SPATIAL PARAMETER SIGNALLING

Abstract

An apparatus comprising means for: obtaining at least one audio signal; obtaining at least one parameter respectively for each of at least two frequency bands associated with the at least one audio signal; and selecting a frequency band of the at least two frequency bands based on comparing at least one further respective parameter for each of the at least two frequency bands wherein the at least one further respective parameter is determined from each of the at least two frequency bands; generating an output comprising a selection of the at least one parameter associated with the selected frequency band of the at least two frequency bands, such that the selection of the at least one parameter associated with the selected frequency band is configured to reduce a bitrate or size of the output.

Inventors: PIHLAJAKUJA; Tapani (Vantaa, FI), LAITINEN; Mikko-Ville (Espoo, FI)

Applicant: NOKIA TECHNOLOGIES OY (Espoo, FI)

Family ID: 63920928

Assignee: NOKIA TECHNOLOGIES OY (Espoo, FI)

Appl. No.: 19/196641

Filed: May 01, 2025

Foreign Application Priority Data

GB	1814227.3	Aug. 31, 2018

Related U.S. Application Data

parent US continuation 17270354 20210222 parent-grant-document US 12327569 WO continuation PCT/FI2019/050581 20190808 child US 19196641

Background/Summary

CROSS-REFERENCE TO RELATED APPLICATIONS [0001] This application is a continuation of U.S. application Ser. No. 17/270,354, filed Feb. 22, 2021, which is national phase entry of International Application No. PCT/FI2019/050581, filed Aug. 8, 2019, which claims priority to Great Britain Patent Application No. 1814227.3, filed Aug. 31, 2018, the entire contents of which are incorporated herein by reference.

FIELD
[0002] The present application relates to apparatus and methods for spatial parameter signalling, but not exclusively for spatial parameter signalling within and between spatial audio encoders and decoders.
BACKGROUND
[0003] Parametric spatial audio processing is a field of audio signal processing where the spatial aspect of the sound is described using a set of parameters. For example, in parametric spatial audio capture from microphone arrays, it is a typical and an effective choice to estimate from the microphone array signals a set of parameters such as directions of the sound in frequency bands, and the ratios between the directional and non-directional parts of the captured sound in frequency bands. These parameters are known to well describe the perceptual spatial properties of the captured sound at the position of the microphone array. These parameters can be utilized in synthesis of the spatial sound accordingly, for headphones binaurally, for loudspeakers, or to other formats, such as Ambisonics.
SUMMARY
[0004] There is provided according to a first aspect an apparatus comprising means for: obtaining at least one audio signal; obtaining at least one parameter respectively for each of at least two frequency bands associated with the at least one audio signal; and selecting a frequency band of the at least two frequency bands based on comparing at least one further respective parameter for each of the at least two frequency bands wherein the at least one further respective parameter is determined from each of the at least two frequency bands; generating an output comprising a selection of the at least one parameter associated with the selected frequency band of the at least two frequency bands, such that the selection of the at least one parameter associated with the selected frequency band is configured to reduce a bitrate or size of the output and wherein the at least one parameter of the selected frequency band is configured to represent respective parameters of the at least two frequency bands.
[0005] The means for obtaining at least one parameter respectively for each of at least two frequency bands associated with the at least one audio signal may be further for obtaining a direction and energy respectively for each of the at least two frequency bands associated with the at least one audio signal, and wherein the means for selecting a frequency band of the at least two frequency bands based on comparing at least one further respective parameter for each of the at least two frequency bands may be further for: determining a directional energy weight factor for each of the at least two frequency bands based on the direction and energy for each of the at least

two frequency bands, wherein the directional energy weight factor is the at least one further respective parameter for each of the at least two frequency bands; determining a weight limit factor based on an averaged energy; comparing the directional energy weight factor for each of the at least two frequency bands to the weight limit factor; and selecting a highest frequency band where the directional energy weight factor is greater than the weight limit factor.

[0006] The energy may be a normalized energy.

[0007] The means for selecting a frequency band of the at least two frequency bands based on comparing at least one further respective parameters for each of the at least two frequency bands may be further for selecting the highest frequency band of the at least two frequency bands.

[0008] The means for obtaining respectively at least one parameter for at least two frequency bands associated with the at least one audio signal may be further for obtaining at least one of: a directional parameter; a distance parameter; an energy parameter; and an energy ratio parameter.

[0009] The means for selecting a frequency band of the at least two frequency bands based on comparing at least one further respective parameter for each of the at least two frequency bands may be further for: saving the at least one parameter for one of the at least two frequency bands; and discarding any other of the at least one parameter for the at least two frequency bands, wherein the means for generating an output comprising a selection of the at least one parameter associated with the selected frequency band of the at least two frequency bands may be further for generating an output comprising the saved at least one parameter for one of the at least two frequency bands and not the discarded other of the at least one parameter for the at least two frequency bands.

[0010] The means for selecting a frequency band of the at least two frequency bands based on comparing at least one further respective parameters for each of the at least two frequency bands may be further for: saving the at least one parameter for one of the at least two frequency bands; and determining a difference between any other of the at least one parameter for the at least two frequency bands and the at least one parameter for one of the at least two frequency bands, wherein the means for generating an output comprising a selection of the at least one parameter associated with the selected frequency band of the at least two frequency bands may be further for generating an output further comprising the difference between any other of the at least one parameter for the at least two frequency bands and the at least one parameter for one of the at least two frequency bands.

[0011] The means are further for generating at least one transport signal based on the at least one audio signal and wherein the means for generating an output comprising a selection of the at least one parameter associated with the selected frequency band of the at least two frequency bands may be further for generating a datastream for storing/transmission based on a combination of the at least one parameter and the at least one transport signal.

[0012] The means for generating a datastream for storing/transmission based on a combination of the at least one parameter and the at least one transport signal may be further for: encoding the at least one transport signal; encoding the at least one parameter associated with the selected frequency band of the at least two frequency bands; and combining the encoded transport signal and the encoded at least one parameter associated with the selected frequency band of the at least two frequency bands.

[0013] The means for generating at least one transport signal based on the at least one audio signal may be further for at least one of: downmixing the at least one audio signal; selecting at least one audio signal from the at least one audio signal, when the at least one audio signal comprises two or more audio signals; generating directional signals directed to different directions, when the at least one audio signal comprises first order ambisonic audio signals; generating cardioid signals directed to different directions, when the at least one audio signal comprises first order ambisonic audio signals; generating cardioid signals directed at opposite directions, when the at least one audio signal comprises first order ambisonic audio signals; and passing at least one transport audio signal, when the at least one audio signal comprises at least one transport audio signal.

[0014] According to a second aspect there is provided an apparatus comprising means for: obtaining at least one signal, the at least one signal comprising at least one parameter associated with a selected frequency band from at least two frequency bands and at least one transport signal; replicating, based on the at least one parameter for one of the at least two frequency bands and a transport signal, at least one parameter for at least one other of the at least two frequency bands; and synthesising at least two audio signals based on the at least one parameter associated with the selected frequency band from at least two frequency bands and at least one replicated parameter for the at least one other of the at least two frequency bands and the transport signal, wherein the at least two audio signals are configured to provide spatial audio reproduction.

[0015] The means for obtaining at least one signal, the at least one signal comprising at least one parameter associated with a selected frequency band from at least two frequency bands may be further for obtaining at least one of: a directional parameter; a distance parameter; an energy parameter; and an energy ratio parameter.

[0016] The means for replicating, based on the at least one parameter for one of the at least two frequency bands, at least one parameter for at least one other of the at least two frequency bands may be further for copying the at least one parameter for one of the at least two frequency bands as the at least one other of the at least two frequency bands.

[0017] The at least one signal may further comprise at least one parameter associated with a difference between at least one other of the at least two frequency bands and the at least one parameter for one of the at least two frequency bands, wherein the means for replicating, based on the at least one parameter for one of the at least two frequency bands, at least one parameter for at least one other of the at least two frequency bands may be further for replicating the at least one parameter for at least one other of the at least two frequency bands based on a combination of the at least one parameter for one of the at least two frequency bands and the at least one parameter associated with a difference between at least one other of the at least two frequency bands and the at least one parameter for one of the at least two frequency bands.

[0018] According to a third aspect there is provided a method comprising: obtaining at least one audio signal; obtaining at least one parameter respectively for each of at least two frequency bands associated with the at least one audio signal; and selecting a frequency band of the at least two frequency bands based on comparing at least one further respective parameter for each of the at least two frequency bands wherein the at least one further respective parameter is determined from each of the at least two frequency bands; generating an output comprising a selection of the at least one parameter associated with the selected frequency band of the at least two frequency bands, such that the selection of the at least one parameter associated with the selected frequency band is configured to reduce a bitrate or size of the output and wherein the at least one parameter of the selected frequency band is configured to represent respective parameters of the at least two frequency bands.

[0019] Obtaining at least one parameter respectively for each of at least two frequency bands associated with the at least one audio signal may comprise obtaining a direction and energy respectively for each of the at least two frequency bands associated with the at least one audio signal, and wherein selecting a frequency band of the at least two frequency bands based on comparing at least one further respective parameter for each of the at least two frequency bands may comprise: determining a directional energy weight factor for each of the at least two frequency bands based on the direction and energy for each of the at least two frequency bands, wherein the directional energy weight factor is the at least one further respective parameter for each of the at least two frequency bands; determining a weight limit factor based on an averaged energy; comparing the directional energy weight factor for each of the at least two frequency bands to the weight limit factor; and selecting a highest frequency band where the directional energy weight factor is greater than the weight limit factor.

[0020] The energy may be a normalized energy.

