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

(11) EP 2 934 025 A1

(12) **EUROPEAN PATENT APPLICATION**

(43) Date of publication:
21.10.2015 Bulletin 2015/43

(51) Int Cl.:
H04S 3/00 (2006.01) ***G10L 19/008*** (2013.01)

(21) Application number: **14305559.8**

(22) Date of filing: **15.04.2014**

(84) Designated Contracting States:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR
Designated Extension States:
BA ME

(71) Applicant: **Thomson Licensing**
92130 Issy-les-Moulineaux (FR)

(72) Inventors:
- **Boehm, Johannes**
  37081 Göttingen (DE)
- **Keiler, Florian**
  30161 Hannover (DE)

(74) Representative: **König, Uwe**
Deutsche Thomson OHG
Karl-Wiechert-Allee 74
30625 Hannover (DE)

(54) **Method and device for applying dynamic range compression to a higher order ambisonics signal**

(57) Dynamic Range Control (DRC) cannot be simply applied to Higher Order Ambisonics (HOA) based signals. A method for performing DRC on a HOA signal comprises transforming the HOA signal to the spatial domain, analyzing the transformed HOA signal, and obtaining, from results of said analyzing, gain factors that are usable for dynamic compression. The gain factors can be transmitted together with the HOA signal. When applying the DRC, the HOA signal is transformed to the spatial domain, the gain factors are extracted and multiplied with the transformed HOA signal in the spatial domain, wherein a gain compensated transformed HOA signal is obtained. The gain compensated transformed HOA signal is transformed back into the HOA domain, wherein a gain compensated HOA signal is obtained.

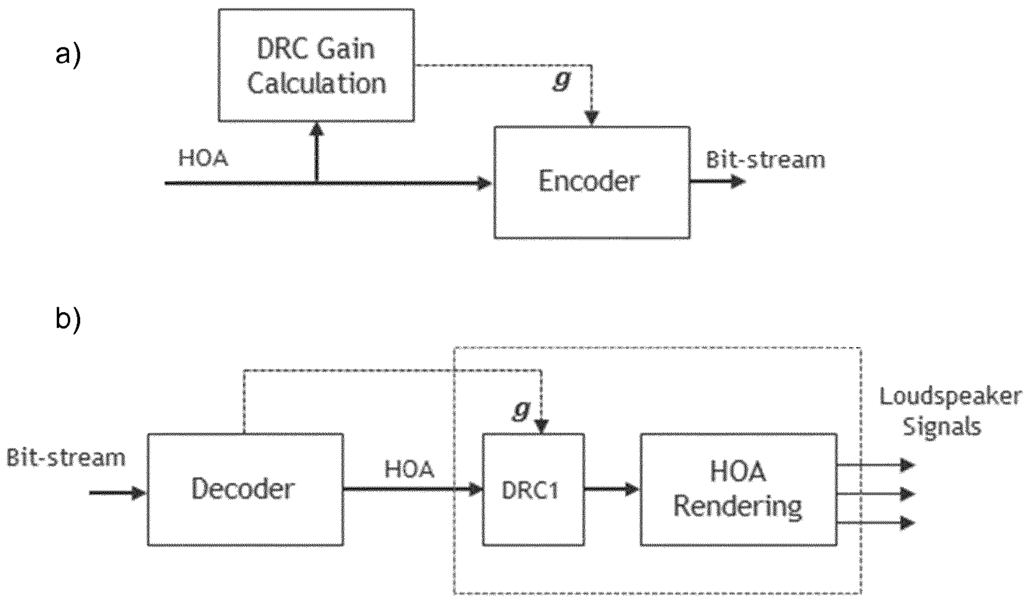

Fig.2

EP 2 934 025 A1

Description

Field of the invention

[0001] This invention relates to a method and a device for performing Dynamic Range Compression (DRC) to an Ambisonics signal, and in particular to a Higher Order Ambisonics (HOA) signal.

Background

[0002] The purpose of Dynamic Range Compression (DRC) is to reduce the dynamic range of an audio signal. A time-varying gain factor is applied to the audio signal. Typically this gain factor is dependent on the amplitude envelope of the signal used for controlling the gain. The mapping is in general non-linear. Large amplitudes are mapped to smaller ones while faint sounds are often amplified. Scenarios are noisy environments, late night listening, small speakers or mobile headphone listening.

[0003] A common concept for streaming or broadcasting Audio is to generate the DRC gains before transmission and apply these gains after receiving and decoding. The principle of using DRC is shown in Fig.1 a). Fig.1 a) shows how DRC is applied to an audio signal. The signal level, usually the signal envelope, is detected and a related gain is computed. The time-varying gain is used to change the amplitude of the audio signal. Fig.1 b) shows the principle of using DRC for encoding/decoding, wherein gain factors are transmitted together with the coded audio signal. On the decoder side, the gains are applied to the decoded audio signal in order to reduce the dynamic range.

[0004] For 3D audio, different gains can be applied to loudspeaker channels, which represent different spatial positions. These positions then need to be known at the sending side, in order to be able to generate a matching set of gains. This is usually only possible for idealized conditions, while in a realistic case the number of speakers and its placement varies in many ways. This is more influenced from practical considerations than from specifications. Higher Order Ambisonics (HOA) allows for flexible rendering. A HOA signal is composed of coefficient channels that do not directly represent sound levels. Therefore, DRC cannot be simply applied to HOA based signals.

Summary of the Invention

[0005] The present invention describes how DRC can be applied to HOA signals. A HOA signal is analysed in order to obtain one or more gain coefficients. In one embodiment, at least two gain coefficients are obtained and the analysis of the HOA signal comprises a transformation into the spatial domain (iDSHT). The one or more gain coefficients are transmitted together with the original HOA signal. A special indication can be transmitted to indicate if all gain coefficients are equal. This is the case in a so-called simplified mode, whereas at least two different gain coefficients are used in a non-simplified mode. At the decoder, the one or more gains can (but need not) be applied to the HOA signal. The user has a choice whether or not to apply the one or more gains. An advantage of the simplified mode is that it requires considerably less computations, since only one gain factor is used, and since the gain factor can be applied to the coefficient channels of the HOA signal directly in the HOA domain, so that the transform into the spatial domain and subsequent transform back into the HOA domain can be skipped. In the simplified mode, the gain factor is obtained by analysis of only the zeroth order coefficient channel of the HOA signal.

[0006] According to one embodiment of the invention, a method for performing DRC on a HOA signal comprises transforming the HOA signal to the spatial domain (by an inverse DSHT), analyzing the transformed HOA signal and obtaining, from results of said analyzing, gain factors that are usable for dynamic range compression. In further steps, the obtained gain factors are multiplied (in the spatial domain) with the transformed HOA signal, wherein a gain compressed transformed HOA signal is obtained. Finally, the gain compressed transformed HOA signal is transformed back into the HOA domain (by a DSHT), i.e. coefficient domain, wherein a gain compressed HOA signal is obtained.

[0007] Further, according to one embodiment of the invention, a method for performing DRC in a simplified mode on a HOA signal comprises analyzing the HOA signal and obtaining from results of said analyzing a gain factor that is usable for dynamic range compression. In further steps, upon evaluation of the indication, the obtained gain factor is multiplied with coefficient channels of the HOA signal (in the HOA domain), wherein a gain compressed HOA signal is obtained. Also upon evaluation of the indication, it can be determined that a transformation of the HOA signal can be skipped. The indication to indicate simplified mode, i.e. that only one gain factor is used, can be set implicitly, e.g. if only simplified mode can be used due to hardware or other restrictions, or explicitly, e.g. upon user selection of either simplified or non-simplified mode.

[0008] Further, according to one embodiment of the invention, a method for applying DRC gain factors to a HOA signal comprises receiving a HOA signal, an indication and gain factors (together with the HOA signal or separately), determining that the indication indicates non-simplified mode, transforming the HOA signal into the spatial domain (using an inverse DSHT), wherein a transformed HOA signal is obtained, multiplying the gain factors with the transformed HOA signal, wherein a dynamic range compressed transformed HOA signal is obtained, and transforming the dynamic range compressed transformed HOA signal back into the HOA domain (i.e. coefficient domain) (using DSHT), wherein a dynamic range compressed HOA signal is obtained.

[0009] Further, according to one embodiment of the invention, a method for applying a DRC gain factor to a HOA signal comprises receiving a HOA signal, an indication and a gain factor (together with the HOA signal or separately), determining that the indication indicates simplified mode, and upon said determining multiplying the gain factor with the HOA signal, wherein a dynamic range compressed HOA signal is obtained.

[0010] An apparatus for performing DRC on a HOA signal is disclosed in claim 12.
An apparatus for applying DRC gain factors to a HOA signal is disclosed in claim 13.

[0011] In one embodiment, the invention provides a computer readable medium having executable instructions to cause a computer to perform a method comprising steps described above.

[0012] Advantageous embodiments of the invention are disclosed in the dependent claims, the following description and the figures.

Brief description of the drawings

[0013] Exemplary embodiments of the invention are described with reference to the accompanying drawings, which show in

Fig.1 the general principle of DRC applied to audio;
Fig.2 a general approach for applying DRC to HOA based signals according to the invention;
Fig.3 Spherical speaker grids for N=1 to N=6;
Fig.4 Creation of DRC gains for HOA;
Fig.5 Applying DRC to HOA signals;
Fig.6 Dynamic Range Compression processing at the decoder side;
Fig.7 DRC for HOA in QMF domain combined with rendering step; and
Fig.8 DRC for HOA in QMF domain combined with rendering step for the simple case of a single DRC gain group.

Detailed description of the invention

[0014] The present invention describes how DRC can be applied to HOA. Fig.2 depicts the principle of the approach. On the encoding or transmitting side, as shown in Fig.2 a), HOA signals are analyzed, DRC gains ***g*** are calculated from the analysis of the HOA signal, and the DRC gains are coded and transmitted along with a coded representation of the HOA content. This may be a multiplexed bitstream or two or more separate bitstreams.

[0015] On the decoding or receiving side, as shown in Fig.2 b), the gains ***g*** are extracted from such bitstream or bitstreams. After decoding of the bitstream or bitstreams in a Decoder, the gains ***g*** are applied to the HOA signal as described below. By this, the gains are applied to the HOA signal, i.e. in general a dynamic range reduced HOA signal is obtained. Finally, the dynamic range adjusted HOA signal is rendered in a HOA renderer.

[0016] In the following, used assumptions and definitions are explained.
Assumptions are that the HOA renderer is energy preserving, i.e. N3D normalized Spherical Harmonics are used, and the energy of a single directional signal coded inside the HOA representation is maintained after rendering. It is described e.g. in EP13306042 (PD030040) how to achieve this energy preserving HOA rendering.

[0017] Definitions of used terms are as follows. $\boldsymbol{B} \in \mathbb{R}^{(N+1)^2 \mathrm{x}\, \tau}$ denotes a block of τ HOA samples, $\boldsymbol{B} = [\boldsymbol{b}(1), \boldsymbol{b}(2), .., \boldsymbol{b}(t), .., \boldsymbol{b}(\tau)]$, with vector $\boldsymbol{b}(t) = \left[b_1, b_2, \dots b_o, \dots b_{(N+1)^2}\right]^T = [B_0^0, B_1^{-1}, \dots\ B_n^m, \dots B_N^N,]^T$ which contains the Ambisonics coefficients in ACN order (vector index $o = n^2 + n + m + 1$, with coefficient order index n and coefficient degree index m). N denotes the HOA truncation order. The number of higher order coefficients in $\boldsymbol{b}$ is $(N + 1)^2$. The sample index for one block of data is t. τ may range from usually one sample to 64 samples or more.

