Claims

1. 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.

2. The encoder of claim 1, wherein the at least one codebook associates parts of tensors to indexes, so that the quantizer converts the current tensor onto a plurality of indexes.

3. The encoder of claim 1 or 2, 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.

4. The encoder of any of the preceding claims, configured to split the current tensor into a plurality of subtensors, so as to quantize each subtensor.

5. The encoder of the preceding claims, configured to decompose the current tensor among a main portion and at least one residual portion, so as to quantize the main portion and the at least one residual portion.

6. The encoder of the preceding claims, configured to transmit the bitstream to a receiver through a communication channel.

7. The encoder of claim 6, configured to monitor the payload 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,

wherein the encoder is 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.

8. The encoder of claim 7, 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.

9. The encoder of claim 7 or 8, 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.

10. The encoder of any of the preceding claims, 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.

11. 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.

12. The decoder claim 11, wherein the quantization index converter is configured to convert a plurality of indexes onto a respective plurality of subtensors, then put together the plurality of subtensors to obtain the at least one current code.

13. The decoder of claim 11 or 12, wherein the quantization index converter is configured to convert a plurality of indexes onto at least one main portion of code and at least one residual portion of code, then to compose the main portion of code and the at least one residual portion of code to obtain the at least one current code.

14. The audio signal representation decoder of any of claims 11-13, 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.

15. The audio signal representation decoder of any of claims 11-14, 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 subcode, 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.

16. 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.

17. The audio signal representation decoder of claim 16, wherein the quantization index converter is configured to convert a plurality of indexes onto a respective plurality of subtensors, then put together the plurality of subtensors to obtain the at least one current code.

18. The audio signal representation decoder of claim 16 or 17, wherein the quantization index converter is configured to convert a plurality of indexes onto at least one main portion of code and at least one residual portion of code, then to compose the main portion of code and the at least one residual portion of code to obtain the at least one current code.

19. The audio signal representation decoder of any of claims 16-18, 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.

20. The audio signal representation decoder of any of claims 16-19, 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 subcode, 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.

21. The audio signal representation decoder of any of claims 16-20, 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.

22. An audio generator for generating an audio signal from a bitstream, comprising the audio signal representation decoder according to any of claims 11-21,

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

23. The audio generator of claim 22 comprising:

a first data provisioner configured to provide, for a given frame, first data derived from an input signal;

a first processing block, configured, for the given frame, to receive the first data and to output first output data in the given frame,

further comprising:

at least one conditioning learnable layer configured, for the given frame, to process target data, from the decoded audio signal representation, to obtain conditioning feature parameters; and

a styling element, configured to apply the conditioning feature parameters to the first data or normalized first data.

24. The audio generator of claim 22 or 23, configured to obtain the input signal from noise.

25. The audio generator of claim 22 or 23 or 24, further comprising at least one preconditioning learnable layer configured to receive the audio signal representation and output target data representing the audio signal.

26. The audio generator of claim 25, wherein the at least one preconditioning learnable layer is configured to provide the target data as a spectrogram or a decoded spectrogram.

27. The audio generator of any of claims 16-26, wherein a first convolution layer is configured to convolute the target data or up-sampled target data to obtain first convoluted data using a first activation function.

28. The audio generator of any of claims 16-27, further comprising a normalizing element, which is configured to normalize the first data.

29. The audio generator of any of claims 16-28, wherein the target data includes a spectrogram.

30. A method comprising:

generating, 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;

converting 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;

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

wherein the method includes writing redundancy information of the current tensor in at least one preceding or following packet of the bitstream different from the current packet, and/or writing, in the current packet, redundancy information of at least one tensor to be written in at least one preceding or following packet of the bitstream different from the current packet.

31. A non-transitory storage unit storing instructions which, when executed by a computer, cause the computer 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;

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;

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

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 writing, in the current packet, redundancy information of at least one tensor to be written in at least one preceding or following packet of the bitstream different from the current packet.