothing comprises convolving the head-related transfer function with a frequency dependent weighting function, the width of which is a function of the width of the ear's critical band.
14. An audio display method according to claim 13 wherein the weighting function has a bandwidth which is a multiple of one, or greater, of the width of the ear's critical band.
- 15 15. An audio display method according to claim 14 wherein said head-related transfer function parameters are extended at low and high frequencies and wherein said bandwidth is wider than a bandwidth proportional to the width of the ear's critical band in said low- and high-frequency regions.
- 20 16. An audio display method according to claim 13 wherein said weighting function has a shape having a higher-order continuity than a rectangularly-shaped window.
17. An audio display method according to claim 1 wherein smoothing frequency components comprises smoothing said frequency components in the frequency domain.
- 25 18. An audio display method according to claim 17 wherein said smoothing comprises convolving said known transfer function $H(f)$ with the frequency response of a weighting function $w_f(f)$ in the frequency domain according to the relationship

30

$$S(f) = \frac{1}{2b_f + 1} \sum_{n=-b_f}^{b_f} W_f(n) \cdot H(f-n)$$

35 where at least the smoothing bandwidth b_f and, optionally, the weighting function shape W_f are a function of frequency.

- 40 19. An audio display method according to claim 1 wherein smoothing frequency components comprises applying a frequency warping function (121) to said known head-related transfer function, transforming the frequency-warped transfer function to the time domain, and time-domain windowing the impulse response of the frequency-warped transfer function,
20. An audio display method according to claim 1 wherein smoothing frequency components comprises applying a frequency warping function (121) to said known head-related transfer function and frequency-domain convolving the frequency-warped transfer function with the frequency response of a constant weighting function.
- 45 21. An audio display method according to claim 19 or claim 20 wherein said frequency warping function maps the transfer function to Bark.
22. An audio display method according to claim 19 or claim 20 further comprising applying a non-linear scaling (51) to said known head-related transfer function prior to said multiplication or said convolving and applying an inverse scaling (56) to the windowed or convolved transfer function.
- 50 23. An audio display method according to claim 1 wherein said filtering is principal-component filtering (15').
24. An audio display method according to claim 1 wherein said head-related transfer function parameters are equalized transfer function parameters and said filtering includes fixed equalization filtering and filtering in response to said equalized transfer function parameters.
- 55 25. An audio display method according to claim 1 wherein said set of head-related transfer functions are derived by

smoothing frequency components of known head-related transfer functions over different bandwidths as a function of the spatial location or directions associated with the transfer function.

- 5
26. An audio display method according to claim 1 wherein said set of head-related transfer functions are derived by smoothing frequency components of known head-related transfer functions over different bandwidths as a function of the complexity of the transfer function.
- 10
27. An audio display method according to claim 1 wherein said set of head-related transfer functions are derived by smoothing frequency components of known head-related transfer functions over different bandwidths as a function of the spatial location or direction associated with the transfer function and as a function of the complexity of the transfer function.
- 15
28. An audio display method according to claim 26 or 27 wherein the bandwidth increases with increasing transfer function complexity.
- 20
29. An audio display method according to claim 1 or claim 28 wherein the bandwidth is selected such that the complexity of the most complex resulting compressed head-related transfer function does not exceed a predetermined complexity.
- 25
30. An audio display method according to claim 1 wherein said set of head-related transfer functions are derived by smoothing frequency components of known head-related transfer functions over different bandwidths as a function of the relative psychoacoustic importance of the transfer function.
- 30
31. An audio display method according to claim 1 wherein said set of head-related transfer functions are derived by smoothing frequency components of known head-related transfer functions over different bandwidths as a function of the spatial location or direction associated with the transfer function and as a function of the relative psychoacoustic importance of the transfer function.
- 35
32. A method according to claim 1, wherein a set of equalized head-related transfer function parameters is generated in response to the spatial location or direction signal (70), wherein fixed equalization filtering parameters and said set of equalized head-related transfer function parameters are selected from or interpolated among parameters derived by splitting a group of known head-related transfer functions into a fixed head related transfer function common to all head-related transfer functions in the group and a variable head-related transfer function associated with each of the known head-related transfer functions, the combination of the fixed and each variable head related transfer function being substantially equal to the respective original known head-related transfer function, and wherein frequency components of each of the variable head-related transfer functions are smoothed over the bandwidth which is a non-constant function of frequency, and wherein the parameters of said fixed head-related transfer function for characterizing said fixed equalization filtering are noted, and the parameters of each of the head-related transfer functions of the resulting variable head-related transfer function for use as said equalized transfer function parameters are noted; and wherein the audio signal is filtered with fixed equalization filtering (74) and in response (75) to said set of equalized head-related transfer function parameters.
- 40
33. An audio display method according to claim 32 wherein the derivation of said fixed equalization filtering parameters and said set of equalized transfer function parameters further includes
- 45
- smoothing frequency components of the fixed transfer function over a bandwidth which is a non-constant function of frequency.
- 50
34. An audio display method according to claim 32 wherein said group of known head-related transfer functions is split into a fixed transfer function and a plurality of variable transfer functions by selecting a fixed transfer function resulting in the least complex variable transfer functions.
- 55
35. An audio display method according to claim 32 wherein said group of known head-related transfer functions is split into a fixed head-related transfer function and a plurality of variable head-related transfer functions by selecting a fixed head-related transfer function representing the diffuse field sound component of the group of known head-related transfer functions.
36. An audio display method according to claim 32 wherein said group of known head-related transfer functions are head-related transfer functions representing a particular direction or range of directions in space.

37. An audio display method according to claim 32 wherein sets of equalized head-related transfer function parameters generated in response to a spatial location or direction signal are generated by principal-component filtering.

38. Three-dimensional virtual audio display apparatus comprising:

means for smoothing frequency components (54, 125) of a known head related transfer function over a bandwidth which is a non-constant function of frequency;
 means for noting the parameters of the transfer function of a resulting compressed transfer function;
 means for generating a set of head-related transfer function parameters (71) in response to a spatial location or direction signal, said set of head-related transfer function parameters being selected from, or interpolated among, said parameters of the transfer function of the resulting compressed transfer function; and
 means for filtering (74, 75) an audio signal in response to said set of head-related transfer function parameters.

