Claims

1. An audio signal representation decoder configured to decode an audio signal representation from a bitstream, the bitstream being divided in a sequence of packets, the audio signal representation decoder comprising:

a bitstream reader, configured to sequentially read the sequence of packets;

a packet loss controller configured to check whether a current packet is well received or is to be considered as lost;

a quantization index converter, configured, in case the packet loss controller has determined that the current packet is well received, to convert at least one index extracted from the current packet onto at least one current code from at least one codebook, thereby forming at least one portion of the audio signal representation; and

wherein the audio signal representation decoder is configured, in case the packet loss controller has determined that the current packet is to be considered as lost, to generate, through at least one learnable predictor layer, at least one current code by prediction from at least one preceding code or index, thereby forming at least one portion of the audio signal representation.

2. The audio signal representation decoder of claim 1, wherein the at least one codebook includes:

a base codebook associating indexes to main portions of codes or high-ranking subcodes; and

at least one low-ranking codebook associating indexes to residual portions of codes or low-ranking subcodes,

wherein the at least one index extracted from the current packet includes at least one high-ranking index and at least one low-ranking index,

wherein the quantization index converter is configured to convert the at least one high-ranking index onto a main portion of the current code or a high-ranking sub-code, and the at least one low-ranking index onto at least one residual portion of the current code or a or a low-ranking subcode, wherein the quantization index converter is further configured to reconstruct the current code by adding the main portion to the at least one residual portion or by putting together the high-ranking subcode with the at least one low-ranking subcode.

3. The audio signal representation decoder of claim 2, configured to predict the current code from at least the at least one high-ranking index of the at least one preceding or following packet, but not from at least one lowest-ranking index of the of the at least one preceding or following packet.

4. The audio signal representation decoder of claim 2 or 3, configured to predict the current code from at least the high-ranking index of the at least one preceding packet and from at least one middle-ranking index, but not from the lowest-ranking index of the of the at least one preceding packet.

5. The audio signal representation decoder of any of claims 1 to 4, configured to store redundancy information written in packets of the bitstream but referring to different packets, the audio signal representation decoder being configured to store the redundancy information in a storage unit,

wherein the audio signal representation decoder is configured, in case the at least one current packet is to be considered as lost, to first search the storage unit, and, in case the redundancy information referring to the at least one current packet is retrieved, to:

retrieve at least one index from the redundancy information referring to the current packet;

cause the quantization index converter to convert the at least one retrieved index from the at least one codebook onto a substitutive code, so as to be part of the audio signal representation.

6. The audio signal representation decoder of claim 5, configured to perform the prediction of the at least one current code only in case of the redundancy information being not retrieved in the storage unit.

7. The audio signal representation decoder of claim 5 or 6 when depending on any of claims 1 to 4, wherein the redundancy information provides at least the high-ranking index(es) of the at least one preceding or following packet, but not at least one of the lower-ranking index(es) of the at least one preceding or following packet.

8. The audio signal representation decoder of any of claims 1 to 7, further comprising at least one learnable predictor configured to perform the prediction, the at least one learnable predictor having at least one learnable predictor layer.

9. The audio signal representation decoder of claim 8, wherein the at least one learnable predictor is trained by sequentially predicting predicted current codes, or respectively current indexes, from preceding and/or following packets, and by comparing the predicted current codes, or the current codes obtained from predicted indexes, with converted codes converted from packets having been well received, so as to learn learnable parameters of the at least one learnable predictor layer which minimize errors of the predicted current codes with respect the converted codes converted from the packets having correct format.

10. The audio signal representation decoder of any of claims 8 or 9, wherein the at least one learnable predictor layer includes at least one recurrent learnable layer.

11. The audio signal representation decoder of any of claims 8 to 10, wherein the at least one learnable predictor layer includes at least one gated recurrent unit.

12. The audio signal representation decoder of any of claims 8 to 11, wherein the at least one learnable predictor layer includes or is part of at least one neural network.