[0021] Selecting a frequency band of the at least two frequency bands based on comparing at least one further respective parameters for each of the at least two frequency bands may comprise selecting the highest frequency band of the at least two frequency bands.

[0022] Obtaining respectively at least one parameter for at least two frequency bands associated with the at least one audio signal may comprise obtaining at least one of: [0023] a directional parameter; a distance parameter; an energy parameter; and an energy ratio parameter.

[0024] Selecting a frequency band of the at least two frequency bands based on comparing at least one further respective parameter for each of the at least two frequency bands may comprise: saving the at least one parameter for one of the at least two frequency bands; and discarding any other of the at least one parameter for the at least two frequency bands, wherein generating an output comprising a selection of the at least one parameter associated with the selected frequency band of the at least two frequency bands may comprise generating an output comprising the saved at least one parameter for one of the at least two frequency bands and not the discarded other of the at least one parameter for the at least two frequency bands.

[0025] Selecting a frequency band of the at least two frequency bands based on comparing at least one further respective parameters for each of the at least two frequency bands may comprise: saving the at least one parameter for one of the at least two frequency bands; and determining a difference between any other of the at least one parameter for the at least two frequency bands and the at least one parameter for one of the at least two frequency bands, wherein generating an output comprising a selection of the at least one parameter associated with the selected frequency band of the at least two frequency bands may comprise generating an output further comprising the difference between any other of the at least one parameter for the at least two frequency bands and the at least one parameter for one of the at least two frequency bands.

[0026] The method may further comprise generating at least one transport signal based on the at least one audio signal and wherein generating an output comprising a selection of the at least one parameter associated with the selected frequency band of the at least two frequency bands may comprise generating a datastream for storing/transmission based on a combination of the at least one parameter and the at least one transport signal.

[0027] Generating a datastream for storing/transmission based on a combination of the at least one parameter and the at least one transport signal may further comprise: encoding the at least one transport signal; encoding the at least one parameter associated with the selected frequency band of the at least two frequency bands; and combining the encoded transport signal and the encoded at least one parameter associated with the selected frequency band of the at least two frequency bands.

[0028] Generating at least one transport signal based on the at least one audio signal may further comprise at least one of: downmixing the at least one audio signal; selecting at least one audio signal from the at least one audio signal, when the at least one audio signal comprises two or more audio signals; generating directional signals directed to different directions, when the at least one audio signal comprises first order ambisonic audio signals; generating cardioid signals directed to different directions, when the at least one audio signal comprises first order ambisonic audio signals; generating cardioid signals directed at opposite directions, when the at least one audio signal comprises first order ambisonic audio signals; and passing at least one transport audio signal, when the at least one audio signal comprises at least one transport audio signal.

[0029] According to a fourth aspect there is provided a method comprising: obtaining at least one signal, the at least one signal comprising at least one parameter associated with a selected frequency band from at least two frequency bands and at least one transport signal; replicating, based on the at least one parameter for one of the at least two frequency bands and a transport signal, at least one parameter for at least one other of the at least two frequency bands; and synthesising at least two audio signals based on the at least one parameter associated with the selected frequency band from at least two frequency bands and at least one replicated parameter for

the at least one other of the at least two frequency bands and the transport signal, wherein the at least two audio signals are configured to provide spatial audio reproduction.

[0030] Obtaining at least one signal, the at least one signal comprising at least one parameter associated with a selected frequency band from at least two frequency bands may further comprise obtaining at least one of: a directional parameter; a distance parameter; an energy parameter; and an energy ratio parameter.

[0031] Replicating, based on the at least one parameter for one of the at least two frequency bands, at least one parameter for at least one other of the at least two frequency bands may be further comprise copying the at least one parameter for one of the at least two frequency bands as the at least one other of the at least two frequency bands.

[0032] The at least one signal may further comprise at least one parameter associated with a difference between at least one other of the at least two frequency bands and the at least one parameter for one of the at least two frequency bands, wherein replicating, based on the at least one parameter for one of the at least two frequency bands, at least one parameter for at least one other of the at least two frequency bands may further comprise replicating the at least one parameter for at least one other of the at least two frequency bands based on a combination of the at least one parameter for one of the at least two frequency bands and the at least one parameter associated with a difference between at least one other of the at least two frequency bands and the at least one parameter for one of the at least two frequency bands. According to a fifth aspect there is provided an apparatus comprising at least one processor and at least one memory including a computer program code, the at least one memory and the computer program code configured to, with the at least one processor, cause the apparatus at least to: obtain at least one audio signal; obtain at least one parameter respectively for each of at least two frequency bands associated with the at least one audio signal; and select a frequency band of the at least two frequency bands based on comparing at least one further respective parameter for each of the at least two frequency bands wherein the at least one further respective parameter is determined from each of the at least two frequency bands; generate an output comprising a selection of the at least one parameter associated with the selected frequency band of the at least two frequency bands, such that the selection of the at least one parameter associated with the selected frequency band is configured to reduce a bitrate or size of the output and wherein the at least one parameter of the selected frequency band is configured to represent respective parameters of the at least two frequency bands.

[0033] The apparatus caused to obtain at least one parameter respectively for each of at least two frequency bands associated with the at least one audio signal may be further be caused to obtain a direction and energy respectively for each of the at least two frequency bands associated with the at least one audio signal, and wherein the apparatus caused to select at least one frequency band of the at least two frequency bands based on comparing at least one further respective parameter for each of the at least two frequency bands may be further be caused to: determine a directional energy weight factor for each of the at least two frequency bands based on the direction and energy for each of the at least two frequency bands, wherein the directional energy weight factor is the at least one further respective parameter for each of the at least two frequency bands; determine a weight limit factor based on an averaged energy; compare the directional energy weight factor for each of the at least two frequency bands to the weight limit factor; and select a highest frequency band where the directional energy weight factor is greater than the weight limit factor.

[0034] The energy may be a normalized energy.

[0035] The apparatus caused to select at least one frequency band of the at least two frequency bands based on comparing at least one further respective parameters for each of the at least two frequency bands may be further caused to select the highest frequency band of the at least two frequency bands.

[0036] The apparatus caused to obtain respectively at least one parameter for at least two frequency bands associated with the at least one audio signal may be further caused to obtain at least one of: a

directional parameter; a distance parameter; an energy parameter; and an energy ratio parameter.

[0037] The apparatus caused to select at least one frequency band of the at least two frequency bands based on comparing at least one further respective parameter for each of the at least two frequency bands may be further caused to: save the at least one parameter for one of the at least two frequency bands; and discard any other of the at least one parameter for the at least two frequency bands, wherein the apparatus caused to generate an output comprising a selection of the at least one parameter associated with the selected frequency band of the at least two frequency bands may be further caused to generate an output comprising the saved at least one parameter for one of the at least two frequency bands and not the discarded other of the at least one parameter for the at least two frequency bands.

[0038] The apparatus caused to select at least one frequency band of the at least two frequency bands based on comparing at least one further respective parameters for each of the at least two frequency bands may be further cause to: save the at least one parameter for one of the at least two frequency bands; and determine a difference between any other of the at least one parameter for the at least two frequency bands and the at least one parameter for one of the at least two frequency bands, wherein the apparatus caused to generate an output comprising a selection of the at least one parameter associated with the selected frequency band of the at least two frequency bands may be further caused to generate an output further comprising the difference between any other of the at least one parameter for the at least two frequency bands and the at least one parameter for one of the at least two frequency bands.

[0039] The apparatus may be further caused to generate at least one transport signal based on the at least one audio signal and wherein the apparatus caused to generate an output comprising a selection of the at least one parameter associated with the selected frequency band of the at least two frequency bands may be further caused to generate a datastream for storing/transmission based on a combination of the at least one parameter and the at least one transport signal.

[0040] The apparatus caused to generate a datastream for storing/transmission based on a combination of the at least one parameter and the at least one transport signal may be further caused to: encode the at least one transport signal; encode the at least one parameter associated with the selected frequency band of the at least two frequency bands; and combine the encoded transport signal and the encoded at least one parameter associated with the selected frequency band of the at least two frequency bands.

[0041] The apparatus caused to generate at least one transport signal based on the at least one audio signal may be further caused to perform least one of: downmix the at least one audio signal; selecting at least one audio signal from the at least one audio signal, when the at least one audio signal comprises two or more audio signals; generate directional signals directed to different directions, when the at least one audio signal comprises first order ambisonic audio signals; generate cardioid signals directed to different directions, when the at least one audio signal comprises first order ambisonic audio signals; generate cardioid signals directed at opposite directions, when the at least one audio signal comprises first order ambisonic audio signals; and pass at least one transport audio signal, when the at least one audio signal comprises at least one transport audio signal.

[0042] According to a sixth aspect there is provided an apparatus comprising at least one processor and at least one memory including a computer program code, the at least one memory and the computer program code configured to, with the at least one processor, cause the apparatus at least to: obtain at least one signal, the at least one signal comprising at least one parameter associated with a selected frequency band from at least two frequency bands and at least one transport signal; replicate, based on the at least one parameter for one of the at least two frequency bands and a transport signal, at least one parameter for at least one other of the at least two frequency bands; and synthesise at least two audio signals based on the at least one parameter associated with the selected frequency band from at least two frequency bands and at least one replicated parameter for

the at least one other of the at least two frequency bands and the transport signal, wherein the at least two audio signals are configured to provide spatial audio reproduction.

[0043] The apparatus caused to obtain at least one signal, the at least one signal comprising at least one parameter associated with a selected frequency band from at least two frequency bands may be further caused to obtain at least one of: a directional parameter; a distance parameter; an energy parameter; and an energy ratio parameter.

