Claims

1. Multi-mode audio decoder for providing a decoded representation of audio content on the basis of an encoded bitstream, the multi-mode audio decoder configured to

decode a global gain value per frame of the encoded bitstream, wherein a first subset of the frames being coded in a first coding mode and a second subset of the frames being coded in a second coding mode, with each frame of the second subset being composed of more than one sub-frames,

decode, per sub-frame of at least a subset of the sub-frames of the second subset of frames, a corresponding bitstream element differentially to the global gain value of the respective frame, and

complete decoding the bitstream using the global gain value and the corresponding bitstream element in decoding the sub-frames of the at least subset of the sub-frames of the second subset of frames and the global gain value in decoding the first subset of frames,

wherein the multi-mode audio decoder is configured such that a change of the global gain value of the frames within the encoded bitstream results in an adjustment of an output level of the decoded representation of the audio content.

2. Multi-mode audio decoder according to claim 1, wherein the first coding mode is a frequency domain coding mode, and the second coding mode is a linear prediction coding mode.

3. Multi-mode audio decoder according to claim 2, wherein the multi-mode audio decoder is configured to, in completing the decoding of the encoded bitstream, decode the sub-frames of the at least subset of the sub-frames of the second subset of frames by using transformed excitation linear prediction decoding, and decode a disjoined subset of the sub-frames of the second subset of the frames by use of CELP.

4. Multi-mode audio decoder according to any of claims 1 to 3, wherein the multi-mode audio decoder is configured to decode, per frame of the second subset of the frames, a further bitstream element revealing a decomposition of the respective frame into one or more sub-frames.

5. Multi-mode audio decoder according to any of the previous claims, wherein the frames of the second subset are of equal length, and the at least subset of the sub-frames of the second subset of frames have a varying sample length selected from the group consisting of 256, 512 and 1024 samples, and a disjoined subset of the sub-frames have a sample length of 256 samples.

6. Multi-mode audio decoder according to any of the previous claims, wherein the multi-mode audio decoder is configured to decode the global gain value on fixed number of bits and the bitstream element on a variable number of bits, the number depending on a sample length of the respective sub-frame.

7. Multi-mode audio decoder according to any of claims 1 to 5, wherein the multi-mode audio decoder is configured to decode the global gain value on fixed number of bits and to decode the bitstream element on fixed number of bits.

8. Multi-mode audio decoder for providing a decoded representation of an audio content on the basis of an encoded bitstream, a first subset of frames of which is CELP coded and a second subset of frames of which is transform coded, the multi-mode audio decoder comprising:

a CELP decoder configured to decode a current frame of the first subset, the CELP decoder comprising:

an excitation generator configured to generate a current excitation of the current frame of the first subset by constructing an codebook excitation based on a past excitation and an codebook index of the current frame of the first subset within the encoded bitstream, and setting a gain of the codebook excitation based on a global gain value within the encoded bitstream; and

a linear prediction synthesis filter configured to filter the current excitation based on linear prediction filter coefficients for the current frame of the first subset within the encoded bitstream;

a transform decoder configured to decode a current frame of the second subset by

constructing spectral information for the current frame of the second subset from the encoded bitstream and performing a spectral-to-time-domain transformation onto the spectral information to obtain a time-domain signal such that a level of the time-domain signal depends on the global gain value.

9. Multi-mode audio decoder according to claim 8, wherein the excitation generator is configured to, in generating the current excitation of the current frame of the first subset,

construct an adaptive codebook excitation based on a past excitation and an adaptive codebook index of the current frame of the first subset within the encoded bitstream;

construct an innovation codebook excitation based on an innovation codebook index for the current frame of the first subset within the encoded bitstream;

set, as the gain of the codebook excitation, a gain of the innovation codebook excitation based on the global gain value within the encoded bitstream; and

combine the adaptive codebook excitation and the innovation codebook excitation to obtain the current excitation of the current frame of the first subset.

