Claims

1. Apparatus for synthesizing a spatially extended sound source (SESS), comprising:

a storage for storing one or more rendering data items for different limited spatial sectors, wherein the different limited spatial sectors are located in a rendering range for a listener, wherein the one or more rendering data items for a limited spatial sector comprises at least one of a left variance data item, a right variance data item, and a left-right covariance data item related to the left head related function data and the right head related function data;

a sector identification processor for identifying one or more limited spatial sectors for the spatially extended sound source within the rendering range for the listener based on spatially extended sound source data;

a target data calculator for calculating target rendering data from the stored left variance data, the stored right variance data, or the stored covariance data; and

an audio processor for processing an audio signal representing the spatially extended sound source using the target rendering data.

2. Apparatus of claim 1,

wherein the storage is configured to store the variance data items or the covariance data item related to head related transfer function data, or binaural room impulse response data, or binaural room transfer function data, or head related impulse response data.

3. Apparatus of claim 1 or 2, wherein the one or more rendering data items comprise variance or covariance data item values for different frequencies.

4. Apparatus of one of the preceding claims, wherein the storage is configured to store, for each limited spatial sector, a frequency dependent representation of the left variance data item, a frequency dependent representation of the right

variance data item, and a frequency dependent representation of the covariance data item.

5. Apparatus of one of the preceding claims,

wherein the target data calculator is configured for calculating, as the target rendering data, at least one of an inter-aural or inter-channel coherence cue, an inter-aural or inter-channel level difference cue, an inter-aural or inter-channel phase difference cue, a first side gain, and a second side gain as the target rendering data, and

wherein the audio processor is configured to perform at least one of an inter-channel or inter-aural coherence adjustment, an inter-aural or inter-channel phase difference adjustment, or an inter-aural or inter-channel level difference adjustment using corresponding cues as the target rendering data.

6. Apparatus of claim 5, wherein the target data calculator is configured

to calculate the inter-aural or inter-channel coherence cue based on the left variance data item, the right variance data item and the covariance data item, or

to calculate the inter-channel or inter aural phase difference cue based on the left variance data item, and the right variance data item, or

to calculate the inter-channel or inter-aural phase difference cue based on the covariance data item, or

to calculate the left or right side gain using the left or right variance data item and an information related to a signal power of the audio signal.

7. Apparatus of claim 6,

wherein the target data calculator is configured to calculate the inter-aural or inter-channel coherence cue, so that a value of the inter-aural or inter-channel coherence cue is within a range of +/- 20% of a value obtained by the following equation for the inter-aural or inter-channel coherence cue

$$\mathrm{IACC}\,(\omega) = \frac{\sum_{n=1}^{N} \mathrm{IACC}_{grid,n}\,(\omega)}{N_{weighted}},$$

wherein the target data calculator is configured to calculate the inter-aural or inter-channel level difference cue so that a value of the inter-aural or inter-channel level difference cue is within a range of +/- 20% of a value obtained by the following equation for the inter-aural or inter-channel level difference cue, or

wherein the target data calculator is configured to calculate the inter-aural or inter-channel phase difference cue so that a value of the inter-aural or inter-channel phase difference cue is within a range of +/- 20% of a value obtained by the following equation for the inter-aural or inter-channel phase difference cue,

$$\begin{aligned}\mathrm{IALD}\,(\omega) &= 10\log_{10}\frac{\mathbf{E}\{|Y_l(\omega)|^2\}}{\mathbf{E}\{|Y_r(\omega)|^2\}} \\ &= 10\log_{10}\frac{\sum_{n=1}^{N} A_{l,n}^2}{\sum_{n=1}^{N} A_{r,n}^2},\end{aligned} \tag{24}$$

$$\begin{aligned}\mathrm{IAPD}\,(\omega) &= \angle\left(\mathbf{E}\{Y_l(\omega)\cdot Y_r^*(\omega)\}\right) \\ &= \angle\left(\mathrm{IACC}(\omega)\right) \\ &= \angle\left(\sum_{n=1}^{N} A_{l,n}\cdot A_{r,n}\cdot e^{j(\phi_{l,n}-\phi_{r,n})}\right).\end{aligned} \tag{25}$$

or

wherein the target data calculator is configured to calculate the first or the second side gain so that a value of the first or the second side gain is within a range of +/- 20% of a value obtained by the following equations for the left side gain or the right side gain:

$$G_l\,(\omega) = \frac{\sum_{n=1}^{N} G_{grid,l,n}}{N_{weighted}}, \text{ or}$$

$$G_r\,(\omega) = \frac{\sum_{n=1}^{N} G_{grid,r,n}}{N_{weighted}},$$

wherein ω is a frequency variable, wherein $\mathrm{N}_{\mathrm{weighted}}$ is a normalization parameter, wherein $\sum_{\mathrm{n}=1}^{\mathrm{N}} \mathrm{G}_{\mathrm{grid,l,n}}$ is a sum over the stored and potentially modified left variance data items for the identified limited spatial sectors, wherein $\sum_{\mathrm{n}=1}^{\mathrm{N}} \mathrm{G}_{\mathrm{grid,r,n}}$ is a sum over the stored and potentially modified right variance data items for the identified limited spatial sectors, wherein N is the number of identified limited spatial sectors, wherein $\mathrm{G}_{\mathrm{l}}(\omega)$ is the left side gain, wherein $\mathrm{G}_{\mathrm{r}}(\omega)$ is the right side gain, wherein $\mathrm{IACC}(\omega)$ is the inter-aural or inter-channel coherence cue, wherein $\sum_{\mathrm{n}=1}^{\mathrm{N}} \mathrm{IACC}_{\mathrm{grid,n}}(\omega)$ is a sum over the stored and potentially modified covariance data items for the identified limited spatial sectors, wherein $IALD(\omega)$ is the inter-aural or inter-channel level difference cue, wherein $IAPD(\omega)$ is the inter-aural or inter-channel phase difference cue, wherein $A^2{}_{\mathrm{l,n}}$ is a squared amplitude of a left side transfer function in the identified limited spatial sector, wherein $A^2{}_{\mathrm{r,n}}$ is a squared amplitude of a right side transfer function in the identified limited spatial sector, wherein $\varphi_{\mathrm{l,n}}$ is a phase of a left side transfer function in the identified limited spatial sector, wherein $\varphi_{\mathrm{l,n}}$ is a phase of a right side transfer function in the identified limited spatial sector, or

wherein the inter-aural or inter-channel phase difference cue is a phase extracted from an (e.g. head related) transfer function point in a the center of the limited spatial sector or from an elementary spatial sector in a center of an extent defined by an identified set of elementary spatial sectors.

8. Apparatus of one of the preceding claims, wherein $A^2{}_{\mathrm{l,n}}$ is a squared amplitude of a left side transfer function in the identified limited spatial sector, wherein $A^2{}_{\mathrm{r,n}}$ is a squared amplitude of a right side transfer function in the identified limited spatial sector

wherein the sector identification processor is configured to apply a projection algorithm or a ray tracing analysis to determine the one or more limited spatial sectors as a set of elementary spatial sectors, or

to use, as the listener data, a listener position or a listener orientation, or to use, as the spatially extended sound source (SESS) data, an SESS orientation, an SESS position, or information on a geometry of the SESS.

9. Apparatus of one of the preceding claims, wherein the rendering range comprises a sphere or a portion of a sphere around the listener, wherein the rendering range is tied to the listener position or the listener orientation, and wherein the one or more limited spatial sector has an azimuth size and an elevation size.

10. Apparatus of claim 9, wherein the azimuth size and the elevation size of the different limited spatial sectors are different from each other, so that an azimuth size is finer for a limited spatial sector directly in front of the listener compared to an azimuth size of a limited spatial sector more to the side of the listener, or wherein the azimuth size decreases towards a side of the listener, or wherein an elevation size of a limited spatial sector is smaller than an azimuth size of this sector.

11. Apparatus of one of the preceding claims, wherein the sector identification processor is configured to determine a set of elementary spatial sectors as the one or more limited spatial sectors, wherein, for each elementary spatial sector, at least one of the left variance data item, the right variance data item, and the covariance data item is stored.

12. Apparatus of claim 11, wherein the sector identification processor is configured

to receive, from a description of an audio scene, occluding information on a potentially occluding object, and

to determine, based on the occlusion information, a specific spatial sector of the set of elementary spatial sectors as an occluding sector, and

wherein the target data calculator is configured to apply an occlusion function to the rendering data items stored for the occluding sector to obtain modified data, and to use the modified data for calculating the target rendering data.

13. Apparatus of claim 12, wherein the occlusion function is a low pass function having different attenuation values for different frequencies, and wherein the rendering data items are data items for different frequencies, and

wherein the target data calculator is configured to weight, for several frequencies, a data item for a certain frequency with the attenuation value for the certain frequency to obtain the modified rendering data.