[0044] The apparatus caused to replicate, based on the at least one parameter for one of the at least two frequency bands, at least one parameter for at least one other of the at least two frequency bands may be further caused to copy the at least one parameter for one of the at least two frequency bands as the at least one other of the at least two frequency bands.

[0045] The at least one signal may further comprise at least one parameter associated with a difference between at least one other of the at least two frequency bands and the at least one parameter for one of the at least two frequency bands, wherein the apparatus caused to replicate, based on the at least one parameter for one of the at least two frequency bands, at least one parameter for at least one other of the at least two frequency bands may be further caused to replicate the at least one parameter for at least one other of the at least two frequency bands based on a combination of the at least one parameter for one of the at least two frequency bands and the at least one parameter associated with a difference between at least one other of the at least two frequency bands and the at least one parameter for one of the at least two frequency bands.

[0046] According to a seventh aspect there is provided a computer program comprising instructions [or a computer readable medium comprising program instructions] for causing an apparatus to perform at least the following: obtaining at least one audio signal; obtaining at least one parameter respectively for each of at least two frequency bands associated with the at least one audio signal; and selecting a frequency band of the at least two frequency bands based on comparing at least one further respective parameter for each of the at least two frequency bands wherein the at least one further respective parameter is determined from each of the at least two frequency bands; generating an output comprising a selection of the at least one parameter associated with the selected frequency band of the at least two frequency bands, such that the selection of the at least one parameter associated with the selected frequency band is configured to reduce a bitrate or size of the output and wherein the at least one parameter of the selected frequency band is configured to represent respective parameters of the at least two frequency bands.

[0047] According to an eighth aspect there is provided a computer program comprising instructions [or a computer readable medium comprising program instructions] for causing an apparatus to perform at least the following: obtaining at least one signal, the at least one signal comprising at least one parameter associated with a selected frequency band from at least two frequency bands and at least one transport signal; replicating, based on the at least one parameter for one of the at least two frequency bands and a transport signal, at least one parameter for at least one other of the at least two frequency bands; and synthesising at least two audio signals based on the at least one parameter associated with the selected frequency band from at least two frequency bands and at least one replicated parameter for the at least one other of the at least two frequency bands and the transport signal, wherein the at least two audio signals are configured to provide spatial audio reproduction.

[0048] According to a ninth aspect there is provided a non-transitory computer readable medium comprising program instructions for causing an apparatus to perform at least the following: obtaining at least one audio signal; obtaining at least one parameter respectively for each of at least two frequency bands associated with the at least one audio signal; and selecting a frequency band of the at least two frequency bands based on comparing at least one further respective parameter for each of the at least two frequency bands wherein the at least one further respective parameter is determined from each of the at least two frequency bands; generating an output comprising a selection of the at least one parameter associated with the selected frequency band of the at least

two frequency bands, such that the selection of the at least one parameter associated with the selected frequency band is configured to reduce a bitrate or size of the output and wherein the at least one parameter of the selected frequency band is configured to represent respective parameters of the at least two frequency bands.

[0049] According to a tenth aspect there is provided a non-transitory computer readable medium comprising program instructions for causing an apparatus to perform at least the following: obtaining at least one signal, the at least one signal comprising at least one parameter associated with a selected frequency band from at least two frequency bands and at least one transport signal; replicating, based on the at least one parameter for one of the at least two frequency bands and a transport signal, at least one parameter for at least one other of the at least two frequency bands; and synthesising at least two audio signals based on the at least one parameter associated with the selected frequency band from at least two frequency bands and at least one replicated parameter for the at least one other of the at least two frequency bands and the transport signal, wherein the at least two audio signals are configured to provide spatial audio reproduction.

[0050] According to an eleventh aspect there is provided a computer readable medium comprising program instructions for causing an apparatus to perform at least the following: obtaining at least one audio signal; obtaining at least one parameter respectively for each of at least two frequency bands associated with the at least one audio signal; and selecting a frequency band of the at least two frequency bands based on comparing at least one further respective parameter for each of the at least two frequency bands wherein the at least one further respective parameter is determined from each of the at least two frequency bands; generating an output comprising a selection of the at least one parameter associated with the selected frequency band of the at least two frequency bands, such that the selection of the at least one parameter associated with the selected frequency band is configured to reduce a bitrate or size of the output and wherein the at least one parameter of the selected frequency band is configured to represent respective parameters of the at least two frequency bands.

[0051] According to a twelfth aspect there is provided a computer readable medium comprising program instructions for causing an apparatus to perform at least the following: obtaining at least one signal, the at least one signal comprising at least one parameter associated with a selected frequency band from at least two frequency bands and at least one transport signal; replicating, based on the at least one parameter for one of the at least two frequency bands and a transport signal, at least one parameter for at least one other of the at least two frequency bands; and synthesising at least two audio signals based on the at least one parameter associated with the selected frequency band from at least two frequency bands and at least one replicated parameter for the at least one other of the at least two frequency bands and the transport signal, wherein the at least two audio signals are configured to provide spatial audio reproduction.

[0052] According to a thirteenth aspect there is provided a computer readable medium comprising program instructions for causing an apparatus to perform at least the following: obtaining at least one audio signal; obtaining at least one parameter respectively for each of at least two frequency bands associated with the at least one audio signal; and selecting a frequency band of the at least two frequency bands based on comparing at least one further respective parameter for each of the at least two frequency bands wherein the at least one further respective parameter is determined from each of the at least two frequency bands; generating an output comprising a selection of the at least one parameter associated with the selected frequency band of the at least two frequency bands, such that the selection of the at least one parameter associated with the selected frequency band is configured to reduce a bitrate or size of the output and wherein the at least one parameter of the selected frequency band is configured to represent respective parameters of the at least two frequency bands.

[0053] According to a fourteenth aspect there is provided a computer readable medium comprising program instructions for causing an apparatus to perform at least the following: obtaining at least

one signal, the at least one signal comprising at least one parameter associated with a selected frequency band from at least two frequency bands and at least one transport signal; replicating, based on the at least one parameter for one of the at least two frequency bands and a transport signal, at least one parameter for at least one other of the at least two frequency bands; and synthesising at least two audio signals based on the at least one parameter associated with the selected frequency band from at least two frequency bands and at least one replicated parameter for the at least one other of the at least two frequency bands and the transport signal, wherein the at least two audio signals are configured to provide spatial audio reproduction.

[0054] According to a fifteenth aspect there is provided an apparatus comprising: audio signal obtaining circuitry configured to obtain at least one audio signal; parameter obtaining circuitry configured to obtain at least one parameter respectively for each of at least two frequency bands associated with the at least one audio signal; and selecting circuitry configured to select a frequency band of the at least two frequency bands based on comparing at least one further respective parameter for each of the at least two frequency bands wherein the at least one further respective parameter is determined from each of the at least two frequency bands; output generating circuitry configured to generate an output comprising a selection of the at least one parameter associated with the selected frequency band of the at least two frequency bands, such that the selection of the at least one parameter associated with the selected frequency band is configured to reduce a bitrate or size of the output and wherein the at least one parameter of the selected frequency band is configured to represent respective parameters of the at least two frequency bands.

[0055] According to a sixteenth aspect there is provided an apparatus comprising: signal obtaining circuitry configured to obtain at least one signal, the at least one signal comprising at least one parameter associated with a selected frequency band from at least two frequency bands and at least one transport signal; replicating circuitry configured to replicate, based on the at least one parameter for one of the at least two frequency bands and a transport signal, at least one parameter for at least one other of the at least two frequency bands; and synthesising circuitry configured to synthesise at least two audio signals based on the at least one parameter associated with the selected frequency band from at least two frequency bands and at least one replicated parameter for the at least one other of the at least two frequency bands and the transport signal, wherein the at least two audio signals are configured to provide spatial audio reproduction.

[0056] An apparatus comprising means for performing the actions of the method as described above.

[0057] An apparatus configured to perform the actions of the method as described above.

[0058] A computer program comprising program instructions for causing a computer to perform the method as described above.

[0059] A computer program product stored on a medium may cause an apparatus to perform the method as described herein.

[0060] An electronic device may comprise apparatus as described herein. A chipset may comprise apparatus as described herein.

[0061] Embodiments of the present application aim to address problems associated with the state of the art.

Description

SUMMARY OF THE FIGURES

[0062] For a better understanding of the present application, reference will now be made by way of example to the accompanying drawings in which:

[0063] FIG. **1** shows schematically a system of apparatus suitable for implementing some

embodiments;

[0064] FIG. **2** shows a flow diagram of the operation of the system as shown in FIG. **1** according to some embodiments;

[0065] FIG. **3** shows schematically capture/encoding apparatus;

[0066] FIG. **4** shows a flow diagram of the operation of capture/encoding apparatus as shown in FIG. **3**;

[0067] FIG. **5** shows schematically capture/encoding apparatus according to some embodiments;

[0068] FIG. **6** shows a flow diagram of the operation of capture/encoding apparatus as shown in FIG. **5** according to some embodiments;

[0069] FIG. **7** shows a flow diagram of the operation of encoding apparatus encoding obtained transport signals and metadata according to some embodiments;

[0070] FIG. **8** shows a flow diagram of the band selection operation of capture/encoding apparatus as shown in FIG. **5** according to some embodiments; and

[0071] FIG. **9** shows schematically shows schematically an example device suitable for implementing the apparatus shown.

EMBODIMENTS OF THE APPLICATION

[0072] The following describes in further detail suitable apparatus and possible mechanisms for the provision of effective spatial analysis derived metadata parameters associated with energy ratios for microphone array and other input format audio signals.