10. Multi-mode audio decoder according to claim 8 or 9, wherein the transform decoder is configured such that the spectral information relates to a current excitation of the current frame of the second subset, and the transform decoder is further configured to, in decoding the current frame of the second subset, spectrally form the current excitation of the current frame of the second subset according to a linear

prediction synthesis filter transfer function defined by linear prediction filter coefficients for the current frame of the second subset within the encoded bitstream so that the performance of the spectral-to-time-domain transformation onto the spectral information results in the decoded representation of the audio content.

11. Multi-mode audio decoder according to claim 10, wherein the transform decoder is configured to perform the spectral formation by converting the linear prediction filter coefficients into a linear prediction spectrum and weighting the spectral information of the current excitation with the linear prediction spectrum.

12. Multi-mode audio decoder according to any of claims 8 to 11, wherein the transform decoder is configured to scale the spectral information with the global gain value.

13. Multi-mode audio decoder according to claim 8 or 9, wherein the transform decoder is configured to construct the spectral information for the current frame of the second subset by use of spectral transform coefficients within the encoded bitstream, and scale factors within the encoded bitstream for scaling the spectral transform coefficients in a spectral granularity of scale factor bands, with scaling the scale factors based on the global gain value, so as to obtain the decoded representation of the audio content.

14. CELP decoder comprising:

an excitation generator configured to generate a current excitation for a current frame of a bitstream by

constructing an adaptive codebook excitation based on a past excitation and an adaptive codebook index for the current frame within the bitstream;

constructing an innovation codebook excitation based on an innovation codebook index for the current frame within the bitstream;

computing an estimate of an energy of the innovation codebook excitation spectrally weighted by a weighted linear prediction synthesis filter constructed from linear prediction filter coefficients within the bitstream;

setting a gain of the innovation codebook excitation based on a ratio between a global gain value within the bitstream and the estimated energy; and

combining the adaptive codebook excitation and the innovation codebook excitation to obtain the current excitation; and

a linear prediction synthesis filter configured to filter the current excitation based on the linear prediction filter coefficients.

15. CELP decoder according to claim 14, wherein the excitation generator is configured to, in constructing the adaptive codebook excitation, filter the past excitation with a filter depending on the adaptive codebook index.

16. CELP decoder according to claim 14 or 15, wherein the excitation generator is configured to construct the innovation codebook excitation such that the latter comprises a zero vector with a number of non-zero pulses, the number and positions of the non-zero pulses being indicated by the innovation codebook index.

17. CELP decoder according to any of claims 14 to 16, wherein the excitation generator is configured to, in computing the estimate of the energy of the innovation codebook excitation, filter the innovation codebook excitation with

$$\frac{\hat{W}(z)}{\hat{A}(z)\,H_{emph}(z)},$$

wherein the linear prediction synthesis filter is configured to filter the current excitation according to $1/\hat{A}(z)$, wherein $\hat{W}(z) = \hat{A}(z/\gamma)$ and γ is a perceptual weighting factor, $H_{emph} = 1 - \alpha\, z^{-1}$ and α is a high-frequency-emphasis factor, wherein the excitation generator is further configured to compute a quadratic sum of samples of the filtered innovation codebook excitation to obtain the estimate of the energy.

18. CELP decoder according to any of claims 14 to 17, wherein the excitation generator is configured to, in combining the adaptive codebook excitation and the

innovation codebook excitation, form a weighted sum of the adaptive codebook excitation weighted with a weighting factor depending on the adaptive codebook index, and the innovation codebook excitation weighted with the gain.

19. SBR decoder comprising a core decoder for decoding core-coder portion of a bitstream to obtain a core band signal according to any of the previous claims, the SBR decoder configured to decode envelope energies for a spectral band to be replicated, from an SBR portion of the bitstream, and scaling the envelope energies according to an energy of the core band signal.

20. Multi-mode audio encoder configured to encode an audio content into an encoded bitstream with encoding a first subset of frames in a first coding mode and a second subset of frames in a second coding mode, wherein the second subset of frames is respectively composed of one or more sub-frames, wherein the multi-mode audio encoder is configured to determine and encode a global gain value per frame, and determine and encode, per sub-frames of at least a subset of the sub-frames of the second subset, a corresponding bitstream element differentially to the global gain value of the respective frame, wherein the multi-mode audio encoder is configured such that a change of the global gain value of the frames within the encoded bitstream results in an adjustment of an output level of a decoded representation of the audio content at the decoding side.

