(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 999 999 B1

(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention of the grant of the patent:
02.11.2011 Bulletin 2011/44

(21) Application number: **06777145.1**

(22) Date of filing: **01.09.2006**

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

(86) International application number:
PCT/EP2006/008566

(87) International publication number:
WO 2007/110103 (04.10.2007 Gazette 2007/40)

(54) **GENERATION OF SPATIAL DOWNMIXES FROM PARAMETRIC REPRESENTATIONS OF MULTI CHANNEL SIGNALS**

ERZEUGUNG RÄUMLICHER HERUNTERMISCHUNGEN AUS PARAMETRISCHEN DARSTELLUNGEN MEHRKANALIGER SIGNALE

PROCÉDÉ DE PRODUCTION DE MIXAGES RÉDUCTEURS SPATIAUX À PARTIR DE REPRÉSENTATIONS PARAMÉTRIQUES DE SIGNAUX MULTICANAL

(84) Designated Contracting States:
AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IS IT LI LT LU LV MC NL PL PT RO SE SI SK TR

(30) Priority: **24.03.2006 SE 0600674**
10.04.2006 US 744555 P

(43) Date of publication of application:
10.12.2008 Bulletin 2008/50

(73) Proprietors:
- **Dolby Sweden AB**
  113 30 Stockholm (SE)
- **Koninklijke Philips Electronics N.V.**
  5621 BA Eindhoven (NL)

(72) Inventors:
- **VILLEMOES, Lars**
  S-175 56 Järfälla (SE)
- **KJÖRLING, Kristofer**
  S-170 75 Solna (SE)
- **BREEBAART, Jeroen**
  NL-5621 BA Eindhoven (NL)

(74) Representative: **Zinkler, Franz et al**
Schoppe, Zimmermann, Stöckeler & Zinkler
Patentanwälte
Postfach 246
82043 Pullach bei München (DE)

(56) References cited:
WO-A-2006/008683 **US-A1- 2003 035 553**
US-A1- 2006 045 274

- **FALLER C ET AL: "Binaural Cue Coding -Part II: Schemes and Applications" IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING, IEEE SERVICE CENTER, NEW YORK, NY, US, vol. 11, no. 6, 6 October 2003 (2003-10-06), pages 520-531, XP002338415 ISSN: 1063-6676**

EP 1 999 999 B1

Description

Field of the Invention

[0001] The present invention relates to decoding of encoded multi-channel audio signals based on a parametric multichannel representation end in particular to the generation of 2-channel downmixes providing a spatial listening experience as for example a headphone compatible down mix or a spatial downmix for 2 speaker setups.

Background of the Invention in Prior Art

[0002] Recent development in audio coding has made available the ability to recreate a multi-channel representation of an audio signal based on a stereo (or mono) signal and corresponding control data. These methods differ substantially from older matrix based solutions such as Dolby Prologic, since additional control data is transmitted to control the recreation, also referred to as up-mix, of the surround channels based on the transmitted mono or stereo channels.

[0003] Hence, such a parametric multi-channel audio decoder, e.g. MPEG Surround, reconstructs N channels based on M transmitted channels, where N > M, and the additional control data. The additional control data represents a significant lower data rate than transmitting the all N channels, making the coding very efficient while at the same time ensuring compatibility with both M channel devices and N channel devices.

[0004] These parametric surround coding methods usually comprise a parameterization of the surround signal based on IID (Inter channel Intensity Difference) or CLD (Channel Level Difference) and ICC (Inter Channel Coherence). These parameters describe power ratios and correlations, between channel pairs in the up-mix process. Further parameters also used in prior art comprise prediction parameters used to predict intermediate or output channels during the up-mix procedure.

[0005] Other developments in reproduction of multi-channel audio content have provided means to obtain a spatial listening impression using stereo headphones. To achieve a spatial listening experience using only the two speakers of the headphones, multi-channel signals are down mixed to stereo signals using HRTF (head related transfer functions), intended to take- into account the extremely complex transmission characteristics of a human head for providing the spatial listening experience.

[0006] Another related approach is to use a conventional 2-channel playback environment and to filter the channels of a multi-channel audio- signal, with appropriate filters to achieve a listening experience close to that of the playback with the original number of speakers. The processing of the signals is similar as in the case of headphone playback to create an appropriate "spatial stereo down mix" having the desired properties. Contrary to the headphone- case, the signal of both speakers directly reaches both ears of a listener, causing undesired "crosstalk effects". As this has to be taken into account for optimal reproduction quality, the filters used for signal processing are commonly called crosstalk-cancellation filters. Generally, the aim of this technique is to extend the possible range of sound sources outside the stereo speaker base by cancellation of inherent crosstalk using complex crosstalk-cancellation filters.

[0007] Because of the complex filtering, HRTF filters are very long, i.e. they may comprise several hundreds of filter taps each. For the same reason, it is hardly possible to find a parameterization of the filters that works well enough not to degrade the perceptual quality when used instead of the actual filter.

[0008] Thus, on the one- hand, bit saving parametric representations of multi-channel signals do exist that allow for an_ efficient transport of an encoded multi-channel signal. On the other hand, elegant ways to create- a-spatial listening experience for a multi-channel signal when using stereo headphones or stereo speakers only are known. However, these require the full number of channels of the multi-channel signal as input for the application of the head related transfer functions that create the headphone down mix- signal. Thus, either the full set of multi-channels signals has to be transmitted- or a parametric representation has to be fully reconstructed before applying the head related transfer functions or the crosstalk-cancellation filters and thus either the transmission bandwidth or the computational complexity is unacceptably high.

[0009] The US application 2006/0045274 relates to the generation of a sound signal by the application of two head-related transfer functions to one transmitted monophonic sound signal. Each- of the head-related transfer functions is derived adding two other head-related transfer functions.

[0010] The international application W02006/008683 describes a method and a device for processing a stereo signal obtained from an encoder encoding n-channel audio signals into spatial parameters and a stereo downmix.

[0011] Faller C et al: "Binaural Cue Coding - Part II: Schemes and Applications" introduce a coding scheme intended to transmit multiple channels in a bit rate efficient manner.

[0012] The US application 2003/0035553 relates to backwards compatible perceptual coding of spatial cues to convert two or more audio signals into a combined audio signal, which is embedded with two or more sets.of one or more auditory scene parameters, wherein each set of auditory scene parameters (e.g. one or more spatial cues such as ILD, ITD or head-related transfer functions) corresponds to a different frequency band in the combined audio signal.

Summary of the invention

[0013] It is the object of the present invention to provide a concept allowing for a more efficient reconstruction of a 2-channel signal providing a spatial listening experience using parametric representations of multi-channel signals.

[0014] In accordance with a first aspect of the present invention, this object is achieved by a decoder according to claim 1 or 19.

[0015] In accordance with a second aspect of the present. invention, this object is achieved by a binaural decoder according to claim 18.

[0016] In accordance with a third aspect of the present invention, this object is achieved a method of deriving a headphone down mix signal according to claim 20.

[0017] In accordance with a fourth aspect of the present invention, this object is achieved by a receiver or audio player according to claim 21.

[0018] In accordance with a fifth aspect of the present invention, this object is achieved by method of receiving or audio playing according to claim 22. In accordance with a sixth aspect of the present invention this object is achieved by a computer program according to claim 23.

[0019] The present invention is based on the finding that a headphone down mix signal can be derived from a parametric drown mix of a multi-channel signal, when a filter calculator is used for deriving modified HRTFs (head related transfer functions) from original HRTFs of the multi-channel signal and when the filter converter uses a level parameter having information on a level relation between two channels of the multi-channel signal such that modified HRTFs are stronger influenced by the HRTF of a channel having a higher level than by the HRTF of a channel having a lower level. Modified HRTFs are derived during the decoding process taking into account the relative strength of the channels associated to the HRTFs. The original HRTFs are modified such, that a down mix signal of a parametric representation of a multi-channel signal can be directly used to synthesize the headphone down mix signal without the need of a full parametric multi-channel reconstruction of the parametric down mix signal.

[0020] In one embodiment of the present invention, an inventive decoder is used implementing a parametric multi-channel reconstruction as well as an inventive binaural reconstruction of a transmitted parametric down mix of an original multi-channel signal. According to the present, invention, a full reconstruction of the multi-channel signal prior to binaural down mixing is not required, having the obvious great advantage of a strongly reduced computational complexity. This allows, for example, mobile devices having only limited energy reservoirs to extend the playback length significantly. A further advantage is that the same device can serve as provider for complete multi-channel signals (for example 5.1, 7.1, 7.2 signals) as well as for a binaural down mix of the signal having- a spatial listening experience even when using only two-speaker headphones. This might, for example, be extremely advantageous in home-entertainment configurations.

[0021] In a further embodiment of the present invention a filter calculator is used for deriving modified HRTFs not only operative to combine the HRTFs of two channels by applying individual weighting factors to the HRTF but by introducing additional phase factors for each HRTF to be combined. The introduction of the phase factor has the advantage of achieving a delay compensation of two filters prior to their superposition or combination. This leads to a combined response that models a main delay time corresponding to an intermediate position between the front and the back speakers.

[0022] A second advantage is that a gain factor, which has to be applied during the combination of the filters to ensure energy conservation, is- much more stable with respect- to-its behavior with frequency than without the introduction of the phase factor. This is particular relevant for the inventive concept, as according to an embodiment of the present invention a representation of- a down mix of a multi-channel signal is processed within a filterbank domain to derive the headphone down mix signal. As such, different frequency bands of the representation of the down mix signal are to be processed separately and therefore, a smooth behavior of the individually applied gain functions is vital.

[0023] In a further embodiment of the present invention the head-related transfer functions are converted to subband-filters for the subband domains such that the total number of modified HRTFs used in the subband domain is smaller than-the total number of original HRTFs. This has the evident advantage that the computational complexity for deriving headphone down mixed signals is even decreased compared to the down mixing using standard HRTF filters.

[0024] Implementing the inventive concept allows for the use of extremely long HRTFs and thus allows for the reconstruction of headphone down mix signals based on a representation of a parametric down mix of a multi-chanael signal with excellent perceptual quality.

[0025] Furthermore, using the inventive concept on crosstalk-cancellation filters allows for the generation of a spatial stereo down mix to be used with a standard 2 speaker setup based on a representation of a parametric down mix of a multi-channel signal with excellent perceptual quality.

[0026] One further big advantage of the inventive decoding concept is that a single inventive binaural decoder implementing the inventive concept may be used to derive a binaural downmix as well as a multi-channel reconstruction of a transmitted down mix taking into account the additionally transmitted spatial parameters.

[0027] In one embodiment of the present invention an inventive binaural decoder is having an analysis filterbank for deriving the representation of the down mix of the multi-channel signal in a subband domain and an inventive decoder implementing the calculation of the modified HRTFs. The decoder further comprises a synthesis filterbank to finally derive a time domain representation of a headphone down mix signal, which is ready to be played back by any conventional audio playback equipment.

[0028] In the following paragraphs, prior art parametric multi-channel decoding schemes and binaural decoding schemes are explained in more detail referencing the accompanying drawings, to more clearly outline the great advantages of the inventive concept.

[0029] Most of the embodiments of the present invention detailed below describe the inventive concept using HRTFs. As previously noted, HRTF processing is similar to the use of crosstalk-cancellation filters. Therefore, all of the embodiments are to be understood as to refer to HRTF processing as well as to crosstalk-cancellation filters. In other words, all HRTE Filters could be replaced by crosstalk-cancellation filters below to apply the inventive concept to the use of crosstalk-cancellation filters.

