Claims

1. Encoder for encoding an audio signal, the encoder comprising:

an analyzer (120; 320) configured for deriving prediction coefficients (122; 322) and a residual signal from an unvoiced frame of the audio signal (102);

a gain parameter calculator (550; 550') configured for calculating a first gain parameter (g_c) information for defining a first excitation signal (c(n)) related to a deterministic codebook and for calculating a second gain parameter (g_n) information for defining a second excitation signal (n(n)) related to a noise-like signal for the unvoiced frame; and

a bitstream former (690) configured for forming an output signal (692) based on an information (142) related to a voiced signal frame, the first gain parameter (g_c) information and the second gain parameter (g_n) information;

wherein the encoder comprises an LTP memory (350n) and a signal generator (850) for generating an adaptive excitation signal for the voiced frame; and

wherein the encoder is configured for not transmitting LTP parameters for the unvoiced frame to save bits, wherein the adaptive excitation signal is set to zero for the unvoiced frame, and wherein the deterministic codebook is configured to code more pulses for a same bit-rate using the saved bits.

2. Encoder according to claim 1, wherein the gain parameter calculator (550; 550') is configured for calculating a first gain parameter (g_c) and a second gain parameter (g_n) and wherein the bitstream former (690) is configured for forming the output signal (692) based on the first gain parameter (g_c) and the second gain parameter(g_n); or

wherein the gain parameter calculator (550; 550') comprises a quantizer (170-1, 170-2) configured for quantizing the first gain parameter (g_c) for obtaining a first quantized gain parameter ($\hat{g}_c$) and for quantizing the second gain parameter (g_n) for obtaining a second quantized gain parameter ($\hat{g}_n$) and wherein the bitstream former (690) is configured for forming the output signal (692) based on the first quantized gain parameter ($\hat{g}_c$) and the second quantized gain parameter ($\hat{g}_n$).

3. Encoder according to claim 1 or 2, further comprising a formant information calculator (160) configured for calculating a speech related spectral shaping information (162) from the prediction coefficients (122; 322) and wherein the gain parameter calculator (550; 550') is configured to calculate the first gain parameter information (g_c) and the second gain parameter information (g_n) based on the speech related spectral shaping information (162).

4. Encoder according to one of previous claims, wherein the gain parameter calculator (550') comprises:

a first amplifier (550e) configured for amplifying the first excitation signal (c(n)) by applying the first gain parameter g_c to obtain a first amplified excitation signal (550f);

a second amplifier (350e; 550g) configured for amplifying the second excitation signal (n(n)) different from the first excitation signal (c(n)) by applying the second gain parameter (g_n) to obtain a second amplified excitation signal (350g; 550h);

a combiner (550i) configured for combining the first amplified excitation signal (550f) and the second amplified excitation signal (350g; 550h) to obtain a combined excitation signal (550k; 550k');

a controller (550n) configured for filtering the combined excitation signal (550k; 550k') with a synthesis filter to obtain a synthesized signal (350l'), for comparing the synthesized signal (350l') and the audio signal frame (102) to obtain a comparison result, to adapt the first gain parameter (g_c) or the second gain parameter (g_n) based on the comparison result; and

wherein the bitstream former (690) is configured for forming the output signal (692) based on an information ($\hat{g}_c$; $\hat{g}_n$) related to the first gain parameter (g_c) and the second gain parameter (g_n).

5. Encoder according to one of previous claims, wherein the gain parameter controller (550; 550') further comprises at least one shaper (350; 550b) configured for spectrally shaping the first excitation signal (c(n)) or a signal derived thereof or the second excitation signal (n(n)) or a signal derived thereof based on a spectral shaping information (162).

6. Encoder according to one of previous claims, wherein the encoder is configured for encoding the audio signal (102) framewise in a sequence of frames and wherein the gain parameter calculator (550; 550') is configured for determining the first gain parameter (g_c) and the second gain parameter (g_n) for each of a plurality of subframes of a processed frame and wherein the gain parameter controller (550; 550') is configured for determining an average energy value associated to the processed frame.

7. Encoder according to one of previous claims, further comprising:

a formant information calculator (160) configured for calculating at least a first a speech related spectral shaping information from the prediction coefficients (122; 322);

a decider (130) configured for determining if the residual signal was determined from an unvoiced signal audio frame.

8. Encoder according to one of previous claims, wherein the gain parameter controller (550; 550') comprises a controller (550n) configured for determining the first gain parameter (g_c) based on:

$$g_c = \frac{\sum_{n=0}^{Lsf-1} xw(n) \cdot cw(n)}{\sum_{n=0}^{Lsf-1} cw(n) \cdot cw(n)}$$

wherein cw(n) is a filtered excitation signal of an innovative codebook and xw(n) is a perceptual target excitation computed in CELP encoder;

wherein the controller (550n) is configured to determine the quantized noise gain ($\widehat{g_n}$) based on quantized value of the first gain parameter ($\widehat{g_c}$) and the root square energy ratio between the first excitation and the second excitation:

$$\frac{\sum_{n=0}^{Lsf-1} \sqrt{c(n) \cdot c(n)}}{\sum_{n=0}^{Lsf-1} \sqrt{n(n) \cdot n(n)}}$$

wherein Lsf is the size in samples of a subframe.