The zeroth order signal $\vartheta_0 = [b_1(1), b_1(2), \dots, b_1(\tau)]$ is the first row of $\boldsymbol{B}$. $\boldsymbol{D} \in \mathbb{R}^{\mathrm{L\,x}\,(N+1)^2}$ denotes an energy preserving rendering matrix that renders a block of HOA samples to a block of L loudspeaker channel in spatial domain: $\boldsymbol{W} = \boldsymbol{DB}$, with $\boldsymbol{W} \in \mathbb{R}^{L \,\mathrm{x}\, \tau}$. This is the assumed procedure of the HOA renderer in Fig.2 b) (HOA rendering).

$\boldsymbol{D}_L \in \mathbb{R}^{(N+1)^2 \,\mathrm{x}\, (N+1)^2}$ denotes a rendering matrix related to $L_L = (N + 1)^2$ channels which are positioned on a sphere in a very regular manner, in a way that all neighboring positions share the same distance. $\boldsymbol{D}_L$ is well-conditioned and its inverse $\boldsymbol{D}_L^{-1}$ exists. Thus both define a pair of transformation matrices (DSHT - Discrete Spherical Harmonics Transform):

$$\boldsymbol{W}_L = \boldsymbol{D}_L \boldsymbol{B} \text{ , } \boldsymbol{B} = \boldsymbol{D}_L^{-1} \boldsymbol{W}_L$$

[0018] $\boldsymbol{g}$ is a vector of $L_L = (N + 1)^2$ gain DRC values. Gain values are assumed to be applied to a block of τ samples and are assumed to be smooth from block to block. For transmission gain values that share the same values can be combined to gain-groups. A single gain-group means a single DRC gain value, here indicated by g_1, is applied to all speaker channel τ samples.

[0019] For every HOA truncation order N, an ideal $L_L = (N + 1)^2$ virtual speaker grid and related rendering matrix $\boldsymbol{D}_L$ are defined. The virtual speaker positions sample spatial areas surrounding a virtual listener. The grids for N=1 to 6 are shown in Fig.3, where areas related to a speaker are shaded cells. One sampling position is always related to a central speaker position (azimuth = 0, inclination = $\pi/2$; azimuth is measured from frontal direction related to the listening position). The sampling positions, $\boldsymbol{D}_L$, $\boldsymbol{D}_L^{-1}$ are known at the encoder side when the DRC gains are created. At the decoder side, $\boldsymbol{D}_L$ and $\boldsymbol{D}_L^{-1}$ need to be known for applying the gain values.

[0020] Creation of DRC gains for HOA works as follows.

The HOA signal is converted to the spatial domain by $\boldsymbol{W}_L = \boldsymbol{D}_L\boldsymbol{B}$. Up to $L_L = (N + 1)^2$ DRC gains g_l are created by analyzing these signals. If the content is a combination of HOA and Audio Objects (AO), AO signals such as e.g. dialog tracks may be used for side chaining. This is shown in Fig.4b). When creating different DRC gain values related to different spatial areas, care needs to be taken that these gains do not influence the spatial image stability at the decoder side. To avoid this, a single gain may be assigned to all L channels, in the simplest case (simplified mode). This can be done by analyzing all spatial signals **W** or by analyzing the zeroth order HOA coefficient sample block (ϑ_0) and the transformation to spatial domain is not needed (Fig.4a). The latter is identical to analyzing the downmix signal of $\boldsymbol{W}$. Further details are given below.

[0021] In Fig.4, creation of DRC gains for HOA is shown. In Fig.4a) is depicted how a single gain (for a single gain group) can be derived from the zeroth HOA order component ϑ_0 (optional with side chaining from AOs). In Fig.4b) is depicted how two or more DRC gains are created by transforming the HOA representation into a spatial domain. As an example, sounds from the back (e.g. background sound) might get more attenuation than sounds originating from front and side directions. This would lead to $(N + 1)^2$ gain values in $\boldsymbol{g}$ which could be transmitted within two gain groups for this example. Optional, here also is side chaining by Audio Objects wave forms and their directional information. Distracting sounds in the HOA mix sharing the same spatial source areas with the AO foreground sounds can get stronger attenuation gains than spatially distant sounds.

[0022] The gain values are transmitted to a receiver or decoder side.

A variable number of 1 to $L_L = (N + 1)^2$ gain values related to a block of τ samples is transmitted. Gain values can be assigned to channel groups for transmission. In an embodiment, all equal gains are combined in one channel group to minimize transmission data. If a single gain is transmitted, it is related to all L_L channels. Transmitted are the number of channel groups gain values g_{l_g}. Channel groups are signaled.

[0023] The gain values are applied as follows.

The receiver/decoder can determine the number of transmitted gain values, decode related information and assign the gains to $L_L = (N + 1)^2$ channels.

If only one gain value (one channel group) is transmitted, it can be directly applied to the *H*OA signal ($B_{DRC} = g_1$ $\boldsymbol{B}$), as shown in Fig.5 a). This has an advantage because the decoding is much simpler and requires considerably less processing. The reason is that no matrix operations are required; instead, the gain values can be applied directly, e.g. multiplied with the HOA coefficients. For further details see below.

If two or more gains are transmitted, the channel group gains are assigned to L channel gains $\boldsymbol{g} = [g_1, ..., g_L]$.

[0024] For the virtual regular loudspeaker grid, the loudspeaker signals with the DRC gains applied are computed by

$$\widehat{\boldsymbol{W}}_L = diag(\boldsymbol{g}) \cdot \boldsymbol{W}_L.$$

The resulting modified HOA representation is then computed by

$$\boldsymbol{B}_{DRC} = \boldsymbol{D}_L^{-1}\widehat{\boldsymbol{W}}_L.$$

[0025] This can be simplified, as shown in Fig.5 b). Instead of transforming the HOA signal into the spatial domain, applying the gains and transforming the result back to the HOA domain, the gain vector is transformed to the HOA domain by:

$$\boldsymbol{G} = \boldsymbol{D}_L^{-1}\, diag(\boldsymbol{g})\, \boldsymbol{D}_L,$$

with $\boldsymbol{G} \epsilon \mathbb{R}^{(N+1)^2 \times (N+1)^2}$. The gain matrix is applied directly to the HOA coefficients:

$$\boldsymbol{B}_{DRC} = \boldsymbol{GB}.$$

This is more efficient in terms of computational operations needed for $(N + 1)^2 < \tau$. That is, this solution has an advantage because the decoding is much simpler and requires considerably less processing. The reason is that no matrix operations are required; instead, the gain values can be applied directly, e.g. multiplied with the HOA coefficients.

[0026] An even more efficient way of applying the gain matrix is to manipulate the Renderer matrix by $\hat{D} = DG$, apply the DRC and render in one step: $\mathbf{W} = \hat{\mathbf{D}}\mathbf{B}$. This is shown in Fig.5 c). This is beneficial if $L < \tau$.

[0027] In Fig.5, applying DRC to HOA signals is shown. In Fig.5 a), a single channel group gain is transmitted and decoded and applied directly onto the HOA coefficients. In Fig. 5 b), more than one channel group gains are transmitted, decoded and a gain vector $\boldsymbol{g}$ of $(N + 1)^2$ gain values is decoded. A gain matrix $\boldsymbol{G}$ is created and applied to a block of HOA samples. In Fig. 5 c), instead of applying the gain matrix / gain value to the HOA signal directly, it is applied directly onto the renderer's matrix. This is computationally beneficial if the DRC block size τ is larger than the number of output channels L.

[0028] In the following, calculation of ideal DSHT (Discrete Spherical Harmonics Transform) matrices for DRC is described.

The requirements for the ideal rendering and encoding matrices $\mathbf{D}_L$ and $\boldsymbol{D}_L^{-1}$ related to an ideal spherical layout are derived below. Even for ideal rendering layouts, requirement 2 and 3 seem to be in contradiction to each other. Either one or the other can be fulfilled without error, but with errors exceeding 3dB for the other one. This is considered to lead to audible artifacts. A method to overcome this is described in the following.

[0029] First, an ideal spherical layout with $L = (N + 1)^2$ is selected. The L directions of the (virtual) speaker positions are given by Ω_l and the related mode matrix is denoted as $\psi_L = [\varphi(\Omega_1), ..., \varphi(\Omega_l), \varphi(\Omega_L]$. Each $\varphi(\Omega_l)$ is a mode vector containing the spherical harmonics of the direction Ω_l. L Quadrature gains related to the spherical layout positions are assembled in vector $\boldsymbol{q}$. These quadrature gains rate the spherical area of such a position and all sum up to a value of 4π related to the surface of a sphere with radius one. A first prototype rendering matrix $\tilde{D}_L$ is derived by

$$\widetilde{\boldsymbol{D}}_L = diag(\boldsymbol{q})\frac{\boldsymbol{\Psi_L}}{L}.$$

Note that the division by L can be omitted due to a later normalization step (see below).

[0030] Second, a compact singular value decomposition is performed: $\widetilde{\boldsymbol{D}}_L = \boldsymbol{USV}^T$ and a second prototype matrix is derived by

$$\widehat{\widetilde{\boldsymbol{D}}}_L = \boldsymbol{UV}^T.$$

[0031] Third, the prototype matrix is normalized:

$$\breve{\boldsymbol{D}}_L = \frac{\widehat{\widehat{\boldsymbol{D}}}_L}{\left\|\widehat{\widehat{\boldsymbol{D}}}_L\right\|_k},$$

where k denotes the matrix norm type. Two matrix norm types show equally good performance. Either the k = 1 norm or the Frobenius norm should be used. This matrix fulfills the requirement 3 (energy preservation).

[0032] Fourth, in the last step the Amplitude error to fulfill requirement 2 is substituted:

Row-vector e is calculated by $\boldsymbol{e} = -\frac{\mathbf{1}_L^T \breve{\boldsymbol{D}}_L - [1,0,0,..,0]}{\mathrm{L}}$, where [1,0,0,..,0] is a row vector of $(N + 1)^2$ all zero elements except for the first element with a value of one. $\mathbf{1}_L^T \breve{\boldsymbol{D}}_L$ denotes the sum of rows vectors of $\breve{\boldsymbol{D}}_L$. The rendering matrix D_L is now derived by substituting the amplitude error:

$$\boldsymbol{D}_L = \breve{\boldsymbol{D}}_L + [\,\boldsymbol{e}^T, \boldsymbol{e}^T, \boldsymbol{e}^T, ..]^T,$$

where vector e is added to every row of $\breve{\boldsymbol{D}}_L$. This matrix fulfills requirement 2 and requirement 3. The first row elements of $\boldsymbol{D}_L^{-1}$ all become one.

[0033] In the following, detailed requirements for DRC are explained.

First, L_L identical gains with a value of g_1 applied in spatial domain is equal to apply the gain g_1 to the HOA coefficients:

$$\boldsymbol{D}_L^{-1}\, \boldsymbol{g}\, \boldsymbol{W}_L = \boldsymbol{D}_L^{-1}\, g_1 \boldsymbol{I}\, \boldsymbol{D}_L\, \boldsymbol{B} = g_1 \boldsymbol{D}_L^{-1}\, \boldsymbol{D}_L\, \boldsymbol{B} = g_1 \boldsymbol{B}$$

[0034] This leads to the requirement: $\boldsymbol{D}_L^{-1}\, \boldsymbol{D}_L = \boldsymbol{I}$, which means that $L = (N + 1)^2$ and $\boldsymbol{D}_L^{-1}$ needs to exist (trivial).