Brief Description of the Drawings

[0030] Preferred embodiments of the present invention are subsequently described by referring to the enclosed drawings, wherein:

Fig. 1 shows a conventional binaural synthesis using HRTFs;

Fig. 1b shows a conventional use of crosstalk-cancellation filters;

Fig. 2 shows an example of a multi-channel spatial encoder;

Fig. 3 shows an example for prior art spatial/binaural-decoders;

Fig. 4 shows an example of a parametric multi-channel encoder;

Fig. 5 shows an example of a parametric multi-channel decoder;

Fig. 6 shaws an example of an inventive decoder;

Fig. 7 shows a block diagram illustrating the concept of transforming filters into the subband domain;

Fig. 8 shows an example of an inventive decoder;

Fig. 9 shows a further example of an inventive decoder; and

Fig. 10 shows an example for an inventive receiver or audio player.

Detailed Description of Preferred Embodiments

[0031] The below-described embodiments are merely illustrative for the principles of the present invention for Binaural

[0032] Decoding of Multi-Channel Signals By Morphed HRTF Filtering. It is understood that modifications and variations of the arrangements and the details described herein will be apparent to others skilled in the art. It is the intent, therefore, to be limited only by the scope of the impending patent claims and not by the specific details presented by way of description and explanation of the embodiments herein.

[0033] In order to better outline the features and advantages of the present invention a more elaborate description of prior art will be given now.

[0034] A conventional binaural synthesis algorithm is outlined in Fig. 1. A set of input channels (left front (LF), right front (RF), left surround (LS), right surround (RS) and center (C)), 10a, 10b, 10c, 10d and 10e is filtered by a set of HRTFs 12a to 12j. Each input signal is split into two signals (a left "L" and a right "R" component) wherein each of these signal components is subsequently filtered by an HRTF corresponding to the desired sound position. Finally, all left ear signals are summed by a summer 14a to generate the left binaural output signal L and the right-ear signals are summed by a summer 14b to generate the right binaural output signal R. It may be noted that HRTF convolution can principally be performed in the- time domain, but it is often preferred to perform filtering in the frequency domain due to the increased computational efficiency. That means that, the summation shown in Fig. 1 is also performed in the frequency domain and a subsequent transformation into a time domain is additionally required.

[0035] Fig. 1b illustrates crosstalk cancellation processing intended to achieve a spatial listening impression using only two speakers of a standard stereo playback environment.

[0036] The aim is reproduction of a multi-channel signal by means of a stereo playback system having only two speakers 16a and 16b such that a listener 18 experiences a spatial listening experience. Am major difference with respect to headphone reproduction is that signals of both speakers 16a and 16b directly reach both ears of listener 18. The signals indicated by dashed lines (crosstalk) therefore have to be taken into account additionally.

[0037] For- ease- of explanation- only a 3 channel input, signal having 3 sources 20a to 20c is illustrated in Fig. 1b. It goes without saying that the scenario can in principle be extended to arbitrary number of channels.

[0038] To derive the stereo signal to be played back, each input source is processed by 2 of the crosstalk cancellation filters 21a to 21f, one filter for each channel of the playback signal. Finally, all filtered signals for the left playback channel 16a and the right playback channel 16b are summed up for playback. It is evident that the crosstalk cancellation filters will in general be different for each source 20a and 20b (depending on its desired perceived position) and that they could furthermore even depend on the listener.

[0039] Owing to the high flexibility of the inventive concept, one benefits from high flexibility in the design and application of the crosstalk cancellation filters such that filters can be optimized for each application or playback device individually. One further advantage is that the method is computationally extremely efficient, since only 2 synthesis filterbanks are required.

[0040] A principle sketch of a spatial audio encoder is shown in Fig. 2. In such a basic encoding scenario, a spatial audio decoder 40 comprises a spatial encoder 42, a down mix encoder 44 and a multiplexer 46.

[0041] A multi-channel input signal 50 is analyzed by the spatial encoder 42, extracting spatial parameters describing spatial properties of the multi-channel input signal that have to be transmitted to the decoder side. The down mixed signal generated by the spatial encoder 42 may for example be a monophonic or a stereo signal depending on different encoding scenarios. The down mix encoder 44 may then encode the monophonic or stereo down mix signal using any conventional mono or stereo audio coding scheme. The multiplexer 46 creates an output bit stream hy combining the spatial parameters and the encoded down mix signal into the output bit stream.

[0042] Fig. 3 shows a possible direct combination of a multi-channel decoder corresponding to the encoder of Fig. 2 and a binaural synthesis method as, for example, outlined in Fig. 1. As can be seen, the prior art approach of combining the features is simple and straight forward. The set-up comprises a de-multiplexer 60, a down mix decoder 62, a spatial decoder 64 and a binaural synthesizer 66. An input bit stream 68 is de-multiplexed resulting in spatial parameters 70 and a down mix signal bit: stream The latter down-mix signal bit stream is decoded by the down mix decoder 62 using a conventional mono or stereo decoder. The decoded down mix is input, together with the spatial parameters 70, into the spatial decoder 64 that generates a multi-channel output signal 72 having the spatial properties indicated by the spatial parameters 70. Having a multi-channel signal 72 completely reconstructed, the approach of simply adding a binaural synthesizer 66 to implement the binaural synthesis concept of Fig. 1 is straight-forward. Therefore, the multi-channel, output signal 72 is used as an input for the binaural synthesizer 66 which processes the multi-channel output signal to derive the resulting binaural output signal 74. The approach shown in Fig. 3 has at, least three disadvantages: a complete multi-channel signal representation has to be computed as an intermediate step, followed by HRTF convolution and down mixing in the binaural synthesis. Although HRTF convolution should be performed on a per channel basis, given the fact that each audio channel can have a different spatial position, this is an undesirable situation from a complexity point of view. Thus, computational complexity is high and energy is wasted.

[0043] The spatial decoder operates in a filterbank (QMF) domain. HRTF convolution, on the other hand, is-typically applied in the FFT domain. Therefor, a cascade of a multi-channel QMF synthesis-filterbank, a multi-channel DFT transform, and a stereo inverse DFT transform is necessary, resulting in a system with high computational demands.

[0044] Coding artefacts created by the spatial decoder to create a multi-channel reconstruction will be audible, and possibly enhanced in the (stereo) binaural output.

[0045] An even more detailed description of multi-channel encoding and decoding is given in Figs. 4 and 5.

[0046] The spatial encoder 100 shown in Fig. 4 comprises a first OTT (1-to-2-encoder) 102a, a second OTT 102b and a TTT box (3-to-2-encoder) 104. A multi-channel input signal 106 consisting of LF, LS, C, RF, RS (left-front, left-surround, center, right-front and right-surround) channels is processed by the spatial encoder 100. The OTT boxes receive two input audio channels each, and derive a single monophonic audio output channel and associated spatial parameters the parameters having information on the spatial properties of the original channels with respect to one another or with respect to the output channel (for example CLD, ICC, parameters). In the encoder 100, the LF and the LS channels are processed by OTT encoder 102a and the RF and RS channels are processed by the OTT encoder 102b. Two signals, L and R are generated, the one only having information on the left side and the other only having information on the right side. The signals L, R and C are further processed by the TTT encoder 104, generating a stereo down mix and additional parameters.

[0047] The parameters resulting from the TTT encoder typically consist of a pair of prediction coefficients for each parameter band-, or a pair of level differences to describe the energy ratios of the three input signals. The parameters of the 'OTT' encoders consist of level differences and coherence or cross-correlation values-between the input signals for each frequency band.

[0048] It may be noted that although the schematic sketch of the spatial encoder 100 points to a sequential processing of the individual channels of the down mix signal during the encoding, it is also possible to implement the complete down mixing process of the encoder 100 within one single matrix operation.

[0049] Fig. 5 shows a corresponding spatial decoder, receiving as an input the down mix signals as provided by the encoder of Fig. 4 and the corresponding spatial parameters.

[0050] The spatial decoder 120 comprises a 2-to-3-decoder 122 and 1-to-2-decoders 124a to 124c. The down mix signals L_0 and R_0 are input into the 2-to-3-decoder 122 that recreates a center channel C, a right channel R and a left channel L. These three channels are further processed by the OTT-decoders 124a to 124c yielding six output channels. It may be noted that the derivation of a low-frequency enhancement channel LFE is not mandatory and can be omitted such that one single OTT-encoder may be saved within the surround decoder 120 shown in Fig. 5.

[0051] According to one embodiment of the present invention the inventive concept is applied in a decoder as shown in Fig. 6. The inventive decoder 200 comprises a 2-to-3-decoder 104 and six HRTF-filters 106a to 106f. A stereo input signal (L_0, R_0) is processed by the TTT-decoder 104, deriving three signals L, C and R. It may be noted, that the stereo input signal is assumed to be delivered within a subband domain, since the TTT-encoder may be the same encoder as shown in Fig. 5 and hence adapted to be operative on subband signals. The signals L, R and C are subject to HRTF parameter processing by the HRTF filters 106a to 106f.

[0052] The resulting 6 channels are summed to generate the stereo binaural output pair (L_b, R_b).

[0053] The TTT decoder, 106, can be described as the following matrix operation:

$$\begin{bmatrix} L \\ R \\ C \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \\ m_{31} & m_{32} \end{bmatrix} \begin{bmatrix} L_0 \\ R_0 \end{bmatrix},$$

with matrix entries m_{xy} dependent on the spatial parameters. The relation of spatial parameters and matrix entries is identical to those relations as in the 5.1-multichannel MPEG surround decoder. Each of the three resulting signals *L, R,* and *C* are split in two and processed with HRTF parameters corresponding to the desired (perceived) position of these sound sources. For the center channel (*C*), the spatiaL parameters of the sound source position can be applied directly, resulting in two output signals for the center, $L_B(C)$ and $R_B(C)$:

$$\begin{bmatrix} L_B(C) \\ R_B(C) \end{bmatrix} = \begin{bmatrix} H_L(C) \\ H_R(C) \end{bmatrix} C .$$

[0054] For the left (*L*) channel, the HRTF parameters from the left-front and left-surround channels are combined into a single HRTF parameter set, using the weights w_{lf} and w_{rf}. The resulting 'composite' HRTF parameters simulate the effect of both the front and surround channels in a statistical sense. The following equations are used to generate the binaural output pair (L_B, R_B) for the left channel:

$$\begin{bmatrix} L_B(L) \\ R_B(L) \end{bmatrix} = \begin{bmatrix} H_L(L) \\ H_R(L) \end{bmatrix} L ,$$

[0055] In a similar fashion, the binaural output for the right channel is obtained according to:

$$\begin{bmatrix} L_B(R) \\ R_B(R) \end{bmatrix} = \begin{bmatrix} H_L(R) \\ H_R(R) \end{bmatrix} R,$$

[0056] Given the above definitions of $L_B(C)$, $R_B(C)$, $L_B(L)$, $R_B(L)$, $L_B(R)$ and $R_B(R)$, the complete L_B and R_B signals can be derived from a single 2 by 2 matrix given the stereo input signal:

$$\begin{bmatrix} L_B \\ R_B \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} L_0 \\ R_0 \end{bmatrix},$$

with

$$h_{11} = m_{11} H_L(L) + m_{21} H_L(R) + m_{31} H_L(C),$$
$$h_{12} = m_{12} H_L(L) + m_{22} H_L(R) + m_{32} H_L(C),$$
$$h_{21} = m_{11} H_R(L) + m_{21} H_R(R) + m_{31} H_R(C),$$
$$h_{22} = m_{12} H_R(L) + m_{22} H_R(R) + m_{32} H_R(C).$$

[0057] In the above it was assumed that the $H_Y(X)$ elements, for $Y=L_0,R_0$ and $X=L,R,C,$ were complex scalars. However, the present invention teaches how to extend the approach of a 2 by 2 matrix binaural decoder to handle arbitrary length HRTF filters. In order to achieve thins, the present invention comprises the following steps:

- Transform the HRTF filter responses to a filterbank domain;
- Overall delay difference or phase difference extraction from HRTF filter pairs;
- Morph the responses of the HRTF filter pair as a function of the CLD parameters
- Gain adjustment

[0058] This is achieved by replacing the six complex gains $H_Y(X)$ for, $Y=L_0,R_0$ and $X=L,R,C$ with six filters. These filters are derived from the ten filters $H_Y(X)$ for $Y=L_0,R_0$ and $X=Lf,Ls,Rf,Rs,C,$ which describe the given HRTF filter responses in the QMF domain. These QMF representations can be achieved according to the method described-in one of the subsequent paragraphs.

