Claims

1. Audio data converter, comprising:

an input interface for receiving a plurality of audio object descriptions of a plurality of audio objects, an audio object description having audio object metadata;

a metadata converter for converting the audio object metadata of the plurality of audio object descriptions into combined DirAC (Directional Audio Coding) metadata; and

an output interface for transmitting or storing the combined DirAC metadata,
wherein the metadata converter is configured

to convert each audio object metadata of the plurality of audio object descriptions into an individual DirAC data description comprising individual DirAC metadata for each audio object description, and

to combine the individual DirAC metadata by individually combining direction of arrival metadata from the individual DirAC metadata for the audio object descriptions by a weighted addition, wherein a weighting of the weighted addition is done in accordance with energies of associated pressure signal energies to obtain the combined DirAC metadata,

or

wherein the metadata converter is configured

to convert each audio object metadata of the plurality of audio object descriptions into an individual DirAC data description comprising individual DirAC metadata for each audio object description,

to convert DirAC parameters in the individual DirAC metadata into individual pressure vectors and individual velocity vectors,

to sum the individual pressure vectors to obtain a combined pressure vector and to sum the individual velocity vectors to obtain a combined individual velocity vector; and

to apply a DirAC analysis to the combined pressure vector and the combined velocity vector to obtain the combined DirAC metadata,

or

wherein the metadata converter is configured

to convert each audio object metadata of the plurality of audio object descriptions into an individual DirAC data description comprising individual DirAC metadata for each audio object description, wherein each individual DirAC metadata comprises direction of arrival metadata, and

to select a direction of arrival value among a first direction of arrival value of first individual DirAC metadata for a first audio object description of the audio object descriptions and a second direction of arrival value of second individual DirAC metadata for a second audio object description of the audio object descriptions, wherein the selected direction of arrival value is associated with a higher energy of an associated pressure signal energy and represents the combined DirAC metadata.

2. Audio data converter of claim 1, in which the audio object metadata has an object position, and wherein the DirAC metadata has a direction of arrival with respect to a reference position.

3. Audio data converter in accordance with claim 1, wherein the individual DirAC metadata additionally comprises diffuseness data, and wherein the metadata converter is configured to combine the diffuseness metadata from the individual DirAC metadata for the audio object descriptions by a weighted addition, wherein a weighting of the weighted addition is done in accordance with energies of associated pressure signal energies.

4. Audio data converter in accordance with claim 1, wherein each individual DirAC metadata comprises additionally comprises diffuseness metadata, and wherein the

metadata converter is configured to select a diffuseness value among a first diffuseness value of first individual DirAC metadata for the first audio object description and a second diffuseness value of the second individual DirAC metadata for the second audio object description, wherein the selected diffuseness value is associated with a higher energy of an associated pressure signal energy, and wherein the selected diffuseness value represents a combined diffuseness value of the combined DirAC metadata.

5. Audio data converter is accordance with one of claims 1 to 4,

wherein the input interface is configured to receive, for each audio object of the plurality of audio objects, an audio object wave form signal in addition to the audio object metadata,

wherein the audio data converter further comprises a downmixer for downmixing the audio object wave form signals of the plurality of audio objects into one or more transport channels, and

wherein the output interface is configured to transmit or store the one or more transport channels in association with the DirAC metadata.

6. Method for performing an audio data conversion, comprising:

receiving a plurality of audio object descriptions of a plurality of audio objects, an audio object description having audio object metadata,

metadata converting the audio object metadata into combined DirAC metadata (Directional Audio Coding); and

transmitting or storing the combined DirAC metadata,

wherein the metadata converting comprises:

converting each audio object metadata of the plurality of audio object descriptions into an individual DirAC data description comprising individual DirAC metadata for each audio object description, and

combining the individual DirAC metadata by individually combining direction of arrival metadata from the individual DirAC metadata for the audio object descriptions by a weighted addition, wherein a weighting of the weighted addition is done in accordance with energies of associated pressure signal energies to obtain the combined DirAC metadata,

or

wherein the metadata converting comprises:

converting each audio object metadata of the plurality of audio object descriptions into an individual DirAC data description comprising individual DirAC metadata for each audio object description,

converting DirAC parameters in the individual DirAC metadata into individual pressure vectors and individual velocity vectors,

summing the individual pressure vectors to obtain a combined pressure vector and summing the individual velocity vectors to obtain a combined individual velocity vector; and

applying a DirAC analysis to the combined pressure vector and the combined velocity vector to obtain the combined DirAC metadata,

or

wherein the metadata converting comprises:

converting each audio object metadata of the plurality of audio object descriptions into an individual DirAC data description comprising individual DirAC metadata for each audio object description, wherein each individual DirAC metadata comprises direction of arrival metadata, and

selecting a direction of arrival value among a first direction of arrival value of first individual DirAC metadata for a first audio object description of the audio object descriptions and a second direction of arrival value of second individual DirAC metadata for a second audio object description of the audio object descriptions, wherein the selected direction of arrival value is associated with

a higher energy of an associated pressure signal energy and represents the combined DirAC metadata.

7. Computer program for performing, when running on a computer or a processor, the method of claim 6.

8. Audio scene encoder, comprising:

an input interface for receiving a DirAC description of an audio scene having DirAC (Directional Audio Coding) metadata and for receiving an object signal having object metadata;

a metadata generator for generating a combined metadata description comprising the DirAC metadata and the object metadata, wherein the DirAC metadata comprises a direction of arrival for individual time-frequency tiles and the object metadata comprises a direction or additionally a distance or a diffuseness of an individual object, wherein the metadata generator comprises a metadata converter as described in any of the claims 1 to 4.

9. Audio scene encoder of claim 8, wherein the input interface is configured for receiving a transport signal associated with the DirAC description of the audio scene and wherein the input interface is configured for receiving an object wave form signal associated with the object signal, and

wherein the audio scene encoder further comprises a transport signal encoder for encoding the transport signal and the object wave form signal.

10. An audio scene encoder of one of claims 8 to 9,

wherein the metadata generator is configured to generate, for the object metadata, a single broadband direction per time and wherein the metadata generator is configured to refresh the single broadband direction per time less frequently than the DirAC metadata.

11. Method of encoding an audio scene, comprising:

receiving a DirAC (Directional Audio Coding) description of an audio scene having DirAC metadata and receiving an object signal having audio object metadata; and

generating a combined metadata description comprising the DirAC metadata and the object metadata, wherein the DirAC metadata comprises a direction of arrival for individual time-frequency tiles and wherein the object metadata comprises a direction or, additionally, a distance or a diffuseness of an individual object, wherein the generating comprises using a metadata generator comprising a metadata converter as described in any of the claims 1 to 5.

12 Computer program for performing, when running on a computer or a processor, the method of claim 11.