Claims

1. Apparatus for transforming an audio stream with more than one channel into another representation, the apparatus being on an encoder side and comprising:

means for deriving one or more parameters describing an acoustic or psychoacoustic model of the audio stream on the encoder side, wherein the means for deriving are configured to calculate prediction coefficients as the one or more parameters, wherein the prediction coefficients are calculated based on a covariance matrix by the means for deriving, wherein the covariance matrix is calculated based on direction of arrival parameters;

means for transforming the audio stream in a signal-adaptive way dependent on the prediction coefficients; and

wherein the one or more parameters comprise at least an information on at least one direction of arrival,

wherein the means for transforming are configured to perform a downmixing of the audio stream on the encoder side.

2. Apparatus for transforming an audio stream with more than one channel into another representation, the apparatus being on a decoder side and comprising:

means for receiving one or more parameters describing an audio scene with an acoustic or psychoacoustic model on the decoder side, wherein the one or more parameters comprise prediction coefficients;

means for transforming the audio stream in a signal-adaptive way dependent on the prediction coefficients;

wherein the prediction coefficients are calculated based on a covariance matrix,

wherein the covariance matrix is calculated based on direction of arrival parameters, and

wherein the means for transforming are configured to perform upmix generation of the audio stream on the decoder side.

3. Apparatus according to one of the previous claims, wherein prediction coefficients are calculated based on $Y_{l,m}$, especially beads on the formula

$$\boldsymbol{P} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -C_{x,w}/\,C_{w,w} & 1 & 0 & 0 \\ -C_{y,w}/C_{w,w} & 0 & 1 & 0 \\ -C_{z,w}/C_{w,w} & 0 & 0 & 1 \end{pmatrix}, \tag{4}$$

with the matrix elements

$$C_{x/y/z,w} = E^{\mathrm{dir}} Y_{0,0}(\theta_D) Y_{1,-1/0/1}(\theta_D)$$

and

$$C_{w,w} = E^{\mathrm{dir}} Y_{0,0}(\theta_D) Y_{0,0}(\theta_D) + \; E_w^{\mathrm{diff}} \tag{13}$$

where $Y_{l,m}$ are real spherical harmonics with degree and index l and m.

4. Apparatus according to one of the previous claims, wherein the one or more parameters further comprise at least an information on a diffuseness factor or on one or more DOAs or on energy ratios, and/or wherein the one or more parameters are derived from the audio stream.

5. Apparatus according to claim 1, wherein the means for deriving are configured to calculate a covariance matrix or a covariance matrix from the acoustic or psychoacoustic model.

6. Apparatus according to one of the previous claims, wherein the means for deriving are configured to calculate a covariance matrix based on the DoA and a diffuseness factor or an energy ratio.

7. Apparatus according to claim 6, wherein the means for deriving are configured to calculate a covariance matrix based on an information about diffuseness, spherical harmonics and a time-dependent scalar-valued signal, especially based on the formula

$$C_{x/y/z,w} = \int dt\, s^2(t) Y_{0,0}(\theta_D) Y_{1,-1/0/1}(\theta_D)$$

where $Y_{l,m}$ is a spherical harmonic with the degree and index l and m and where s(t) is a time-dependent scalar-valued signal; and/or

based on a signal energy, especially based on the following formula

$$C_{x/y/z,w} = (1 - \Psi)\, E\, Y_{0,0}(\theta_D) Y_{1,-1/0/1}(\theta_D)$$

where ψ describes the diffuseness and where E describes the signal energy for the audio stream; and/or based on the formula

$$C_{w,w} = (1 - \Psi)\, E\, Y_{0,0}(\theta_D) Y_{0,0}(\theta_D) + \Psi\, E$$

where E is the signal energy; and/or based on the formula

$$C_{x,x} = (1 - \Psi)\, E\, Y_{1,-1}(\theta_D) Y_{1,-1}(\theta_D) + \frac{\Psi}{3}\, E$$

and for the y and z channels analogously.

8. Apparatus according to claims 7, wherein the signal energy E is directly calculated from the audio stream; or

wherein the signal energy E is estimated from the model of the audio stream.

9. Apparatus according to one of the previous claims, wherein the audio stream is preprocessed by a parameter estimator or wherein the audio stream is preprocessed by a parameter estimator comprising a metadata encoder or metadata decoder and/or wherein the audio stream is preprocessed by an analysis filterbank.