PCT/EP 2014/071 769

9. Encoder according to one of previous claims, further comprising a quantizer (170-1, 170-2) configured for quantizing the first gain parameter (g_c) to obtain a quantized first gain parameter ($\widehat{g_c}$), wherein the gain parameter controller (550n) is configured for determining the first gain parameter (g_c) as a based on:

$$g_c = \frac{\sum_{n=0}^{Lsf-1} xw(n) \cdot cw(n)}{\sum_{n=0}^{Lsf-1} cw(n) \cdot cw(n)}$$

wherein gc is the first gain parameter, Lsfis the size of the subframe in samples, cw(n) denotes the first shaped excitation signal, xw(n) denotes a Code Excited Linear Prediction encoding signal.

wherein the gain parameter controller (550n) or the quantizer (170-1, 170-2) is further configured for normalizing the first gain parameter (g_c) to obtain a normalized first gain parameter based on:

$$g_{nc} = g_c \cdot \frac{\sum_{n=0}^{Lsf-1} \sqrt{c(n) \cdot c(n)}}{Lsf \cdot 10^{\widehat{nrg}/20}}$$

wherein g_{nc} denotes the normalized fist gain parameter and $\widehat{nrg}$ is a measure for an average energy of the unvoiced residual signal over the whole frame; and

wherein the quantizer (170-1, 170-2) is configured for quantizing the normalized first gain parameter to obtain the quantized first gain parameter ($\widehat{g_c}$).

10. Encoder according to claim 9, wherein the quantizer (170-1, 170-2) is configured for quantizing the second gain parameter (g_n) to obtain a quantized second gain parameter ($\widehat{g}_n$) wherein the gain parameter controller (550; 550') is configured to determine the second gain parameter (g_n) by determining an error value based on:

$$\frac{1}{Lsf}\left|\sum_{n=0}^{Lsf-1} k \cdot xw^2(n) - \sum_{n=0}^{Lsf-1} (\widehat{g_c} \cdot cw(n) + g_n nw(n))^2\right|$$

wherein k is a variable attenuation factor in a range between 0.5 and 1, Lsf corresponds to the size of a subframe of a processed audio frame, cw(n) denotes the first shaped excitation signal (c(n)), xw(n) denotes a Code Excited Linear Prediction encoding signal, gn denotes the second gain parameter and $\widehat{g_c}$ denotes a quantized first gain parameter;

wherein the gain parameter controller (550; 550') is configured for determining the error for the current subframe and wherein the quantizer (170-1, 170-2) is configured for determining the quantized second gain ($\widehat{g_n}$) which minimizes the error and for obtaining the quantized second gain ($\widehat{g_n}$) based on:

$$\widehat{g_n} = Q(index_n) \cdot \widehat{g_c} \cdot \frac{\sum_{n=0}^{Lsf-1} \sqrt{c(n) \cdot c(n)}}{\sum_{n=0}^{Lsf-1} \sqrt{n(n) \cdot n(n)}}$$

where $Q(index_n)$ denotes a scalar value from a finite set a possible values.

11. Encoder according to claim 10, wherein the combiner (550i) is configured for combining the first gain parameter (g_c) and the second gain parameter (g_n) to obtain a combines excitation signal (e(n)) based on:

$$e(n) = \widehat{g_c} \cdot c(n) + \widehat{g_n} \cdot n(n)$$

12. Encoder according to claim 10 or 11, wherein the quantizer (170-2) is configured for determining the error value based on an energy mismatch between the first shaped excitation signal (c(n)) and the second excitation signal, wherein the quantizer (170-1) is configured for determining the first gain parameter (g_c) based on a mean squared error or mean squared root error.

13. Encoder according to one of previous claims, wherein the encoder is configured for not transmitting long term prediction parameters for unvoiced frames.

14. Decoder (1000) for decoding a received audio signal (1002) comprising an information related to prediction coefficients (122), the decoder (1000) comprising:

a first signal generator (1010) configured for generating a first excitation signal (1012) from a deterministic codebook for a portion of a synthesized signal (1062);

a second signal generator (1020) configured for generating a second excitation signal (1022) from a noise-like signal for the portion of the synthesized signal (1062);

a combiner (1050) configured for combining the first excitation signal (1012) and the second excitation signal (1022) for generating a combined excitation signal (1052) for the portion of the synthesized signal (1062); and

a synthesizer (1060) configured for synthesizing the portion of the synthesized signal (1062) from the combined excitation signal (1052) and the prediction coefficients (122);

wherein the decoder comprises an LTP memory (350n) and a signal generator (850) for generating an adaptive excitation signal for the voiced frame; and

wherein the received audio signal does not comprise LTP parameters for the unvoiced frame, wherein the decoder is configured for setting to zero the adaptive excitation signal for the unvoiced frame, and wherein the deterministic codebook is configured to provide more pulses for a same bit-rate due to bits saved because of the lack of LTP parameters for the unvoiced frame.