[0073] Apparatus has been designed to transmit a spatial audio modelling of a sound field using Q (which is typically 2) transport audio signals and spatial metadata. The transport audio signals are typically compressed with a suitable audio encoding scheme (for example advanced audio coding—AAC or enhanced voice services—EVS codecs). The spatial metadata may contain parameters such as Direction (for example azimuth, elevation) in time-frequency domain.

[0074] Furthermore other parameters which may be determined and signalled to a renderer or receiver is one or more direct-to-total energy ratios (in the time-frequency domain) which represents the distribution of energy between each specific direction and the total audio energy. Another parameter may be one (or more where practical) diffuse-to-total energy ratio (in the time-frequency domain) which represents distribution of energy between ambient or diffuse signal (i.e., non-directional signal such as reverberation) and total energy.

[0075] The parametric spatial audio signals may be represented as Q channels+metadata. This format can be compressed in encoding to efficiently store it for later retrieval or transmit it over a suitable transmission channel. Various methods can be used depending on how the channels are configured and what the metadata contains. A common procedure is to define a constant bitrate budget for the whole bitstream that contains audio channels and the metadata. This bitrate budget can then be divided statically or adaptively (dynamically) between audio channels and metadata.

[0076] For example, a bitrate budget of 64 kb/s for 2-channels+metadata could be used in various ways. Using the full 64 kb/s for the 2 audio channels would offer very good quality for encoding the stereo signal (for example using an EVS codec), but in this example the metadata would not be transmitted. In using 56 kb/s for the audio and 8 kb/s for metadata would usually provide a higher overall quality as the difference in audio coding quality is not large but the signalled metadata can provide full 3d surround reproduction.

[0077] With lower bitrates, dividing the bitrate budget becomes even more difficult. For example with a 16 kb/s budget, there may be the following coding modes: [0078] One channel audio 16 kb/s [0079] One channel audio 15 kb/s+1 kb/s metadata [0080] One channel audio 11 kb/s+5 kb/s metadata [0081] Two channel audio 16 kb/s [0082] Two channel audio 15 kb/s+1 kb/s metadata [0083] Two channel audio 11 kb/s+5 kb/s metadata

[0084] Optimizing between these example modes may require listening experiments. However, previous experiments have shown that with such low bitrates offering more bitrate to the raw audio quality over multiple channels tends to offer better perceived quality. The effect of metadata bitrate

budgeting is that reducing the metadata bitrate such that the audio signal receives at least 90% of the total bitrate budget is believed to be a good target.

[0085] However the amount of metadata generated and therefore the amount of data defining spatial parameters is frequency band related. For example, for B (e.g., 5, 10, 20, or 30) frequency bands and two parameters (direction and energy ratio) for each time frame, there may be at minimum 2*B*K (K is number of bits per parameter) bits of metadata per time frame. Assuming the common number of 50 frames per second, B=5, and K=10 there may be 5 kb/s metadata generated. With low bitrate applications (such as IVAS) the total target bitrate with audio can be so low as 14 kb/s so the metadata would take a big portion of the bitrate budget even after entropy coding (which may reduce the bitrate to half of the generated total).

[0086] Currently, attempts to reduce the generated include reducing bit accuracy per parameter or even removing less important parameters when the bitrate budget is low. Another approach is to reduce the number of frequency bands for metadata, for example generating just one parameter per timeframe and thus producing a reduction of generated metadata by B. One method for achieving this is to perform a wideband analysis (in other words assume only one frequency band for the full audible frequency range) and encode this wideband group.

[0087] The concept as discussed herein attempts to improve on these methods and in particular, instead of wideband analysis, attempts to: [0088] require a single analysis system for different bitrates (and therefore not one band analysis for low bitrates, multiple band analysis for high bitrates); and [0089] improve for the sound scene time-frequency resolution in a practical manner suitable for the human hearing range.

[0090] Thus, the concept as discussed in further detail in the embodiments herein implements an analysis system with multiple bands and then selects the best frequency band to represent the current time frame.

[0091] The embodiments discussed herein therefore attempt to reduce the bitrate by selecting one frequency band from the analysed metadata to represent all frequency bands. This reduces bitrate usage by factor of B (where B is the original number of frequency bands). The selection process in some embodiments may thus relate to audio encoding and decoding using a sound-field related parametrization (e.g., direction(s) and direct-to-total energy ratio(s) in frequency bands) where a solution is provided for automatically reducing the bitrate of the direction parameters by transmitting only one direction value for all frequency bands and where the transmitted one direction value is determined by: [0092] obtaining audio signals; [0093] determining (spatial parameters) directions and direct-to-total energy ratios in frequency bands; [0094] determining normalized energy in frequency bands; [0095] determining directional energy weight factor (e.g., energy multiplied by direct-to-total energy ratio); [0096] determining the highest frequency band with directional energy weight factor above a threshold; [0097] encoding/storing/transmitting only the direction of the determined band. [0098] This may be further expanded or detailed as analysis apparatus configured to: Obtain multichannel audio signals (for example Capture spatial audio signals); Apply time-frequency transform to the multichannel audio signals; [0099] Perform spatial analysis for the transformed signal; [0100] Calculate normalized energy for each frequency band for the transformed signal; [0101] Calculate frequency band weight factor for each band (energy multiplied with energy ratio) for the transformed signal; [0102] Choose or select a highest band that has a weight factor over defined limit (e.g., 0.5); [0103] Discard other metadata and save only the metadata for the chosen frequency band; [0104] Create transport signals; [0105] Encode and transmit/store transport signals and metadata.

[0106] With respect to the synthesis apparatus it is then configured to: [0107] Obtain (receive/retrieve) the transmitted/stored transport signals and metadata; [0108] replicate the selected/chosen metadata to all frequency bands; and Synthesize output using transport signals and replicated metadata.

[0109] The directions and the direct-to-total energy ratios can be estimated using any suitable

method (e.g., SPAC), and depends on the type of the audio signals (e.g., microphone-array, Ambisonics, multichannel audio signals).

[0110] The normalized energy can be estimated as discussed in the embodiments herein in a suitable manner. For example by computing the sum of squares of the frequency-domain samples and dividing with the largest energy.

[0111] The threshold value may in some embodiments be determined for example by multiplying the average normalized energy by a factor.

[0112] In addition to the direction, also all other parameters (e.g., direct-to-total energy ratios) may be encoded using the same scheme. In other words transmitting only one parameter value for all frequency bands. The value to be transmitted can be selected using the same procedure.

[0113] The decoding can be performed using any suitable method for example by using the same parameter value at all frequency bands.

[0114] In some embodiments, in encoding, the selected frequency band can be used as a reference band and a very low bitrate difference coding related to it determined for other bands.

[0115] With respect to FIG. **1** an example apparatus and system for implementing embodiments of the application are shown. The system **171** is shown with an 'analysis' part **121** and a 'synthesis' part **131**. The 'analysis' part **121** is the part from receiving the input (multichannel loudspeaker, microphone array, ambisonics, or mobile device capture) audio signals **100** up to an encoding of the metadata and transport signal **102** which may be transmitted or stored **104**. The 'synthesis' part **131** may be the part from a decoding of the encoded metadata and transport signal **104** to the presentation of the synthesized signal (for example in multi-channel loudspeaker form **106** via loudspeakers **107** or binaural or ambisonic formats).

[0116] The input to the system **171** and the 'analysis' part **121** is therefore audio signals **100**. These may be suitable input multichannel loudspeaker audio signals, microphone array audio signals, ambisonic audio signals, or mobile captured audio signals.

[0117] The input audio signals **100** may be passed to an analysis processor **101**. The analysis processor **101** may be configured to receive the input audio signals and generate a suitable data stream **104** comprising suitable transport signals. The transport audio signals may also be known as associated audio signals and be based on the audio signals. For example in some embodiments the transport signal generator **103** is configured to downmix or otherwise select or combine, for example, by beamforming techniques the input audio signals to a determined number of channels and output these as transport signals. In some embodiments the analysis processor is configured to generate a 2-audio-channel output of the microphone array audio signals. The determined number of channels may be two or any suitable number of channels.

[0118] In some embodiments the analysis processor is configured to pass the received input audio signals **100** unprocessed to an encoder in the same manner as the transport signals. In some embodiments the analysis processor **101** is configured to select one or more of the microphone audio signals and output the selection as the transport signals **104**. In some embodiments the analysis processor **101** is configured to apply any suitable encoding or quantization to the transport audio signals.

[0119] In some embodiments the analysis processor **101** is also configured to analyse the input audio signals **100** to produce metadata associated with the input audio signals (and thus associated with the transport signals). The analysis processor **101** can, for example, be a computer (running suitable software stored on memory and on at least one processor), mobile device, or alternatively a specific device utilizing, for example, FPGAs or ASICs. As shown herein in further detail the metadata may comprise, for each time-frequency analysis interval, at least one direction parameter and at least one energy ratio parameter. The at least one direction parameter and the at least one energy ratio parameter may in some embodiments be considered to be spatial audio parameters. In other words the spatial audio parameters comprise parameters which aim to characterize the sound-field of the input audio signals.

[0120] In some embodiments the parameters generated may differ from frequency band to frequency band and may be dependent on the transmission bit rate. Thus for example in band X all of the parameters are generated and transmitted, whereas in band Y only one of the parameters is generated and transmitted, and furthermore in band Z any other number of parameters are generated or transmitted. A practical example of this may be that for some frequency bands such as the highest band some of the parameters are not required for perceptual reasons.

[0121] The transport signals and the metadata **102** may be transmitted or stored, this is shown in FIG. **1** by the dashed line **104**. Before the transport signals and the metadata are transmitted or stored they may in some embodiments be coded in order to reduce bit rate, and multiplexed to one stream. The encoding and the multiplexing may be implemented using any suitable scheme.