[0059] In other words, the present invention teaches a concept for deriving modified HRTFs as by modifying (morphing) of the front end surround channel filters using a complex linear combination according to

$$H_Y(X) = g w_f \exp(-j\phi_{XY} w_s^2) H_Y(Xf) + g w_s \exp(j\phi_{XY} w_f^2) H_Y(Xs).$$

[0060] As it can be seen from the above formula, deriving of the modified HRTFs, is a weighted superposition of the original HRTFs, additionally applying phase factors. The-weights w_s, w_f depend on the CLD parameters intended to be used by the OTT decoders 124a and 124b of Fig. 5.

[0061] The weights, w_{lf} and w_{ls} depends on the CLD parameter of the 'OTT' box for Lf and Ls:

$$w_{lf}^2 = \frac{10^{CLD_l/10}}{1+10^{CLD_l/10}},$$

$$w_{ls}^2 = \frac{1}{1+10^{CLD_l/10}}.$$

[0062] The weights w_{rf} and w_{rs} depend on the CLD parameter of the 'OTT' box for Rf and Rs:

$$w_{rf}^2 = \frac{10^{CLD_r/10}}{1+10^{CLD_r/10}},$$

$$w_{rs}^2 = \frac{1}{1+10^{CLD_r/10}}.$$

[0063] The phase parameter ϕ_{XY} can be derived from the main delay time difference τ_{XY} between the front and back HRTF filters and the subband index n of the QMF bank:

$$\phi_{XY} = \frac{\pi(n+\frac{1}{2})}{64}\tau_{XY}.$$

[0064] The role of this phase parameter in the morphing of filters is twofold. First, it realizes a delay compensation of the two filters prior to superposition which leads to a combined response which models a main delay time corresponding to a source position between the front and the back speakers. Second, it makes the necessary gain compensation factor g much more stable and slowly varying over frequency than in the case of simple superposition with $\phi_{XY} = 0$.

[0065] The gain factor g is derermined by the incoherent addition power rule,

$$P_Y(X)^2 = w_f^2 P_Y(Xf)^2 + w_s^2 P_Y(Xs)^2,$$

where

$$P_Y(X)^2 = g^2\left(w_f^2 P_Y(Xf)^2 + w_s^2 P_Y(Xs)^2 + 2w_f w_s P_Y(Xf) P_Y(Xs)\rho_{XY}\right)$$

and ρ_{XY} is the real value of the normalized complex cross correlation between the filters

$\exp(-j\phi_{XY})H_Y(Xf)$ and $H_Y(Xs)$.

[0066] For the above equations, P denotes a parameter describing an average level per frequency band for the impulse response of the filter specified by the indexes. This mean intensity is of course easily derived, once the filter response function are known.

[0067] In the case of simple superposition with ϕ_{XY} =0, the value of ρ_{XY} varies in an erratic and oscillatory manner as a function of frequency, which leads to the need for extensive gain adjustment. In practical implementation it is necessary to limit the value of the gaingand a remaining spectral colorization of the signal cannot be avoided.

[0068] In contrast, the use of morphing with a delay based phase compensation as taught by the present invention leads to a smooth behaviour of ρ_{XY} as a function of frequency. This value is often even close to one for natural HRTF derived filter pairs since they differ mainly in delay and amplitude, and the purpose of the phase parameter is to take the delay difference into account in the QMF filterbank domain.

[0069] An alternative beneficial choice of phase parameter ϕ_{XY} taught by the present invention is given by the phase angle of the normalized complex cross correlation between the filters

$H_Y(Xf)$ and $H_Y(X_s)$,

and unwrapping the phase values with standard unwrapping techniques as a function of the subband index n of the QMF bank. This choice has the consequence that ρ_{XY} is never negative and hence the compensation gain g satisfies $1/\sqrt{2} \leq g \leq 1$ for all subbands. Moreover this choice of phase parameter enables the morphing of the front and surround channel filters in situations where a main delay time difference τ_{XY} is not available.

[0070] For the embodiment of the present invention as described above, it is taught to accurately transform the HRTFs into an efficient representation of the HRTF filters within the QMF domain.

[0071] Fig. 7 gives a principle sketch of the concept to accurately transform time-domain filters into filters within the subband domain having the same net effect on a reconstructed signal. Fig. 7 shows a complex analysis bank 300, a synthesis bank 302 corresponding to the analysis bank 300, a filter converter 304 and a subband filter 306.

[0072] An input signal 310 is provided for which a filter 312 is known having desired properties. The aim of the implementation of the filter converter 304 is- that the output signal 314 has the same characteristics after analysis by the analysis filterbank 300, subsequent subband filtering 306 and synthesis 302 as if it would have when filtered by filter 312 in the time domain. The task of providing a number of subband filters corresponding to the number of subbands used is fulfilled by filter converter 304.

[0073] The following description outlines a method for implementing a given FIR filter $h(v)$ in the complex QMF subband domain. The principle of operation is shown in Figure 7.

[0074] Here, the subband filtering is simply the application of one complex valued FIR filter for each subband, n=0,1,..., L-1 to transform the original indices c_n into their filtered counterparts d_n according to the following formula:

$$d_n(k) = \sum_l g_n(l) c_n(k-l).$$

[0075] Observe that this is different from well known methods developed for critically sampled filterbanks, since those methods require multiband filtering with longer responses. The key component is the filter converter, which converts any time domain. FIR filter into the complex subband domain filters. Since the complex QMF subband domain is over-sampled, there is no canonical set of subband filters for a given time domain filter. Different subband filters can have the same net effect of the time domain signal. What will be described here is a particularly attractive approximate solution, which is obtained by restricting the filter converter to be a complex analysis bank similar to the QMF.

[0076] Assuming that the filter converter prototype is of length $64K_Q$, a real $64K_H$ tap FIR filter is transformed into a set of 64 complex K_H+K_Q-1 tap subband- filters. For K_Q= 3, a FIR filter of 1024 taps is converted into 18 tap subband filtering with an approximation quality of 50 dB.

[0077] The subband filter taps are computed from the formula

$$g_n(k) = \sum_{v=-\infty}^{\infty} h(v+kL) q(v) \exp\left(-i\frac{\pi}{L}\left(n+\tfrac{1}{2}\right)v\right),$$

where $q(v)$ is a FIR prototype filter derived from the QMF prototype filter. As it can be seen, this is just a complex filterbank analysis of the given filter $h(v)$.

[0078] In the following, the inventive concept will be outlined for a further embodiment of the present invention, where a multi-channel parametric representation for a multi-channel signal having five channels is available. Please note that in this particular embodiment of the present invention, original 10 HRTF filters $V_{Y,X}$ (as for example given by a QMF representation of the filters 12a to 12j of Fig 1) are morphed into six filters $h_{v,x}$ for Y = L, R and X = L, R, C.

[0079] The ten filters $v_{Y,X}$ for $Y=L,R$ and- $X=FL,BL,FR,BR,C$ describe the given HRTF filter responses in a hybrid QMF domain.

[0080] The combination of the front and surround channel filters is performed with a complex linear combination according to

$$\mathbf{h}_{L,C} = \mathbf{v}_{L,C}$$

$$\mathbf{h}_{R,C} = \mathbf{v}_{R,C}$$

$$\mathbf{h}_{L,L} = g_{L,L}\sigma_{FL}\exp\left(-j\phi^{L}_{FL,BL}\sigma^{2}_{BL}\right)\mathbf{v}_{L,FL} + g_{L,L}\sigma_{BL}\exp\left(j\phi^{L}_{FL,BL}\sigma^{2}_{FL}\right)\mathbf{v}_{L,BL}$$

$$\mathbf{h}_{L,R} = g_{L,R}\sigma_{FR}\exp\left(-j\phi^{L}_{FR,BR}\sigma^{2}_{BR}\right)\mathbf{v}_{L,FR} + g_{L,R}\sigma_{BR}\exp\left(j\phi^{L}_{FR,BR}\sigma^{2}_{FR}\right)\mathbf{v}_{L,BR}$$

$$\mathbf{h}_{R,L} = g_{R,L}\sigma_{FL}\exp\left(-j\phi^{R}_{FL,BL}\sigma^{2}_{BL}\right)\mathbf{v}_{R,FL} + g_{R,L}\sigma_{BL}\exp\left(j\phi^{R}_{FL,BL}\sigma^{2}_{FL}\right)\mathbf{v}_{R,BL}$$

$$\mathbf{h}_{R,R} = g_{R,R}\sigma_{FR}\exp\left(-j\phi^{R}_{FR,BR}\sigma^{2}_{BR}\right)\mathbf{v}_{R,FR} + g_{R,R}\sigma_{BR}\exp\left(j\phi^{R}_{FR,BR}\sigma^{2}_{FR}\right)\mathbf{v}_{R,BR}$$

[0081] The gain factors. $g_{L,L}, g_{L,R}, g_{R,L}, g_{R,R}$ are determined by

$$g_{Y,X} = \left(\frac{\sigma^{2}_{FX}CFB^{2}_{Y,X} + \sigma^{2}_{BX}}{\sigma^{2}_{FX}CFB^{2}_{Y,X} + \sigma^{2}_{BX} + 2\sigma_{FX}\sigma_{BX}CFB_{Y,X}ICCFB^{\phi}_{Y,X}}\right)^{1/2}$$

[0082] The parameters $CFB_{Y,X}, ICCFB^{\phi}_{Y,X}$ and the phase parameters ϕ are defined as follows:

[0083] An average front/back level quotient per hybrid band for the HRTF filters is defined for $Y = L,R$ and $X = L,R$ by

$$\left(CFB_{Y,X}\right)_{k} = \left(\frac{\sum_{l=0}^{L_{q}-1}\left|\left(v_{Y,FX}\right)_{k}(l)\right|^{2}}{\sum_{l=0}^{L_{q}-1}\left|\left(v_{Y,BX}\right)_{k}(l)\right|^{2}}\right)^{1/2}.$$