Patentansprüche

1. Dreidimensionales, virtuelles Audioausgabeverfahren, umfassend:

Erzeugen eines Satzes kopfbezogener Übertragungsfunktionsparameter ansprechend auf ein Raumortsignal oder ein Raumrichtungssignal, wobei der Satz kopfbezogener Übertragungsfunktionsparameter aus kopfbezogenen Übertragungsfunktionsparametern ausgewählt wird oder aus diesen interpoliert wird, die durch Glätten (54, 125) von Frequenzanteilen (50) einer bekannten kopfbezogenen Übertragungsfunktion über eine Bandbreite, die eine nicht konstante Funktion der Frequenz ist, und Festhalten der kopfbezogenen Übertragungsfunktionsparameter einer resultierenden, komprimierten kopfbezogenen Übertragungsfunktion (57, 130) hergeleitet werden; und
 Filtern eines Audiosignals ansprechend auf den Satz kopfbezogener Übertragungsfunktionsparameter.

2. Audioausgabeverfahren nach Anspruch 1, wobei die Bandbreite eine Funktion der Breite des kritischen Bandes des Ohres ist.

3. Audioausgabeverfahren nach Anspruch 2, wobei das Glätten (54) für jeden Frequenzanteil in zumindest einem Teil eines Audiobandes der Ausgabe das Anwenden einer Mittelungsfunktion bei den Frequenzanteilen innerhalb der Bandbreite, die den Frequenzanteil enthält, aufweist.

4. Audioausgabeverfahren nach Anspruch 3, wobei die Mittelungsfunktion eine Funktion der Amplitude der Frequenzanteile ist.

5. Audioausgabeverfahren nach Anspruch 3, wobei die Mittelungsfunktion eine Funktion der Leistung der Frequenzanteile ist.

6. Audioausgabeverfahren nach Anspruch 4 oder Anspruch 5, wobei die Mittelungsfunktion den Medianwert bestimmt.

7. Audioausgabeverfahren nach Anspruch 4 oder Anspruch 5, wobei die Mittelungsfunktion das gewichtete arithmetische Mittel bestimmt.

8. Audioausgabeverfahren nach Anspruch 4 oder Anspruch 5, wobei die Mittelungsfunktion den gewichteten geometrischen Mittelwert bestimmt.

9. Audioausgabeverfahren nach Anspruch 4 oder Anspruch 5, wobei die Mittelungsfunktion einen gestutzten Mittelwert bestimmt.

10. Audioausgabeverfahren nach Anspruch 2, wobei das Glätten eine Faltung der kopfbezogenen Übertragungsfunktion mit einer frequenzabhängigen Wichtungsfunktion aufweist, wobei die Wichtungsfunktion eine rechteckige Form hat.

11. Audioausgabeverfahren nach Anspruch 1, wobei die Bandbreite proportional zur Breite des kritischen Bandes des Ohres ist.

12. Audioausgabeverfahren nach Anspruch 11, wobei die Parameter der kopfbezogenen Übertragungsfunktionen bei

niedrigen und hohen Frequenzen erweitert werden und wobei die Bandbreite breiter als eine Bandbreite ist, die proportional zur Breite des kritischen Bandes des Ohrs in den Bereichen mit niedrigen und hohen Frequenzen ist.

- 5 13. Audioausgabeverfahren nach Anspruch 1, wobei das Glätten eine Faltung der kopfbezogenen Übertragungsfunktion mit einer frequenzabhängigen Wichtungsfunktion aufweist, wobei deren Breite eine Funktion der Breite des kritischen Bandes des Ohrs ist.
- 10 14. Audioausgabeverfahren nach Anspruch 13, wobei die Wichtungsfunktion eine Bandbreite aufweist, die ein Vielfaches von eins oder höher der Breite des kritischen Bandes des Ohrs ist.
- 15 15. Audioausgabeverfahren nach Anspruch 14, wobei die Parameter der kopfbezogenen Übertragungsfunktionen bei niedrigen und hohen Frequenzen erweitert werden und wobei die Bandbreite breiter als eine Bandbreite proportional zur Breite des kritischen Bandes des Ohrs in Bereichen mit niedrigen und hohen Frequenzen ist.
- 20 16. Audioausgabeverfahren nach Anspruch 13, wobei die Wichtungsfunktion eine Form mit einer Stetigkeit höherer Ordnung als das rechteckförmige Fenster aufweist.
17. Audioausgabeverfahren nach Anspruch 1, wobei die Glättungsfrequenzanteile ein Glätten der Frequenzanteile in der Frequenzdomäne aufweisen.
- 25 18. Audioausgabeverfahren nach Anspruch 17, wobei das Glätten eine Faltung der bekannten Übertragungsfunktion $H(f)$ mit der Frequenzantwort einer Wichtungsfunktion $w_f(i)$ in der Frequenzdomäne gemäß der Beziehung