[0035] Second, analyzing the sum signal in spatial domain is equal to analyzing the zeroth order HOA component. DRC analyzers use the signals' energy as well as its amplitude. Thus the sum signal is related to amplitude and energy. The signal model of HOA: $\boldsymbol{B} = \boldsymbol{\Psi}_e \boldsymbol{X}_s,\ \boldsymbol{X}_s \in \mathbb{R}^{S \times \tau}$ is a matrix of S directional signals; $\psi_e = [\varphi(\Omega_1), ..., \varphi(\Omega_s), \varphi(\Omega_S)]$ is a N3D mode matrix related to the directions $\Omega_1, .., \Omega_s$. The mode vector $\boldsymbol{\varphi}(\boldsymbol{\Omega}_s) = [Y_0^0(\boldsymbol{\Omega}_s),\ Y_1^{-1}(\boldsymbol{\Omega}_s), ... Y_N^N(\boldsymbol{\Omega}_s)]^T$ is assembled out of Spherical Harmonics. In N3D notation the zeroth order component $Y_0^0(\boldsymbol{\Omega}_s) = 1$ is independent of the direction.

The zeroth order component HOA signal needs to become the sum of the directional signals $\boldsymbol{b}_o = [b_1(1), b_1(2), ..., b_1(T)] = \mathbf{1}_S^T \boldsymbol{X}_s$ to reflect the correct amplitude of the summation signal. $\mathbf{1}_S$ is a vector assembled out of S elements with a value of 1.

The energy of the directional signals is preserved in this mix because $\boldsymbol{b}_o \boldsymbol{b}_o^T = \mathbf{1}_S^T \boldsymbol{X}_s \boldsymbol{X}_s^T\, \mathbf{1}_S$. This would simplify to $\sum_{s=1}^{S} \sum_{t=1}^{\tau} X_{s,t}^2 = \left\|\boldsymbol{X}_s\right\|_{fro}^2$ if the signals X_s are not correlated.

[0036] The sum of amplitudes in spatial domain is given by $\mathbf{1}_L^T \boldsymbol{W}_L = \mathbf{1}_L^T \boldsymbol{D}_L\, \boldsymbol{\Psi}_e\, \boldsymbol{X}_s = \mathbf{1}_L^T\, \boldsymbol{M}_L \boldsymbol{X}_s$ with HOA panning matrix $\boldsymbol{M}_L = \boldsymbol{D}_L\, \psi_e$.

This becomes $\boldsymbol{b}_o = \mathbf{1}_S^T \boldsymbol{X}_s$ for $\mathbf{1}_L^T\, \boldsymbol{M}_L = \mathbf{1}_L^T \boldsymbol{D}_L\, \boldsymbol{\Psi}_e = \mathbf{1}_S^T$. The latter requirement can be compared to the sum of amplitudes requirement sometimes used in panning like VBAP. Empirically it can be seen that this can be achieved in good approximation for very symmetric spherical speaker setups with $\boldsymbol{D}_L = \boldsymbol{\Psi}_e^{-1}$, because there we find: $\mathbf{1}_L^T \boldsymbol{D}_L \approx [1,0,0,..,0] \implies \mathbf{1}_L^T \boldsymbol{D}_L\, \boldsymbol{\Psi}_e \approx [Y_0^0(\boldsymbol{\Omega}_1), ... Y_0^0(\boldsymbol{\Omega}_s)] = \mathbf{1}_S^T$. The Amplitude requirement can then be reached within necessary accuracy.

This also ensures that the energy requirement for the sum signal can be met:

The energy sum in spatial domain is given by: $\mathbf{1}_L^T \boldsymbol{W}_L \, \boldsymbol{W}_L^T \, \mathbf{1}_L = \ \mathbf{1}_L^T \, \boldsymbol{M}_L \boldsymbol{X}_s \, \boldsymbol{X}_s^T \, \boldsymbol{M}_L \mathbf{1}_L$ which would become in good approximation $\mathbf{1}_S^T \boldsymbol{X}_s \boldsymbol{X}_s^T \, \mathbf{1}_S$, the existence of an ideal symmetric speaker setup required.

This leads to the requirement: $\mathbf{1}_L^T \boldsymbol{D}_L \simeq [1{,}0{,}0{,}\ldots{,}0]$ and in addition from the signal model we can conclude that the top row of $\boldsymbol{D}_L^{-1}$ needs to be [1,1,1,1,..], i.e. a vector of length L with "one" elements) in order that the re-encoded order zero signal maintains amplitude and energy.

[0037] Third, energy preservation is a prerequisite: The energy of signal $\boldsymbol{x}_s \in \mathbb{R}^{1 \boldsymbol{x} \tau}$ should be preserved after conversion to HOA and spatial rendering to loud speakers independent of the signal's direction Ω_S. This leads to $||\boldsymbol{D}_L \, \boldsymbol{\varphi}(\boldsymbol{\Omega}_S)||_2^2 = 1.$ This can be achieved by modelling $\boldsymbol{D}_L$ from rotation matrices and a diagonal gain matrix: $\boldsymbol{D}_L = UV^T \, diag(a)$ (the dependency on the direction (Ω_s) was removed for clarity): $||\boldsymbol{D}_L \, \boldsymbol{\varphi}||_2^2 = \boldsymbol{\varphi}^T \boldsymbol{D}_L^T \boldsymbol{D}_L \boldsymbol{\varphi} =$

$$\boldsymbol{\varphi}^T diag(\boldsymbol{a}) \boldsymbol{V} \boldsymbol{U}^T \boldsymbol{U} \boldsymbol{V}^T diag(\boldsymbol{a}) \boldsymbol{\varphi} = \ \boldsymbol{\varphi}^T diag(\boldsymbol{a})^2 \boldsymbol{\varphi} = \ \sum_{\boldsymbol{o}=\mathbf{1}}^{(N+1)^2} a_o^2 \, \varphi_o^2 \equiv 1$$

For Spherical harmonics $\varphi_o^2 = Y_n^{m^2}(\boldsymbol{\Omega}_s) \ = 1,$ so all gains a_o^2 related to $||\boldsymbol{D}_L||_{fro}^2 = \sum_{o=1}^{(N+1)^2} a_o^2 = 1$ would satisfy the equation. If all gains are selected equal, this leads to $a_o^2 = (N+1)^{-2}$.

The requirement $VV^T = 1$ can be achieved for $L \geq (N+1)^2$ and only be approximated for $L < (N+1)^2$.)

[0038] This leads to the requirement: $\boldsymbol{D}_L^T \, \boldsymbol{D}_L = \ diag(\boldsymbol{a})^2$ with $\sum_{o=1}^{(N+1)^2} a_o^2 = 1.$

[0039] As an example, a case with ideal spherical positions (for HOA orders N=1 to N=3) is described in the following (Tabs.1-3). Ideal spherical positions for further HOA orders (N=4 to N=6) are described further below (Tabs.4-6). All the below-mentioned positions are derived from modified positions published by Jörg Fliege in "Integration nodes for the sphere", http://www.mathematik.uni-dortmund.de/lsx/research/projects/fliege/nodes/-nodes.html, 2010. Online, accessed 2010-10-05. The method to derive these positions and related quadrature/cubature gains was published in Jörg Fliege and Ulrike Maier. "A two-stage approach for computing cubature formulae for the sphere", Technical report, Fachbereich Mathematik, Universität Dortmund, 1999. In these tables, the azimuth is measured counter-clockwise from frontal direction related to the listening position and the inclination is measured from the z-axis with an inclination of 0 being above the listening position.

[0040] The term numerical quadrature is often abbreviated to quadrature and is quite a synonym for *numerical integration,* especially as applied to 1-dimensional integrals. Numerical integration over more than one dimension is called cubature herein.

N=1 Positions

[0041]

Spherical position Ω_l		q_l
Inclination θ / rad	Azimuth ϕ / rad	Quadrature gains
0.33983655	3.14159265	3.14159271
1.57079667	0.00000000	3.14159267
2.06167886	1.95839324	3.14159262
2.06167892	-1.95839316	3.14159262

a)

$\boldsymbol{D}_L$:

0.2500 -0.0000 0.4082 -0.1443
0.2500 0.0000 -0.0000 0.4330

0.2500 0.3536 -0.2041 -0.1443
0.2500 -0.3536 -0.2041 -0.1443

b)
Tab.1: a) Spherical positions of virtual loudspeakers for HOA order N=1, and b) resulting rendering matrix for spatial transform (DSHT)

N=2 Positions

[0042]

Spherical position Ω_l		q_l
Inclination θ / rad	Azimuth ϕ / rad	Quadrature gains
1.57079633	0.00000000	1.41002219
2.35131567	3.14159265	1.36874571
1.21127801	-1.18149779	1.36874584
1.21127606	1.18149755	1.36874598
1.31812905	-2.45289512	1.41002213
0.00975782	-0.00009218	1.41002214
1.31812792	2.45289621	1.41002230
2.41880319	1.19514740	1.41002223
2.41880555	-1.19514441	1.41002209

a)
D_L:

0.1117	0.0000	0.0067	0.2001	0.0000	-0.0000	-0.0931	-0.0078	0.2235
0.1099	-0.0000	-0.1237	-0.1249	-0.0000	0.0000	0.0486	0.2399	0.0889
0.1099	-0.1523	0.0619	0.0625	-0.1278	-0.1266	-0.0850	0.0841	-0.1455
0.1099	0.1523	0.0619	0.0625	0.1278	0.1266	-0.0850	0.0841	-0.1455
0.1117	-0.1272	0.0450	-0.1479	0.1938	-0.0427	-0.0898	-0.1001	0.0350
0.1117	-0.0000	0.2001	0.0086	0.0000	-0.0000	0.2402	-0.0040	0.0310
0.1117	0.1272	0.0450	-0.1479	-0.1938	0.0427	-0.0898	-0.1001	0.0350
0.1117	0.1272	-0.1484	0.0436	0.0408	-0.1942	0.0769	-0.0982	-0.0612
0.1117	-0.1272	-0.1484	0.0436	-0.0408	0.1942	0.0769	-0.0982	-0.0612

b)
Tab.2: a) Spherical positions of virtual loudspeakers for HOA order N=2 and b) resulting rendering matrix for spatial transform (DSHT)

N=3 Positions

[0043]

Spherical position Ω_l		q_l
Inclination θ / rad	Azimuth ϕ / rad	Quadrature gains
0.49220083	0.00000000	0.75567412
1.12054210	-0.87303924	0.75567398

(continued)

Spherical position Ω_l		q_l
Inclination θ / rad	Azimuth ϕ / rad	Quadrature gains
2.52370429	-0.05517088	0.75567401
2.49233024	-2.15479457	0.87457076
1.57082248	0.00000000	0.87457075
2.02713647	1.01643753	0.75567388
1.61486095	-2.60674413	0.75567396
2.02713675	-1.01643766	0.75567398
1.08936018	2.89490077	0.75567412
1.18114721	0.89523032	0.75567399
0.65554353	1.89029902	0.75567382
1.60934762	1.91089719	0.87457082
2.68498672	2.02012831	0.75567392
1.46575084	-1.76455426	0.75567402
0.58248614	-2.22170415	0.87457060
2.00306837	2.81329239	0.75567389

a)
$\boldsymbol{D}_L$:

