Claims

1. Apparatus for processing an encoded audio scene representing a sound field related to a virtual listener position, the encoded audio scene comprising information on a transport signal and a first set of parameters related to the virtual listener position, the apparatus comprising:

a parameter converter for converting the first set of parameters into a second set of parameters related to a channel representation comprising two or more channels for a reproduction at predefined spatial positions for the two or more channels,

wherein the first set of parameters comprises, for each input time frame of a plurality of input time frames and for each input frequency band of a plurality of input frequency bands, at least one DirAC parameter, and

wherein the parameter converter is configured to calculate the second set of parameters as parametric stereo or multichannel parameters; and

an output interface for generating a processed audio scene using the second set of parameters and the information on the transport signal.

2. Apparatus of claim 1,

wherein the output interface is configured for upmixing the transport signal using the second set of parameters into an upmix signal comprising the two or more channels.

3. Apparatus of claim 1, wherein the output interface is configured to generate the processed audio scene by combining the transport signal or the information on the transport signal and the second set of parameters to obtain a transcoded audio scene as the processed audio scene.

4. Apparatus of one of the preceding claims, wherein the at least one DirAC parameter comprises at least one of a direction of arrival parameter, a diffuseness parameter, a direction information parameter related to a sphere with the virtual listening position as an origin of the sphere, and a distance parameter, and

wherein the parametric stereo or multichannel parameters comprise at least one of a side gain parameter, a residual prediction gain parameter, an inter-channel level difference parameter, an inter-channel time difference parameter, an inter-channel phase difference parameter and an inter-channel coherence parameter.

5. Apparatus of one of the preceding claims, wherein the input time frame, to which the first set of parameters is related, comprises two or more input time subframes, and wherein an output time frame, to which the second set of parameters is related, is smaller than the input time frame and longer than an input time subframe of the two or more input time subframes, and

wherein the parameter converter is configured to calculate, from the at least one DirAC parameter for each one of the two or more input time subframes being time-subsequent, a raw parameter of the second set of parameters for each one of the two or more input time subframes being time-subsequent and to combine the at least two raw parameters for the two or more input time subframes being time-subsequent to derive a parameter of the second set of parameters related to the output time frame.

6. Apparatus of claim 5, wherein the parameter converter is configured to perform, in the combining the at least two raw parameters for the two or more input time subframes being time-subsequent, a weighted combination of the at least two raw parameters, wherein weighting factors for the weighted combination are derived based on an amplitude-related measure of the transport signal in the corresponding input time subframe wherein the amplitude-related measure is, an energy or a power of the transport signal in the corresponding input time subframe, and wherein a weighting factor for an input time subframe is greater in case of a higher energy or higher power of the transport signal in the corresponding input time subframe compared to a weighting factor for an input time subframe having a lower energy or lower power of the transport signal in the corresponding input time subframe.

7. Apparatus of one of the preceding claims, wherein the parameter converter is configured to calculate at least one raw parameter for each output time frame using at least one parameter of the first set of parameters for the input time frame,

wherein the parameter converter is configured to calculate a smoothing factor for each raw parameter in accordance with a smoothing rule, and

wherein the parameter converter is configured to apply a corresponding smoothing factor to the corresponding raw parameter to derive the parameter of the second set of parameters for the output time frame.

8. Apparatus of claim 7, wherein the parameter converter is configured

to calculate a long-term average over an amplitude-related measure of a first time portion of the transport signal, and

to calculate a short-term average over an amplitude-related measure of a second time portion of the transport signal, wherein the second time portion is shorter than the first time portion, and

to calculate a smoothing factor based on a ratio between the long-term average and the short-term average.

9. Apparatus of one of the preceding claims,

wherein the input time frame, to which the first set of parameters is related, comprises two or more input time subframes, and wherein an output time frame, to which the second set of parameters is related, is smaller than the input time frame and longer than an input time subframe of the two or more input time subframes,

wherein the parameter converter is configured to calculate, from the parameters for each one of the two or more input time subframes being time-subsequent, a raw parameter of the second set of parameters for each one of the two or more input time subframes being time-subsequent and to combine the at least two raw parameters for the two or more input time subframes being time-subsequent to derive a parameter of the second set of parameters related to the output time frame, and

wherein the output interface is configured

to perform a conversion of a time portion of the transport signal corresponding to the output time frame into a spectral representation, wherein the time portion of the transport signal is shorter than the input time frame, in which the parameters of the first set of parameters are organized,

to perform an upmixing operation of the spectral representation using the second set of parameters to obtain the two or more channels in the spectral representation; and

to convert each channel of the two or more channels in the spectral representation into a time representation.