[0122] In the decoder side **131**, the received or retrieved data (stream) may be input to a synthesis processor **105**. The synthesis processor **105** may be configured to demultiplex the data (stream) to coded transport and metadata. The synthesis processor **105** may then decode any encoded streams in order to obtain the transport signals and the metadata.

[0123] The synthesis processor **105** may then be configured to receive the transport signals and the metadata and create a suitable multi-channel audio signal output **106** (which may be any suitable output format such as binaural, multi-channel loudspeaker or Ambisonics signals, depending on the use case) based on the transport signals and the metadata. In some embodiments with headphone or loudspeaker reproduction, an actual physical sound field is reproduced (using the output device **107** for example loudspeakers/headphones etc.) having the desired perceptual properties. In other embodiments, the reproduction of a sound field may be understood to refer to reproducing perceptual properties of a sound field by other means than reproducing an actual physical sound field in a space. For example, the desired perceptual properties of a sound field can be reproduced over headphones using the binaural reproduction methods as described herein. In another example, the perceptual properties of a sound field could be reproduced as an Ambisonic output signal, and these Ambisonic signals can be reproduced with Ambisonic decoding methods to provide for example a binaural output with the desired perceptual properties.

[0124] The synthesis processor **105** can in some embodiments be a computer (running suitable software stored on memory and on at least one processor), mobile device, or alternatively a specific device utilizing, for example, FPGAs or ASICs.

[0125] With respect to FIG. **2** an example flow diagram of the overview shown in FIG. **1** is shown.

[0126] First the system (analysis part) is configured to receive input audio signals or suitable multichannel input as shown in FIG. **2** by step **201**.

[0127] Then the system (analysis part) is configured to generate a transport signal channels or transport signals (for example downmix/selection/beamforming based on the multichannel input audio signals) as shown in FIG. **2** by step **203**.

[0128] Also the system (analysis part) is configured to analyse the audio signals to generate metadata: Directions; Energy ratios as shown in FIG. **2** by step **205**.

[0129] The system is then configured to (optionally) encode for storage/transmission the transport signals and metadata as shown in FIG. **2** by step **207**.

[0130] After this the system may store/transmit the transport signals and metadata as shown in FIG. **2** by step **209**.

[0131] The system may retrieve/receive the transport signals and metadata as shown in FIG. **2** by step **211**.

[0132] Then the system is configured to extract from the transport signals and metadata as shown in FIG. **2** by step **213**.

[0133] The system (synthesis part) is configured to synthesize an output spatial audio signals (which as discussed earlier may be any suitable output format such as binaural, multi-channel loudspeaker or Ambisonics signals, depending on the use case) based on extracted audio signals and metadata as shown in FIG. **2** by step **215**.

[0134] With respect to FIG. **3** an example analysis processor **101** is shown where the input audio signal is provided from an audio source **301** which in this example is a spatial capture device configured to generate multichannel audio signals from multiple microphones. The multichannel audio signals in this example are passed to a transport (audio) signal generator **311**. The transport signal generator **311** is configured to generate the transport audio signals according to any of the options described previously. For example the transport signals may be downmixed from the input signals. The number of the transport audio signals may be any number and may be 2 or more or fewer than 2.

[0135] In the example shown in FIG. **3** the multichannel audio signals are also input to a time frequency transform **303**. The time frequency transform **303** may be configured to generate suitable time-frequency representations of the multichannel audio signals and pass these to a frequency band processor **307**.

[0136] The frequency band processor **305** is configured to generate spatial metadata outputs such as shown as the directions, direct-to-total energy ratios, and in some embodiments other types of energy ratios such as diffuse-to-total energy ratio(s) and remainder-to-total energy ratio(s).

[0137] The implementation of the analysis may be any suitable implementation that produces the described metadata outputs.

[0138] Thus for example in some embodiments the frequency band processor **305** comprises a direction analyser **307** configured to generate the direction metadata and an energy ratio analyser **309** configured to generate the energy ratio metadata.

[0139] The direction and energy ratio metadata for all of the analysed frequency bands may then be passed to a transmission/storage encoder **313**. The transmission/storage encoder **313** may be configured to combine and encode the transport signals, the directions, and the energy ratios to generate the data stream **102**.

[0140] For example in some embodiments the transmission/storage encoder **313** may comprise a suitable transport signal compressor/encoder configured to compress the audio signals using a suitable codec (e.g., AAC or EVS).

[0141] With respect to FIG. **4** is shown a flow diagram of the operation of the analysis processor.

[0142] The first operation is one of receiving the (multichannel loudspeaker or other) audio signals as shown in FIG. **4** by step **401**.

[0143] In some embodiments the audio signals are processed in some form to generate the transport audio signals as shown in FIG. **4** by step **403**.

[0144] The following operation may be one of spatially analysing the (multichannel loudspeaker) signals in order to determine direction metadata as shown in FIG. **4** by step **405**.

[0145] Then the energy ratios (for example the direct, diffuse and remainder energy ratios) are determined as shown in FIG. **4** by step **407**.

[0146] In some embodiments the metadata and transport audio signals are processed (compressed/encoded). For example the number of the directions and ratios are furthermore controlled (and may be selected and/or combined). The processing of the metadata/transport audio signals is shown in FIG. **4** by step **409**.

[0147] The processed transport audio signals and the metadata may then be furthermore be combined to generate a suitable data stream as shown in FIG. **4** by step **411**.

[0148] With respect to FIG. **5** there is shown an example analysis processor **101** suitable for implementing some embodiments with additions over the example provided in FIG. **3**.

[0149] The example analysis processor **101** is shown again with the input audio signal provided from an audio source **301** which also in this example is a spatial capture device configured to generate multichannel audio signals from multiple microphones. For example capturing a spatial audio signal can be performed with any known capture device. For example an Eigenmike or Nokia 8 mobile phone are suitable. As described previously the multichannel (spatial) audio signal may be any format such as mixed content (e.g., a multichannel audio format such as 5.1) and Ambisonics

content that may produce the relevant spatial audio parameters.

[0150] The multichannel audio signals in this example are passed to a transport (audio) signal generator **311**.

[0151] The transport signal generator **311** similar to the example in FIG. **3** is configured to generate the transport audio signals according to any of the options described previously. For example the transport signals may be downmixed from the input signals. The number of the transport audio signals may be any number and may be 2 or more or fewer than 2.

[0152] In the example shown in FIG. **5** the multichannel audio signals are also input to a time frequency transform **303**. The time frequency transform **303** may be configured to generate suitable time-frequency representations of the multichannel audio signals and pass these to a frequency band processor **505**.

[0153] The frequency band processor **505** is configured to generate spatial metadata outputs such as shown as the directions, direct-to-total energy ratios, and in some embodiments other types of energy ratios such as diffuse-to-total energy ratio(s) and remainder-to-total energy ratio(s).

[0154] The implementation of the analysis may be any suitable implementation that produces the described metadata outputs.

[0155] Thus for example in some embodiments the frequency band processor **505** comprises a direction analyser **307** configured to generate the direction metadata and an energy ratio analyser **309** configured to generate the energy ratio metadata.

[0156] These may be determined by performing spatial analysis on the time-frequency transformed multichannel audio signal.

[0157] An example of spatial analysis may be for example DirAC (Directional Audio Coding) spatial analysis.

[0158] DirAC may estimate the directions and diffuseness ratios (equivalent information to a direct-to-total ratio parameter) from a first-order Ambisonic (FOA) signal, or its variant the B-format signal.

$$[00001] \mathrm{FOA}_i(t) = \begin{bmatrix} w_i(t) \\ x_i(t) \\ y_i(t) \\ z_i(t) \end{bmatrix}$$

[0159] The signals of $\Sigma.\mathrm{sub}.i=1.\mathrm{sup}.\mathrm{NUM\ CRFOA.sub}.i(t)$ are transformed into frequency bands for example by STFT, resulting in time-frequency signals $w(k,n)$, $x(k,n)$, $y(k,n)$, $z(k,n)$ where k is the frequency bin index and n is the time index. DirAC estimates the intensity vector by

$$[00002] I(k,n) = -\mathrm{Re}\left\{w(k,n) .\mathrm{Math.} \begin{bmatrix} x(k,n) \\ y(k,n) \\ z(k,n) \end{bmatrix}\right\},$$ [0160] where Re means real part, and asterisk * means complex conjugate.

[0161] The direction parameter is opposite of the direction of the real part of the intensity vector. The intensity vector may be averaged over several time and/or frequency indices prior to the determination of the direction parameter.

[0162] DirAC determines the diffuseness as

$$[00003] \quad (k,n) = 1 - \frac{.\mathrm{Math.}\ E[I(k,n)]\ .\mathrm{Math.}}{E[0.5(w^2(k,n) + x^2(k,n) + y^2(k,n) + z^2(k,n))]}$$

[0163] Diffuseness is a ratio value that is 1 when the sound is fully ambient, and 0 when the sound is fully directional. Again, all parameters in the equation are typically averaged over time and/or frequency. The expectation operator E[ ] can be replaced with an average operator in practical systems.

[0164] When averaged, the diffuseness (and direction) parameters typically are determined in

frequency bands combining several frequency bins k, for example, approximating the Bark frequency resolution.

[0165] DirAC, as determined above, is only one of the options to determine the directional and ratio metadata, and clearly one may utilize other methods to determine the metadata, for example, using a spatial audio capture (SPAC) algorithm with microphone-array signals (real or simulated). Furthermore, there are also many variants of DirAC analysis in the literature. For example where the input content is not FOA, a suitable modification can be done to convert the signal into FOA-format to perform analysis. Other analysis methods are also applicable as long as they produce the directional and energy ratio metadata.