13. The audio signal representation decoder of any of claims 8 to 12, wherein the at least one learnable predictor layer has at least one state,

the at least one learnable predictor layer being iteratively instantiated, along a sequential plurality of learnable predictor layer instantiations,

in such a way that, to predict the at least one current code, a current learnable predictor layer instantiation receives a state from at least one preceding learnable predictor layer instantiation which has predicted at least one preceding code for at least one preceding packet.

14. The audio signal representation decoder of claim 13, wherein, to predict the at least one current code, the current learnable predictor layer instantiation receives in input:

the at least one preceding converted code in case the at least one preceding packet is considered well received; and

the at least one preceding predicted code) in case the at least one preceding packet is considered as lost.

15. The audio signal representation decoder of claim 14, wherein, to predict the current code, the current learnable predictor layer instantiation receives the state from the at least one preceding iteration both in case the at least one preceding packet is considered well received and in case the at least one preceding packet is considered as lost.

16. The audio signal representation decoder of any of claims 13 to 15, wherein the at least one learnable predictor layer is configured to predict the current code and/or to receive the state from the at least one preceding learnable predictor layer instantiation both in case the at least one preceding packet is considered well received and in case the at least one preceding packet is considered as lost, so as to provide the predicted code and/or to output the state to at least one subsequent learnable predictor layer instantiation.

17. The audio signal representation decoder of any of claims 13 to 16, wherein the current learnable predictor layer instantiation includes at least one learnable convolutional unit.

18. The audio signal representation decoder of any of claims 13 to 17, wherein the current learnable predictor layer instantiation includes at least one learnable recurrent unit.

19. The audio signal representation decoder of claim 18, wherein the at least one recurrent unit of the current learnable layer is inputted with a state from a correspondent at least one recurrent unit from the at least one preceding learnable predictor layer instantiation, and outputs a state to a corresponding at least one recurrent unit of at least one subsequent learnable predictor layer instantiation.

20. The audio signal representation decoder of any of claims 13 to 19, wherein the current learnable predictor layer instantiation has a series of learnable layers.

21. The audio signal representation decoder of claim 20, wherein for the current learnable predictor layer instantiation, the series of learnable layers includes at least one dimension-reducing learnable layer and at least one dimension-increasing learnable layer subsequent to the at least one dimension-reducing learnable layer.

22. The audio signal representation decoder of claim 21, wherein the at least one dimension-reducing learnable layer includes at least one learnable layer with a state.

23. The audio signal representation decoder of any of claims 21 to 22, wherein the at least one dimension-increasing learnable layer includes at least one learnable layer without a state.

24. The audio signal representation decoder of any of claims 20 to 22, wherein the series of learnable layers is gated.

25. The audio signal representation decoder of any of claims 20 to 24, wherein the wherein the series of learnable layers is gated through a softmax activation function.

26. An audio signal representation decoder configured to decode an audio signal representation from a bitstream, the bitstream being divided in a sequence of packets, the audio signal representation decoder comprising:

a bitstream reader, configured to sequentially read the sequence of packets, and to extract, from the at least one current packet:

at least one index of the at least one current packet; and

redundancy information on at least one preceding or following packet, the redundancy information permitting to reconstruct at least one index, or information provided by the index, within the at least one preceding or following packet;

a packet loss controller, PLC, configured to check whether the at least one current packet is well received or is to be considered as lost;

a quantization index converter, configured to convert the at least one index of the at least one current packet onto at least one current converted code at least one codebook, thereby forming a portion of the audio signal representation;

a redundancy information storage unit, configured to store the redundancy information and to provide the stored redundancy information on the at least one current packet in case the PLC has determined that the at least one current packet is to be considered as lost, to form a portion of the audio signal representation through the redundancy information.

27. The audio signal representation decoder of claim 26, configured to read a signalling indicating a packet offset between the current packet and the at least one preceding or following packet having the redundant information at least in function of the payload of the communication channel, so that the redundancy information storage unit, or another component of the audio signal representation decoder, reconstructs the packet to which the redundancy information refers and stores the redundancy information associated with the packet to which the redundancy information refers.