$$S(f) = \frac{1}{2b_f + 1} \sum_{n=-b_f}^{b_f} W_f(n) \cdot H(f - n),$$

- 30 aufweist, wobei zumindest die Glättungsbandbreite (b_f) und optional die Form der Wichtungsfunktion W_f eine Funktion der Frequenz sind.
- 35 19. Audioausgabeverfahren nach Anspruch 1, wobei die Glättungsfrequenzanteile ein Anwenden einer Frequenzwölbungsfunktion (121) bei der bekannten kopfbezogenen Übertragungsfunktion, ein Umwandeln der frequenzgewölbten Übertragungsfunktion in die Zeitdomäne und das Anwenden einer Fenstertechnik in der Zeitdomäne auf die Impulsantwort der frequenzgewölbten Übertragungsfunktion aufweist.
- 40 20. Audioausgabeverfahren nach Anspruch 1, wobei das Glätten der Frequenzanteile das Anwenden einer Frequenzwölbungsfunktion (121) der bekannten kopfbezogenen Übertragungsfunktion und eine Faltung in der Frequenzdomäne der frequenzgewölbten Übertragungsfunktion mit der Frequenzantwort einer konstanten Wichtungsfunktion umfasst.
- 45 21. Audioausgabeverfahren nach Anspruch 19 oder Anspruch 20, wobei die Frequenzwölbungsfunktion die Übertragungsfunktion auf die Bark-Funktion abbildet.
- 50 22. Audioausgabeverfahren nach Anspruch 19 oder Anspruch 20, ferner aufweisend das Anwenden einer nicht linearen Skalierung (51) bei der bekannten kopfbezogenen Übertragungsfunktion vor der Multiplikation oder dem Falten und das Anwenden einer inversen Skalierung (56) bei der gefensterten oder gefalteten Übertragungsfunktion.
- 55 23. Audioausgabeverfahren nach Anspruch 1, wobei das Filtern ein Filtern des Hauptanteils (15') ist.
24. Audioausgabeverfahren nach Anspruch 1, wobei die kopfbezogenen Übertragungsfunktionsparameter entzerrte Übertragungsfunktionsparameter sind und das Filtern ein festgelegtes Entzerrungsfiltern und ein Filtern in Reaktion auf die Parameter der entzerrten Übertragungsfunktion umfasst.
25. Audioausgabeverfahren nach Anspruch 1, wobei der Satz kopfbezogener Übertragungsfunktionen durch Glätten der Frequenzanteile bekannter kopfbezogener Übertragungsfunktionen über unterschiedliche Bandbreiten als eine Funktion der der Übertragungsfunktion zugeordneten Raumorts oder der der Übertragungsfunktion zugeordneten

Raumrichtungen hergeleitet wird.

- 5 **26.** Audioausgabefunktion nach Anspruch 1, wobei der Satz kopfbezogener Übertragungsfunktionen durch Glätten der Frequenzanteile bekannter kopfbezogener Übertragungsfunktionen über unterschiedliche Bandbreiten als eine Funktion der Komplexität der Übertragungsfunktion hergeleitet wird.
- 10 **27.** Audioausgabeverfahren nach Anspruch 1, wobei der Satz kopfbezogener Übertragungsfunktionen durch Glätten der Frequenzanteile bekannter kopfbezogener Übertragungsfunktionen über unterschiedliche Bandbreiten als eine Funktion des der Übertragungsfunktion zugeordneten Raumorts oder der der Übertragungsfunktion zugeordneten Raumrichtung und als eine Funktion der Komplexität der Übertragungsfunktion hergeleitet wird.
- 15 **28.** Audioausgabeverfahren nach Anspruch 26 oder 27, wobei die Bandbreite mit der Zunahme der Komplexität der Übertragungsfunktion zunimmt.
- 20 **29.** Audioausgabeverfahren nach Anspruch 1 oder Anspruch 28, wobei die Bandbreite so ausgewählt ist, dass die Komplexität der resultierenden komplexesten komprimierten kopfbezogenen Übertragungsfunktion nicht eine vorbestimmte Komplexität überschreitet.
- 25 **30.** Audioausgabeverfahren nach Anspruch 1, wobei der Satz kopfbezogener Übertragungsfunktionen durch Glätten der Frequenzanteile bekannter kopfbezogener Übertragungsfunktionen über unterschiedliche Bandbreiten als eine Funktion der relativen psychoakustischen Wichtigkeit der Übertragungsfunktion hergeleitet wird.
- 30 **31.** Audioausgabeverfahren nach Anspruch 1, wobei der Satz kopfbezogener Übertragungsfunktionen durch Glätten der Frequenzanteile bekannter kopfbezogener Übertragungsfunktionen über unterschiedliche Bandbreiten als eine Funktion des der Übertragungsfunktion zugeordneten Raumortes oder der der Übertragungsfunktion zugeordneten Raumrichtung und als eine Funktion der relativen psychoakustischen Wichtigkeit der Übertragungsfunktion hergeleitet werden.
- 35 **32.** Verfahren nach Anspruch 1, wobei ein Satz entzerrter kopfbezogener Übertragungsfunktionsparameter in Reaktion auf das Raumortssignal oder das Raumrichtungssignal (70) erzeugt wird, wobei festgelegte Entzerrungsfilterparameter und der Satz entzerrter kopfbezogener Übertragungsfunktionsparameter aus Parametern ausgewählt werden oder aus diesen interpoliert werden, die durch Aufteilen einer Gruppe bekannter kopfbezogener Übertragungsfunktionen in eine festgelegte kopfbezogene Übertragungsfunktion, die alle kopfbezogene Übertragungsfunktionen in der Gruppe gemeinsam haben, und in eine variable kopfbezogene Übertragungsfunktion hergeleitet werden, die jeder der bekannten kopfbezogenen Übertragungsfunktionen zugeordnet ist, wobei die Kombination aus der festgelegten und jeder variablen kopfbezogenen Übertragungsfunktion im Wesentlichen gleich der entsprechenden ursprünglichen bekannten kopfbezogenen Übertragungsfunktionen ist und wobei die Frequenzanteile jeder der variablen kopfbezogenen Übertragungsfunktionen über die Bandbreite mit einer nicht konstanten Funktion der Frequenz geglättet werden und wobei die Parameter der festgelegten kopfbezogenen Übertragungsfunktion zum Charakterisieren der festgelegten Entzerrungsfilterung festgehalten werden und die Parameter jeder der kopfbezogenen Übertragungsfunktionen der resultierenden variablen kopfbezogenen Übertragungsfunktion zur Verwendung als entzerrte Übertragungsfunktionsparameter festgehalten werden und wobei das Audiosignal mit einer festgelegten Entzerrungsfilterung (74) und in Reaktion (75) auf den Satz entzerrter kopfbezogener Übertragungsfunktionsparameter gefiltert wird.
- 40 **33.** Audioausgabeverfahren nach Anspruch 32, wobei das Herleiten der festgelegten Entzerrungsfilterparameter und des Satzes entzerrter Übertragungsfunktionsparameter ferner Folgendes umfasst:
- 45 das Glätten der Frequenzanteile der festgelegten Übertragungsfunktion über eine Bandbreite, die eine nicht konstante Funktion der Frequenz ist.
- 50 **34.** Audioausgabeverfahren nach Anspruch 32, wobei die Gruppe bekannter kopfbezogener Übertragungsfunktionen in eine festgelegte Übertragungsfunktion und eine Mehrzahl variabler Übertragungsfunktionen durch Auswahl einer festgelegten Übertragungsfunktion, die in den am wenigsten komplexen variablen Übertragungsfunktionen resultiert, aufgeteilt wird.
- 55 **35.** Audioausgabeverfahren nach Anspruch 32, wobei die Gruppe bekannter kopfbezogener Übertragungsfunktionen in eine festgelegte kopfbezogene Übertragungsfunktion und eine Mehrzahl variabler kopfbezogener Übertragungs-