[0084] Furthermore, phase parameters $\phi^{L}_{FL,BL}, \phi^{L}_{FR,BR}, \phi^{R}_{FL,BL}, \phi^{R}_{FR,BR}$ are then defined for Y=L,R and X=L,R by

$$\left(CIC_{Y,X}\right)_{k} = \left|\left(CIC_{Y,X}\right)_{k}\right|\exp\left(j\left(\phi^{Y}_{FX,BX}\right)_{k}\right).$$

where the complex cross correlations $(CIC_{Y,X})_k$ are defined by

$$\left(CIC_{Y,X}\right)_{k} = \frac{\sum_{l=0}^{L_{q}-1}\left(v_{Y,FX}\right)_{k}(l)\left(v_{Y,BX}\right)^{*}_{k}(l)}{\left(\sum_{l=0}^{L_{q}-1}\left|\left(v_{Y,FX}\right)_{k}(l)\right|^{2}\right)^{1/2}\left(\sum_{l=0}^{L_{q}-1}\left|\left(v_{Y,BX}\right)_{k}(l)\right|^{2}\right)^{1/2}}.$$

[0085] A phase unwrapping is applied to the phase parameters along, the subband index k, such that the absolute value of the phase increment from subband k to subband k+1 is smaller or equal to π for k = 0,1,... . In cases where there are two choices, $\pm\pi$,for the increment, the sign of the increment for a phase measurement in the interval $]-\pi,\pi]$ is chosen. Finally, normalized phase compensated cross correlations are defined for Y=L,R and X=L,R by

$$\left(ICCFB^{\phi}_{Y,X}\right)_k = \left|\left(CIC_{Y,X}\right)_k\right|.$$

[0086] Please note that in the case where the multi-channel processing is performed within a hybrid subband domain, i.e. in a. domain where subbands are further decomposed into different frequency bands, a mapping of the HRTF responses to the hybrid band filters may for example be performed as follows:

[0087] As in the case without an hybrid filterbank, the ten given HRTF impulse responses from source $X = FL,BL,FR,BR,C$ to target $Y = L,R$ are all converted into QMF subband filters according to the method outlined above. The result is the ten subband filters $\hat{v}_{Y,X}$ with components

$$\left(\hat{v}_{Y,X}\right)_m (l)$$

for QMF subband m = 0,1,...,63 and QMF time slot $l = 0,1,...,L_q - 1$. Let the index mapping from the hybrid band k to QMF band m be denoted by $m = Q(k)$.

[0088] Then the HRTF filters $v_{Y,X}$ in the hybrid band domain are defined by

$$\left(v_{Y,X}\right)_k (l) = \left(\hat{v}_{Y,X}\right)_{Q(k)} (l).$$

[0089] For the specific embodiment described in the previous paragraphs, the filter conversion of HRTF filters into the QMF domain can be implemented as follows, given a FIR filter $h(v)$ of length N_k to be transferred to the complex QMF subband domain:

[0090] The subband filtering consists of the separate application of one complex valued FIR filter $h_m(l)$ for each QMF subband, m = 0,1,...,63. The key component is the filter converter, which converts the given time domain FIR filter $h(v)$ into the complex subband domain filters $h_m(l)$. The filter converter is a complex analysis bank similar to the QMF analysis bank. Its prototype filter $q(v)$ is of length 192. An extension with zeros of the time domain FIR filter is defined by

$$\tilde{h}(\nu) = \begin{cases} h(\nu), & \nu = 0,1,\ldots,N_h - 1; \\ 0, & \text{otherwise}, \end{cases}$$

[0091] The subband domain- filters of length, $L_q = K_h + 2$ where $K_h = [N_h/64]$ is then given for m = 0,1,...,63 and $l = 0,1,...,K_h + 1$ by

$$h_m(l) = \sum_{\nu=0}^{191} \tilde{h}\left(\nu + 64(l-2)\right) q(\nu) \exp\left(-j\frac{\pi}{64}\left(m + \tfrac{1}{2}\right)(\nu - 95)\right) .$$

[0092] Although the inventive concept has been detailed with respect to a down mix signal having two channels, i.e. a transmitted stereo signal, the application of the inventive concept is by no means restricted to a scenario having a stereo-down mix signal.

[0093] Summarizing, the present invention relates to the problem of using long HRTF or crosstalk cancellation filters for binaural rendering of parametric multi-channel signals. The invention teaches new ways to extend the parametric HRTF approach to arbitrary length of HRTF filters.

[0094] The present invention comprises the following features:

- Multiplying the stereo down mix signal, by a 2 by 2 matrix where every matrix element is a FIR filter or arbitrary length (as given by the HRTF filter);

- Deriving the filters in the 2 by 2 matrix by morphing the original HRTF filters based on the transmitted multi-channel parameters;

- Calculation of the morphing of the HRTF filters so that the correct spectral envelope and overall energy is obtained.

[0095] Fig. 8 shows an example for an inventive decoder 300 for deriving a headphone down- mix signal. The decoder comprises-a filter calculator 302 and a synthesizer 304. The filter calculator receives as a first input level parameters 306 and as a second input HRTFs (head-related transfer functions) 308 to derive codified. HRTFs 310 that have the same net effect on a signal when applied to the signal in the subband domain than the head-related transfer functions 308 applied in the time domain. The modified HRTFs 310 serve as first input to the synthesizer 304 that receives as a second input a representation of a down-mix signal 312 within a subband domain. The representation of the down-mix signal 312 is derived by a parametric multi-channel encoder. and intended to be used as a basis for reconstruction of a full multi-channel signal by a multi-channel decoder. The synthesizer 404 is thus able to derive a headphone down-mix signal 314 using- the modified HRTEs 310 and the representation of the down-mix signal 312.

[0096] It may be noted, that the HRTFs could be provided in any possible parametric representation, for example as the transfers function associated to the filter, as the impulse response of the filter or as a series of tap coefficients for an FIR-filter.

[0097] The previous examples assume, that the representation of the down-mix signal is already supplied as a filterbank representation, i.e. as samples derived by a filterbank. In practical applications, however, a time-domain down-mix signal is typically supplied and transmitted to allow also for a direct playback of the submitted signal in simple playback environments. Therefore, in Fig. 9 in a further embodiment of the present invention, where a binaural. compatible decoder 400 comprises an analysis filterbank 402 and a synthesis filterbank 404 and an inventive decoder, which could, for example, be the decoder 300 of Fig. 8. Decoder functionalities and their descriptions are applicable in Fig. 9 as well as in Fig. 8 and the description of the decoder 300 will be omitted within the following paragraph.

[0098] The analysis filterbank 402 receives a downmix of a multi-channel signal 406 as created by a multi-channel parametric encoder. The analysis filterbank 402 derives the filterbank representation of the received down mix signal 406 which is then input into decoder 300 that derives a headphone downmix signal 408, still within the filterbank domain. That is, the down mix is represented by a multitude of samples or coefficients within the frequency bands introduced by the analysis filterbank 402. Therefore, to provide a final headphone down mix signal 410 in the time domain the headphone downmix signal 408 is input into synthesis filterbank 404 that derives the headphone down mix signal 410, which is ready to be played back by stereo reproduction equipment.

[0099] Fig. 10 shows an inventive receiver or audio player 500, having an inventive audio decode 501, a bit stream input 502, and an audio output 504.

[0100] A bit stream- can be input at the input 502 of the inventive receiver/audio player 500. The bit- stream then is decoded by the decoder 501 and the decoded signal is output or played at the output 504 of the inventive receiver/audio player 500.

[0101] Although examples have been derived in the preceding paragraphs to implement the inventive concept relying on a transmitted stereo down mix, the inventive concept may also be applied in configurations based on a single monophonic down mix channel or on more than two down mix channels.

[0102] One particular implementation of the transfer of head-related transfer functions into the subband domain is given in the description of the present invention. However, other techniques of deriving the subband filters may also be used without limiting the inventive concept.

[0103] The phase factors introduced in the derivation of the modified HRTFs can be derived also by other computations than the ones previously presented.

[0104] Even as the inventive concept is shown particularly for HRTF and crosstalk cancellation filters, it can be used for other filters defined for one or more individual channels of a multi channel signal to allow for a computationally efficient generation of a high quality stereo playback signal. The filters are furthermore not only restricted to filters intended to model a listening environment. Even filters adding "artificial" components" to a signal can be used, such as for example reverberation or other distortion filters.

[0105] Depending on certain implementation requirements of the inventive methods, the inventive methods can be implemented in hardware or in software. The implementation can be-performed using a digital storage medium, in particular a disk, DVD or a CD having electronically readable, control signals stored thereon, which cooperate with a programmable computer system such that the inventive methods are performed. Generally, the present invention is, therefore, a computer program product with a program code stored on a machine readable carrier, the program code being operative for performing the inventive methods when the computer program product runs on a computer. In other words, the inventive methods are, therefore, a computer program having a program code for performing at least one of the inventive methods when the computer program runs on a computer.

[0106] While the foregoing has been particularly shown and described with reference to particular embodiments thereof, it will be understood by those skilled in the art that various other changes in the form and details may be made.

Claims

1. Decoder for deriving a headphone down mix signal (314) using a representation of a down mix of a multi-channel signal (312) and using a level parameter (306) having information on a level relation between two channels of the multi-channel signal and using head-related transfer functions (308) related to the two channels of the multi-channel signal, wherein a first channel of the two channels is a front channel of the left or the right side of the multi-channel signal and a second channel of the two channels is a back channel of the same side, comprising:

   a filter calculator (302) for deriving a modified head-related transfer function $H_Y(X)$ (310) by weighting the front channel head-related transfer function $H_Y(Xf)$ and the back channel head-related transfer function $H_Y(Xs)$ using the level parameter (306) such that the modified head-related transfer function $H_Y(X)$ (310) is stronger influenced by the head-related transfer function (308) of a channel having a higher level than by the head-related transfer function (308) of a channel having a lower level by using the following complex linear combination:

   $$H_Y(X) = gw_f \exp(-j\phi_{XY} w_s^2) H_Y(Xf) + gw_s \exp(j\phi_{XY} w_f^2) H_Y(Xs),$$ wherein

   Φ_{XY} is a phase parameter, w_s and w_f are weighting factors derived using the level parameter (306) and g is a common gain factor derived using the level parameter (306); and

   a synthesizer (304) for deriving the headphone down mix signal (314) using the modified head-related transfer function (310) and the representation of the down mix signal (312).

2. Decoder in accordance with claim 1 in which the filter calculator (302) is operative such that the number of modified head-related transfer functions (310) derived is smaller than the number of associated head-related transfer functions (308) of the two channels.

3. Decoder in accordance with claim 1 in which the filter calculator (302) is operative to derive a modified head-related transfer function (310) adapted to be applied to a filterbank representation of the down mix signal.

4. Decoder in accordance with claim 1, adapted to use a representation of the down mix signal derived in a filterbank domain.

5. Decoder in accordance with claim 1, in which the filter calculator (302) is operative to derive the modified head-related transfer function (310) using head-related transfer functions (308) **characterized by** more than three parameters.

6. Decoder in accordance with claim 1, in which the filter calculator (302) is operative to derive the weighting factors for the head-related transfer functions (308) of the two channels using the same level parameter (306).

7. Decoder in accordance with claim 6, in which the filter calculator (302) is operative to derive a first weighting factor w_f for a first channel f and a second weighting factor w_s for a second channel s using the level parameter CLD_1 according to the following formulas:

   $$w_f^2 = \frac{10^{CLD_l/10}}{1+10^{CLD_l/10}},$$

$$w_s^2 = \frac{1}{1+10^{CLD_l/10}} \text{ .}$$

8. Decoder in accordance with claim 1, in which the filter calculator (302) is operative to derive the modified head-related transfer function (310) applying a common gain factor g to the head-related transfer function (308) of the two channels such that energy is preserved when deriving the modified head-related transfer functions (310).