[0166] The direction and energy ratio metadata for all of the analysed frequency bands may then be passed to a metadata selector **521**.

[0167] Furthermore the output of the energy ratio analyser **309** is output to a weight factor determiner **517**.

[0168] Furthermore the frequency band processor **505** comprises a normalised energy determiner **515** configured to generate a normalised energy determination and pass this to a weight factor determiner **517** and to a weight limit determiner **519**.

[0169] In some embodiments the normalised energy determination may be performed as a two step operation. A first step being to calculate the average energy for each frequency band in this time instant for example with the following equation:

$$[00004] E_{\text{avg}} = \frac{1}{NKI} \sum_{n=1}^{N} \sum_{k=K_b}^{K_r} \sum_{i=1}^{I} |S(i,k,n)|^2$$

[0170] where N is number of time samples in this time frame, K_b and K_t are the current frequency band bottom and top frequency bins, and I is the number of input channels in the signal. $S(i,k,n)$ is the time-frequency domain representation of the transport signal.

[0171] The second step may be to normalize the average energies of each frequency band so that the largest energy of any frequency band is found and then divide all [0172] energies with the largest energy value. This may be seen as the largest energy of a frequency band is (always) 1 and other frequency bands have less energy or represented as an equation as:

$$[00005] E_{\text{norm}}(i) = \frac{E_{\text{avg}}(i)}{\max(E_{\text{avg}})}$$

[0173] In some embodiments any suitable alternative normalization methods may be employed (e.g., normalizing with total energy instead of largest energy) and can be used but the limit parameter (as discussed hereafter) is appropriately tuned. In addition, in some embodiments unnormalized energy may be employed but the limit parameter requires even more careful tuning.

[0174] The frequency band processor **505** in some embodiments further comprises a weight factor determiner **517** configured to receive the normalised energy and the energy ratios and determine at least one weighting factor which is output to the metadata selector **521**.

[0175] With the normalized energy known, the weight factor may be determined by based on the product of energy ratio and the normalized energy in the frequency band. The weight factor may therefore be determined by the equation:

$$w = r E_{\text{norm}}$$

[0176] where r is the energy ratio parameter.

[0177] This weight factor is a number between 0 and 1. It will be a very high value when there is a directional impulsive onset present in the scene as both energy ratio and normalized energy will be high. Likewise, if there is no onset present, these values tend to be lower for higher frequencies. The use of the product ensures that, for example, high normalized energy but low energy ratio (i.e., loud reverberation) does not produce high weight values as the direction and the metadata in this case is not the best representative.

[0178] In some embodiments, this weight factor can be any other suitable weight factor such as only the energy ratio parameter r.

[0179] The analysis processor **101** in some embodiments comprises a weight limit determiner **519** configured to receive the normalised energy determination and output a weight limit value to the metadata selector **521**.

[0180] The weight limit can be a constant value (e.g., 0.5) or it can be based on the average normalized energy of all frequency bands in the time frame (e.g., average [0181] normalized energy multiplied with a constant like 0.5). The latter option is preferred and is formed as:

[00006] $w_{\text{thr}} = \frac{c}{B} . \underset{i=1}{\overset{B}{\text{Math.}}} E_{\text{norm}}(i)$ [0182] where c is tuned threshold constant such as 0.5 and B is the total number of frequency bands.

[0183] In some embodiments, this weight limit can be any other suitable value.

[0184] The analysis processor **101** in some embodiments comprises a metadata selector **521** configured to receive the output of the direction analyser **307** (direction metadata for each band), energy ratio analyser **309** (energy ratio metadata for each band), weight factor determiner **517** (weight factors) and weight limit determiner **519**. The metadata selector **521** is then configured to select one of the directions and energy ratios based on the weight factor and weight factor limit and pass the selected metadata to a transmission/storage encoder **513**.

[0185] The metadata selector may be configured to choose or select the highest frequency band that has a weight factor over the weight limit. If for some reason no band has weight over the limit, the metadata selector in some embodiments is configured to select the lowest frequency band.

[0186] In some embodiments once the metadata selector determines the selected frequency band, it may be configured to discard metadata associated with the other bands.

[0187] In some embodiments the metadata selector is configured to prioritize and only discard part of the metadata. For example, in some embodiments the direction information for the other bands are discarded but the energy ratio parameters are kept for all frequency bands.

[0188] In some embodiments, two or more frequency bands (but fewer than the total number of frequency bands) are selected to represent the other frequency bands. For example, two frequency bands can be selected such that two (or N where N is less than the total number of frequency bands) highest frequency bands with weights over the threshold (or weight limit) are selected. The parameters associated with the selected higher frequency band is then used to represent parameters for frequency bands above it, and parameters associated with the lower frequency band is used to represent parameters for frequency bands below it, and both are used to represent frequency bands between them.

[0189] In some embodiments the 'best' frequency band is selected but a difference coding technique is employed to represent the other frequency bands.

[0190] For example for each frequency band: [0191] Direction may be coded separately for azimuth and elevation [0192] Azimuth has 2 bits and represents offsets of 0°, 90°, 180°, or 270° from the chosen band azimuth [0193] Elevation has 2 bits and represents offsets of 0°, 45°, and −45° (one value not used) [0194] Each ratio parameter has 2 bits and represents offsets of 0, 0.25, −0.25, −0.5

[0195] In some embodiments, a few bits are used to signal which frequency band is the reference band for the difference coding. Using this method still significantly reduces the bitrate but offers more accurate representation.

[0196] In some embodiments the highest frequency band is selected and the metadata associated with the highest frequency band is used to 'represent' all frequency bands. This is less optimal in quality but is computationally more efficient to implement.

[0197] The analysis processor **101** may further comprise a transmission/storage encoder **513**. The transmission/storage encoder **513** may be configured to combine and encode the transport signals, the selected direction, and the energy ratio to generate the data stream **102**.

[0198] For example in some embodiments the transmission/storage encoder **513** may comprise a suitable transport signal compressor/encoder configured to compress the audio signals using a

suitable codec (e.g., AAC or EVS) and encoding metadata using entropy coding methods (e.g., codebook coding).

[0199] With respect to FIG. **6** is shown a flow diagram of the operation of the analysis processor shown in FIG. **5** (and additionally the synthesis processor shown in FIG. **1**).

[0200] The first operation is one of obtaining the (multichannel loudspeaker or other) audio signals as shown in FIG. **6** by step **601**.

[0201] The audio signals may be processed by the application of a time-frequency transform as shown in FIG. **6** by step **603**.

[0202] In some embodiments the time-frequency domain audio signals are processed in some form to generate the transport signals as shown in FIG. **6** by step **617**.

[0203] Furthermore in some embodiments the time-frequency domain audio signals are processed and spatial analysis performed to determine parameters such as direction(s) (and/or distance) and energy ratio(s) for each band as shown in FIG. **6** by step **607**.

[0204] Additionally in some embodiments the time-frequency domain audio signals are processed and a normalised energy per band calculated as shown in FIG. **6** by step **605**.

[0205] Having determined the normalised energy per band and spatial analysis then in some embodiments the weight factor per band is formed or determined as shown in FIG. **6** by step **609**.

[0206] Also having determined the normalised energy per band in some embodiments the weight factor limit is formed or determined as shown in FIG. **6** by step **611**.

[0207] Based on the weight factor per band and the weight factor limit a highest band with a weight over the limit is chosen as shown in FIG. **6** by step **613**.

[0208] The other metadata is then discarded and the chosen band metadata saved as shown in FIG. **6** by step **615**.

[0209] The selected metadata and transport signals are then compressed/encoded (and combined) before being stored and/or transmitted as shown in FIG. **6** by step **619**.

[0210] With respect to the synthesis processor operations the transmitted/retrieved signal is decoded and metadata replicated for all frequency bands as shown in FIG. **6** by step **621**.

[0211] Then a suitable spatial synthesis is performed as shown in FIG. **6** by step **623**.

[0212] As described previously the audio signal input format may be any suitable format. For example with respect to FIG. **7** is shown a flow diagram of the operation of an encoder suitable to encoding an obtained transport audio signal and metadata. In such an embodiment the frequency band processor may comprise only the normalised energy determiner and weight factor determiner as the direction and energy ratios have been determined.

[0213] The first operation is one of obtaining the transport audio signals and metadata as shown in FIG. **7** by step **701**.

[0214] In this example the parameters such as direction(s) (and/or distance) and energy ratio(s) for each band have been obtained and a normalised energy per band calculated as shown in FIG. **7** by step **705**.

[0215] Having determined the normalised energy per band and spatial analysis then in some embodiments the weight factor per band is formed or determined as shown in FIG. **7** by step **709**.

[0216] Also having determined the normalised energy per band in some embodiments the weight factor limit is formed or determined as shown in FIG. **7** by step **711**.

[0217] Based on the weight factor per band and the weight factor limit a highest band with a weight over the limit is chosen as shown in FIG. **7** by step **713**.

[0218] The other metadata is then discarded and the chosen band metadata saved as shown in FIG. **7** by step **715**.

[0219] The selected metadata and transport signals are then compressed/encoded (and combined) before being stored and/or transmitted as shown in FIG. **7** by step **719**.

[0220] With respect to the synthesis processor operations the transmitted/retrieved signal is decoded and metadata replicated for all frequency bands as shown in FIG. **7** by step **721**. Then a

suitable spatial synthesis is performed as shown in FIG. **7** by step **723**.

[0221] With respect to FIG. **8** an example operation of the metadata selector is shown in further detail. The first operation is to start and receive the inputs such as weight factors, weight limits, and parameters as shown in FIG. **8** by step **801**.