21. Multi-mode audio encoder for encoding an audio content into an encoded bitstream by CELP encoding a first subset of frames of the audio content and transform encoding a second subset of the frames, the multi-mode audio encoder comprising:

a CELP encoder configured to encode a current frame of the first subset, the CELP encoder comprising

a linear prediction analyzer configured to generate linear prediction filter coefficients for the current frame of the first subset and encode same into the encoded bitstream; and

an excitation generator configured to determine a current excitation of the current frame of the first subset, which, when filtered by a linear

prediction synthesis filter based on the linear prediction filter coefficients within the encoded bitstream , recovers the current frame of the first subset, defined by a past excitation and a codebook index for the current frame of the first subset and encoding the codebook index into the encoded bitstream; and

a transform encoder configured to encode a current frame of the second subset by performing a time-to-spectral-domain transformation onto a time-domain signal for the current frame of the second subset to obtain spectral information and encode the spectral information into the encoded bitstream,

wherein the multi-mode audio encoder is configured to encode a global gain value into the encoded bitstream, the global gain value depending on an energy of a version of the audio content of the current frame of the first subset, filtered with the linear prediction analysis filter depending on the linear prediction coefficients, or an energy of the time-domain signal.

22. CELP encoder comprising

a linear prediction analyzer configured to generate linear prediction filter coefficients for a current frame of an audio content and encode the linear prediction filter coefficients into a bitstream;

an excitation generator configured to determine a current excitation of the current frame as a combination of an adaptive codebook excitation and an innovation codebook excitation, which, when filtered by a linear prediction synthesis filter based on the linear prediction filter coefficients, recovers the current frame, by

constructing the adaptive codebook excitation defined by a past excitation and an adaptive codebook index for the current frame and encoding the adaptive codebook index into the bitstream; and

constructing the innovation codebook excitation defined by an innovation codebook index for the current frame and encoding the innovation codebook index into the bitstream; and

an energy determiner configured to determine an energy of a version of the audio content of the current frame filtered a weighting filter, to obtain a global gain value and encoding the global gain value into the bitstream, the weighting filter construed from the linear prediction filter coefficients.

23. CELP encoder according to claim 22, wherein the linear prediction analyzer is configured to determine the linear prediction filter coefficients by linear prediction analysis applied onto a windowed and, according to a predetermined pre-emphasis filter, pre-emphasized version of the audio content.

24. CELP encoder according to claim 22 or 23, wherein the excitation generator is configured to, in constructing the adaptive codebook excitation and the innovation codebook excitation, minimize a perceptual weighted distortion measure relative to the audio content.

25. CELP encoder according to any of claims 22 to 24, wherein the excitation generator is configured to, in constructing the adaptive codebook excitation and the innovation codebook excitation, minimize a perceptual weighted distortion measure relative to the audio content using a perceptual weighting filter

$$W(z) = A(z/\gamma),$$

wherein γ is a perceptual weighting factor and A(z) is 1/H(z), wherein H(z) is the linear prediction synthesis filter, and wherein the energy determiner is configured to use the perceptual weighting filter as a weighting filter.

26. CELP encoder according to any of claims 22 to 25, wherein the excitation generator is configured to perform an excitation update to obtain a past excitation of a next frame, by

estimating an innovation codebook excitation energy estimate by filtering an innovation codebook vector defined by first information contained within the innovation codebook index with

$$\frac{\hat{W}(z)}{\hat{A}(z)\,H_{emph}(z)},$$

and determining an energy of the result filtering result, wherein $1/\hat{A}(z)$ is the linear prediction synthesis filter and depends on the linear prediction filter coefficients, $\hat{W}(z) = \hat{A}(z/\gamma)$ and γ is a perceptual weighting factor, $H_{emph} = 1 - \alpha z^{-1}$ and α is a high-frequency-emphasis factor;

forming a ratio between the innovation codebook excitation energy estimate and an energy determined by the global gain value in order to obtain a prediction gain;

multiplying the prediction gain with an innovation codebook correction factor contained within the innovation codebook index as a second information thereof, to yield an actual innovation codebook gain; and

actually generating the past excitation for the next frame by combining the adaptive codebook excitation and the innovation codebook excitation with weighting the latter with the actual innovation codebook gain.