10. Apparatus for transforming an audio stream in a directional audio coding system, being on an encoder side and comprising:

means for deriving one or more acoustic model parameters of a model of the audio stream, wherein the one or more acoustic model parameters are transmitted to enable restoring all channels of the audio stream and comprise at least an information on direction of arrival (DoA),

means for transforming the audio stream in a signal-adaptive way dependent on the one or more acoustic model parameters; where all or a subset of the channels of the audio stream are transformed;
wherein the means for transforming are configured to perform a downmixing of the audio stream on the encoder side.

11. Apparatus according to one of the previous claims for transforming an audio stream in a directional audio coding system, wherein one or more acoustic model parameters are received to restore all channels of the audio stream and comprise at least an information on direction of arrival, or

wherein the means for transforming are configured to perform upmix generation of the audio stream on the decoder side.

12. Apparatus according to one of the previous claims, wherein the one or more parameters are quantized prior to a transmission.

13. Apparatus according to one of the previous claims, wherein the one or more parameters are dequantized after a transmission.

14. Apparatus according to one of the previous claims, wherein the parameters are smoothed over time.

15. Apparatus according to one of the previous claims, wherein a transform is computed such that correlations between transport channels are reduced by use of Karhunen-Loève transform or prediction matrix.

16. Apparatus according to one of the previous claims, wherein an inter-channel covariance matrix of the audio stream is estimated from the model or the acoustic or psychoacoustic model of the audio stream.

17. Apparatus according to one of the previous claims, wherein a transform matrix is derived from a covariance matrix of the model or the acoustic or psychoacoustic model of the audio stream.

18. Apparatus according to one of the previous claims, wherein a transform matrix is calculated using the covariance matrix from the acoustic or psychoacoustic model for one or more frequency bands and a different method to calculate the covariance matrix for one or more other frequency bands

19. Apparatus according to one of the previous claims, wherein at least one of transform methods used by the means for transforming is multiplication of a vector of audio channels by a constant matrix.

20. Apparatus according to one of the previous claims, wherein at least one of transform methods used by the means for transforming uses prediction based on the inter-channel covariance matrix of a vector of audio channels.

21. Apparatus according to one of the previous claims, wherein at least one of transform methods used by the means for transforming uses prediction based on inter-channel covariance matrix based on the DOA and an additional diffuseness factor or an energy ratio.

22. Apparatus according to one of the previous claims, wherein the means for deriving the one or more parameters are configured to process all or a subset of the channels of a first-order or higher-order Ambisonics input signal of the audio stream.

23. Apparatus according to one of claims 10-22, wherein a sound scene of the audio stream is rotatable in such a way that:

- an audio signal in the spherical-harmonics domain resulting from a transform is pre-multiplied by a rotation matrix;
- model parameters and/or prediction coefficients are transformed in accordance with the transform of a transport channel signal; and
- non-transport channels of an output signal are reconstructed using the transformed model and/or prediction coefficients parameters.

24. Encoder comprising an apparatus according to one of the claims 1, 3-22 having backreference to claim 1.

25. Decoder comprising an apparatus according to one of the claims 2 or 3-22 having backreference to claim 2.

26. A system comprising an encoder according to claim 24 and a decoder according to 25, wherein the encoder is configured to calculate a prediction matrix and/or a downmix and wherein decoder is configured to calculate an upmix matrix from estimated parameters or the one or more parameters of the acoustic model independently of each other.

27. Method for transforming an audio stream with more than one channel into another representation, performed on an encoder side and comprising the following steps:

deriving the one or more parameters describing an acoustic or psychoacoustic model of an audio stream from the audio stream, wherein deriving comprises calculating prediction coefficients as the one or more parameters, wherein the covariance matrix is calculated based on direction of arrival parameters and wherein the prediction coefficients calculated are calculated based on a covariance matrix and wherein the one or more parameters comprise at least an information on direction of arrival; and

transforming the audio stream in a signal-adaptive way dependent on the prediction coefficients; wherein transforming comprises a downmixing of the audio stream on the encoder side.

28. Method for transforming an audio stream with more than one channel into another representation, performed on a decoder side and comprising the following steps:

receiving one or more parameters describing an audio scene with an acoustic or psychoacoustic model on the decoder side, wherein the one or more parameters comprise prediction coefficients; and

transforming the audio stream in a signal-adaptive way dependent on the prediction coefficients; wherein the prediction coefficients are calculated based on a

covariance matrix and wherein the covariance matrix is calculated based on direction of arrival parameters; wherein transforming comprises upmixing of the audio stream on the decoder side.

29. Computer program for performing, when running on a computer, the method according to claims 27, or 28.