[0222] The next operation is setting an index i=B as shown in FIG. **8** by step **803**. The next operation is testing the index weight factor Wi against the weight limit W.sub.thr as shown in FIG. **8** by step **803**.

[0223] If W.sub.i>W.sub.thr then the next operation is determining i is the selected frequency band as shown in FIG. **8** by step **809** and then ending the operation as shown in FIG. **8** by step **813**.

[0224] If W.sub.i is not >W.sub.thr then the next operation is decrementing i by 1 as shown in FIG. **8** by step **807**.

[0225] Having decremented i by 1 then the next operation is checking whether i=1 as shown in FIG. **8** by step **811**.

[0226] Where i=1 then the next operation is determining i is the selected frequency band as shown in FIG. **8** by step **809** and then ending the operation as shown in FIG. **8** by step **813**.

[0227] Where i is not=1 then the operation may then test the new index, index weight factor Wi against the weight limit W.sub.thr as shown in FIG. **8** by step **803** and the process may continue until W.sub.i>W.sub.thr for the index or the index=1.

[0228] The above assumes that frequency band indexing starts from 1. The above can be modified to accommodate any other indexing system (such as starting from 0).

[0229] With respect to the synthesis processor the single band metadata values may be obtained and then replicated for all frequency bands. This results in a normal full set of metadata that can be used in further synthesis.

[0230] The synthesis operation may then use the transport signals and replicated metadata to generate a suitable rendering of the audio signals. This procedure can be performed using any suitable means, for example, with methods such as DirAC based spatial audio signal synthesis. An example procedure for synthesising audio signals for loudspeakers is that the directions are synthesized into specific directions using 3D panning techniques such as vector-base amplitude panning (VBAP) multiplied with $\sqrt{\text{square root over } (r)}$, and non-directional ambient signal is decorrelated with a phase-scrambling filter and reproduced to all directions multiplied with [00007]$\sqrt{\frac{r}{C}}$, where r is the energy ratio parameter and C is the number of loudspeaker channels. In such a manner some embodiments may be implemented which reduce bitrate usage while offering quality that is in many cases at least reasonable and can be almost transparent (and in many cases is) to full metadata transmission with many signals. Bitrate reduction with the primary method is by factor of B, where B is the original number of frequency bands. I.e., if original metadata bitrate is 5 kb/sand B=5 then this method achieves bitrate of 1 kb/s.

[0231] Furthermore such embodiments may be able to produce a signal which is at least as good or better than using single wideband parametric analysis.

[0232] Additionally in some embodiments the implementation is computationally efficient method to reduce bitrate as it only requires a determination of the energies (this is often part of the analysis already) and weight factors and then discard data.

[0233] In some embodiments spatial sound transmission storage can be achieved even at very low bitrates.

[0234] For example a teleconference system may use a parametric spatial audio, e.g., DirAC, as the main analysis and synthesis method. Spatial capture may be obtained with an Eigenmike that produces first-order Ambisonics for this use. The spatial audio is analysed in time-frequency (20 ms frame and 30 frequency bands) domain and produces direction parameters as azimuth and elevation, and energy ratio parameter in form of diffuseness. Rather than encoding these parameters using a determined number of bits per parameter, i.e., 8 bits, to produce metadata at a bitrate of 36

kb/s (before other compression) the application of some embodiments may result in a bitrate of just 1.2 kb/s for the metadata (before other compression). This leaves more bits to use for the coding of the audio signal which directly results in better perceived audio quality.

[0235] A further example would be using time-frequency resolution such as 10 ms time frame and 12 frequency bands would result in following comparison bitrates. 24 kb/s compared to 2.4 kb/s according to some embodiments.

[0236] As the reduction in bitrate of metadata is quite large, it especially benefits the use case where the bitrate budget is very low. For example, 24 kb/s is usually in the domain of mono downmix or very compressed stereo if only raw audio encoding is used. If spatial metadata is introduced using, for example, the second time-frequency resolution above, the full spatial metadata would be hard to fit to the bitrate budget even after expected 50% entropy coding for it (metadata would take 12 kb/s of 24 kb/s available). However, using the presented embodiments it may be possible to reduce the metadata down to a fifth and in this case we achieve very reasonable division of bitrate after entropy coding (1.2 kb/s for metadata, 22.8 kb/s for audio) thus offering full spatial audio even at low bitrates instead of mono or stereo. This means that at low bitrates, it may be possible to achieve a significant sound quality increase compared to sending full metadata.

[0237] With respect to FIG. **9** an example electronic device which may be used as the analysis or synthesis device is shown. The device may be any suitable electronics device or apparatus. For example in some embodiments the device **1900** is a mobile device, user equipment, tablet computer, computer, audio playback apparatus, etc.

[0238] In some embodiments the device **1900** comprises at least one processor or central processing unit **1907**. The processor **1907** can be configured to execute various program codes such as the methods such as described herein.

[0239] In some embodiments the device **1900** comprises a memory **1911**. In some embodiments the at least one processor **1907** is coupled to the memory **1911**. The memory **1911** can be any suitable storage means. In some embodiments the memory **1911** comprises a program code section for storing program codes implementable upon the processor **1907**. Furthermore in some embodiments the memory **1911** can further comprise a stored data section for storing data, for example data that has been processed or to be processed in accordance with the embodiments as described herein. The implemented program code stored within the program code section and the data stored within the stored data section can be retrieved by the processor **1907** whenever needed via the memory-processor coupling.

[0240] In some embodiments the device **1900** comprises a user interface **1905**. The user interface **1905** can be coupled in some embodiments to the processor **1907**. In some embodiments the processor **1907** can control the operation of the user interface **1905** and receive inputs from the user interface **1905**. In some embodiments the user interface **1905** can enable a user to input commands to the device **1900**, for example via a keypad. In some embodiments the user interface **1905** can enable the user to obtain information from the device **1900**. For example the user interface **1905** may comprise a display configured to display information from the device **1900** to the user. The user interface **1905** can in some embodiments comprise a touch screen or touch interface capable of both enabling information to be entered to the device **1900** and further displaying information to the user of the device **1900**.

[0241] In some embodiments the device **1900** comprises an input/output port **1909**. The input/output port **1909** in some embodiments comprises a transceiver. The transceiver in such embodiments can be coupled to the processor **1907** and configured to enable a communication with other apparatus or electronic devices, for example via a wireless communications network. The transceiver or any suitable transceiver or transmitter and/or receiver means can in some embodiments be configured to communicate with other electronic devices or apparatus via a wire or wired coupling.

[0242] The transceiver can communicate with further apparatus by any suitable known

communications protocol. For example in some embodiments the transceiver or transceiver means can use a suitable universal mobile telecommunications system (UMTS) protocol, a wireless local area network (WLAN) protocol such as for example IEEE 802.X, a suitable short-range radio frequency communication protocol such as Bluetooth, or infrared data communication pathway (IRDA).

[0243] The transceiver input/output port **1909** may be configured to receive the loudspeaker signals (or other input format audio signals) and in some embodiments determine the parameters as described herein by using the processor **1907** executing suitable code. Furthermore the device may generate a suitable transport signal and parameter output to be transmitted to the synthesis device.

[0244] In some embodiments the device **1900** may be employed as at least part of the synthesis device. As such the input/output port **1909** may be configured to receive the transport signals and in some embodiments the parameters determined at the capture device or processing device as described herein, and generate a suitable audio signal format output by using the processor **1907** executing suitable code. The input/output port **1909** may be coupled to any suitable audio output for example to a multichannel speaker system and/or headphones or similar.

[0245] In general, the various embodiments of the invention may be implemented in hardware or special purpose circuits, software, logic or any combination thereof. For example, some aspects may be implemented in hardware, while other aspects may be implemented in firmware or software which may be executed by a controller, microprocessor or other computing device, although the invention is not limited thereto. While various aspects of the invention may be illustrated and described as block diagrams, flow charts, or using some other pictorial representation, it is well understood that these blocks, apparatus, systems, techniques or methods described herein may be implemented in, as non-limiting examples, hardware, software, firmware, special purpose circuits or logic, general purpose hardware or controller or other computing devices, or some combination thereof.

[0246] The embodiments of this invention may be implemented by computer software executable by a data processor of the mobile device, such as in the processor entity, or by hardware, or by a combination of software and hardware. Further in this regard it should be noted that any blocks of the logic flow as in the Figures may represent program steps, or interconnected logic circuits, blocks and functions, or a combination of program steps and logic circuits, blocks and functions. The software may be stored on such physical media as memory chips, or memory blocks implemented within the processor, magnetic media such as hard disk or floppy disks, and optical media such as for example DVD and the data variants thereof, CD.

[0247] The memory may be of any type suitable to the local technical environment and may be implemented using any suitable data storage technology, such as semiconductor-based memory devices, magnetic memory devices and systems, optical memory devices and systems, fixed memory and removable memory. The data processors may be of any type suitable to the local technical environment, and may include one or more of general purpose computers, special purpose computers, microprocessors, digital signal processors (DSPs), application specific integrated circuits (ASIC), gate level circuits and processors based on multi-core processor architecture, as non-limiting examples.

[0248] Embodiments of the inventions may be practiced in various components such as integrated circuit modules. The design of integrated circuits is by and large a highly automated process. Complex and powerful software tools are available for converting a logic level design into a semiconductor circuit design ready to be etched and formed on a semiconductor substrate.

[0249] Programs, such as those provided by Synopsys, Inc. of Mountain View, California and Cadence Design, of San Jose, California automatically route conductors and locate components on a semiconductor chip using well established rules of design as well as libraries of pre-stored design modules. Once the design for a semiconductor circuit has been completed, the resultant design, in a standardized electronic format (e.g., Opus, GDSII, or the like) may be transmitted to a

semiconductor fabrication facility or "fab" for fabrication.