0.061457	-0.000075	0.093499	0.050400	-0.000027	0.000060	0.091035	0.098988	0.026750	0.019405	0.001461	0.003133	0.065741	0.124248	0.086602	0.029345
0.061457	-0.073257	0.046432	0.061316	-0.094748	-0.071487	-0.029426	0.059688	-0.016892	-0.055360	-0.097812	-0.010980	-0.082425	-0.007027	-0.048502	-0.080998
0.061457	-0.003584	-0.086661	0.061312	-0.004319	0.006362	0.068273	-0.111895	0.039506	0.008330	0.001142	-0.027428	-0.044323	0.125349	-0.097700	0.021534
0.065628	-0.057573	-0.090918	-0.038050	0.042921	0.102558	0.066570	0.067780	-0.018289	0.008866	-0.087449	-0.104655	-0.011720	-0.061567	0.025778	0.023749
0.065628	-0.000000	-0.000003	0.114142	-0.000000	0.000000	-0.073690	-0.000007	0.127634	0.002742	0.000000	0.010620	0.012464	-0.093807	0.009642	0.121106
0.061457	0.081011	-0.046687	0.050396	0.085735	-0.079893	-0.028706	-0.049469	-0.042390	0.016897	-0.101358	0.003784	0.101201	-0.012537	0.040833	-0.076613
0.061457	-0.054202	-0.004471	-0.091238	0.104013	0.005102	-0.068089	0.008829	0.056943	-0.149185	0.004553	0.050065	0.007556	0.060425	-0.003395	-0.002394
0.061457	-0.080936	-0.046816	0.050396	-0.085707	0.079834	-0.028795	-0.049516	-0.042442	-0.030388	0.099898	0.015986	0.082103	-0.014540	0.065488	-0.078162
0.061457	0.023227	0.049179	-0.091237	-0.044356	0.023858	-0.024641	-0.094498	0.082023	0.072649	-0.042376	-0.007211	-0.082403	0.008618	0.112746	-0.042512
0.061457	0.076842	0.040224	0.061316	0.099067	0.065125	-0.038969	0.052207	-0.022402	0.028674	0.096668	-0.032684	-0.098253	-0.008594	-0.028068	-0.082210
0.061457	0.061293	0.084298	-0.020472	-0.026210	0.108838	0.060891	-0.036183	-0.035381	-0.026726	-0.058661	0.111083	0.035312	-0.053574	-0.087737	0.014123
0.065628	0.107524	-0.004399	-0.038047	-0.080156	-0.009268	-0.073361	0.003280	-0.099081	-0.064714	0.014164	-0.085660	-0.004839	0.038775	0.016889	0.101473
0.061457	0.042357	-0.095230	-0.020477	-0.018235	-0.084766	0.096995	0.040799	-0.014532	-0.025100	0.058531	0.110659	-0.076710	-0.053780	0.056883	0.013978
0.061457	-0.103651	0.010933	-0.020474	0.044445	-0.024073	-0.066259	-0.004608	-0.108789	0.127480	0.000140	0.071265	-0.019816	0.026559	-0.016573	0.076201
0.065628	-0.049951	0.095320	-0.038045	0.037235	-0.093290	0.080481	-0.071053	-0.010264	-0.018490	0.073275	-0.097597	0.032029	-0.080959	-0.030699	0.008722
0.061457	0.030975	-0.044701	-0.091239	-0.059658	-0.028961	-0.032307	0.085658	0.077606	0.084920	0.037824	-0.010382	0.084083	0.002412	-0.102187	-0.047341

b)

Tab.3: a) Spherical positions of virtual loudspeakers for HOA order N=3 and b) resulting rendering matrix for spatial transform (DSHT)

[0044] Typical application scenarios to apply DRC gains to HOA signals are shown in Fig.5. For mixed content applications like HOA plus Audio Objects, DRC gain application can be realized in at least two ways for flexible rendering. Fig.6 shows exemplarily Dynamic Range Compression (DRC) processing at the decoder side. In Fig.6 a), DRC is applied before rendering and mixing. In Fig.6 b), DRC is applied to the loudspeaker signals, i.e. after rendering and mixing.

[0045] In Fig.6a), DRC gains are applied to Audio Objects and HOA separately: DRC gains are applied to Audio Objects in an Audio Object DRC block 610, and DRC gains are applied to HOA in a HOA DRC block 615. Here the realization of the block HOA DRC block 615 matches one of those in Fig.5. In Fig.6b), a single gain is applied to all channels of the mixture signal of the rendered HOA and rendered Audio Object signal. Here no spatial emphasis and attenuation is possible. The related DRC gain cannot be created by analyzing the sum signal of the rendered mix, because the speaker layout of the consumer site is not known at the time of creation at the broadcast or content creation site. The DRC gain can be derived analyzing $\boldsymbol{y}_m \epsilon \mathbb{R}^{1\mathrm{x}\,\tau}$ where y_m is mix of the zeroth order HOA signal $\boldsymbol{b}_{\mathrm{w}}$ and the mono downmix of S Audio Objects:

$$\boldsymbol{y}_{\mathrm{m}} = \boldsymbol{b}_o + \sum_{\mathrm{s}=1}^{\mathrm{S}} \boldsymbol{x}_{\mathrm{s}} \,.$$

[0046] In the following, further details of the disclosed solution are described.

DRC for HOA Content

[0047] DRC is applied to the HOA signal before rendering and may be combined with rendering. DRC for HOA can be applied in time domain or in QMF-filter bank domain.

DRC in Time Domain

[0048] The DRC decoder shall provide $(\mathrm{N} + 1)^2$ gain values $\boldsymbol{g}_{drc} = \left[g_1, \ldots, g_{(N+1)^2}\right]^T$ according to the number of HOA coefficient channels of the HOA signal c. N is the HOA order. Application of DRC gains to the HOA signals:

$$\boldsymbol{c_{drc}} = \boldsymbol{D}_{\boldsymbol{L}}^{\boldsymbol{-1}} diag(\boldsymbol{g}_{drc}) \boldsymbol{D}_{DSHT}\, \boldsymbol{c}$$

where c is a vector of one time sample of HOA coefficients $(\boldsymbol{c} \,\epsilon\, \mathbb{R}^{(N+1)^2 \mathrm{x}\, 1})$, and $\boldsymbol{D}_{DSHT} \,\epsilon\, \mathbb{R}^{(N+1)^2 \mathrm{x}\, (N+1)^2}$ and its inverse $\boldsymbol{D}_{DSHT}^{\boldsymbol{-1}}$ are matrices related to a Discrete Spherical Harmonics Transform (DSHT) optimized for DRC purposes.

In one embodiment, to decrease the computational load by $(N + 1)^4$ operations per sample, it can be advantageous to include the rendering step and calculate the loudspeaker signals directly by: $\boldsymbol{w_{drc}} = \left(\boldsymbol{D}\, \boldsymbol{D_{DSHT}^{-1}}\right)\ \left(diag(\boldsymbol{g}_{drc}) \boldsymbol{D}_{DSHT}\right) \boldsymbol{c}$, where D is the rendering matrix and $\left(\boldsymbol{D}\, \boldsymbol{D}_{\boldsymbol{L}}^{\boldsymbol{-1}}\right)$ can be precomputed.

If all gains $\mathbf{g}_1, \ldots, \mathbf{g}_{(\mathrm{N}+1)2}$ have the same value of $\mathbf{g}_{\mathbf{drc}}$, as in the simplified mode, a single gain group has been used to transmit the coder DRC gains. This case can be flagged by the DRC decoder, because in this case the calculation in the spatial filter is not needed, so that the calculation simplifies to:

$$\mathbf{c}_{\mathrm{drc}} = \mathrm{g}_{\mathrm{drc}}\, \mathbf{c}.$$

Calculation of DSHT matrices for DRC

[0049] In the following, $\boldsymbol{D}_{\mathrm{L}}$ is renamed to $\boldsymbol{D}_{\mathrm{DSHT}}$. The matrices to determine the spatial filter $\boldsymbol{D}_{\mathrm{DSHT}}$ and its inverse $\boldsymbol{D}_{DSHT}^{-1}$ are calculated as follows:

A set of spherical positions $\boldsymbol{\mathfrak{D}}_{DSHT} = [\,\boldsymbol{\Omega}_1, \boldsymbol{\Omega}_{\mathrm{l}}, \ldots, \boldsymbol{\Omega}_{(\mathrm{N}+1)^2}]$ with $\boldsymbol{\Omega}_{\mathrm{l}} = [\theta_l, \phi_l]^T$ and related quadrature (cubature) gains $\boldsymbol{q} \in \mathbb{R}^{(N+1)^2 \mathrm{x}\, 1}$ are selected indexed by the HOA order N from Tables 1-4. A mode matrix ψ_{DSHT} related to these positions is calculated (see above). A first prototype matrix is calculated by $\widetilde{\boldsymbol{D}}_1 = diag(\boldsymbol{q}) \frac{\boldsymbol{\Psi}_{DSHT}}{(N+1)^2}$ (the division by $(N+1)^2$ can be skipped due to a subsequent normalization). A compact singular value decomposition is performed $\tilde{D}_1 = USV^T$ and a new prototype matrix is calculated by: $\hat{\widetilde{\boldsymbol{D}}}_2 = \boldsymbol{U}\boldsymbol{V}^T$. This matrix is normalized by: $\breve{\boldsymbol{D}}_2 = \frac{\hat{\tilde{\boldsymbol{D}}}_2}{\left\|\hat{\tilde{\boldsymbol{D}}}_2\right\|_{fro}}$. A row-vector e is calculated by $\boldsymbol{e} = -\frac{\mathbf{1}_L^T \breve{\boldsymbol{D}}_2 - [1,0,0,..,0]}{(\mathrm{N}+1)^2}$, where [1,0,0,..,0] is a row vector of $(N+1)^2$ all zero elements except for the first element with a value of one. $\mathbf{1}_L^T \breve{\boldsymbol{D}}_2$ denotes the sum of rows of $\breve{\boldsymbol{D}}_2$. The optimized DSHT matrix $\boldsymbol{D}_{DSHT}$ is now derived by: $\boldsymbol{D}_{\mathrm{DSHT}}$. It has been found that, if erroneously $-\boldsymbol{e}$ is used instead of e, the invention provides slightly worse but still usable results.

[0050] For DRC in QMF-filter bank domain, the following applies.

[0051] The DRC decoder provides a gain value $g_{ch}(n, m)$ for every time frequency tile n, m for $(N + 1)^2$ spatial channels. The gains for time slot n and frequency band m are arranged in $\boldsymbol{g}(n,m) \in \mathbb{R}^{(N+1)^2 \mathrm{x}\, 1}$.

Multiband DRC is applied in QMF Filter bank domain. The processing steps are shown in Fig.7. The reconstructed HOA signal is transformed into spatial domain by (inverse DSHT): $\boldsymbol{W}_{DSHT} = \boldsymbol{D}_{DSHT}\boldsymbol{C}$, where $\boldsymbol{C} \in \mathbb{R}^{(N+1)^2 \mathrm{x}\, \tau}$ is a block of T HOA samples and $\boldsymbol{W}_{DSHT} \in \mathbb{R}^{(N+1)^2 \mathrm{x}\, \tau}$ is a block of spatial samples matching the input time granularity of the QMF filter bank. Then the QMF analysis filter bank is applied. Let $\hat{\boldsymbol{w}}_{DSHT}(n,m) \in \mathbb{C}^{(N+1)^2 \mathrm{x}\, 1}$ denote the a vector of spatial channels per time frequency tile (*n*, m). Then the DRC gains are applied: $\breve{\boldsymbol{w}}_{DRC}(n,m) = diag(\boldsymbol{g}(n,m))\, \hat{\boldsymbol{w}}_{DSHT}(n,m)$.

To minimize the computational complexity, the DSHT and rendering to loudspeaker channels are combined: $\boldsymbol{w}(n,m) = \boldsymbol{D}\, \boldsymbol{D}_{DSHT}^{-1}\, \breve{\boldsymbol{w}}_{DRC}(n,m)$, where **D** denotes the HOA rendering matrix. The QMF signals then can be fed to the mixer for further processing.

Fig.7 shows DRC for HOA in the QMF domain combined with a rendering step.

If only a single gain group for DRC has been used this should be flagged by the DRC decoder because again computational simplifications are possible. In this case the gains in vector g(n, m) all share the same value of $g_{DRC}(n, m)$. The QMF filter bank can be directly applied to the HOA signal and the gain $g_{DRC}(n, m)$ can be multiplied in filter bank domain.

[0052] Fig.8 shows DRC for HOA in the QMF domain (a filter domain of a Quadrature Mirror Filter) combined with a rendering step, with computational simplifications for the simple case of a single DRC gain group.