funktionen durch Auswahl einer festgelegten kopfbezogenen Übertragungsfunktion, die den diffusen Feldschallanteil der Gruppe bekannter kopfbezogener Übertragungsfunktionen darstellt, aufgeteilt wird.

5 36. Audioanzeigeverfahren nach Anspruch 32, wobei die Gruppe bekannter kopfbezogener Übertragungsfunktionen kopfbezogene Übertragungsfunktionen sind, die eine spezielle Richtung oder einen Bereich von Richtungen im Raum darstellen.

10 37. Audioausgabeverfahren nach Anspruch 32, wobei die Sätze entzerrter kopfbezogener Übertragungsfunktionsparameter in Reaktion auf ein Raumortsignal oder ein Raumrichtungssignal durch Filtern des Hauptanteils erzeugt werden.

38. Dreidimensionale virtuelle Audioausgabevorrichtung, mit:

15 einer Einrichtung zum Glätten von Frequenzanteilen (54, 125) einer bekannten kopfbezogenen Übertragungsfunktion über eine Bandbreite, die eine nicht konstante Funktion der Frequenz ist;
 einer Einrichtung zum Festhalten der Parameter der Übertragungsfunktion einer resultierenden komprimierten Übertragungsfunktion;
 einer Einrichtung zum Erzeugen eines Satzes kopfbezogener Übertragungsfunktionsparameter (71) in Reaktion auf ein Raumortsignal oder ein Raumrichtungssignal, wobei der Satz kopfbezogener Übertragungsfunktionsparameter aus den Parametern der Übertragungsfunktion der resultierenden komprimierten Übertragungsfunktion ausgewählt wird oder aus diesen interpoliert wird; und
 20 einer Einrichtung zum Filtern (74, 75) eines Audiosignals in Reaktion auf den Satz kopfbezogener Übertragungsfunktionsparameter.

25

Revendications

1. Procédé d'affichage audio virtuel tridimensionnel consistant à :

30 générer un ensemble de paramètres de fonction de transfert asservie aux mouvements de la tête en réponse à un signal de situation ou de direction dans l'espace, dans lequel ledit ensemble de paramètres de fonction de transfert asservie aux mouvements de la tête sont sélectionnés, ou interpolés, parmi les paramètres de fonction de transfert asservie aux mouvements de la tête dérivés en lissant (54, 125) les composants de fréquence (50) d'une fonction de transfert asservie aux mouvements de la tête sur une largeur de bande qui est
 35 une fonction non constante de la fréquence, et en notant les paramètres de fonction de transfert asservie aux mouvements de la tête d'une fonction de transfert asservie aux mouvements de la tête comprimée résultante (57, 130) ; et
 filtrer un signal audio en réponse audit ensemble de paramètres de fonction de transfert asservie aux mouvements de la tête.

40

2. Procédé d'affichage audio selon la revendication 1, dans lequel la largeur de bande est fonction de la largeur de la bande critique de l'ouïe.

45 3. Procédé d'affichage audio selon la revendication 2, dans lequel le lissage (54) consiste, pour chaque composante de fréquence dans au moins une partie de la bande audio de l'affichage, à appliquer une fonction de moyenne aux composantes de fréquence à l'intérieur de la largeur de bande contenant la composante de fréquence.

50 4. Procédé d'affichage audio selon la revendication 3, dans lequel la fonction de moyenne est fonction de l'amplitude des composantes de fréquence.

5. Procédé d'affichage audio selon la revendication 3, dans lequel la fonction de moyenne est fonction de la puissance des composantes de fréquence.

55 6. Procédé d'affichage audio selon la revendication 4 ou la revendication 5, dans lequel ladite fonction de moyenne détermine la médiane.

7. Procédé d'affichage audio selon la revendication 4 ou la revendication 5, dans lequel ladite fonction de moyenne détermine la moyenne arithmétique pondérée.