9. Decoder in accordance with claim 8, in which the common gain factor is within the interval $[1/\sqrt{2}, 1]$.

10. Decoder in accordance with claim 1, in which the filter calculator (302) is operative to derive the phase parameter using a delay time between impulse responses of head-related transfer functions (308) of the two channels.

11. Decoder in accordance with claim 10, in which the filter calculator (302) is operative in a filterbank domain having n frequency bands and to derive individual phase parameters for each frequency band using the delay time.

12. Decoder in accordance with claim 10, in which the filter calculator (302) is operative in a filterbank domain having more than 2 frequency bands and to derive individual phase parameters Φ_{XY} for each frequency band using the delay time τ_{XY} according to the following formula:

$$\phi_{XY} = \frac{\pi(n+\frac{1}{2})}{64}\tau_{XY} \text{ .}$$

13. Decoder in accordance with claim 1, in which the filter calculator (302) is operative to derive the phase parameter using the phase angle of the normalized complex cross correlation between the impulse responses of head-related transfer functions (308) of the first and the second channel.

14. Decoder in accordance with claim 1, adapted to use a representation of a down mix signal (312) having a left and a right channel derived from a multi-channel signal having a left-front, a left-surround, a right-front, a right-surround and a center channel.

15. Decoder in accordance with claim 1, in which the synthesizer is operative to derive channels of the headphone down mix signal (314) applying a linear combination of the modified head-related transfer functions (310) to the representation of the down mix (312) of the multi-channel signal.

16. Decoder in accordance with claim 15, in which the synthesizer is operative to use coefficients for the linear combination of the modified head-related transfer functions (310) depending on the level parameter (306).

17. Decoder in accordance with claim 15, in which the synthesizer (304) is operative to use coefficients for the linear combination depending on additional multi-channel parameters related to additional spatial properties of the multi-channel signal.

18. Binaural decoder, comprising:

a decoder in accordance with claim 1;
an analysis filterbank (300) for deriving the representation of the down mix of the multi-channel signal (312) by subband filtering the downmix of the multi-channel signal; and
a synthesis filterbank (302) for deriving a time-domain headphone signal by synthesizing the headphone down mix signal (314).

19. Decoder for deriving a spatial stereo down mix signal using a representation of a down mix of a multi-channel signal (312) and using a level parameter (306) having information on a level relation between two channels of the multi-channel signal and using crosstalk cancellation filters related to the two channels of the multi-channel signal, wherein a first channel of the two channels is a front channel of the left or the right side of the multi-channel signal and a second channel of the two channels is a back channel of the same side, comprising:

a filter calculator (302) for deriving a modified crosstalk cancellation filter $H_Y(X)$ by weighting the front channel head-related transfer function $H_Y(Xf)$ and the back channel head-related transfer function $H_Y(Xs)$ of the two channels using the level parameter (306) such that the modified crosstalk cancellation filter $H_Y(X)$ is stronger influenced by the crosstalk cancellation filter of a channel having a higher level than by the crosstalk cancellation filter of a channel having a lower level by using the following complex linear combination:

$$H_Y(X) = gw_f \exp(-j\phi_{XY} w_s^2) H_Y(Xf) + gw_s \exp(j\phi_{XY} w_f^2) H_Y(Xs),$$, wherein

Φ_{XY} is a phase parameter, w_s and w_f are weighting factors derived using the level parameter (306) and g is a common gain factor derived using the level parameter (306); and
a synthesizer (304) for deriving the spatial stereo down mix signal using the modified crosstalk cancellation filters and the representation of the down mix signal (312).

20. Method of deriving a headphone down mix signal (314) using a representation of a down mix of a multi-channel signal (312) and using a level parameter (306) having information on a level relation between two channels of the multi-channel signal and using head-related transfer functions (308) related to the two channels of the multi-channel signal wherein a first channel of the two channels is a front channel of the left or the right side of the multi-channel signal and a second channel of the two channels is a back channel of the same side, the method comprising:

deriving, using the level parameter (306), a modified head-related transfer functions $H_Y(X)$ (310) by weighting the front channel head-related transfer function $H_Y(Xf)$ and the back channel head-related transfer function $H_Y(Xs)$ using the level parameter (306) such that the modified head-related transfer function $H_Y(X)$ is stronger influenced by the head-related transfer function of a channel having a higher level than by the head-related transfer function of a channel having a lower level by using the following complex linear combination:

$$H_Y(X) = gw_f \exp(-j\phi_{XY} w_s^2) H_Y(Xf) + gw_s \exp(j\phi_{XY} w_f^2) H_Y(Xs),$$, wherein

Φ_{XY} is a phase parameter, w_s and w_f are weighting factors derived using the level parameter (306) and g is a common gain factor derived using the level parameter (306); and deriving the headphone down mix signal (314) using the modified head-related transfer functions (310) and the representation of the down mix signal.

21. Receiver or audio player having a decoder for deriving a headphone down mix signal (314) according to claims 1 to 17.

22. Method of receiving or audio playing, the method having a method for deriving a headphone down mix signal (314) according to claim 20.

23. Computer program having a program code for performing, when running on a computer, one of the methods of claims 20 or 22.

Patentansprüche

1. Decodierer zum Ableiten eines Kopfhörer-Abwärtsmischsignals (314) unter Verwendung einer Darstellung einer Abwärtsmischung eines Mehrkanalsignals (312) und unter Verwendung eines Pegelparameters (306), der Informationen über eine Pegelbeziehung zwischen zwei Kanälen des Mehrkanalsignals aufweist, und unter Verwendung kopfbezogener Transferfunktionen (308), die auf die zwei Kanäle des Mehrkanalsignals bezogen sind, wobei ein erster Kanal der zwei Kanäle ein vorderer Kanal der linken oder der rechten Seite des Mehrkanalsignals ist und ein zweiter Kanal der zwei Kanäle ein hinterer Kanal derselben Seite ist, mit folgenden Merkmalen:

einer Filterberechnungseinrichtung (302) zum Ableiten einer modifizierten kopfbezogenen Transferfunktion $H_Y(X)$ (310) durch Gewichten der kopfbezogenen Transferfunktion $H_Y(Xf)$ des vorderen Kanals und der kopfbezogenen Transferfunktion $H_Y(Xs)$ des hinteren Kanals unter Verwendung des Pegelparameters (306) derart, dass die modifizierte kopfbezogene Transferfunktion $H_Y(X)$ (310) stärker durch die kopfbezogene Transferfunktion (308) eines Kanals mit einem höheren Pegel beeinflusst wird als durch die kopfbezogenen Transferfunktion (308) eines Kanals mit einem niedrigeren Pegel, indem die folgende komplexe lineare Kombination verwendet wird: $H_Y(X) = gw_f \exp(-j\phi_{XY} w_s^2) H_Y(Xf) + gw_s \exp(j\phi_{XY} w_f^2) H_Y(Xs)$,, wobei Φ_{XY} ein Phasenparameter ist, w_s und w_f Gewichtungsfaktoren sind, die unter Verwendung des Pegelparameters (306) abgeleitet werden, und g ein gemeinsamer Gewinnfaktor ist, der unter Verwendung des Pegelparameters (306) abgeleitet wird; und

einem Synthetisierer (304) zum Ableiten des Kopfhörer-Abwärtsmischsignals (314) unter Verwendung der modifizierten kopfbezogenen Transferfunktion (310) und der Darstellung des Abwärtsmischsignals (312).

2. Decodierer gemäß Anspruch 1, bei dem die Filterberechnungseinrichtung (302) derart wirksam ist, dass die Anzahl abgeleiteter modifizierter kopfbezogener Transferfunktionen (310) geringer ist als die Anzahl zugeordneter kopfbezogener Transferfunktionen (308) der zwei Kanäle.

3. Decodierer gemäß Anspruch 1, bei dem die Filterberechnungseinrichtung (302) dahin gehend wirksam ist, eine modifizierte kopfbezogene Transferfunktion (310) abzuleiten, die dahin gehend angepasst ist, auf eine Filterbankdarstellung des Abwärtsmischsignals angewendet zu werden.

4. Decodierer gemäß Anspruch 1, der dazu angepasst ist, eine Darstellung des abgeleiteten Abwärtsmischsignals in einem Filterbankbereich zu verwenden.

5. Decodierer gemäß Anspruch 1, bei dem die Filterberechnungseinrichtung (302) dahin gehend wirksam ist, die modifizierte kopfbezogene Transferfunktion (310) unter Verwendung kopfbezogener Transferfunktionen (308), die durch mehr als drei Parameter gekennzeichnet sind, abzuleiten.

6. Decodierer gemäß Anspruch 1, bei dem die Filterberechnungseinrichtung (302) dahin gehend wirksam ist, die Gewichtungsfaktoren für die kopfbezogenen Transferfunktionen (308) der zwei Kanäle unter Verwendung desselben Pegelparameters (306) abzuleiten.

7. Decodierer gemäß Anspruch 6, bei dem die Filterberechnungseinrichtung (302) dahin gehend wirksam ist, einen ersten Gewichtungsfaktor w_f für einen ersten Kanal f und einen zweiten Gewichtungsfaktor W_s für einen zweiten Kanal s unter Verwendung des Pegelparameters CLD_1 gemäß den folgenden Formeln abzuleiten:

$$w_f^2 = \frac{10^{CLD_l/10}}{1+10^{CLD_l/10}},$$

$$w_s^2 = \frac{1}{1+10^{CLD_l/10}}.$$

8. Decodierer gemäß Anspruch 1, bei dem die Filterberechnungseinrichtung (302) dahin gehend wirksam ist, die modifizierte kopfbezogene Transferfunktion (310) abzuleiten, wobei ein gemeinsamer Gewinnfaktor g an die kopfbezogenen Transferfunktion (308) der zwei Kanäle angelegt wird, derart, dass Energie bewahrt wird, wenn die modifizierten kopfbezogenen Transferfunktionen (310) abgeleitet werden.

9. Decodierer gemäß Anspruch 8, bei dem der gemeinsame Gewinnfaktor innerhalb des Intervalls $[1/\sqrt{2},1]$ liegt.

10. Decodierer gemäß Anspruch 1, bei dem die Filterberechnungseinrichtung (302) dahin gehend wirksam ist, den Phasenparameter unter Verwendung einer Verzögerungszeit zwischen Pulsantworten kopfbezogener Transferfunktionen (308) der zwei Kanäle abzuleiten.

11. Decodierer gemäß Anspruch 10, bei dem die Filterberechnungseinrichtung (302) in einem Filterbankbereich wirksam ist, der n Frequenzbänder aufweist, und dahin gehend wirksam ist, einzelne Phasenparameter für jedes Frequenzband unter Verwendung der Verzögerungszeit abzuleiten.

12. Decodierer gemäß Anspruch 10, bei dem die Filterberechnungseinrichtung (302) in einem Filterbankbereich wirksam ist, der mehr als 2 Frequenzbänder aufweist, und dahin gehend wirksam ist, einzelne Phasenparameter Φ_{XY} für jedes Frequenzband unter Verwendung der Verzögerungszeit τ_{XY} gemäß der folgenden Formel abzuleiten:

$$\phi_{XY} = \frac{\pi(n+\frac{1}{2})}{64}\tau_{XY} .$$

13. Decodierer gemäß Anspruch 1, bei dem die Filterberechnungseinrichtung (302) dahin gehend wirksam ist, den Phasenparameter unter Verwendung des Phasenwinkels der normierten komplexen Kreuzkorrelation zwischen den Pulsantworten kopfbezogener Transferfunktionen (308) des ersten und des zweiten Kanals abzuleiten.