10. Apparatus of claim 9, wherein the output interface is configured to perform the upmixing operation based on the following equations:

$$\tilde{L}_{t,k} = \frac{(\tilde{M}_{t,k}(1+\tilde{g}_{t,b})+\tilde{r}_{t,b}g_{norm}\tilde{\rho}_{t,k})}{\sqrt{2}}$$

and

$$\tilde{R}_{t,k} = \frac{(\tilde{M}_{t,k}(1-\tilde{g}_{t,b})-\tilde{r}_{t,b}g_{norm}\tilde{\rho}_{t,k})}{\sqrt{2}},$$

wherein $\tilde{M}_{t,k}$ is the transport signal for a frame t and a frequency bin k, wherein $\tilde{L}_{t,k}$ is a first channel of the two or more channels in the spectral representation for the frame t and the frequency bin k, wherein $\tilde{R}_{t,k}$ is a second channel of the two or more channels in the spectral representation for the frame t and the frequency bin k, wherein $\tilde{g}_{t,b}$ is a side gain parameter for the frame t and a subband b, wherein $\tilde{r}_{t,b}$ is a residual prediction gain parameter for the frame t and the suband b, wherein g_{norm} is an energy adjusting factor that can be there or not, and wherein $\tilde{\rho}_{t,k}$ is a raw residual signal for the frame t and the frequency bin k.

11. Apparatus of one of the preceding claims,

wherein the first set of parameters comprises a direction of arrival parameter for the input frequency band, and wherein the second set of parameters comprises a side gain parameter per input frequency band, and

wherein the parameter converter is configured to calculate the side gain parameter for an output frequency band using the following equation:

$$sidegain[b] = \sin\left(azimuth[b] * \frac{\pi}{180}\right)\cos\left(elevation[b] * \frac{\pi}{180}\right),$$

wherein b is the output frequency band, wherein sidegain is the side gain parameter, wherein azimuth is an azimuth component of the direction of arrival parameter, and wherein elevation is an elevation component of the direction of arrival parameter.

12. Apparatus of claim 11,

wherein the first set of parameters additionally comprises a diffuseness parameter for the input frequency band, and wherein the parameter converter is configured to calculate the side gain parameter for the output frequency band using the following equation:

$$sidegain[b] = \sqrt{1 - diff(b)}\sin\left(azimuth[b] * \frac{\pi}{180}\right)\cos\left(elevation[b] * \frac{\pi}{180}\right)$$

wherein *diff(b)* is the diffuseness parameter for the input frequency band b.

13. Apparatus of one of the preceding claims,

wherein the first set of parameters comprises a diffuseness parameter for the input frequency band, and

wherein the second set of parameters comprises a residual prediction gain parameter for an output frequency band, and

wherein the parameter converter is configured to use, as the residual prediction gain parameter for the output frequency band, the diffuseness parameter from the input frequency band, when the input frequency band and the output frequency band are equal to each other, or to derive, from the diffuseness parameter for the input frequency band, a diffuseness parameter for the output frequency band, and to then

use the diffuseness parameter for the output frequency band as the residual prediction gain parameter for the output frequency band.

14. Method of processing an encoded audio scene representing a sound field related to a virtual listener position, the encoded audio scene comprising information on a transport signal and a first set of parameters related to the virtual listener position, the method comprising:

converting the first set of parameters into a second set of parameters related to a channel representation comprising two or more channels for a reproduction at predefined spatial positions for the two or more channels,

wherein the first set of parameters comprises, for each input time frame of a plurality of input time frames and for each input frequency band of a plurality of input frequency bands, at least one DirAC parameter, and

wherein the converting comprises calculating the second set of parameters as parametric stereo or multichannel parameters; and

generating a processed audio scene using the second set of parameters and the information on the transport signal.

15. Computer program for performing, when running on a computer or a processor, the method of claim 14.