8. Procédé d'affichage audio selon la revendication 4 ou la revendication 5, dans lequel ladite fonction de moyenne détermine la moyenne géométrique pondérée.
- 5 9. Procédé d'affichage audio selon la revendication 4 ou la revendication 5, dans lequel ladite fonction de moyenne détermine une moyenne compensée.
- 10 10. Procédé d'affichage audio selon la revendication 2, dans lequel le lissage consiste à convolutionner la fonction de transfert asservie aux mouvements de la tête avec une fonction de pondération liée à la fréquence, ladite fonction de pondération ayant une forme rectangulaire.
11. Procédé d'affichage audio selon la revendication 1, dans lequel la largeur de bande est proportionnelle à la largeur de la bande critique de l'ouïe.
- 15 12. Procédé d'affichage audio selon la revendication 11, dans lequel lesdits paramètres de fonction de transfert asservie aux mouvements de la tête sont étendus à basse et haute fréquences et dans lequel ladite largeur de bande est plus large qu'une largeur de bande proportionnelle à la largeur de la bande critique de l'ouïe dans lesdites régions de basse et haute fréquences.
- 20 13. Procédé d'affichage audio selon la revendication 1, dans lequel le lissage consiste à convolutionner la fonction de transfert asservie aux mouvements de la tête avec une fonction de pondération liée à la fréquence, dont la largeur est fonction de la largeur de la bande critique de l'ouïe.
- 25 14. Procédé d'affichage audio selon la revendication 13, dans lequel la fonction de pondération a une largeur de bande qui est un multiple de un, ou plus, de la largeur de la bande critique de l'ouïe.
- 30 15. Procédé d'affichage audio selon la revendication 14, dans lequel lesdits paramètres de fonction de transfert asservie aux mouvements de la tête sont étendus à basse et haute fréquences et dans lequel ladite largeur de bande est plus large qu'une largeur de bande proportionnelle à la largeur de la bande critique de l'ouïe dans lesdites régions de basse et haute fréquences.
- 35 16. Procédé d'affichage audio selon la revendication 13, dans lequel ladite fonction de pondération a une forme ayant une continuité d'ordre plus élevé qu'une fenêtre de forme rectangulaire.
- 40 17. Procédé d'affichage audio selon la revendication 1, dans lequel le lissage des composantes de fréquence consiste à lisser lesdites composantes de fréquence dans le domaine fréquentiel.
18. Procédé d'affichage audio selon la revendication 17, dans lequel ledit lissage consiste à convolutionner ladite fonction de transfert connue $H(f)$ avec la réponse en fréquence d'une fonction de pondération $w_f(i)$ dans le domaine fréquentiel selon la relation

$$S(f) = \frac{1}{2b_f + 1} \sum_{n=b_f}^{b_f} W_f(n) \cdot H(f-n)$$

où au moins la largeur de bande de lissage b_f et, éventuellement, la forme de la fonction de pondération W_f sont fonction de la fréquence.

- 50 19. Procédé d'affichage audio selon la revendication 1, dans lequel le lissage des composantes de fréquence consiste à appliquer une fonction de déformation de fréquence (121) à ladite fonction de transfert asservie aux mouvements de la tête connue, à transformer la fonction de transfert déformée en fréquence en domaine temporel, et à assurer le fenêtrage dans le domaine temporel de la réponse d'impulsion de la fonction de transfert déformée en fréquence.
- 55 20. Procédé d'affichage audio selon la revendication 1, dans lequel le lissage des composantes de fréquence consiste à appliquer une fonction de déformation de fréquence (121) à ladite fonction de transfert asservie aux mouvements de la tête connue et à convolutionner dans le domaine fréquentiel la fonction de transfert déformée en fréquence avec la réponse en fréquence d'une fonction de pondération constante.

21. Procédé d'affichage audio selon la revendication 19 ou la revendication 20, dans lequel ladite fonction de déformation de fréquence mappe la fonction de transfert sur Bark.
- 5 22. Procédé d'affichage audio selon la revendication 19 ou la revendication 20, consistant en outre à appliquer une échelle non linéaire (51) à ladite fonction de transfert asservie aux mouvements de la tête connue avant ladite multiplication ou ladite convolution et à appliquer une échelle inverse (56) à la fonction de transfert en fenêtre et convolutée.
- 10 23. Procédé d'affichage audio selon la revendication 1, dans lequel ledit filtrage est un filtrage de composante principale (15').
- 15 24. Procédé d'affichage audio selon la revendication 1, dans lequel lesdits paramètres de fonction de transfert asservie aux mouvements de la tête sont des paramètres de fonction de transfert égalisés et ledit filtrage comprend le filtrage d'égalisation fixe et le filtrage en réponse auxdits paramètres de fonction de transfert égalisés.
- 20 25. Procédé d'affichage audio selon la revendication 1, dans lequel ledit ensemble de fonctions de transfert asservies aux mouvements de la tête est dérivé en lissant les composantes de fréquence de fonctions de transfert asservies aux mouvements de la tête connues sur différentes largeurs de bande en fonction de la situation spatiale ou de directions associées à la fonction de transfert.
- 25 26. Procédé d'affichage audio selon la revendication 1, dans lequel ledit ensemble de fonctions de transfert asservies aux mouvements de la tête est dérivé en lissant les composantes de fréquence de fonctions de transfert asservies aux mouvements de la tête connues sur différentes largeurs de bande en fonction de la complexité de la fonction de transfert.
- 30 27. Procédé d'affichage audio selon la revendication 1, dans lequel ledit ensemble de fonctions de transfert asservies aux mouvements de la tête est dérivé en lissant les composantes de fréquence de fonctions de transfert asservies aux mouvements de la tête connues sur différentes largeurs de bande en fonction de la situation ou de la direction dans l'espace associée à la fonction de transfert et en fonction de la complexité de la fonction de transfert.
- 35 28. Procédé d'affichage audio selon la revendication 26 ou 27, dans lequel la largeur de bande augmente avec l'augmentation de la complexité de la fonction de transfert.
- 40 29. Procédé d'affichage audio selon la revendication 1 ou la revendication 28, dans lequel la largeur de bande est choisie de telle sorte que la complexité de la fonction de transfert asservie aux mouvements de la tête comprimée résultante la plus complexe ne dépasse pas une complexité prédéterminée.
- 45 30. Procédé d'affichage audio selon la revendication 1, dans lequel ledit ensemble de fonctions de transfert asservies aux mouvements de la tête est dérivé en lissant les composantes de fréquence de fonctions de transfert asservies aux mouvements de la tête connues sur différentes largeurs de bande en fonction de l'importance psychoacoustique relative de la fonction de transfert.
- 50 31. Procédé d'affichage audio selon la revendication 1, dans lequel ledit ensemble de fonctions de transfert asservies aux mouvements de la tête est dérivé en lissant les composantes de fréquence de fonctions de transfert asservies aux mouvements de la tête connues sur différentes largeurs de bande en fonction de la situation ou de la direction dans l'espace associée à la fonction de transfert et en fonction de l'importance psychoacoustique relative de la fonction de transfert.
- 55 32. Procédé selon la revendication 1, dans lequel un ensemble de paramètres de fonction de transfert asservie aux mouvements de la tête est généré en réponse au signal de situation ou de direction dans l'espace (70), dans lequel des paramètres de filtrage d'égalisation fixes et ledit ensemble de paramètres de fonction de transfert asservie aux mouvements de la tête égalisés sont sélectionnés ou interpolés parmi les paramètres dérivés en séparant un groupe de fonctions de transfert asservies aux mouvements de la tête en une fonction de transfert asservie aux mouvements de la tête fixe commune à l'ensemble des fonctions de transfert asservies aux mouvements de la tête dans le groupe et une fonction de transfert asservie aux mouvements de la tête variable associée à chacune des fonctions de transfert asservies aux mouvements de la tête connues, la combinaison de la fonction de transfert asservie aux mouvements de la tête fixe et de chaque fonction de transfert variable étant sensiblement égale à la fonction de transfert asservie aux mouvements de la tête connue originale respective, et dans lequel les composantes de