[0053] As has become apparent in view of the above, in one embodiment the invention relates to a method for performing DRC on a HOA signal, the method comprising steps of setting or determining a mode, the mode being either a simplified mode or a non-simplified mode, in the non-simplified mode, transforming the HOA signal to the spatial domain, wherein an inverse DSHT is used, in the non-simplified mode, analyzing the transformed HOA signal, and in the simplified mode, analyzing the HOA signal, obtaining, from results of said analyzing, one or more gain factors that are usable for dynamic range compression, wherein only one gain factor is obtained in the simplified mode and wherein two or more different gain factors are obtained in the non-simplified mode, in the simplified mode multiplying the obtained gain factor with the HOA signal, wherein a gain compressed HOA signal is obtained, in the non-simplified mode, multiplying the obtained gain factors with the transformed HOA signal, wherein a gain compressed transformed HOA signal is obtained, and transforming the gain compressed transformed HOA signal back into the HOA domain, wherein a gain compressed HOA signal is obtained.

[0054] In one embodiment, the method further comprises before said multiplying the obtained factors, transmitting the HOA signals together with the obtained gain factor or gain factors.

[0055] In one embodiment, the HOA signal is divided into frequency subbands, and the steps of analysing the HOA signal (or transformed HOA signal), obtaining one or more gain factors, multiplying the obtained gain factor(s) with the HOA signal (or transformed HOA signal), and transforming the gain compressed transformed HOA signal back into the HOA domain are applied to each frequency subband separately, with individual gains per subband. It is noted that the sequential order of dividing the HOA signal into frequency subbands and transforming the HOA signal to the spatial domain can be swapped, and/or the sequential order of synthesizing the subbands and transforming the gain compressed transformed HOA signals back into the HOA domain can be swapped, independently from each other.

[0056] In one embodiment the invention relates to a method for applying DRC gain factors to a HOA signal, the method comprising steps of receiving a HOA signal together with an indication and one or more gain factors, the indication indicating either a simplified mode or a non-simplified mode, wherein only one gain factor is received if the indication indicates the simplified mode, selecting either a simplified mode or a non-simplified mode according to said indication, in the simplified mode multiplying the gain factor with the HOA signal, wherein a dynamic range compressed HOA signal is obtained, and in the non-simplified mode transforming the HOA signal into the spatial domain, wherein a transformed HOA signal is obtained, multiplying the gain factors with the transformed HOA signals, wherein dynamic range compressed transformed HOA signals are obtained, and transforming the dynamic range compressed transformed HOA signals back into the HOA domain, wherein a dynamic range compressed HOA signal is obtained.

[0057] Further, in one embodiment the invention relates to a device for performing DRC on a HOA signal, the device comprising a processor or one or more processing elements adapted for setting or determining a mode, the mode being either a simplified mode or a non-simplified mode, in the non-simplified mode transforming the HOA signal to the spatial domain, wherein an inverse DSHT is used, in the non-simplified mode analyzing the transformed HOA signal, while in the simplified mode analyzing the HOA signal, obtaining, from results of said analyzing, one or more gain factors that are usable for dynamic range compression, wherein only one gain factor is obtained in the simplified mode and wherein two or more different gain factors are obtained in the non-simplified mode, in the simplified mode multiplying the obtained gain factor with the HOA signal, wherein a gain compressed HOA signal is obtained, and in the non-simplified mode multiplying the obtained gain factors with the transformed HOA signal, wherein a gain compressed transformed HOA signal is obtained, and transforming the gain compressed transformed HOA signal back into the HOA domain, wherein a gain compressed HOA signal is obtained.

[0058] In one embodiment for non-simplified mode only, a device for performing DRC on a HOA signal comprises a processor or one or more processing elements adapted for transforming the HOA signal to the spatial domain, analyzing the transformed HOA signal, obtaining, from results of said analyzing, gain factors that are usable for dynamic range compression, multiplying the obtained factors with the transformed HOA signals, wherein gain compressed transformed HOA signals are obtained, and transforming the gain compressed transformed HOA signals back into the HOA domain, wherein gain compressed HOA signals are obtained. In one embodiment, the device further comprises a transmission unit for transmitting, before multiplying the obtained gain factor or gain factors, the HOA signal together with the obtained gain factor or gain factors.

[0059] Also here it is noted that the sequential order of dividing the HOA signal into frequency subbands and transforming the HOA signal to the spatial domain can be swapped, and the sequential order of synthesizing the subbands and transforming the gain compressed transformed HOA signals back into the HOA domain can be swapped, independently from each other.

[0060] Further, in one embodiment the invention relates to a device for applying DRC gain factors to a HOA signal, the device comprising a processor or one or more processing elements adapted for receiving a HOA signal together with an indication and one or more gain factors, the indication indicating either a simplified mode or a non-simplified mode, wherein only one gain factor is received if the indication indicates the simplified mode, setting the device to either a simplified mode or a non-simplified mode, according to said indication, in the simplified mode, multiplying the gain factor with the HOA signal, wherein a dynamic range compressed HOA signal is obtained; and in the non-simplified mode, transforming the HOA signal into the spatial domain, wherein a transformed HOA signal is obtained, multiplying the gain factors with the transformed HOA signals, wherein dynamic range compressed transformed HOA signals are obtained, and transforming the dynamic range compressed transformed HOA signals back into the HOA domain, wherein a dynamic range compressed HOA signal is obtained.

[0061] In one embodiment, the device further comprises a transmission unit for transmitting, before multiplying the obtained factors, the HOA signals together with the obtained gain factors. In one embodiment, the HOA signal is divided into frequency subbands, and the analysing the transformed HOA signal, obtaining gain factors, multiplying the obtained factors with the transformed HOA signals and transforming the gain compressed transformed HOA signals back into the HOA domain are applied to each frequency subband separately, with individual gains per subband.

[0062] In one embodiment of the device for applying DRC gain factors to a HOA signal, the HOA signal is divided into a plurality of frequency subbands, and obtaining one or more gain factors, multiplying the obtained gain factors with the

HOA signals or the transformed HOA signals, and in the non-simplified mode transforming the gain compressed transformed HOA signals back into the HOA domain are applied to each frequency subband separately, with individual gains per subband.

[0063] Further, in one embodiment where only the non-simplified mode is used, the invention relates to a device for applying DRC gain factors to a HOA signal, the device comprising a processor or one or more processing elements adapted for receiving a HOA signal together with gain factors, transforming the HOA signal into the spatial domain (using iDSHT), wherein a transformed HOA signal is obtained, multiplying the gain factors with the transformed HOA signal, wherein a dynamic range compressed transformed HOA signal is obtained, and transforming the dynamic range compressed transformed HOA signal back into the HOA domain (i.e. coefficient domain) (using DSHT), wherein a dynamic range compressed HOA signal is obtained.

[0064] While there has been shown, described, and pointed out fundamental novel features of the present invention as applied to preferred embodiments thereof, it will be understood that various omissions and substitutions and changes in the apparatus and method described, in the form and details of the devices disclosed, and in their operation, may be made by those skilled in the art without departing from the spirit of the present invention. It is expressly intended that all combinations of those elements that perform substantially the same function in substantially the same way to achieve the same results are within the scope of the invention. Substitutions of elements from one described embodiment to another are also fully intended and contemplated.

[0065] The following tables list spherical positions of virtual loudspeakers for HOA of order N with N=4, 5 or 6.

N=4 Positions

[0066]

Tab.4: Spherical positions of virtual loudspeakers for HOA order N=4

Inclination \rad	Azimuth \rad	Gain g_l
1.57079633	0.00000000	0.52689274
2.39401407	0.00000000	0.48518011
1.14059283	-1.75618245	0.52688432
1.33721851	0.69215601	0.47027816
1.72512898	-1.33340585	0.48037442
1.17406779	-0.79850952	0.51130478
0.69042674	1.07623171	0.50662254
1.47478735	1.43953896	0.52158458
1.67073876	2.25235428	0.52835300
2.52745842	-1.33179653	0.52388165
1.81037110	3.05783641	0.49800736
1.91827560	-2.03351312	0.48516540
0.27992161	2.55302196	0.50663531
0.47981675	-1.18580204	0.50824199
2.37644317	2.52383590	0.45807408
0.98508365	2.03459671	0.47260252
2.18924206	1.58232601	0.49801422
1.49441825	-2.58932194	0.51745117
2.04428895	0.76615262	0.51744164
2.43923726	-2.63989327	0.52146074
1.10308418	2.88498471	0.52158484
0.78489181	-2.54224201	0.47027748
2.96802845	1.25258904	0.52145388
1.91816652	-0.63874484	0.48036020
0.80829458	-0.00991977	0.50824345

N=5 Positions

[0067]

Tab.5: Spherical positions of virtual loudspeakers for HOA orders N= 5

Inclination \rad	Azimuth \rad	Gain g_l
1.57079633	0.00000000	0.34493574
2.68749293	3.14159265	0.35131373
1.92461621	-1.22481468	0.35358151
1.95917092	3.06534485	0.36442231
2.18883411	0.08893301	0.36437350
0.35664531	-2.15475973	0.33953855
1.32915731	-1.05408340	0.35358417
2.21829206	2.45308518	0.33534647
1.00903070	2.31872053	0.34739607
0.99455136	-2.29370294	0.36437101
1.13601102	-0.46303195	0.33534542
0.41863640	0.63541391	0.35131934
1.78596913	-0.56826765	0.34739591
0.56658255	-0.66284593	0.36441956
2.25292410	0.89044754	0.36437098
2.67263757	-1.71236120	0.36442208
0.86753981	-1.50749854	0.34068122
1.38158330	1.72190554	0.35358401
0.98578154	0.23428465	0.35131950
1.45079827	-1.69748851	0.34739437
2.09223697	-1.85025366	0.33534659
2.62854417	1.70110685	0.34494256
1.44817433	-2.83400771	0.33953463
2.37827410	-0.72817212	0.34068529
0.82285875	1.51124182	0.33534531
0.40679748	2.38217051	0.34493552
0.84332549	-3.07860398	0.36437337
1.38947809	2.83246237	0.34068522
1.61795773	-2.27837285	0.34494274
2.17389505	-2.58540735	0.35131361
1.65172710	2.28105193	0.35358166
1.67862104	0.57097606	0.33953819
2.02514031	1.70739195	0.34739443
1.12965858	0.89802542	0.36442004
2.82979093	0.17840931	0.33953488
1.67550339	1.18664952	0.34068114

N=6 Positions

[0068]

Tab.6: Spherical positions of virtual loudspeakers for HOA orders N= 6

Inclination \rad	Azimuth \rad	Gain g_l
1.57079633	0.00000000	0.23821170
2.42144792	0.00000000	0.23821175
0.32919895	2.78993083	0.26169552
1.06225899	1.49243160	0.25534085
1.06225899	1.49243160	0.25534085

(continued)

Inclination \rad	Azimuth \rad	Gain g_l
1.01526896	-2.16495206	0.25092628
1.10570423	-1.59180661	0.25099550
1.47319543	1.14258135	0.26160776
2.15414541	1.88359269	0.24442720
0.20805372	-0.52863458	0.25487678
0.50141101	-2.11057110	0.25619096
1.98041218	0.28912378	0.26288225
0.83752075	-2.81667891	0.25837996
2.44130228	0.81495962	0.26772416
1.21539727	-1.00788022	0.25534092
2.62944184	-1.58354086	0.26437874
1.86884674	-2.40686906	0.25619091
0.68705554	-1.20612227	0.25576026
1.52325470	-1.98940871	0.26169551
2.39097364	-2.37336381	0.25576025
0.98667678	0.86446728	0.26014219
2.27078506	-3.06771779	0.25099551
2.33605400	2.51674567	0.26455002
1.29371004	2.03656562	0.25576032
0.86334494	2.77720222	0.25092620
1.94118355	-0.37820559	0.26772409
2.10323413	-1.28283816	0.24442725
1.87416330	0.80785741	0.23821179
1.63423157	1.65277986	0.26437876
2.06477636	1.31341296	0.25595469
0.82305807	-0.47771423	0.26437883
2.04154780	-1.85106655	0.25487677
0.61285067	0.33640173	0.24442716
1.08029340	0.10986230	0.25595472
1.60164764	-1.43535015	0.26455000
2.66513701	1.69643796	0.26014228
1.35887781	-2.58083733	0.25838000
1.78658555	2.25563014	0.25487674
1.83333508	2.80487382	0.26169549
0.78406009	2.08860099	0.25099560
2.94031615	-0.07888534	0.26160780
1.34658213	2.57400947	0.25619094
1.73906669	-0.87744928	0.26014223
0.50210739	1.33550547	0.26455007
2.38040297	-0.75104092	0.25595462
1.41826790	0.54845193	0.26772418
1.77904107	-2.93136138	0.25092628
1.35746628	-0.47759398	0.26160765
1.31545731	3.12752832	0.25838016
2.81487011	-3.12843671	0.25534100

[0069] It will be understood that the present invention has been described purely by way of example, and modifications of detail can be made without departing from the scope of the invention. Each feature disclosed in the description and (where appropriate) the claims and drawings may be provided independently or in any appropriate combination. Features may, where appropriate be implemented in hardware, software, or a combination of the two.