14. Apparatus of one of claims 11 or 12, wherein the sector identification processor is configured to determine that another elementary spatial sector of the set of elementary spatial sectors determined for the occluding object is not occluded by the potential occluding object, and

wherein the target data calculator is configured to combine the modified data from the occluding sector and the rendering data items of the other sector without a modification using the occluding function or modified by a different modification function to obtain the target rendering data.

15. Apparatus of one of claims 12 to 14, wherein the sector identification processor is configured to determine a first elementary spatial sector of the set of elementary spatial sectors to have a first characteristic and to determine a second elementary spatial sector of the set of elementary spatial sectors to have a second different characteristic, and

wherein the target data calculator is configured to not apply any modification function to the first elementary spatial sector and to apply a modification function to the second elementary spatial sector, or to apply a first modification function to the first elementary spatial sector and to apply a second modification function to the second elementary spatial sector, the second modification function being different from the first modification function.

16. Apparatus of claim 15, wherein the first modification function is frequency selective and the second modification function is constant over frequency, or wherein the first modification function has a first frequency selective characteristic and wherein the second modification function has a second frequency selective characteristic being different from the first frequency selective characteristic, or wherein the first modification function has a first attenuation characteristic and the second modification function has a second different attenuation characteristic, and

wherein the target data calculator is configured to select or adjust the modification function from the first modification function and the second modification function based on a distance between the first elementary spatial sector or the second elementary spatial sector to the listener or based on a characteristic of an object being placed between the listener and the corresponding elementary spatial sector.

17. Apparatus of one of claims 12 to 16, wherein the sector identification processor is configured to classify the set of elementary spatial sectors into different sector classes based on characteristics associated with the elementary spatial sectors,

wherein the target data calculator is configured to combine the rendering data items of the elementary spatial sectors in each class to obtain a combined result for each class, if more than one elementary spatial sectors is in a class, and to apply a specific modification function associated with at least one class to the combined result of this class to obtain a modified combination result for this class, or

to apply the specific modification function associated with at least one class to the one or more data items of the one or more elementary spatial sectors of each class to obtain modified data items and to combine the modified data items of the elementary spatial sectors in each class to obtain a modified combination result for this class,

to combine the combination result or if available the modified combination result for each class to obtain an overall combination result, and

to use the overall combination result as the target rendering data or to calculate the target rendering data from the overall combination result.

18. Apparatus of claim 17,

wherein the characteristic for an elementary spatial sector is determined as being one of a group comprising an occluded elementary spatial sector involving a first occlusion characteristic, an occluded elementary spatial sector involving a second occlusion characteristic being different from the first occlusion

characteristic, an unoccluded elementary spatial sector having a first distance to the listener, and an unoccluded elementary spatial sector having a second distance to the listener, wherein the second distance is different from the first distance.

19. Apparatus of claim 17 or claim 18, wherein the target data calculator is configured to modify or combine frequency dependent variance or covariance parameters as the rendering data items to obtain, as the overall combination result, an overall combined variance or an overall combined covariance parameter, and

to calculate at least one of an inter-aural or inter-channel coherence cue, an inter-aural or inter-channel level difference cue, an inter-aural or inter-channel phase difference cue, a first side gain, or a second side gain as the target rendering data.

20. Apparatus of one of the preceding claims, further comprising an initializer to determine at least one of the left variance data item, the right variance data item, and the covariance data item from pre-stored head related function data, wherein the initializer is configured to calculate the left variance data item, the right variance data item or the covariance data item from a plurality of head related function data for the limited spatial sector, and wherein the limited spatial sector is sized in such a way that at least two left head related function data, at least two right head related function data exist for the limited spatial range.

21. Method of synthesizing a spatially extended sound source (SESS), comprising:

storing one or more rendering data items for different limited spatial sectors, wherein the different limited spatial sectors are located in a rendering range for a listener, wherein the one or more rendering data items for a limited spatial sector comprises at least one of a left variance data item related to left head related function data, a right variance data item related to right head related function data, and a covariance data item related to the left head related function data and the right head related function data;

identifying one or more limited spatial sectors for the spatially extended sound source within the rendering range for the listener based on spatially extended sound source data;

calculating target rendering data from the stored left variance data, the stored right variance data, or the stored covariance data; and

processing an audio signal representing the spatially extended sound source using the target rendering data.

22. Computer program for performing, when running on a computer or a processor, the method of claim 21.