fréquence de chacune des fonctions de transfert asservies aux mouvements de la tête variables sont lissées sur la largeur de bande qui est une fonction non constante de la fréquence, et dans lequel les paramètres de ladite fonction de transfert asservie aux mouvements de la tête fixe pour caractériser ledit filtrage d'égalisation fixe sont notés, et les paramètres de chacune des fonctions de transfert asservies aux mouvements de la tête de la fonction de transfert asservie aux mouvements de la tête variable résultante pour utilisation en tant que paramètres de fonction de transfert égalisés sont notés ; et dans lequel le signal audio est filtré avec un filtrage d'égalisation fixe (74) et en réponse (75) auxdits paramètres de fonction de transfert asservie aux mouvements de la tête égalisés.

33. Procédé d'affichage audio selon la revendication 32, dans lequel la dérivation desdits paramètres de filtrage d'égalisation fixes et dudit ensemble de paramètres de fonction de transfert égalisés consiste en outre à :

lisser les composantes de fréquence de la fonction de transfert fixe sur une largeur de bande qui est une fonction non constante de la fréquence.

34. Procédé d'affichage audio selon la revendication 32, dans lequel ledit groupe de fonctions de transfert asservies aux mouvements de la tête connues est séparé en une fonction de transfert fixe et une pluralité de fonctions de transfert variables en sélectionnant une fonction de transfert fixe résultant en les fonctions de transfert variables les moins complexes.

35. Procédé d'affichage audio selon la revendication 32, dans lequel le groupe de fonctions de transfert asservies aux mouvements de la tête est séparé en une fonction de transfert asservie aux mouvements de la tête fixe et une pluralité de fonctions de transfert asservies aux mouvements de la tête variables en sélectionnant une fonction de transfert asservie aux mouvements de la tête fixe représentant la composante sonore en champ diffus du groupe de fonctions de transfert asservies aux mouvements de la tête connues.

36. Procédé d'affichage audio selon la revendication 32, dans lequel ledit groupe de fonctions de transfert asservies aux mouvements de la tête connues sont des fonctions de transfert asservies aux mouvements de la tête représentant une direction ou plage de directions particulière dans l'espace.

37. Procédé d'affichage audio selon la revendication 32, dans lequel les ensembles de paramètres de fonction de transfert asservie aux mouvements de la tête générés en réponse à un signal de situation ou de direction dans l'espace sont générés par un filtrage de composante principale.

38. Dispositif d'affichage audio virtuel tridimensionnel comprenant :

des moyens pour lisser les composantes de fréquence (54, 125) d'une fonction de transfert asservie aux mouvements de la tête connue sur une largeur de bande qui est une fonction non constante de la fréquence ; des moyens pour noter les paramètres de la fonction de transfert d'une fonction de transfert comprimée résultante ;

des moyens pour générer un ensemble de paramètres de fonction de transfert asservie aux mouvements de la tête (71) en réponse à un signal de situation ou de direction dans l'espace, ledit ensemble de paramètres de fonction de transfert asservie aux mouvements de la tête étant choisi, ou interpolé, parmi lesdits paramètres de la fonction de transfert de la fonction de transfert comprimée résultante ; et

des moyens pour filtrer (74, 75) un signal audio en réponse audit ensemble de paramètres de fonction de transfert asservie aux mouvements de la tête.