Claims

1. A method for performing DRC on a HOA signal, the method comprising steps of

   - setting or determining a mode, the mode being either simplified mode or non-simplified mode;
   - in the non-simplified mode, transforming the HOA signal to the spatial domain, wherein an inverse DSHT is used;
   - in the non-simplified mode, analyzing the transformed HOA signal, and in the simplified mode, analyzing the HOA signal;
   - obtaining, from results of said analyzing, one or more gain factors that are usable for dynamic range compression, wherein only one gain factor is obtained in the simplified mode and wherein two or more different gain factors are obtained in the non-simplified mode;
   - in the simplified mode multiplying the obtained gain factor with the HOA signal, wherein a gain compressed HOA signal is obtained;
   - in the non-simplified mode, multiplying the obtained gain factors with the transformed HOA signal, wherein a gain compressed transformed HOA signal is obtained, and transforming the gain compressed transformed HOA signal back into the HOA domain, wherein a gain compressed HOA signal is obtained.

2. Method of claim 1, further comprising a step of transmitting the HOA signals together with the obtained gain factors before said step of multiplying the obtained factors.

3. Method according to claim 1 or 2, wherein the HOA signal is divided into frequency subbands, and the steps of analysing the transformed HOA signal, obtaining gain factors, multiplying the obtained factors with the transformed HOA signals and transforming the gain compressed transformed HOA signals back into the HOA domain are applied to each frequency subband separately, with individual gains per subband.

4. Method according to claim 3, wherein the sequential order of dividing the HOA signal into frequency subbands and transforming the HOA signal to the spatial domain can be swapped, and the sequential order of synthesizing the subbands and transforming the gain compressed transformed HOA signals back into the HOA domain can be swapped, independently from each other.

5. A method for applying DRC gain factors to a HOA signal, the method comprising

   - receiving a HOA signal and one or more gain factors;
   - transforming the HOA signal into the spatial domain, wherein a transformed HOA signal is obtained;
   - multiplying the gain factors with the transformed HOA signal, wherein a dynamic range compressed transformed HOA signal is obtained; and
   - transforming the dynamic range compressed transformed HOA signal back into the HOA domain (i.e. coefficient domain) (using DSHT), wherein a dynamic range compressed HOA signal is obtained.

6. Method according to claim 5, wherein also an indication is received, the indication indicating either a simplified mode or a non-simplified mode, and wherein only one gain factor is received if the indication indicates the simplified mode, further comprising a step of selecting either a simplified mode or a non-simplified mode according to said indication, wherein the steps of transforming the HOA signal into the spatial domain and transforming the dynamic range compressed transformed HOA signal back into the HOA domain are performed only in the non-simplified mode, and wherein in the simplified mode the gain factors are multiplied with the HOA signal.

7. Method according to claim 1, 5 or 6, wherein the step of transforming the HOA signal into the spatial domain uses a transform matrix according to at least one of Tab.1 b), Tab 2 b) and Tab 3 b).

8. Method according to claim 1, 5 or 6, wherein in the step of transforming the HOA signal into the spatial domain an iDSHT is used with a transform matrix obtained from the spherical positions of virtual loudspeakers and quadrature gains q.

9. Method according to claim 8, wherein the transform matrix is computed from the mode matrix ψ_{DSHT} and corresponding quadrature gains.

10. Method according to claim 8 or 9, wherein spherical positions and corresponding quadrature gains according to at least one of Tab.1 a), Tab.2a), Tab.3a) and Tab.4-6 are used.

11. Method according to claim 8 or 9, wherein the transform matrix is computed according to $\boldsymbol{D}_{DSHT} = \breve{\boldsymbol{D}}_2 + [\,\boldsymbol{e}^T, \boldsymbol{e}^T, \boldsymbol{e}^T, ..]^T$ wherein $\breve{\boldsymbol{D}}_2 = \frac{\hat{\hat{\boldsymbol{D}}}_2}{\left\|\hat{\hat{\boldsymbol{D}}}_2\right\|_{fro}}$ is a normalized version of $\hat{\hat{\boldsymbol{D}}}_2 = \boldsymbol{U}\boldsymbol{V}^T$ with U,V obtained from $\widetilde{\boldsymbol{D}}_1 = \boldsymbol{USV}^T = diag(\boldsymbol{q}) \frac{\boldsymbol{\Psi}_{DSHT}}{(N+1)^2}$ with ψ_{DSHT} being the mode matrix of used spherical positions of virtual loudspeaker, and e^T is a transposed version of $\boldsymbol{e} = -\frac{\boldsymbol{1}_L^T \breve{\boldsymbol{D}}_2 - [1,0,0,..,0]}{(\mathrm{N}+1)^2}$.

12. Device for performing DRC on a HOA signal, the device comprising a processor or one or more processing elements adapted for

- setting or determining a mode, the mode being either a simplified mode or a non-simplified mode;
- in the non-simplified mode, transforming the HOA signal to the spatial domain, wherein an inverse DSHT is used;
- in the non-simplified mode analyzing the transformed HOA signal, while in the simplified mode analyzing the HOA signal;
- obtaining, from results of said analyzing, one or more gain factors that are usable for dynamic range compression, wherein only one gain factor is obtained in the simplified mode and wherein two or more different gain factors are obtained in the non-simplified mode;
- in the simplified mode, multiplying the obtained gain factor with the HOA signal, wherein a gain compressed HOA signal is obtained; and
- in the non-simplified mode, multiplying the obtained gain factors with the transformed HOA signal, wherein a gain compressed transformed HOA signal is obtained, and transforming the gain compressed transformed HOA signal back into the HOA domain, wherein a gain compressed HOA signal is obtained.

13. Device according to claim 12, further comprising a transmission unit for transmitting, before multiplying the obtained gain factor or gain factors, the HOA signal together with the obtained gain factor or gain factors.

14. Device according to claim 12 or 13, wherein the HOA signal is divided into a plurality of frequency subbands, and the analysing the HOA signal or the transformed HOA signal, obtaining one or more gain factors, multiplying the obtained gain factors with the HOA signal or the transformed HOA signal and, in the non-simplified mode, the transforming the gain compressed transformed HOA signal back into the HOA domain are applied to each frequency subband separately, with individual gains per subband.

15. Device for applying DRC gain factors to a HOA signal, the device comprising a processor or one or more processing elements adapted for

- receiving a HOA signal together with an indication and one or more gain factors, the indication indicating either a simplified mode or a non-simplified mode, wherein only one gain factor is received if the indication indicates the simplified mode;
- setting the device to either a simplified mode or a non-simplified mode, according to said indication;
- in the simplified mode, multiplying the gain factor with the HOA signal, wherein a dynamic range compressed HOA signal is obtained; and
- in the non-simplified mode, transforming the HOA signal into the spatial domain, wherein a transformed HOA signal is obtained, multiplying the gain factors with the transformed HOA signals, wherein dynamic range compressed transformed HOA signals are obtained, and transforming the dynamic range compressed transformed HOA signals back into the HOA domain, wherein a dynamic range compressed HOA signal is obtained.

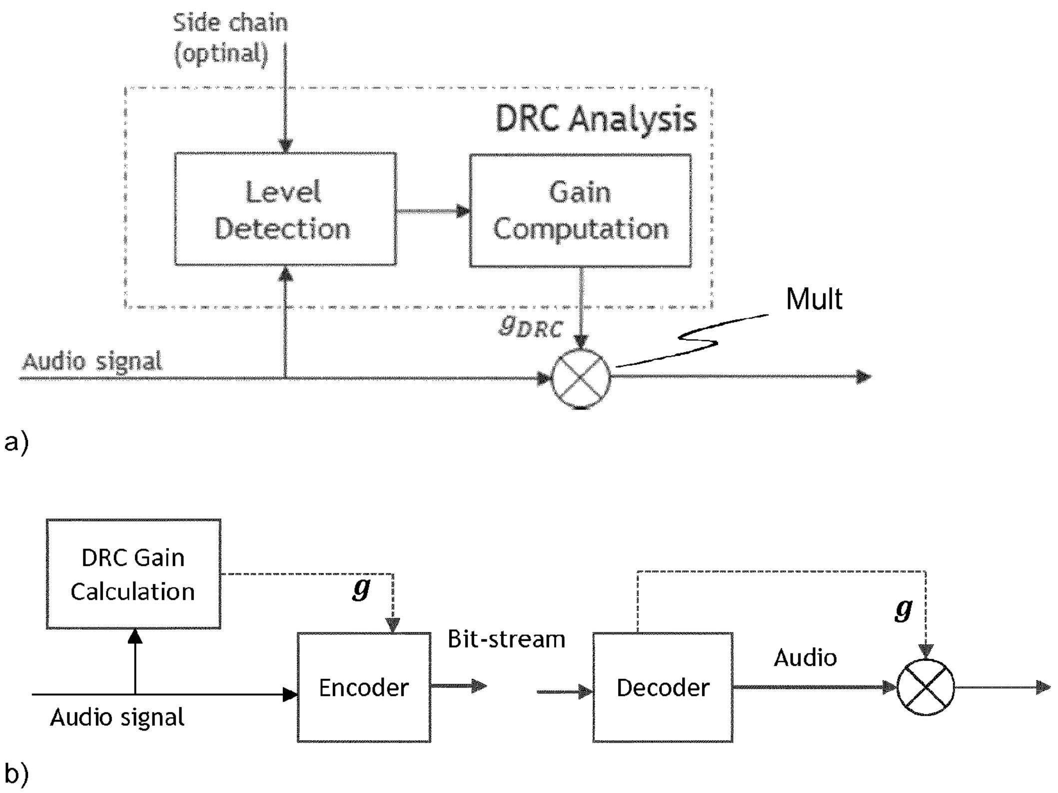
Side chain
(optinal)
DRC Analysis
Level
Detection
Gain
Computation
Mult
g_{DRC}
Audio signal
a)
DRC Gain
Calculation
g
g
Bit-stream
Audio
Encoder
Decoder
Audio signal
b)