14. Decodierer gemäß Anspruch 1, der dazu angepasst ist, eine Darstellung eines Abwärtsmischsignals (312) zu verwenden, das einen linken und einen rechten Kanal aufweist, der von einem Mehrkanalsignal abgeleitet ist, das einen linken vorderen, einen linken Surround-, einen rechten vorderen, einen rechten Surround- und einen Mittelkanal aufweist.

15. Decodierer gemäß Anspruch 1, bei dem der Synthetisierer dahin gehend wirksam ist, Kanäle des Kopfhörer-Abwärtsmischsignals (314) abzuleiten, wobei eine lineare Kombination der modifizierten kopfbezogenen Transferfunktionen (310) auf die Darstellung der Abwärtsmischung (312) des Mehrkanalsignals angewendet wird.

16. Decodierer gemäß Anspruch 15, bei dem der Synthetisierer dahin gehend wirksam ist, Koeffizienten für die lineare Kombination der modifizierten kopfbezogenen Transferfunktionen (310) in Abhängigkeit von dem Pegelparameter (306) zu verwenden.

17. Decodierer gemäß Anspruch 15, bei dem der Synthetisierer (304) dahin gehend wirksam ist, Koeffizienten für die lineare Kombination in Abhängigkeit von zusätzlichen Mehrkanalparametern, die mit zusätzlichen räumlichen Eigenschaften des Mehrkanalsignals zusammenhängen, zu verwenden.

18. Binauraler Decodierer, der folgende Merkmale aufweist:

einen Decodierer gemäß Anspruch 1;
eine Analysefilterbank (300) zum Ableiten der Darstellung der Abwärtsmischung des Mehrkanalsignals (312) durch Teilbandfiltern der Abwärtsmischung des Mehrkanalsignals; und
eine Synthesefilterbank (302) zum Ableiten eines Zeitbereich-Kopfhörersignals durch Synthetisieren des Kopfhörer-Abwärtsmischsignals (314).

19. Decodierer zum Ableiten eines räumlichen Stereo-Abwärtsmischsignals unter Verwendung einer Darstellung einer Abwärtsmischung eines Mehrkanalsignals (312) und unter Verwendung eines Pegelparameters (306), der Informationen über eine Pegelbeziehung zwischen zwei Kanälen des Mehrkanalsignals aufweist, und unter Verwendung von Nebensprechaufhebungsfiltem, die auf die zwei Kanäle des Mehrkanalsignals bezogen sind, wobei ein erster Kanal der zwei Kanäle ein vorderer Kanal der linken oder der rechten Seite des Mehrkanalsignals ist und ein zweiter Kanal der zwei Kanäle ein hinterer Kanal derselben Seite ist, mit folgenden Merkmalen:

einer Filterberechnungseinrichtung (302) zum Ableiten eines modifizierten Nebensprechaufhebungsfilters $H_Y(X)$ durch Gewichten der kopfbezogenen Transferfunktion $H_Y(Xf)$ des vorderen Kanals und der kopfbezogenen Transferfunktion $H_Y(Xs)$ des hinteren Kanals der zwei Kanäle unter Verwendung des Pegelparameters (306) derart, dass das modifizierte Nebensprechaufhebungsfilter $H_Y(X)$ stärker durch das Nebensprechaufhebungsfilter eines Kanals mit einem höheren Pegel beeinflusst wird als durch das Nebensprechaufhebungsfilter eines Kanals mit einem niedrigeren Pegel, indem die folgende komplexe lineare Kombination verwendet wird: $H_Y(X) = gw_f \exp(-j\phi_{XY} w_s^2) H_Y(Xf) + gw_s \exp(j\phi_{XY} w_f^2) H_Y(Xs)$, wobei Φ_{XY} ein Phasenparameter ist, w_s und w_f Gewichtungsfaktoren sind, die unter Verwendung des Pegelparameters (306) abgeleitet werden, und g ein gemeinsamer Gewinnfaktor ist, der unter Verwendung des Pegelparameters (306) abgeleitet

wird; und
einem Synthetisierer (304) zum Ableiten des räumlichen Stereo-Abwärtsmischsignals unter Verwendung des modifizierten Nebensprechaufhebungsfilters und der Darstellung des Abwärtsmischsignals (312).

20. Verfahren zum Ableiten eines Kopfhörer-Abwärtsmischsignals (314) unter Verwendung einer Darstellung einer Abwärtsmischung eines Mehrkanalsignals (312) und unter Verwendung eines Pegelparameters (306), der Informationen über eine Pegelbeziehung zwischen zwei Kanälen des Mehrkanalsignals aufweist, und unter Verwendung kopfbezogener Transferfunktionen (308), die auf die zwei Kanäle des Mehrkanalsignals bezogen sind, wobei ein erster Kanal der zwei Kanäle ein vorderer Kanal der linken oder der rechten Seite des Mehrkanalsignals ist und ein zweiter Kanal der zwei Kanäle ein hinterer Kanal derselben Seite ist, wobei das Verfahren folgende Schritte aufweist:

Ableiten, unter Verwendung des Pegelparameters (306), einer modifizierten kopfbezogenen Transferfunktion $H_Y(X)$ (310) durch Gewichten der kopfbezogenen Transferfunktion $H_Y(Xf)$ des vorderen Kanals und der kopfbezogenen Transferfunktion $H_Y(Xs)$ des hinteren Kanals unter Verwendung des Pegelparameters (306) derart, dass die modifizierte kopfbezogene Transferfunktion $H_Y(X)$ stärker durch die kopfbezogene Transferfunktion eines Kanals mit einem höheren Pegel beeinflusst wird als durch die kopfbezogene Transferfunktion eines Kanals mit einem niedrigeren Pegel, indem die folgende komplexe lineare Kombination verwendet wird:

$$H_Y(X) = gw_f \exp(-j\phi_{XY} w_s^2) H_Y(Xf) + gw_s \exp(j\phi_{XY} w_f^2) H_Y(Xs),$$ wobei

Φ_{XY} ein Phasenparameter ist, w_s und w_f Gewichtungsfaktoren sind, die unter Verwendung des Pegelparameters (306) abgeleitet werden, und g ein gemeinsamer Gewinnfaktor ist, der unter Verwendung des Pegelparameters (306) abgeleitet wird; und
Ableiten des Kopfhörer-Abwärtsmischsignals (314) unter Verwendung der modifizierten kopfbezogenen Transferfunktionen (310) und der Darstellung des Abwärtsmischsignals.

21. Empfänger oder Audio-Abspielgerät, der beziehungsweise das einen Decodierer zum Ableiten eines Kopfhörer-Abwärtsmischsignals (314) gemäß den Ansprüchen 1 bis 17 aufweist.

22. Verfahren zum Empfangen oder Audio-Abspielen, wobei das Verfahren ein Verfahren zum Ableiten eines Kopfhörer-Abwärtsmischsignals (314) gemäß Anspruch 20 aufweist.

23. Computerprogramm, das einen Programmcode zum Durchführen, wenn es auf einem Computer abläuft, eines der Verfahren gemäß den Ansprüchen 20 oder 22 aufweist.

Revendications

1. Décodeur pour dériver un signal de mélange descendant d'écouteur (314) à l'aide d'une représentation d'un mélange descendant d'un signal multicanal (312) et à l'aide d'un paramètre de niveau (306) présentant des informations sur un rapport de niveau entre deux canaux du signal multicanal et à l'aide de fonctions de transfert relatives à la tête (308) relatives aux deux canaux du signal multicanal, dans lequel un premier canal parmi les deux canaux est un canal avant du côté gauche ou droit du signal multicanal et un deuxième canal parmi les deux canaux est un canal arrière du même côté, comprenant:

un calculateur de filtre (302) destiné à dériver une fonction de transfert relative à la tête modifiée $H_Y(X)$ (310) en pondérant la fonction de transfert relative à la tête de canal avant $H_Y(Xf)$ et la fonction de transfert relative à la tête de canal arrière $H_Y(Xs)$ à l'aide du paramètre de niveau (306) de sorte que la fonction de transfert relative à la tête modifiée $H_Y(X)$ (310) soit plus fortement influencée par la fonction de transfert relative à la tête (308) d'un canal présentant un niveau supérieur que par la fonction de transfert relative à la tête (308) d'un canal présentant un niveau inférieur à l'aide de la combinaison linéaire complexe suivante:

$$H_Y(X) = gw_f exp(-j\Phi_{XY} w_s^2) H_Y(Xf) + gw_s exp\left(j\Phi_{XY} w_f^2\right) H_Y(Xs),$$ où

Φ_{XY} est un paramètre de phase, w_s et w_f sont des facteurs de pondération dérivés à l'aide du paramètre de niveau (306) et g est un facteur de gain commun dérivé à l'aide du paramètre de niveau (306); et
un synthétiseur (304) destiné à dériver le signal de mélange descendant d'écouteur (314) à l'aide de la fonction

de transfert relative à la tête modifiée (310) et de la représentation du signal de mélange descendant (312).

2. Décodeur selon la revendication 1, dans lequel le calculateur de filtre (302) est opérationnel de sorte que le nombre de fonctions de transfert relative à la tête modifiées (310) dérivées soit inférieur au nombre de fonctions de transfert relative à la tête associées (308) des deux canaux.

3. Décodeur selon la revendication 1, dans lequel le calculateur de filtre (302) est opérationnel pour dériver une fonction de transfert relative à la tête modifiée (310) adaptée pour être appliquée à une représentation de banc de filtres du signal de mélange descendant.

4. Décodeur selon la revendication 1, adapté pour utiliser une représentation du signal de mélange descendant dérivé dans un domaine de banc de filtres.

5. Décodeur selon la revendication 1, dans lequel le calculateur de filtre (302) est opérationnel pour dériver la fonction de transfert relative à la tête modifiée (310) à l'aide de fonctions de transfert relative à la tête (308) **caractérisées par** plus de trois paramètres.

6. Décodeur selon la revendication 1, dans lequel le calculateur de filtre (302) est opérationnel pour dériver les facteurs de pondération pour les fonctions de transfert relative à la tête (308) des deux canaux à l'aide du même paramètre de niveau (306).

7. Décodeur selon la revendication 6, dans lequel le calculateur de filtre (302) est opérationnel pour dériver un premier facteur de pondération w_f pour un premier canal f et un deuxième facteur de pondération w_s pour un deuxième canal s à l'aide du paramètre de niveau CLD_1 selon les formules suivantes:

$$w_{lf}^2 = \frac{10^{CLD_l/10}}{1+10^{CLD_l/10}},$$

$$w_{ls}^2 = \frac{1}{1+10^{CLD_l/10}}.$$

8. Décodeur selon la revendication 1, dans lequel le calculateur de filtre (302) est opérationnel pour dériver la fonction de transfert relative à la tête modifiée (310) en appliquant un facteur de gain commun g à la fonction de transfert relative à la tête (308) des deux canaux, de sorte que l'énergie soit préservée lors de la dérivation des fonctions de transfert relative à la tête modifiées (310).