27. Multi-mode audio decoding method for providing a decoded representation of audio content on the basis of an encoded bitstream, the method comprising

decoding a global gain value per frame of the encoded bitstream, wherein a first subset of the frames being coded in a first coding mode and a second subset of the frames being coded in a second coding mode, with each frame of the second subset being composed of more than one sub-frames,

decoding, per sub-frame of at least a subset of the sub-frames of the second subset of frames, a corresponding bitstream element differentially to the global gain value of the respective frame, and

completing decoding the bitstream using the global gain value and the corresponding bitstream element in decoding the sub-frames of the at least subset of the sub-frames of the second subset of frames and the global gain value in decoding the first subset of frames,

wherein the multi-mode audio decoding method is performed such that a change of the global gain value of the frames within the encoded bitstream results in an adjustment of an output level of the decoded representation of the audio content.

28. Multi-mode audio decoding method for providing a decoded representation of an audio content on the basis of an encoded bitstream, a first subset of frames of which is CELP coded and a second subset of frames of which is transform coded, the method comprising:

CELP decoding a current frame of the first subset, the CELP decoder comprising:

generating a current excitation of the current frame of the first subset by constructing an codebook excitation based on a past excitation and an codebook index of the current frame of the first subset within the encoded bitstream, and setting a gain of the codebook excitation based on a global gain value within the encoded bitstream; and

filtering the current excitation based on linear prediction filter coefficients for the current frame of the first subset within the encoded bitstream;

transform decoding a current frame of the second subset by

constructing spectral information for the current frame of the second subset from the encoded bitstream and performing a spectral-to-time-domain transformation onto the spectral information to obtain a time-domain signal such that a level of the time-domain signal depends on the global gain value.

29. CELP decoding method comprising:

generating a current excitation for a current frame of a bitstream by

constructing an adaptive codebook excitation based on a past excitation and an adaptive codebook index for the current frame within the bitstream;

constructing an innovation codebook excitation based on an innovation codebook index for the current frame within the bitstream;

computing an estimate of an energy of the innovation codebook excitation spectrally weighted by a weighted linear prediction synthesis filter constructed from linear prediction filter coefficients within the bitstream;

setting a gain of the innovation codebook excitation based on a ratio between a global gain value within the bitstream and the estimated energy; and

combining the adaptive codebook excitation and the innovation codebook excitation to obtain the current excitation; and

filtering the current excitation based on the linear prediction filter coefficients by a linear prediction synthesis filter.

30. Multi-mode audio encoding method comprising encoding an audio content into an encoded bitstream with encoding a first subset of frames in a first coding mode and a second subset of frames in a second coding mode, wherein the second subset of frames is respectively composed of one or more sub-frames, wherein the multi-mode audio encoding method further comprises determining and encoding a global gain value per frame, and determine and encode, per sub-frames of at least a subset of the sub-frames of the second subset, a corresponding bitstream element differentially to the global gain value of the respective frame, wherein the multi-mode audio encoding method is performed such that a change of the global gain value of the frames within the encoded bitstream results in an adjustment of an output level of a decoded representation of the audio content at the decoding side.