[0250] The foregoing description has provided by way of exemplary and non-limiting examples a full and informative description of the exemplary embodiment of this invention. However, various modifications and adaptations may become apparent to those skilled in the relevant arts in view of the foregoing description, when read in conjunction with the accompanying drawings and the appended claims. However, all such and similar modifications of the teachings of this invention will still fall within the scope of this invention as defined in the appended claims.

Claims

1. An apparatus comprising at least one processor and at least one memory including a computer program code, the at least one memory and the computer program code configured to, with the at least one processor, cause the apparatus at least to: obtain at least one audio signal; obtain at least one parameter for a least one of at least two frequency bands associated with the at least one audio signal; select a frequency band of the at least two frequency bands based on comparing at least one further parameter for the at least two frequency bands wherein the at least one further parameter is determined from at least one of the at least two frequency bands; and generate an output comprising a selection of the at least one parameter associated with the selected frequency band of the at least two frequency bands, such that the selection of the at least one parameter associated with the selected frequency band is configured to reduce a bitrate or size of the output.

2. The apparatus as claimed in claim 1, wherein the at least one memory and the computer program code are configured to, with the at least one processor, cause the apparatus to obtain at least one parameter by obtaining an energy ratio and energy respectively for each of the at least two frequency bands, and wherein the at least one memory and the computer program code are configured to, with the at least one processor, cause the apparatus to select the frequency band of the at least two frequency bands by: determining an energy weight factor for each of the at least two frequency bands based on the energy ratio and energy for each of the at least two frequency bands, wherein the energy weight factor is the at least one further respective parameter for each of the at least two frequency bands; determining a weight limit factor based on an averaged energy; comparing the energy weight factor for each of the at least two frequency bands to the weight limit factor; and selecting a highest frequency band where the energy weight factor is greater than the weight limit factor.

3. The apparatus as claimed in claim 2, wherein the energy is a normalized energy.

4. The apparatus as claimed in claim 1, wherein the at least one memory and the computer program code are configured to, with the at least one processor, cause the apparatus to select a frequency band of the at least two frequency bands by selecting the highest frequency band of the at least two frequency bands.

5. The apparatus as claimed in claim 1, wherein the at least one memory and the computer program code are configured to, with the at least one processor, cause the apparatus to obtain at least one parameter by obtaining at least one of: a directional parameter; a distance parameter; an energy parameter; and an energy ratio parameter.

6. The apparatus as claimed in claim 1, wherein the at least one memory and the computer program code are configured to, with the at least one processor, cause the apparatus to select the frequency band by: saving the at least one parameter for one of the at least two frequency bands; and discarding any other of the at least one parameter for the at least two frequency bands, wherein the output comprises the saved at least one parameter for one of the at least two frequency bands and not the discarded other of the at least one parameter for the at least two frequency bands.

7. The apparatus as claimed in claim 1, wherein the at least one memory and the computer program code are configured to, with the at least one processor, cause the apparatus to select the frequency band by: saving the at least one parameter for one of the at least two frequency bands; and

determining a difference between any other of the at least one parameter for the at least two frequency bands and the at least one parameter for one of the at least two frequency bands, wherein the output comprises the difference.

8. The apparatus as claimed in claim 1, wherein the at least one memory and the computer program code are further configured to, with the at least one processor, cause the apparatus to generate at least one transport signal based on the at least one audio signal, wherein the at least one memory and the computer program code are configured to, with the at least one processor, cause the apparatus to generate an output by generating a datastream for storing/transmission based on a combination of the at least one parameter and the at least one transport signal, and wherein the at least one memory and the computer program code are configured to, with the at least one processor, cause the apparatus to generate a datastream for storing/transmission by: encoding the at least one transport signal; encoding the at least one parameter associated with the selected frequency band of the at least two frequency bands; and combining the encoded transport signal and the encoded at least one parameter associated with the selected frequency band of the at least two frequency bands.

9. The apparatus as claimed in claim 1, wherein the at least one memory and the computer program code are further configured to, with the at least one processor, cause the apparatus to generate at least one transport signal based on the at least one audio signal by at least one of: downmixing the at least one audio signal; selecting at least one audio signal from the at least one audio signal, when the at least one audio signal comprises two or more audio signals; generating directional signals directed to different directions, when the at least one audio signal comprises first order ambisonic audio signals; generating cardioid signals directed to different directions, when the at least one audio signal comprises first order ambisonic audio signals; generating cardioid signals directed at opposite directions, when the at least one audio signal comprises first order ambisonic audio signals; and passing at least one transport audio signal, when the at least one audio signal comprises at least one transport audio signal.

10. The apparatus as claimed in claim 1, wherein the apparatus is caused to obtain at least one parameter by obtaining at least one parameter respectively for each of the at least two frequency bands associated with the at least one audio signal and to select a frequency band by selecting a frequency band of the at least two frequency bands based on comparing at least one further respective parameter for each of the at least two frequency bands wherein the at least one further parameter is determined from each of the at least two frequency bands, and wherein the at least one parameter of the selected frequency band is configured to represent respective parameters of the at least two frequency bands.

11. An apparatus comprising at least one processor and at least one memory including a computer program code, the at least one memory and the computer program code configured to, with the at least one processor, cause the apparatus at least to: obtain at least one signal, the at least one signal comprising at least one parameter associated with a selected frequency band from at least two frequency bands and at least one transport signal; replicate, based on the at least one parameter for one of the at least two frequency bands and a transport signal, at least one parameter for at least one other of the at least two frequency bands; and synthesise at least two audio signals based on the at least one parameter associated with the selected frequency band from at least two frequency bands and at least one replicated parameter for the at least one other of the at least two frequency bands and the transport signal, wherein the at least two audio signals are configured to provide spatial audio reproduction.

12. The apparatus as claimed in claim 11, wherein the at least one memory and the computer program code are configured to, with the at least one processor, cause the apparatus to obtain at least one signal by obtaining at least one of: a directional parameter; a distance parameter; an energy parameter; and an energy ratio parameter.

13. The apparatus as claimed in claim 11, wherein the at least one memory and the computer

program code are configured to, with the at least one processor, cause the apparatus to replicate by copying the at least one parameter for one of the at least two frequency bands as the at least one other of the at least two frequency bands.

14. The apparatus as claimed in claim 11, wherein the at least one signal further comprises at least one parameter associated with a difference between at least one other of the at least two frequency bands and the at least one parameter for one of the at least two frequency bands, wherein the at least one memory and the computer program code are configured to, with the at least one processor, cause the apparatus to replicate by replicating the at least one parameter for at least one other of the at least two frequency bands based on a combination of the at least one parameter for one of the at least two frequency bands and the at least one parameter associated with the difference between at least one other of the at least two frequency bands and the at least one parameter for one of the at least two frequency bands.

15. A method comprising: obtaining at least one audio signal; obtaining at least one parameter for at least one of at least two frequency bands associated with the at least one audio signal; selecting a frequency band of the at least two frequency bands based on comparing at least one further parameter for the at least two frequency bands wherein the at least one further respective parameter is determined from at least one of the at least two frequency bands; and generating an output comprising a selection of the at least one parameter associated with the selected frequency band of the at least two frequency bands, such that the selection of the at least one parameter associated with the selected frequency band is configured to reduce a bitrate or size of the output.

16. The method as claimed in claim 15, further comprises generating at least one transport signal based on the at least one audio signal, wherein generating the output comprises generating a datastream for storing/transmission based on a combination of the at least one parameter and the at least one transport signal, and wherein generating the datastream for storing/transmission comprises at least one of: encoding the at least one transport signal; encoding the at least one parameter associated with the selected frequency band of the at least two frequency bands; and combining the encoded transport signal and the encoded at least one parameter associated with the selected frequency band of the at least two frequency bands.

17. The method as claimed in claim 15, further comprises generating at least one transport signal based on the at least one audio signal, wherein generating the at least one transport signal further comprises at least one of: downmixing the at least one audio signal; selecting at least one audio signal from the at least one audio signal, when the at least one audio signal comprises two or more audio signals; generating directional signals directed to different directions, when the at least one audio signal comprises first order ambisonic audio signals; generating cardioid signals directed to different directions, when the at least one audio signal comprises first order ambisonic audio signals; generating cardioid signals directed at opposite directions, when the at least one audio signal comprises first order ambisonic audio signals; and passing at least one transport audio signal, when the at least one audio signal comprises at least one transport audio signal.

18. The method as claimed in claim 15, wherein obtaining at least one parameter comprises obtaining at least one parameter respectively for each of the at least two frequency bands associated with the at least one audio signal, wherein selecting a frequency band comprises selecting a frequency band of the at least two frequency bands based on comparing at least one further respective parameter for each of the at least two frequency bands wherein the at least one further parameter is determined from each of the at least two frequency bands, and wherein the at least one parameter of the selected frequency band is configured to represent respective parameters of the at least two frequency bands.

19. A method comprising: obtaining at least one signal, the at least one signal comprising at least one parameter associated with a selected frequency band from at least two frequency bands and at least one transport signal; replicating, based on the at least one parameter for one of the at least two frequency bands and a transport signal, at least one parameter for at least one other of the at least

two frequency bands; and synthesising at least two audio signals based on the at least one parameter associated with the selected frequency band from at least two frequency bands and at least one replicated parameter for the at least one other of the at least two frequency bands and the transport signal, wherein the at least two audio signals are configured to provide spatial audio reproduction.

20. The method as claimed in claim 19, wherein replicating the at least one parameter further comprises copying the at least one parameters for one of the at least two frequency bands as the at least one other of the at least two frequency bands.