Claims

1. An audio signal representation decoding unit for generating a decompressed ambisonic spatial audio signal representation from a compressed ambisonic spatial audio signal representation representing an audio signal, the compressed ambisonic spatial audio signal representation including at least one transport channel and side information, the side information including sound field parameters, the sound field parameters including, for each spatial sector of a plurality of spatial sectors, directional parameter(s) providing information on a direction of arrival, DoA, in the spatial sector, the sound field parameters including, for at least one spatial sector, sector diffuseness parameter(s) providing information on sector diffuseness of the audio signal in the at least one spatial sector,

the audio signal representation decoding unit including a plurality of sector decoding paths, each sector decoding path being configured to decode a directional sector signal of the decompressed ambisonic spatial audio signal representation in each spatial sector by applying, to the at least one transport channel, or a sector signal derived from the at least one transport channel, the directional parameter(s) and the sector diffuseness parameter(s) of the spatial sector,

the audio signal representation decoding unit including a global diffuseness signal decoding path configured to derive a global diffuseness signal by applying, to the at least one transport channel, a global diffuseness parameter, or other information on the global diffuseness of the audio signal,

the audio signal representation decoding unit including a global diffuseness signal inserter to combine the plurality of decoded directional sector signals and the global diffuseness signal, to output the decompressed ambisonic spatial audio signal representation.

2. The audio signal representation decoding unit of claim 1, configured to apply, to the at least one transport channel or a sector signal derived from the transport channel, the sector diffuseness parameter(s) by weighting the transport channel, in at least one sector decoding path, using a mixing weight derived from the sector diffuseness parameter(s), to thereby derive the directional sector signal.

3. The audio signal representation decoding unit of claim 2, configured to weight the at least one transport channel or sector signal derived from the transport channel using the mixing weight being, or being derived from, a positive coefficient received from, or processed from, the sector diffuseness parameter(s).

4. The audio signal representation decoding unit of claim 2 or 3, configured to weight the at least one transport channel, or sector signal derived from the transport channel, using the mixing weight, for at least one spatial sector,

the mixing weight being, or being derived from, a coefficient indicative of a sector directionality in the specific spatial sector.

5. The audio signal representation decoding unit of any of claims 2-4, configured to weight the at least one transport channel or sector signal derived from the transport channel using the mixing weight, for each spatial sector,

the mixing weight being, or being derived from, a coefficient indicative of the relative directionality of the signal in the specific spatial sector over the relative directionalities of the totality of the spatial sectors.

6. The audio signal representation decoding unit of any of claims 2-5, configured to weight the at least one transport channel or sector signal derived from the transport channel for at least one first spatial sector using a first mixing weight being, or being derived from, a coefficient indicative of the sector directionality in the first spatial sector, and

configured to weight the at least one transport channel or sector signal derived from the transport channel for at least one second spatial sector using a second mixing weight,

the audio signal representation decoding unit being configured to retrieve the second mixing weight being retrieved by complementing, to a predetermined fixed value, the coefficient indicative of the sector directionality in the first spatial sector.

7. The audio signal representation decoding unit of any of claims 2-6, configured to derive each of N-1 mixing weights from parameters written in the side information, and to derive one N-th mixing weight from by complementing the other N-1 mixing weights to a constant positive value, where N is the number of spatial sectors.

8. The decoding unit of any of the preceding claims, configured, in each sector decoding path, to apply, to the at least one sector signal, the directional parameter(s) by multiplying the at least one sector signal by a vector of spherical harmonic functions evaluated along the DoA (Ω_s) in the spatial sector, so as to extend the directional signal for the spatial sector in a higher ambisonics order.

9. The decoding unit of any of the preceding claims, configured to apply a spatial filter to the at least one transport channel or processed version of the at least one transport channel, to limit the at least one transport channel to one spatial sector for each sector decoding path.

10. The decoding unit of any of the preceding claims, configured to compute at least one directional sector signal using

$$\mathbf{x}_s = x_s * \mathbf{Y}(\Omega_s) \quad = [x_s * Y_{00}(\Omega),\ x_s * Y_{1-1}(\Omega), x_s * Y_{10}(\Omega), x_s * Y_{11}(\Omega), \dots],$$

where s indicates the spatial sector, x_s is the transport channel, or processed version thereof, in the specific spatial sector s, Ω_s is the directional parameter for the specific spatial sector s, and **Y**, which is a function of Ω_s, is the vector of spherical harmonic functions given by

$[Y_{00}(\Omega_s), Y_{1-1}(\Omega_s), Y_{10}(\Omega_s), Y_{11}(\Omega_s), \ldots Y_{nm}(\Omega_s)]$, and $Y_{nm}(\Omega_s)$ is a spherical harmonic of order n and degree m.