9. Décodeur selon la revendication 8, dans lequel le facteur de gain commun se situe dans l'intervalle $[1/\sqrt{2}, 1]$.

10. Décodeur selon la revendication 1, dans lequel le calculateur de filtre (302) est opérationnel pour dériver le paramètre de phase à l'aide d'un temps de retard entre les réponses impulsionnelles des fonctions de transfert relative à la tête (308) des deux canaux.

11. Décodeur selon la revendication 10, dans lequel le calculateur de filtre (302) est opérationnel dans un domaine de banc de filtres présentant n bandes de fréquences et pour dériver des paramètres de phase individuels pour chaque bande de fréquences à l'aide du temps de retard.

12. Décodeur selon la revendication 10, dans lequel le calculateur de filtre (302) est opérationnel dans un domaine de banc de filtres présentant plus de 2 bandes de fréquences et pour dériver des paramètres de phase individuels Φ_{XY} pour chaque bande de fréquences à l'aide du temps de retard τ_{XY} selon la formule suivante:

$$\Phi_{XY} = \frac{\pi\left(n+\frac{1}{2}\right)}{64}\tau_{XY}.$$

13. Décodeur selon la revendication 1, dans lequel le calculateur de filtre (302) est opérationnel pour dériver le paramètre de phase à l'aide de l'angle de phase de la corrélation croisée complexe normalisée entre les réponses impulsionnelles des fonctions de transfert relative à la tête (308) du premier et du deuxième canal.

14. Décodeur selon la revendication 1, adapté pour utiliser une représentation d'un signal de mélange descendant (312) présentant un canal gauche et un canal droit dérivés d'un signal multicanal présentant un canal avant gauche, un canal ambiophonique gauche, un canal avant droit, un canal ambiophonique droit et un canal central.

15. Décodeur selon la revendication 1, dans lequel le synthétiseur est opérationnel pour dériver les canaux du signal de mélange descendant de l'écouteur (314) en appliquant une combinaison linéaire des fonctions de transfert relative à la tête modifiées (310) à la représentation du mélange descendant (312) du signal multicanal.

16. Décodeur selon la revendication 15, dans lequel le synthétiseur est opérationnel pour utiliser des coefficients pour la combinaison linéaire des fonctions de transfert relative à la tête modifiées (310) en fonction du paramètre de niveau (306).

17. Décodeur selon la revendication 15, dans lequel le synthétiseur (304) est opérationnel pour utiliser des coefficients pour la combinaison linéaire en fonction de paramètres multicanal additionnels relatifs à des propriétés spatiales additionnelles du signal multicanal.

18. Décodeur binaural, comprenant:

un décodeur selon la revendication 1;
un banc de filtres d'analyse (300) destiné à dériver la représentation du mélange descendant du signal multicanal (312) en filtrant par sous-bande le mélange descendant du signal multicanal; et
un banc de filtres de synthèse (302) destiné à dériver un signal d'écouteur dans le domaine temporel en synthétisant le signal de mélange descendant d'écouteur (314).

19. Décodeur pour dériver un signal de mélange descendant stéréo spatial à l'aide d'une représentation d'un mélange descendant d'un signal multicanal (312) et à l'aide d'un paramètre de niveau (306) présentant des informations sur un rapport de niveau entre deux canaux du signal multicanal et à l'aide de filtres d'annulation de diaphonie relatifs aux deux canaux du signal multicanal, dans lequel un premier canal parmi les deux canaux est un canal avant du côté gauche ou du côté droit du signal multicanal et un deuxième canal parmi les deux canaux est un canal arrière du même côté, comprenant:

un calculateur de filtre (302) destiné à dériver un filtre d'annulation de diaphonie $H_Y(X)$ en pondérant la fonction de transfert relative à la tête de canal avant $H_Y(Xf)$ et la fonction de transfert relative à la tête de canal arrière $H_Y(Xs)$ des deux canaux à l'aide du paramètre de niveau (306) de sorte que le filtre d'annulation de diaphonie modifié $H_Y(X)$ soit plus fortement influencé par le filtre d'annulation de diaphonie d'un canal présentant un niveau supérieur que par le filtre d'annulation de diaphonie d'un canal présentant un niveau inférieur à l'aide de la combinaison linéaire complexe suivante:

$$H_Y(X) = gw_f exp(-j\Phi_{XY}w_s^2)H_Y(Xf) + gw_s exp\left(j\Phi_{XY}w_f^2\right)H_Y(Xs),$$ où Φ_{XY} est un paramètre de phase, w_s et w_f sont des facteurs de pondération dérivés à l'aide du paramètre de niveau (306) et g est un facteur de gain commun dérivé à l'aide du paramètre de niveau (306); et
un synthétiseur (304) destiné à dériver le signal de mélange descendant stéréo spatial à l'aide des filtres d'annulation de diaphonie modifiés et de la représentation du signal de mélange descendant (312).

20. Procédé pour dériver un signal de mélange descendant d'écouteur (314) à l'aide d'une représentation d'un mélange descendant d'un signal multicanal (312) et à l'aide d'un paramètre de niveau (306) présentant des informations sur un rapport de niveau entre deux canaux du signal multicanal et à l'aide de fonctions de transfert relatives à la tête

(308) relatives aux deux canaux du signal multicanal, dans lequel un premier canal parmi les deux canaux est un canal avant du côté gauche ou droit du signal multicanal et un deuxième canal parmi les deux canaux est un canal arrière du même côté, comprenant:

dériver, à l'aide du paramètre de niveau (306), une fonction de transfert relative à la tête modifiée $H_Y(X)$ (310) en pondérant la fonction de transfert relative à la tête de canal avant $H_Y(Xf)$ et la fonction de transfert relative à la tête de canal arrière $H_Y(Xs)$ à l'aide du paramètre de niveau (306) de sorte que la fonction de transfert relative à la tête modifiée $H_Y(X)$ soit plus fortement influencée par la fonction de transfert relative à la tête d'un canal présentant un niveau supérieur que par la fonction de transfert relative à la tête d'un canal présentant un niveau inférieur à l'aide de la combinaison linéaire complexe suivante:

$$H_Y(X) = gw_f exp(-j\Phi_{XY} w_s^2) H_Y(Xf) + gw_s exp\left(j\Phi_{XY} w_f^2\right) H_Y(Xs),\text{où}$$

Φ_{XY} est un paramètre de phase, w_s et w_f sont des facteurs de pondération dérivés à l'aide du paramètre de niveau (306) et g est un facteur de gain commun dérivé à l'aide du paramètre de niveau (306); et

dériver le signal de mélange descendant d'écouteur (314) à l'aide des fonctions de transfert relative à la tête modifiée (310) et de la représentation du signal de mélange descendant.

21. Récepteur ou reproducteur audio présentant un décodeur pour dériver un signal de mélange descendant d'écouteur (314) selon les revendications 1 à 17.

22. Procédé de réception ou de reproduction audio, le procédé présentant un procédé pour dériver un signal de mélange descendant d'écouteur (314) selon la revendication 20.

23. Programme d'ordinateur ayant un code de programme pour réaliser, lorsqu'il est exécuté sur un ordinateur, l'un des procédés des revendications 20 ou 22.

FIG 1

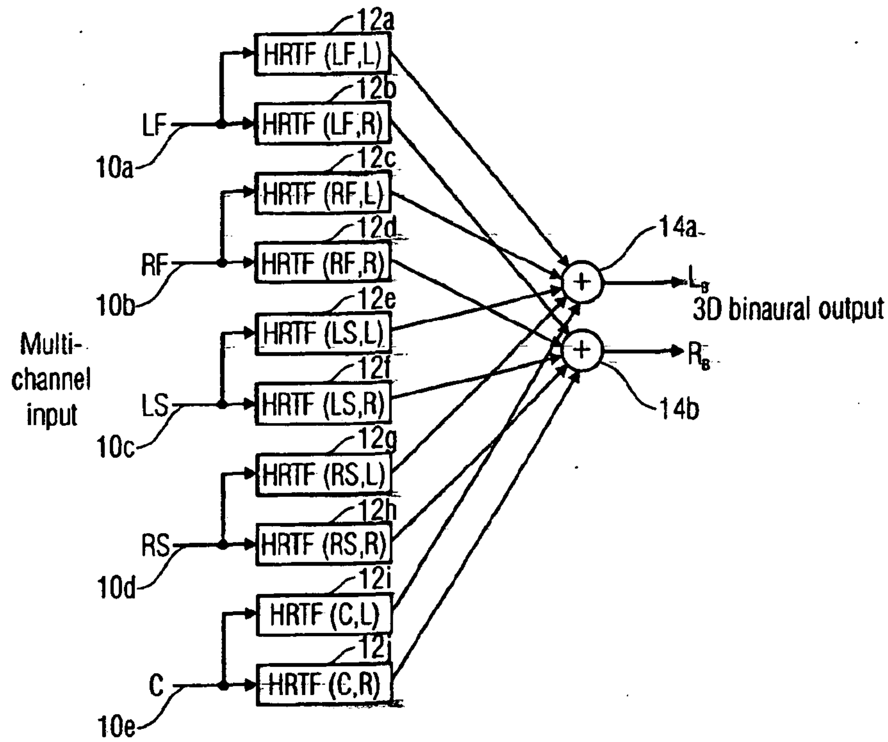

HRTF (LF,L)
12a
HRTF (LF,R)
12b
LF
10a
HRTF (RF,L)
12c
HRTF (RF,R)
12d
RF
10b
HRTF (LS,L)
12e
HRTF (LS,R)
12f
LS
10c
Multi-
channel
input
HRTF (RS,L)
12g
HRTF (RS,R)
12h
RS
10d
HRTF (C,L)
12i
HRTF (C,R)
12j
C
10e
14a
L_B
3D binaural output
R_B
14b

FIG 1b

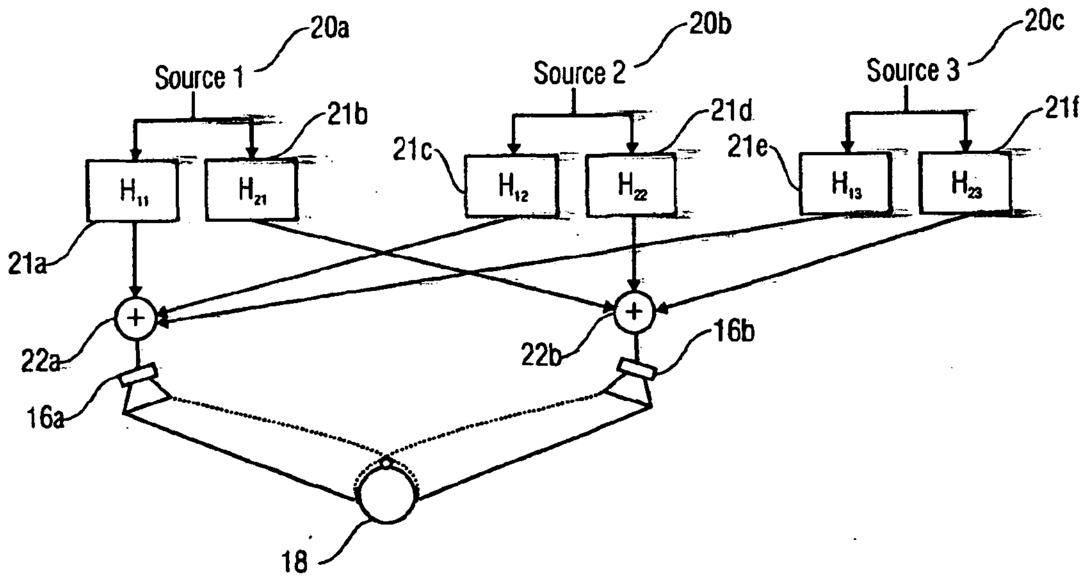

Source 1
20a
Source 2
20b
Source 3
20c
H11
21a
H21
21b
H12
21c
H22
21d
H13
21e
H23
21f
22a
22b
16a
16b
18