31. Multi-mode audio encoding method for encoding an audio content into an encoded bitstream by CELP encoding a first subset of frames of the audio content and transform encoding a second subset of the frames, the multi-mode audio encoding method comprising:

encoding a current frame of the first subset, the CELP encoder comprising

performing linear prediction analysis to generate linear prediction filter coefficients for the current frame of the first subset and encode same into the encoded bitstream; and

determining a current excitation of the current frame of the first subset, which, when filtered by a linear prediction synthesis filter based on the linear prediction filter coefficients within the encoded bitstream, recovers the current frame of the first subset, defined by a past excitation and a codebook index for the current frame of the first subset and encoding the codebook index into the encoded bitstream; and

encoding a current frame of the second subset by performing a time-to-spectral-domain transformation onto a time-domain signal for the current frame of the second subset to obtain spectral information and encode the spectral information into the encoded bitstream,

wherein the multi-mode audio encoding method further comprises encoding a global gain value into the encoded bitstream, the global gain value depending on an energy of a version of the audio content of the current frame of the first subset, filtered with the linear prediction analysis filter depending on the linear prediction coefficients, or an energy of the time-domain signal.

32. CELP encoding method comprising

Performing linear prediction analysis to generate linear prediction filter coefficients for a current frame of an audio content and encode the linear prediction filter coefficients into a bitstream;

determining a current excitation of the current frame as a combination of an adaptive codebook excitation and an innovation codebook excitation, which, when filtered by a linear prediction synthesis filter based on the linear prediction filter coefficients , recovers the current frame, by

constructing the adaptive codebook excitation defined by a past excitation and an adaptive codebook index for the current frame and encoding the adaptive codebook index into the bitstream; and

constructing the innovation codebook excitation defined by an innovation codebook index for the current frame and encoding the innovation codebook index into the bitstream; and

determining an energy of a version of the audio content of the current frame filtered a weighting filter, to obtain a global gain value and encoding the global gain value into the bitstream, the weighting filter construed from the linear prediction filter coefficients.

33. Computer program having a program code for performing, when running on a computer, a method according to any of claims 27 to 32.

Claims

1. Multi-mode audio decoder for providing a decoded representation of audio content on the basis of an encoded bitstream, the multi-mode audio decoder configured to

decode a global gain value per frame of the encoded bitstream, wherein a first subset of the frames being coded in a first coding mode and a second subset of the frames being coded in a second coding mode, with each frame of the second subset being composed of more than one sub-frames,

decode, per sub-frame of at least a subset of the sub-frames of the second subset of frames, a corresponding bitstream element differentially to the global gain value of the respective frame, and

complete decoding the bitstream using the global gain value and the corresponding bitstream element in decoding the sub-frames of the at least subset of the sub-frames of the second subset of frames and the global gain value in decoding the first subset of frames,

wherein the multi-mode audio decoder is configured such that a change of the global gain value of the frames within the encoded bitstream results in an adjustment of an output level of the decoded representation of the audio content.

2. Multi-mode audio decoder according to claim 1, wherein the first coding mode is a frequency domain coding mode, and the second coding mode is a linear prediction coding mode.

3. Multi-mode audio decoder according to claim 2, wherein the multi-mode audio decoder is configured to, in completing the decoding of the encoded bitstream, decode the sub-frames of the at least subset of the sub-frames of the second subset of frames by using transformed excitation linear prediction decoding, and decode a disjoined subset of the sub-frames of the second subset of the frames by use of CELP.

4. Multi-mode audio decoder according to any of claims 1 to 3, wherein the multi-mode audio decoder is configured to decode, per frame of the second subset of the frames, a further bitstream element revealing a decomposition of the respective frame into one or more sub-frames.

5. Multi-mode audio decoder according to any of the previous claims, wherein the frames of the second subset are of equal length, and the at least subset of the sub-frames of the second subset of frames have a varying sample length selected from the group consisting of 256, 512 and 1024 samples, and a disjoined subset of the sub-frames have a sample length of 256 samples.

6. Multi-mode audio decoder according to any of the previous claims, wherein the multi-mode audio decoder is configured to decode the global gain value on fixed number of bits and the bitstream element on a variable number of bits, the number depending on a sample length of the respective sub-frame.

7. Multi-mode audio decoder according to any of claims 1 to 5, wherein the multi-mode audio decoder is configured to decode the global gain value on fixed number of bits and to decode the bitstream element on fixed number of bits.