11. The decoding unit of any of the preceding claims, configured to apply a global diffuseness weight obtained from the global diffuseness parameter, or the information on the global diffuseness of the audio signal, to weight the at least one transport channel, thereby obtaining a global diffuseness signal version to be used in the global diffuseness signal decoding path, and

to apply a second weight, complementary to the global diffuseness weight, to weight the at least one transport channel, thereby obtaining at least one globally non-diffuse signal to be processed in the plurality of sector decoding paths.

12. The decoding unit of any of the preceding claims, configured to derive mixing weight(s) of the global diffuseness signal and the directional sector signals from the global diffuseness parameter, or the information on the global diffuseness of the audio signal.

13. The decoding unit of any of the preceding claims, configured to apply, to the at least one transport channel, a weighting parameter complementary to the global diffuseness parameter used for deriving the global diffuseness signal , so that, for each sector decoding path, the at least transport channel is weighed using the weighting parameter.

14. The audio signal representation decoding unit of any of the preceding claims,

wherein the global diffuseness signal decoding path is configured to weight the at least one transport channel by a global diffuseness gain, which is, or is derived from, the global diffuseness parameter, or the other information on the global diffuseness of the audio signal, and

each of the plurality of sector decoding paths is configured to weight the at least one transport channel by a global directionality gain which is, or is derived from, the global diffuseness parameter, or the other information on the global diffuseness of the audio signal.

15. The audio signal representation decoding unit of claim 14 , wherein the global diffuseness gain is $1 + g(\Psi)$ and is according to

$$1 + g(\Psi) = \sqrt{1 + \Psi * (f_{comp} - 1)} \ ,$$

where Ψ is, or is derived from, the global diffuseness parameter, or the other information on the global diffuseness of the audio signal, and f_{comp} is a diffuse compensation factor.

16. The audio signal representation decoding unit of any of claims 14-15, where the value range of the global diffuseness gain is limited to a certain value range as to prevent too strong deviations from the global diffuseness signal.

17. The audio signal representation decoding unit of any of the preceding claims, configured to switch between:

a low order operation mode, in which, among the plurality of sector decoding paths, at least one of the sector decoding paths is deactivated, while only one of the sector decoding paths is activated, wherein the side information does not contain the sound field parameter(s) for the deactivated at least one of the sector decoding paths; and

a high order operation mode, in which, among the plurality of sector decoding paths, all the plurality sector decoding paths are activated, or at least less sector decoding paths are deactivated in respect to the low order operation mode, wherein the side information also contains the sound field parameter(s) for all the plurality of the sector decoding paths, as well as the global diffuseness parameter.

18. The audio signal representation decoding unit of any of the preceding claims, configured to convert the spatial audio signal representation from an encoded at least one transport channel into a decoded version of the encoded at least one transport channel.

19. The audio signal representation decoding unit of any of the preceding claims, configured to convert the decoded ambisonic spatial audio signal representation from the filterbank domain to the time domain.

20. The audio signal representation decoding unit of any of the preceding claims, further configured to upmix the at least one transport channel from a first number of transport channels to a second number of transport channels greater than the first number.

21. The audio signal representation decoding unit of any of the preceding claims, comprising a mixing-matrix estimator configured to process the sound field parameters, to derive a covariance matrix, or other covariance information, between different transport channels, the mixing-matrix estimator being configured to reconstruct a mixing matrix, or other mixing information, from the covariance matrix, or the other covariance information, and to apply the mixing matrix, or the other mixing information, to the transport channels.