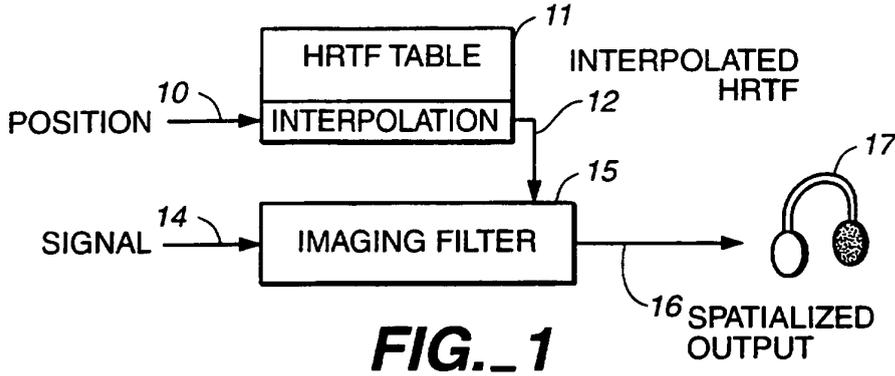


FIG. 1

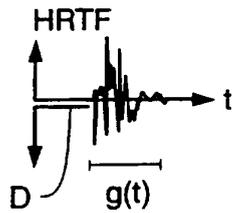


FIG. 2a

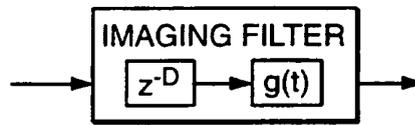


FIG. 2b

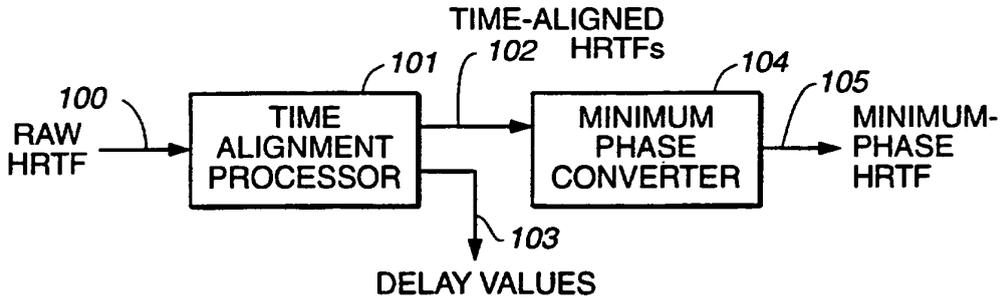


FIG. 3a

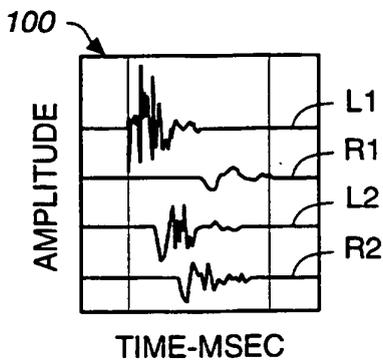


FIG. 3b

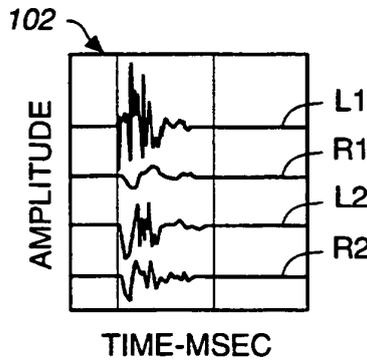


FIG. 3c

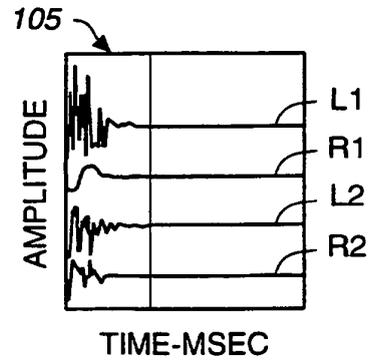


FIG. 3d

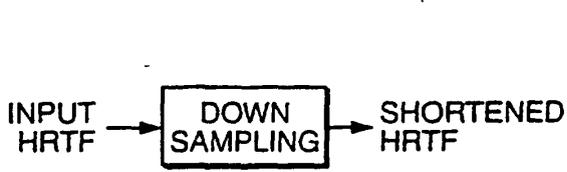


FIG._4a

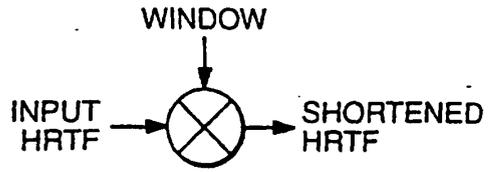


FIG._4b

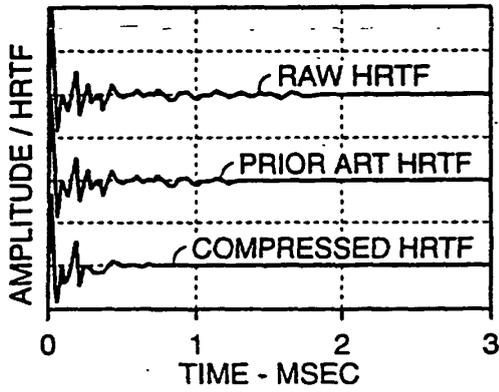


FIG._5a

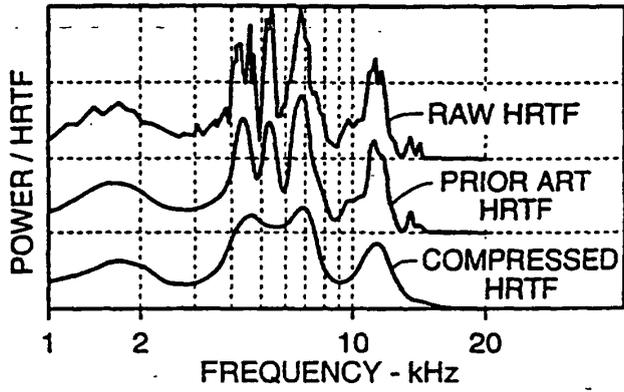


FIG._5b

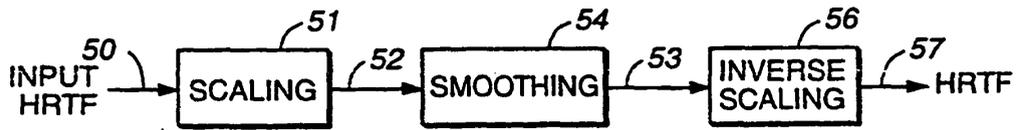


FIG._6a

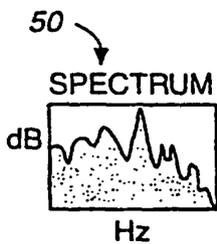


FIG._6b

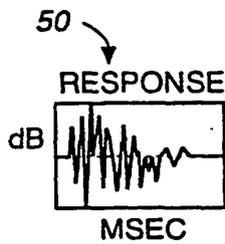


FIG._6c

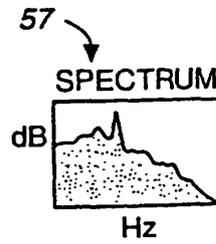


FIG._6d

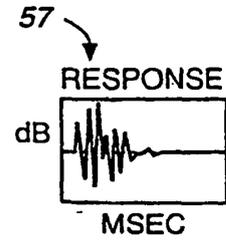