FIG 2

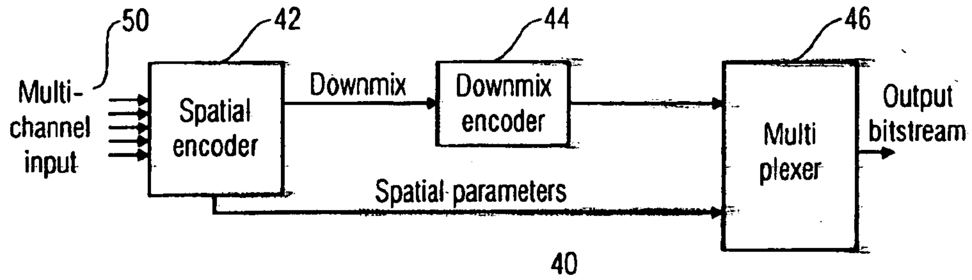


FIG 3

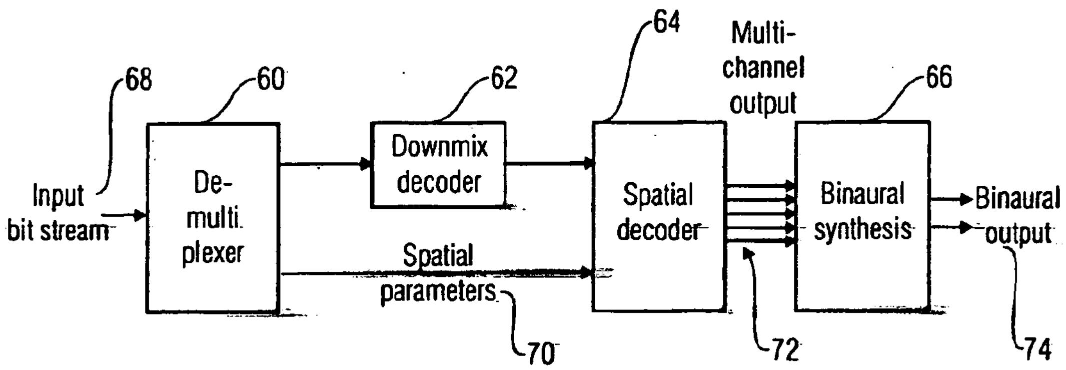


FIG 4

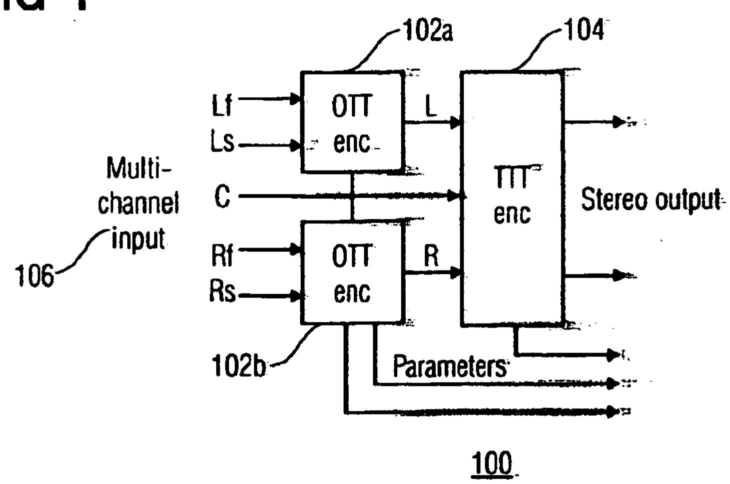
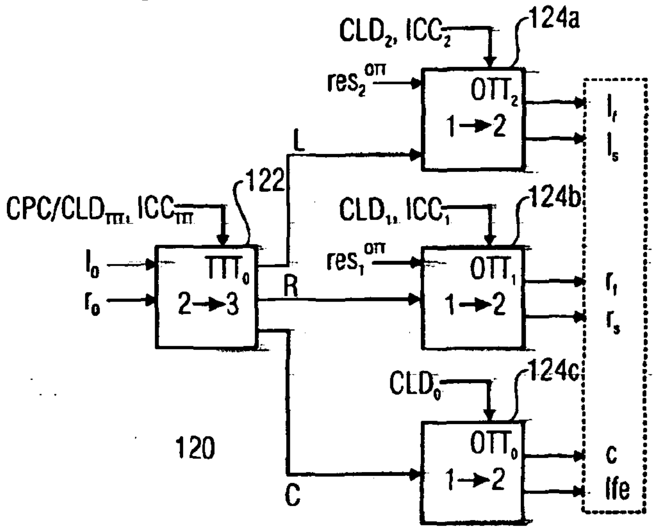
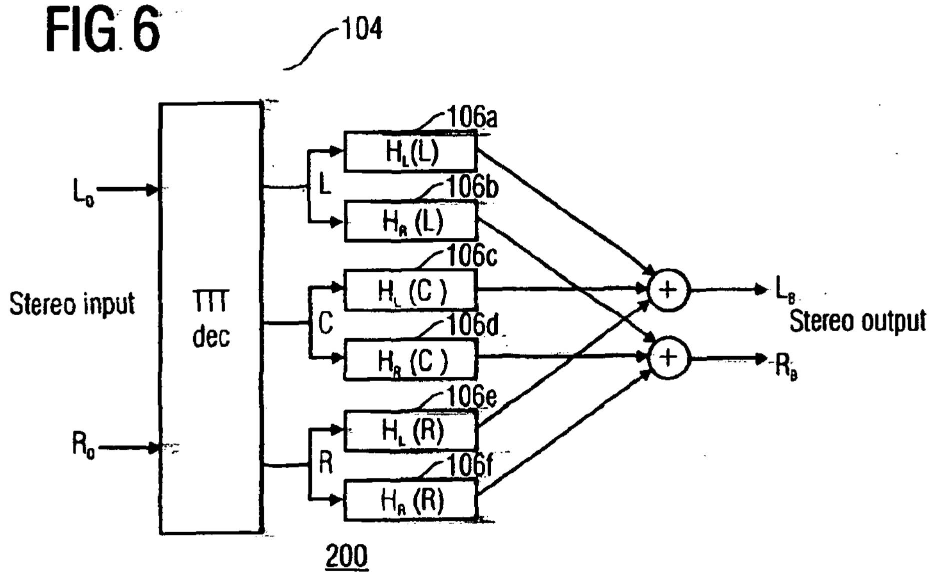


FIG 5



CLD2, ICC2
124a
res2 OTT
OTT2
1→2
lf
ls
L
122
CPC/CLDTTT, ICCTTT
CLD1, ICC1
124b
res1 OTT
l0
r0
TTT0
2→3
R
OTT1
1→2
rf
rs
CLD0
124c
120
OTT0
1→2
c
lfe
C

FIG 6



104
106a
HL(L)
L0
L
106b
HR (L)
106c
Stereo input
TTT
dec
HL (C)
C
LB
Stereo output
106d
HR (C)
RB
106e
R0
HL (R)
R
106f
HR (R)
200

FIG 7

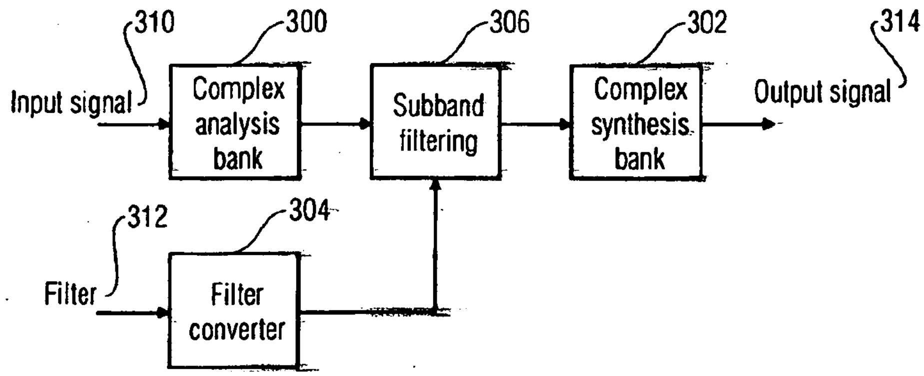


FIG 8

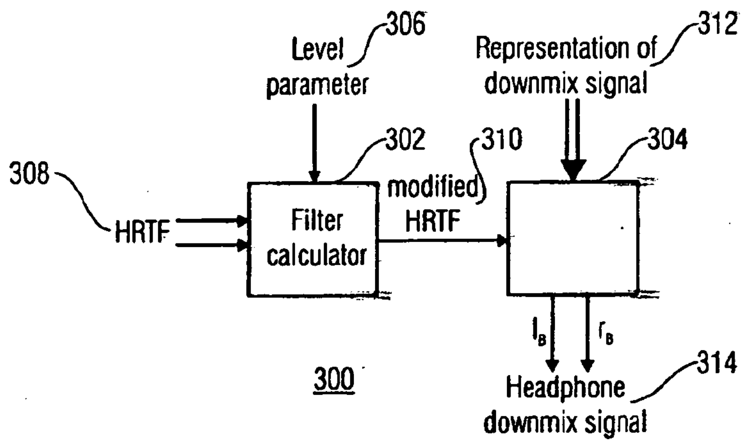


FIG 9

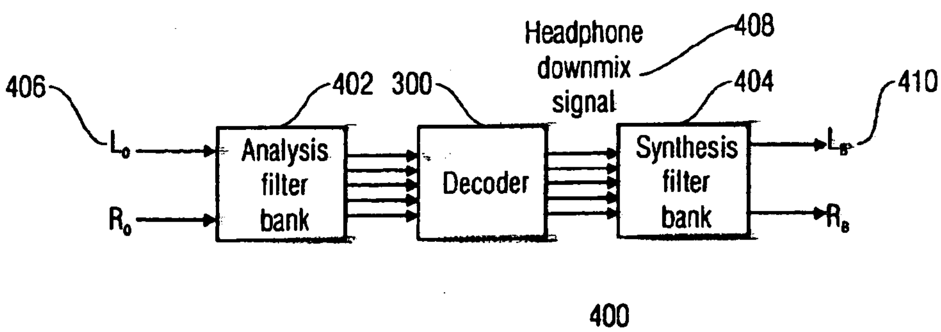
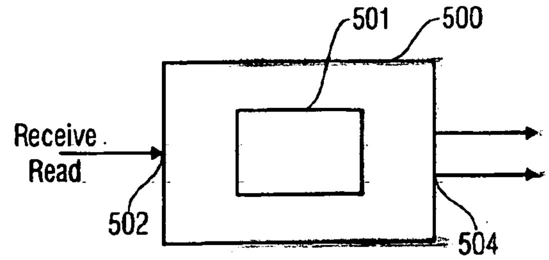


FIG 10



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 20060045274 A **[0009]**
- WO 2006008683 A **[0010]**
- US 20030035553 A **[0012]**