8. Multi-mode audio decoder for providing a decoded representation of an audio content on the basis of an encoded bitstream, a first subset of frames of which is CELP coded and a second subset of frames of which is transform coded, the multi-mode audio decoder comprising:

a CELP decoder configured to decode a current frame of the first subset, the CELP decoder comprising:

an excitation generator configured to generate a current excitation of the current frame of the first subset by constructing an codebook excitation based on a past excitation and an codebook index of the current frame of the first subset within the encoded bitstream, and setting a gain of the codebook excitation based on a global gain value within the encoded bitstream; and

a linear prediction synthesis filter configured to filter the current excitation based on linear prediction filter coefficients for the current frame of the first subset within the encoded bitstream;

a transform decoder configured to decode a current frame of the second subset by

constructing spectral information for the current frame of the second subset from the encoded bitstream and performing a spectral-to-time-domain transformation onto the spectral information to obtain a time-domain signal such that a level of the time-domain signal depends on the global gain value.

9. Multi-mode audio decoder according to claim 8, wherein the excitation generator is configured to, in generating the current excitation of the current frame of the first subset,

construct an adaptive codebook excitation based on a past excitation and an adaptive codebook index of the current frame of the first subset within the encoded bitstream;

construct an innovation codebook excitation based on an innovation codebook index for the current frame of the first subset within the encoded bitstream;

set, as the gain of the codebook excitation, a gain of the innovation codebook excitation based on the global gain value within the encoded bitstream; and

combine the adaptive codebook excitation and the innovation codebook excitation to obtain the current excitation of the current frame of the first subset.

10. Multi-mode audio decoder according to claim 8 or 9, wherein the transform decoder is configured such that the spectral information relates to a current excitation of the current frame of the second subset, and the transform decoder is further configured to, in decoding the current frame of the second subset, spectrally form the current excitation of the current frame of the second subset according to a linear prediction synthesis filter transfer function defined by linear prediction filter coefficients for the current frame of the second subset within the encoded bitstream so that the performance of the spectral-to-time-domain transformation onto the spectral information results in the decoded representation of the audio content.

11. Multi-mode audio decoder according to claim 10, wherein the transform decoder is configured to perform the spectral formation by converting the linear prediction filter coefficients into a linear prediction spectrum and weighting the spectral information of the current excitation with the linear prediction spectrum.

12. Multi-mode audio decoder according to any of claims 8 to 11, wherein the transform decoder is configured to scale the spectral information with the global gain value.

13. Multi-mode audio decoder according to claim 8 or 9, wherein the transform decoder is configured to construct the spectral information for the current frame of the second subset by use of spectral transform coefficients within the encoded bitstream, and scale factors within the encoded bitstream for scaling the spectral transform coefficients in a spectral granularity of scale factor bands, with scaling the scale factors based on the global gain value, so as to obtain the decoded representation of the audio content.

14. Multi-mode audio encoder configured to encode an audio content into an encoded bitstream with encoding a first subset of frames in a first coding mode and a second subset of frames in a second coding mode, wherein the second subset of frames is respectively composed of one or more sub-frames, wherein the multi-mode audio encoder is configured to determine and encode a global gain value per frame, and determine and encode, per sub-frames of at least a subset of the sub-frames of the second subset, a corresponding bitstream element differentially to the global gain value of the respective frame, wherein the multi-mode audio encoder is configured such that a change of the global gain value of the frames within the encoded bitstream results in an adjustment of an output level of a decoded representation of the audio content at the decoding side.

15. Multi-mode audio encoder for encoding an audio content into an encoded bitstream by CELP encoding a first subset of frames of the audio content and transform encoding a second subset of the frames, the multi-mode audio encoder comprising:

a CELP encoder configured to encode a current frame of the first subset, the CELP encoder comprising

a linear prediction analyzer configured to generate linear prediction filter coefficients for the current frame of the first subset and encode same into the encoded bitstream; and

an excitation generator configured to determine a current excitation of the current frame of the first subset, which, when filtered by a linear prediction synthesis filter based on the linear prediction filter coefficients within the

encoded bitstream, recovers the current frame of the first subset, defined by a past excitation and a codebook index for the current frame of the first subset and encoding the codebook index into the encoded bitstream; and

a transform encoder configured to encode a current frame of the second subset by performing a time-to-spectral-domain transformation onto a time-domain signal for the current frame of the second subset to obtain spectral information and encode the spectral information into the encoded bitstream,

wherein the multi-mode audio encoder is configured to encode a global gain value into the encoded bitstream, the global gain value depending on an energy of a version of the audio content of the current frame of the first subset, filtered with the linear prediction analysis filter depending on the linear prediction coefficients, or an energy of the time-domain signal.