Fig.1

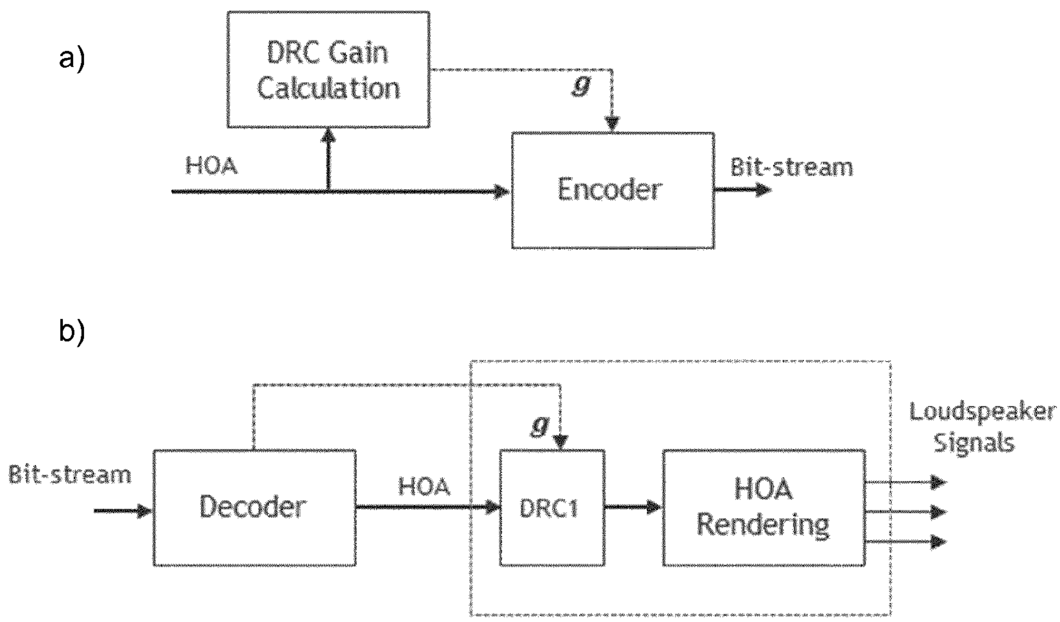
a)
DRC Gain
Calculation
g
HOA
Encoder
Bit-stream
b)
g
Loudspeaker
Signals
Bit-stream
HOA
Decoder
DRC1
HOA
Rendering

Fig.2

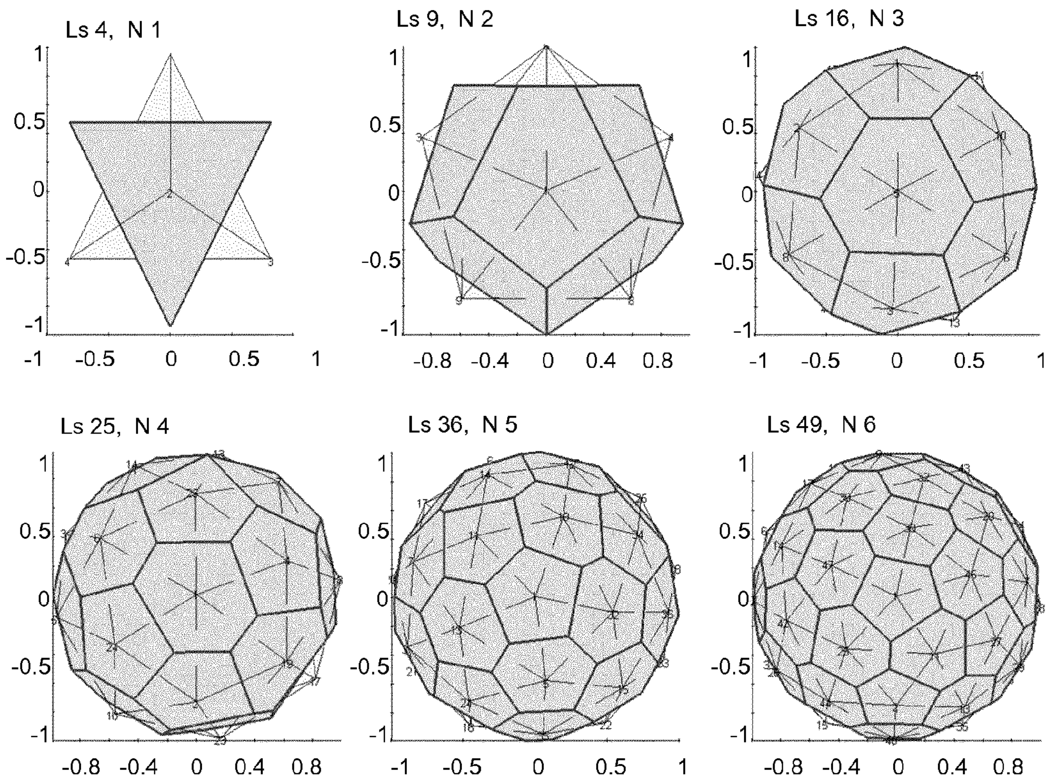
Ls 4, N 1
Ls 9, N 2
Ls 16, N 3
Ls 25, N 4
Ls 36, N 5
Ls 49, N 6

Fig.3

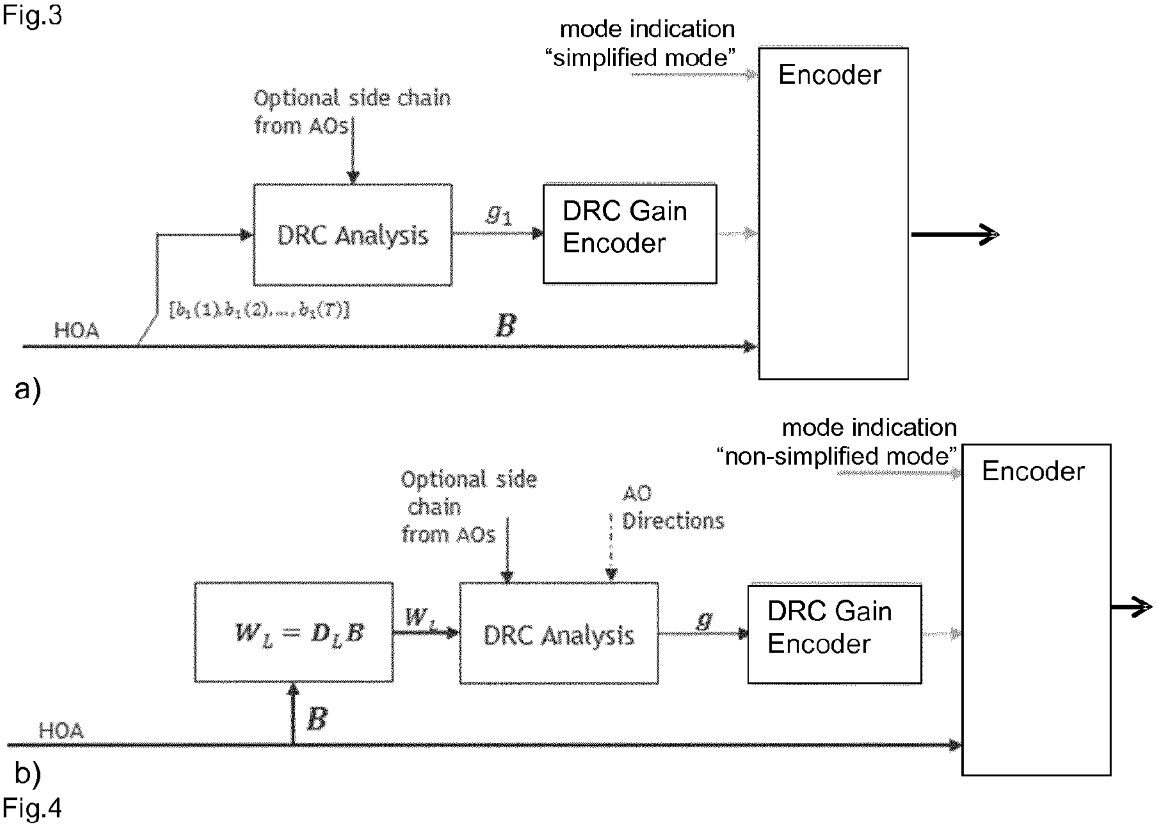
mode indication
"simplified mode"
Encoder
Optional side chain
from AOs
DRC Analysis
g_1
DRC Gain
Encoder
HOA
$[b_1(1), b_1(2), \ldots, b_1(T)]$
B
a)
mode indication
"non-simplified mode"
Encoder
Optional side
chain
from AOs
AO
Directions
$W_L = D_L B$
W_L
DRC Analysis
g
DRC Gain
Encoder
B
HOA
b)

Fig.4

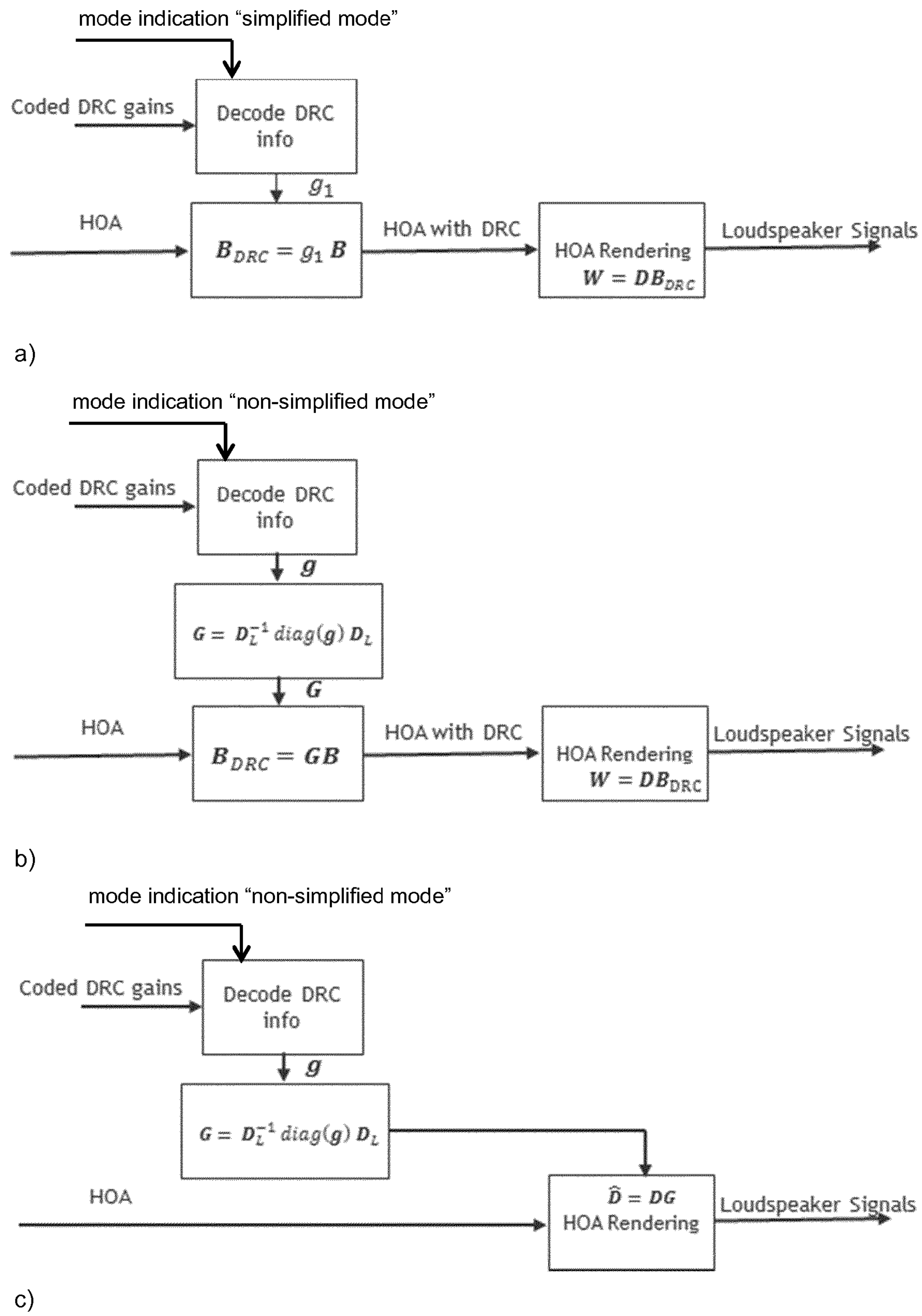
mode indication "simplified mode"
Coded DRC gains
Decode DRC info
g_1
HOA
$B_{DRC} = g_1 B$
HOA with DRC
HOA Rendering $W = DB_{DRC}$
Loudspeaker Signals
a)
mode indication "non-simplified mode"
Coded DRC gains
Decode DRC info
g
$G = D_L^{-1}\, diag(g)\, D_L$
G
HOA
$B_{DRC} = GB$
HOA with DRC
HOA Rendering $W = DB_{DRC}$
Loudspeaker Signals
b)
mode indication "non-simplified mode"
Coded DRC gains
Decode DRC info
g
$G = D_L^{-1}\, diag(g)\, D_L$
HOA
$\tilde{D} = DG$ HOA Rendering
Loudspeaker Signals
c)