22. The audio signal representation decoding unit of claim 21, wherein the covariance-matrix synthesizer is configured to process the sound field parameter(s) including the DoA parameter(s) and sector diffuseness parameter(s) of the plurality of spatial sectors and the global diffuseness parameter, or other information on the global diffuseness, to derive the covariance matrix, or the other covariance information, between different transport channels, the mixing-matrix estimator being configured to reconstruct a mixing matrix, or the other mixing information, from the covariance matrix, or the other

covariance information, so as to employ the sound field parameter(s) to derive the covariance matrix, or the other covariance information, for at least one frequency band, the audio signal representation decoding unit being configured to derive the covariance matrix, or the other covariance information, for at least one other frequency band without using the sound field parameters.

23. The audio signal representation decoding unit of claim 22, configured to derive, for at least one other frequency band, the mixing matrix, or other mixing information, from covariance information which is received from the side information.

24. The audio signal representation decoding unit of claim 18, where the sound field parameters are modified in order to achieve a rotation of the sound field represented by the output ambisonic signal.

25. An apparatus, comprising:

the audio signal representation decoding unit of any of the preceding claims;

a bitstream reader and dequantizer, configured to read a bitstream, in which there is encoded the low order spatial audio signal representation, and to provide the high order spatial audio signal representation to the audio signal representation decoding unit.

26. The apparatus of claim 25, further comprising

a renderer, to render the audio signal from the ambisonic spatial audio signal representation.

27. An audio signal representation encoding unit for encoding an input spatial audio signal representation, representing an audio signal, onto a compressed ambisonic spatial audio signal representation representing the audio signal,

the audio signal representation encoding unit being configured to downmix the input spatial audio signal representation to derive at least one transport channel;

the audio signal representation encoding unit being configured to derive side information, the side information including sound field parameters, the sound field parameters including, for each spatial sector of a plurality of spatial sectors, directional parameter(s) providing information on a direction of arrival, DoA, in the spatial sector, the sound field parameters including sector diffuseness parameter(s) providing information on diffuseness of the audio signal in at least one spatial sector,

the audio signal representation encoding unit including a plurality of sector parameter estimators, each sector parameter estimator being configured to process a specific sector signal of the input spatial audio signal representation in a specific spatial sector of the plurality of spatial sectors, so as to derive the directional parameter(s) and the information on diffuseness of the audio signal in the at least one spatial sector,

the audio signal representation encoding unit including a bitstream writer to encode the at least one transport channel and the side information.

28. The audio signal representation encoding unit of claim 27, further including a global diffuseness parameter estimator to estimate a global diffuseness parameter to be inserted in the side information.

29. The audio signal representation encoding unit of any of claims 27-28, further configured to estimate a relative directionality of each specific spatial sector in respect to the directionalities of the all spatial sectors, and to write the coefficient, or information indicative of the relative directionality, as a sector diffuseness parameter.

30. The audio signal representation encoding unit of claim 29, further configured to estimate the relative directionality as including at least one of a first and a second spatial sector respectively indicated with a_1 and a_2 and satisfies

$$a_1 = \frac{1 - \Psi_1}{(1 - \Psi_1) + (1 - \Psi_2)}$$

and

$$a_2 = 1 - a_1,$$

where Ψ_1 is, or is obtained from, the sector diffuseness information for the first spatial sector and Ψ_2 is, or is obtained from, the sector diffuseness information for the second spatial sector.

31. The audio signal representation encoding unit of any of claims 27-30, further configured to convert the input spatial audio signal representation into the filterbank domain to derive a filterbank version of the input spatial audio signal representation,

further configured to downmix the filterbank domain version of the input spatial audio signal representation to derive the at least one transport channel in the filterbank domain, and

further configured to perform a filterbank synthesis of the at least one transport channel from the filterbank domain to the time domain.

32. The audio signal representation encoding unit of any of claims 27-31 configured to switch between:

a low order operation mode, in which, among a plurality of sector paths, at least one of the sector paths is deactivated, while only one of the sector paths is activated, wherein the side information does not contain the sound field parameter(s) for the deactivated at least one of the sector paths; and

a high order operation mode, in which, among the plurality of sector paths, all the plurality sector paths are activated, or at least less sector paths are deactivated in respect to the low order operation mode, wherein the side information also contains the sound field parameter(s) for all the plurality of the activated sector paths, as well as a global diffuseness parameter.

33. The audio signal representation encoding unit of claim 32, configured to select between the low order operation mode and the high order operation mode based on the bitrate, so as to select the low order operation mode in case of low bitrate, and the high order operation mode in case of bitrate higher than the low bitrate.