28. The audio signal representation decoder of any of claims 26-27, wherein the redundancy information storage unit is configured to store, as redundancy information, at least one index from a preceding or following packet, so as to provide, to the quantization index converter, the stored at least one index in case the PLC has determined that the at least one current packet is to be considered as lost.

29. The audio signal representation decoder of any of claims 26 to 28, wherein the redundancy information storage unit is configured to store, as redundancy information, at least one code, or a part thereof, previously extracted from a preceding or following packet, to bypass the quantization index converter using the stored code in case in case the PLC has determined that the at least one current packet is to be considered as lost.

30. The audio signal representation decoder of any of claims 26 to 29, wherein the at least one codebook associates indexes to codes or parts of codes, so that the quantization index converter converts the at least one index extracted from the current packet onto the at least one converted code, or at least one part of a converted code.

31. The audio signal representation decoder of any of claims 26 to 30, wherein the at least one codebook includes:

a base codebook associating indexes to main portions of codes; and

at least one low-ranking codebook,

wherein the at least one index extracted from the current packet includes at least one high-ranking index and at least one low-ranking index,

wherein the quantization index converter is configured to convert the at least one high-ranking index onto a main portion of the current code, or a high-ranking subcode, and the at least one low-ranking index onto at least one residual portion of the current code, or a high-ranking subcode,

wherein the quantization index converter is further configured to reconstruct the current code by adding the main portion to the at least one residual portion or by putting together the at least one high-ranking subcode with the at least one low-ranking subcode or by putting together the at least one high-ranking subcode with the at least one low-ranking subcode or by putting together the high-ranking subcode with the at least one low-ranking subcode,

the audio signal representation decoder is configured to:

generate or retrieve the at least one current code from at least the at least one high-ranking index of the at least one preceding or following packet, but not from the lowest-ranking index of the at least one preceding or following packet,, and/or generate or retrieve the current code from at least the high-ranking index of the at least one preceding or following packet and from at least one middle-ranking index, but not from the lowest-ranking index of the of the at least one preceding or following packet.

32. An audio generator for generating an audio signal from a bitstream, comprising the audio signal representation decoder of claim 1,

further configured to generate the audio signal by converting the audio signal representation onto the audio signal.

33. The audio generator of claim 32, further configured to render the generated audio signal.

34. An encoder, comprising:

an audio signal representation generator configured to generate, through at least one learnable layer, an audio signal representation as a representation of an audio signal, the audio signal representation including a sequence of tensors;

a quantizer configured to convert each current tensor of the sequence of tensors onto at least one index, wherein each index is obtained from at least one codebook associating a plurality of tensors to a plurality of indexes;

a bitstream writer configured to write packets in the bitstream, so that a current packet includes the at least one index for the current tensor of the sequence of tensors,

wherein the encoder is configured to write redundancy information of the current tensor in at least one preceding or following packet of the bitstream different from the current packet and/or to write, in the current packet, redundancy information of a tensor, different from the current tensor, in the current packet.

35. The encoder of any of claim 34, wherein the at least one codebook includes:

a base codebook associating main portions of tensors to indexes; and

at least one low-ranking codebook associating residual portions of tensors to indexes,

wherein the at least one current tensor has at least one main portion and at least one residual portion,

wherein the quantizer is configured to convert the main portion of the at least one current tensor onto at least one high-ranking index, and the at least one residual portion of the at least one tensor onto at least one low-ranking index,

so that the bitstream writer writes, in the bitstream, both the high-ranking index and the at least one low-ranking index,

wherein the encoder is configured to provide the redundancy information with at least the high-ranking index(es) of the at least one preceding or following packet, but not at least the lowest-ranking low-ranking index(es) of the same at least one preceding or following packet.

36. The encoder of any of claims 34-35, configured to transmit the bitstream to a receiver through a communication channel, and further configured to monitor the pay-

load state of the communication channel, so as, in case the payload state in the communication channel is over a predetermined threshold, to increase the quantity of redundancy information.