FIG._6e

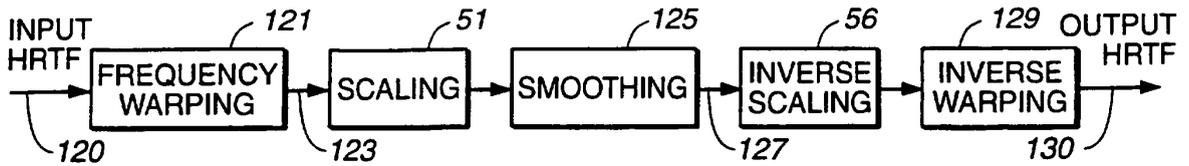


FIG._7a

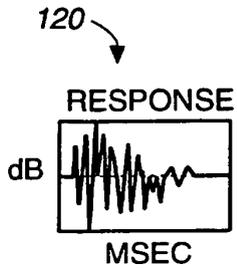


FIG._7b

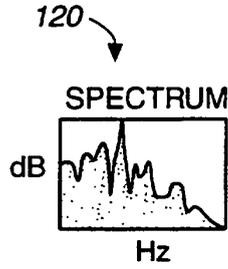


FIG._7c

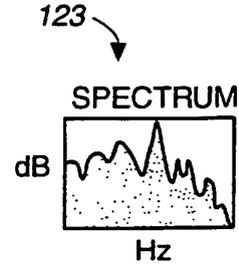


FIG._7d

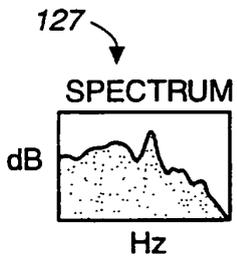


FIG._7e

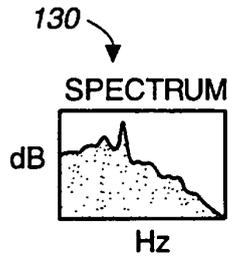


FIG._7f

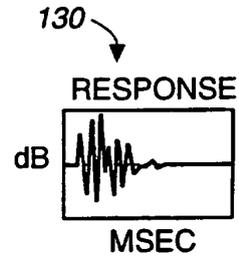


FIG._7g

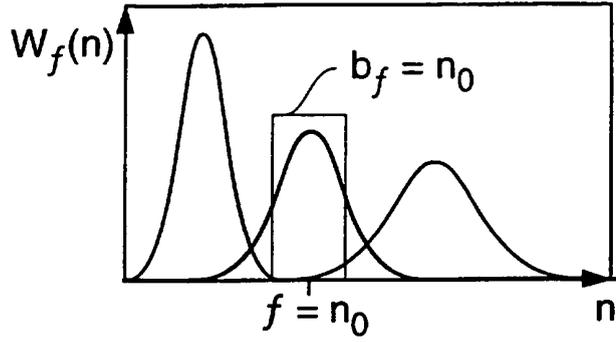


FIG._8

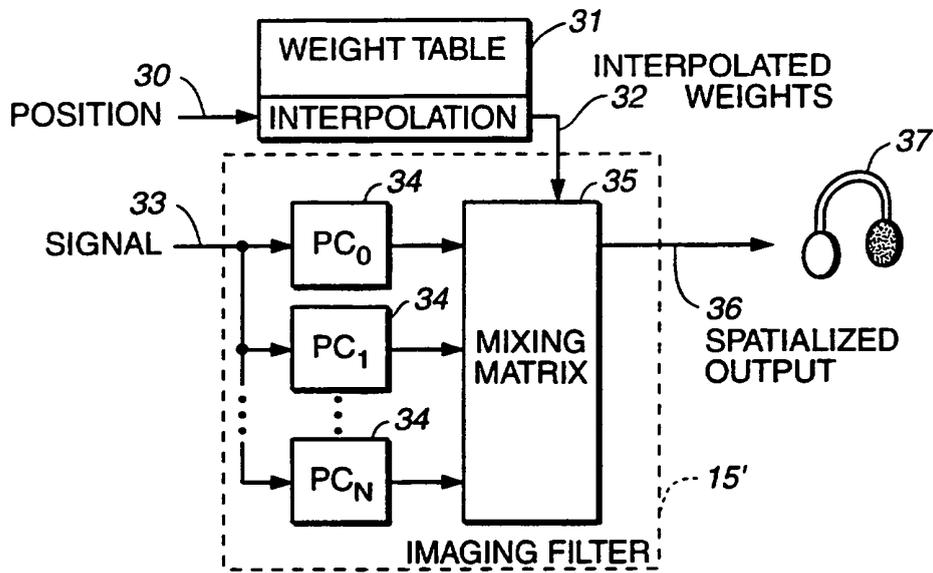


FIG._9

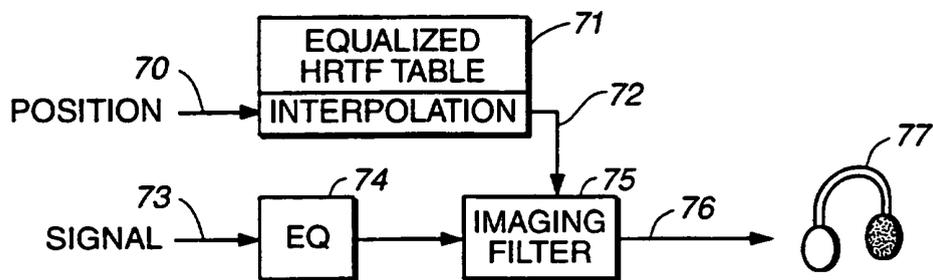


FIG._10

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- US 5105462 A [0016]

Non-patent literature cited in the description

- **WIGHTMAN ; KISTLER.** *J. Acoust. Soc. Am.*, 1989, vol. 85 (2) [0002]
- **ELIZABETH M. WENZEL.** Localization in Virtual Acoustic Displays. *Presence*, 1992, vol. 1 (1) [0003]