Fig.5

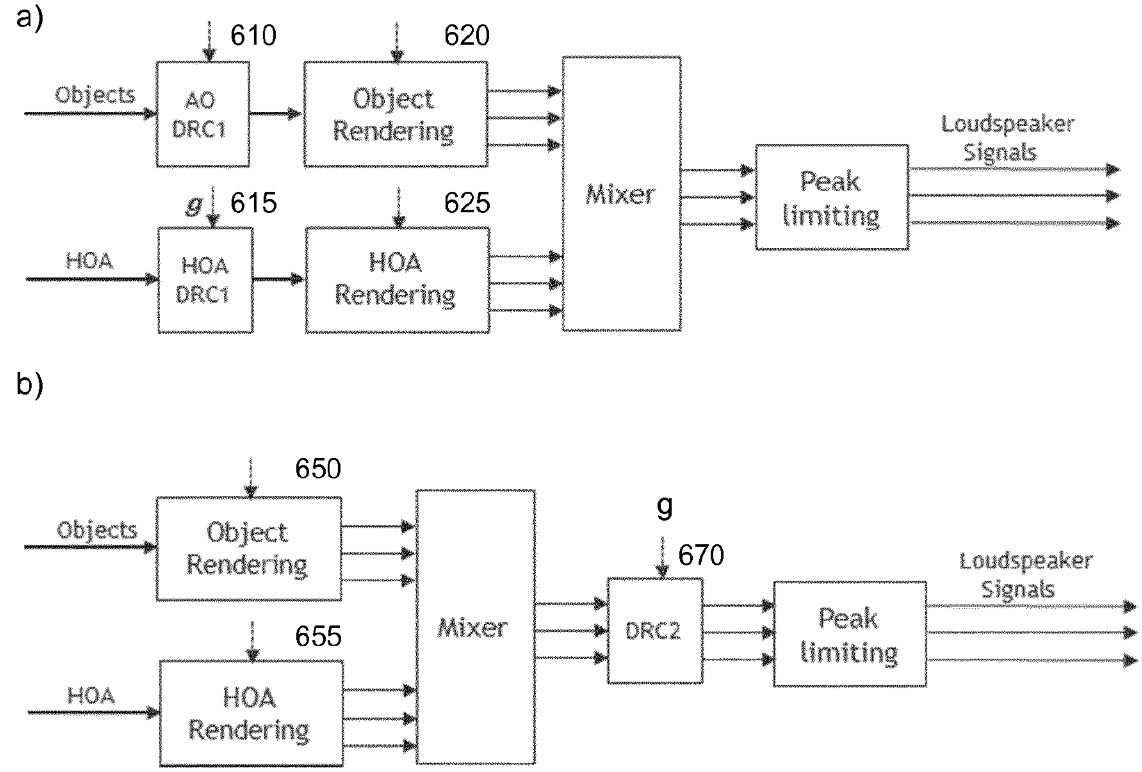
a)
610
620
Objects
AO
DRC1
Object
Rendering
Mixer
Loudspeaker
Signals
Peak
limiting
g
615
625
HOA
HOA
DRC1
HOA
Rendering
b)
650
Objects
Object
Rendering
g
670
Mixer
Loudspeaker
Signals
DRC2
Peak
limiting
655
HOA
HOA
Rendering

Fig.6

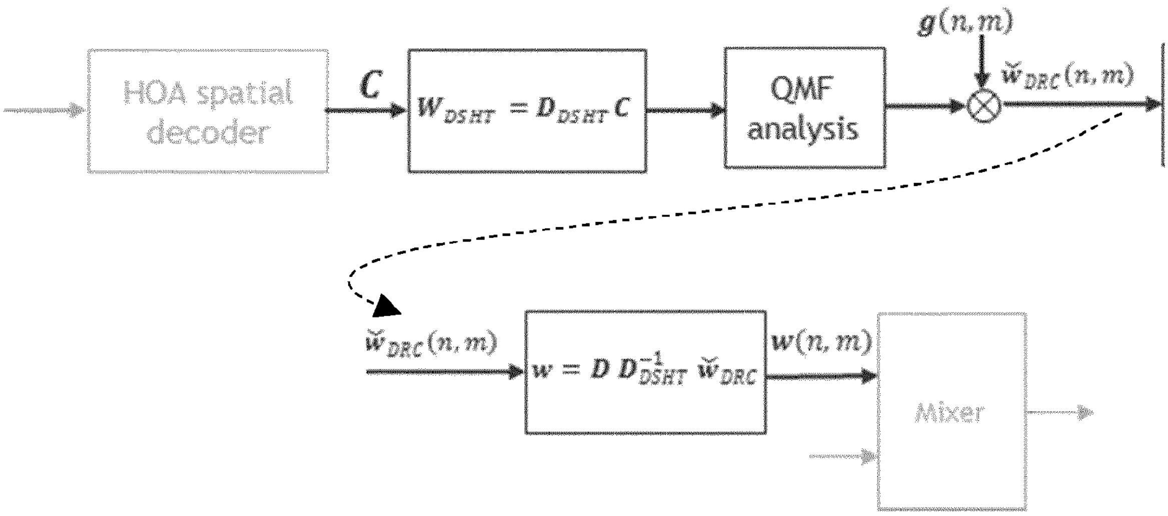
$g(n,m)$
HOA spatial
decoder
C
$W_{DSHT} = D_{DSHT} C$
QMF
analysis
$\breve{w}_{DRC}(n,m)$
$\breve{w}_{DRC}(n,m)$
$w = D\, D_{DSHT}^{-1}\, \breve{w}_{DRC}$
$w(n,m)$
Mixer

Fig.7

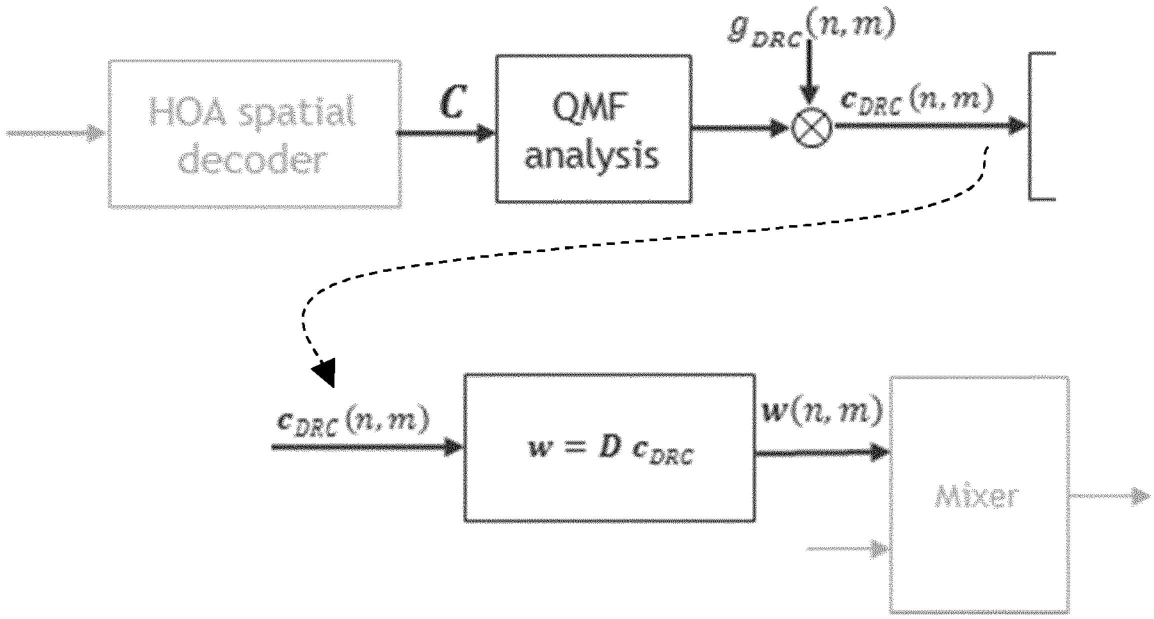
HOA spatial decoder
C
QMF analysis
$g_{DRC}(n,m)$
$c_{DRC}(n,m)$
$c_{DRC}(n,m)$
$w = D\ c_{DRC}$
$w(n,m)$
Mixer

Fig.8

Europäisches Patentamt
European Patent Office
Office européen des brevets

EUROPEAN SEARCH REPORT

Application Number
EP 14 30 5559

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
X	EP 2 688 066 A1 (THOMSON LICENSING [FR]) 22 January 2014 (2014-01-22) * paragraph [0014]; figure 2 * -----	1-15	INV. H04S3/00 G10L19/008
X	BURNETT IAN ET AL: "Encoding Higher Order Ambisonics with AAC", AES CONVENTION 124; MAY 2008, AES, 60 EAST 42ND STREET, ROOM 2520 NEW YORK 10165-2520, USA, 1 May 2008 (2008-05-01), XP040508582, * page 3; figure 1 * -----	1-15	
A	"WD1-HOA Text of MPEG-H 3D Audio", 107. MPEG MEETING;13-1-2014 - 17-1-2014; SAN JOSE; (MOTION PICTURE EXPERT GROUP OR ISO/IEC JTC1/SC29/WG11),, no. N14264, 21 February 2014 (2014-02-21), XP030021001, * page 30; figure 9 * -----	1-15	
A	WO 2013/181115 A1 (DTS INC [US]) 5 December 2013 (2013-12-05) * the whole document * -----	1-15	TECHNICAL FIELDS SEARCHED (IPC) H04S G10L

The present search report has been drawn up for all claims

1

Place of search	Date of completion of the search	Examiner
Munich	17 October 2014	Kunze, Holger

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

ANNEX TO THE EUROPEAN SEARCH REPORT ON EUROPEAN PATENT APPLICATION NO. EP 14 30 5559

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report.
The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

17-10-2014

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 2688066 A1	22-01-2014	EP 2688066 A1	22-01-2014
		TW 201412145 A	16-03-2014
		WO 2014012944 A1	23-01-2014
WO 2013181115 A1	05-12-2013	NONE	

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82

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

- EP 13306042 A **[0016]**

Non-patent literature cited in the description

- **JÖRG FLIEGE.** *Integration nodes for the sphere,* 2010, www.mathematik.uni-dortmund.de/lsx/research/projects/fliege/nodes/-nodes.html **[0039]**
- **JÖRG FLIEGE ; ULRIKE MAIER.** A two-stage approach for computing cubature formulae for the sphere. *Technical report, Fachbereich Mathematik,* 1999 **[0039]**