37. The encoder of claim 36, configured:

in case the payload in the communication channel is below the predetermined threshold, to only transmit, as redundancy information, for each current packet, high-ranking indexes of the at least one preceding or following packet; and/or

in case the payload of the communication channel is over the predetermined threshold, to transmit, as redundancy information, for each current packet, both the high-ranking indexes of the at least one preceding or following packet and at least some low-ranking indexes of the same at least one preceding or following packet.

38. The encoder of claim 36 or 37, configured to compute a packet offset between the current packet and the at least one preceding or following packet having the redundant information at least in function of the payload of the communication channel.

39. The encoder of any of claims 36 to 38, configured to compute a packet offset between the current packet and the at least one preceding or following packet having the redundant information at least in function of the envisioned application.

40. The encoder of any of claims 36 to 39, configured to compute a packet offset between the current packet and the at least one preceding or following packet having the redundant information at least in function of an input provided by the end-user.

41. The encoder of claim 37 or 38, configured to compute a packet offset between the current packet and the at least one preceding or following packet having the redundant information at least in function of the payload of the communication channel, in such a way that the higher the payload in the communication channel, or the higher the error rate in the communication channel, the higher the packet offset.

42. The encoder of any of claims 34 to 41, wherein the at least one codebook includes a redundancy codebook associating a plurality of tensors to a plurality of indexes, wherein the encoder is configured to write the redundancy information of the

current tensor in the at least one preceding or following packet of the bitstream different from the current packet as an index received from the at least one quantization codebook.

43. A method for decoding an audio signal representation from a bitstream, the method including:

reading a sequence of packets comprised in the bitstream, and extracting, from a current packet:

at least one index of the current packet;

redundancy information on at least one preceding or following packet, the redundancy information permitting to reconstruct at least one index within the at least one preceding or following packet;

checking whether the current packet is well received or is to be considered as lost;

converting the at least one index of the current packet onto at least one current converted code from at least one codebook, thereby forming a portion of the audio signal representation;

storing the redundancy information and providing the stored redundancy information on the at least one current packet in case the checking has determined that the at least one current packet is to be considered as lost, to form a portion of the audio signal representation through the redundancy information.

44. A method for decoding an audio signal representation from a bitstream, the bitstream being divided in a sequence of packets, the audio signal representation decoder comprising:

sequentially reading the sequence of packets;

checking whether a current packet is well received or is to be considered as lost;

in case the checking has determined that the current packet is well received, converting at least one index extracted from the current packet onto at least one current code from at least one code-book, thereby forming at least one portion of the audio signal representation; and

in case the packet loss controller has determined that the current packet is to be considered as lost, generating, through at least one learnable predictor layer, at least one current code by prediction from at least one preceding code or index.

45. A non-transitory storage unit storing instructions which, when executed by a computer, cause the computer to:

extract, from a current packet:

at least one index of the current packet; and

redundancy information on at least one preceding or following packet, the redundancy information permitting to reconstruct at least one index within the at least one preceding or following packet;

check whether the current packet is well received or is to be considered as lost;

convert the at least one index of the current packet onto at least one current converted code from at least one codebook, thereby forming a portion of the audio signal representation;

control the storing of the redundancy information and retrieved the stored redundancy information on the at least one current packet in case the checking has determined that the at least one current packet is to be considered as lost, to form a portion of the audio signal representation through the redundancy information.

46. A non-transitory storage unit storing instructions which, when executed by a computer, cause the computer to:

sequentially read a sequence of packets;

check whether a current packet is well received or is to be considered as lost;

in case the check has determined that the current packet is well received, convert at least one index extracted from the current packet onto at least one current code from at least one code-book, thereby forming at least one portion of the audio signal representation; and

in case the check has determined that the current packet is to be considered as lost, generate, through at least one learnable predictor layer, at least one current code by prediction from at least one preceding code or index.