16. Multi-mode audio decoding method for providing a decoded representation of audio content on the basis of an encoded bitstream, the method comprising

decoding a global gain value per frame of the encoded bitstream, wherein a first subset of the frames being coded in a first coding mode and a second subset of the frames being coded in a second coding mode, with each frame of the second subset being composed of more than one sub-frames,

decoding, per sub-frame of at least a subset of the sub-frames of the second subset of frames, a corresponding bitstream element differentially to the global gain value of the respective frame, and

completing decoding the bitstream using the global gain value and the corresponding bitstream element in decoding the sub-frames of the at least subset of the sub-frames of the second subset of frames and the global gain value in decoding the first subset of frames,

wherein the multi-mode audio decoding method is performed such that a change of the global gain value of the frames within the encoded bitstream results in an adjustment of an output level of the decoded representation of the audio content.

17. Multi-mode audio decoding method for providing a decoded representation of an audio content on the basis of an encoded bitstream, a first subset of frames of which is CELP coded and a second subset of frames of which is transform coded, the method comprising:

CELP decoding a current frame of the first subset, the CELP decoder comprising:

generating a current excitation of the current frame of the first subset by constructing an codebook excitation based on a past excitation and an codebook index of the current frame of the first subset within the encoded bitstream, and setting a gain of the codebook excitation based on a global gain value within the encoded bitstream; and

filtering the current excitation based on linear prediction filter coefficients for the current frame of the first subset within the encoded bitstream;

transform decoding a current frame of the second subset by

constructing spectral information for the current frame of the second subset from the encoded bitstream and performing a spectral-to-time-domain transformation onto the spectral information to obtain a time-domain signal such that a level of the time-domain signal depends on the global gain value.

18. Multi-mode audio encoding method comprising encoding an audio content into an encoded bitstream with encoding a first subset of frames in a first coding mode and a second subset of frames in a second coding mode, wherein the second subset of frames is respectively composed of one or more sub-frames, wherein the multi-mode audio encoding method further comprises determining and encoding a global gain value per frame, and determine and encode, per sub-frames of at least a subset of the sub-frames of the second subset, a corresponding bitstream element differentially to the global gain value of the respective frame, wherein the multi-mode audio encoding method is performed such that a change of the global gain value of the frames within the encoded bitstream results in an adjustment of an output level of a decoded representation of the audio content at the decoding side.

19. Multi-mode audio encoding method for encoding an audio content into an encoded bitstream by CELP encoding a first subset of frames of the audio content and

transform encoding a second subset of the frames, the multi-mode audio encoding method comprising:

encoding a current frame of the first subset, the CELP encoder comprising

performing linear prediction analysis to generate linear prediction filter coefficients for the current frame of the first subset and encode same into the encoded bitstream; and

determining a current excitation of the current frame of the first subset, which, when filtered by a linear prediction synthesis filter based on the linear prediction filter coefficients within the encoded bitstream, recovers the current frame of the first subset, defined by a past excitation and a codebook index for the current frame of the first subset and encoding the codebook index into the encoded bitstream; and

encoding a current frame of the second subset by performing a time-to-spectral-domain transformation onto a time-domain signal for the current frame of the second subset to obtain spectral information and encode the spectral information into the encoded bitstream,

wherein the multi-mode audio encoding method further comprises encoding a global gain value into the encoded bitstream, the global gain value depending on an energy of a version of the audio content of the current frame of the first subset, filtered with the linear prediction analysis filter depending on the linear prediction coefficients, or an energy of the time-domain signal.

20. Computer program having a program code for performing, when running on a computer, a method according to any of claims 16 to 19.