15. Decoder according to claim 14, wherein the received audio signal (1002) comprises an information related to a first gain parameter (g_c) and to a second gain parameter (g_n), wherein the decoder further comprises:

a first amplifier (254; 350e; 550e) configured for amplifying the first excitation signal (1012) or a signal derived thereof by applying the first gain parameter (g_c) to obtain a first amplified excitation signal (1012');

a second amplifier (254; 350e; 550e) configured for amplifying the second excitation signal (1022) or a signal derived by applying the second gain parameter to obtain a second amplified excitation signal (1022');

16. Decoder according to claim 14 or 15, further comprising:

a formant information calculator (160; 1090) configured for calculating a first spectral shaping information (1092a) and a second spectral shaping information (1092b) from the prediction coefficients (122; 322);

a first shaper (1070) for spectrally shaping a spectrum of the first excitation signal (1012) or a signal derived thereof using the first spectral shaping information (1092a); and

a second shaper (1080) for spectrally shaping a spectrum of the second excitation signal (1022) or a signal derived thereof using the second shaping information (1092b);

17. Encoded audio signal (692; 1002) comprising an information related to prediction coefficients (122; 322), an information related to a deterministic codebook, an information related to a first gain parameter (g_c) and a second gain parameter (g_n) and an information (142) related to a voiced and an unvoiced signal frame;

wherein the encoded audio signal comprises information related to an adaptive excitation signal for the voiced frame; and

wherein the encoded audio signal is not comprising LTP parameters for the unvoiced frame, wherein the adaptive excitation signal is set to zero for the unvoiced frame.

18. Method (1400) for encoding an audio signal (102) , the method comprising:

deriving (1410) prediction coefficients (122; 322) and a residual signal from an unvoiced frame of the audio signal (102);

calculating (1420) a first gain parameter information ($\widehat{g_c}$) for defining a first excitation signal (c(n)) related to a deterministic codebook and for calculating a second gain parameter information ($\hat{g}_n$) for defining a second excitation signal (n(n)) related to a noise-like signal (n(n)) for the unvoiced frame; and

forming (1430) an output signal (692; 1002) based on an information (142) related to a voiced signal frame, the first gain parameter information ($\widehat{g_c}$) and the second gain parameter information ($\hat{g}_n$);

generating an adaptive excitation signal for the voiced frame using an LTP memory (350n) and a signal generator (850); and

not transmitting LTP parameters for the unvoiced frame to save bits, setting the adaptive excitation signal to zero for the unvoiced frame, and coding more pulses for a same bit-rate using the deterministic codebook and using the saved bits.

19. Method (1500) for decoding a received audio signal (692; 1002) comprising an information related to prediction coefficients (122; 322), the received audio signal not comprise LTP parameters for the unvoiced frame the method comprising:

generating (1510) a first excitation signal (1012, 1012') from a deterministic codebook for a portion of a synthesized signal (1062);

generating (1520) a second excitation signal (1022, 1022') from a noise-like signal (n(n)) for the portion of the synthesized signal (1062);

combining (1530) the first excitation signal (1012, 1012') and the second excitation signal (1022, 1022') for generating a combined excitation signal (1052) for the portion of the synthesized signal (1062); and

synthesizing (1540) the portion of the synthesized signal (1062) from the combined excitation signal (1052) and the prediction coefficients (122; 322);

generating an adaptive excitation signal for the voiced frame using an LTP memory (350n) and a signal generator (850); and

setting to zero the adaptive excitation signal for the unvoiced frame, and providing more pulses for a same bit-rate due to bits saved because of the lack of LTP parameters for the unvoiced frame using the deterministic codebook.

20. Computer program having a program code for executing a method according to claim 18 or 19 when running on a computer.

Claims

1. Encoder for encoding an audio signal, the encoder comprising:

an analyzer configured for deriving prediction coefficients and a residual signal from an unvoiced frame of the audio signal;

a gain parameter calculator configured for calculating a first gain parameter information for defining a first excitation signal related to a deterministic codebook and for calculating a second gain parameter information for defining a second excitation signal related to a noise-like signal for the unvoiced frame; and

a bitstream former configured for forming an output signal based on an information related to a voiced signal frame, the first gain parameter information and the second gain parameter information.

2. Encoder according to claim 1, wherein the gain parameter calculator is configured for calculating a first gain parameter and a second gain parameter and wherein the bitstream former is configured for forming the output signal based on the first gain parameter and the second gain parameter; or

wherein the gain parameter calculator comprises a quantizer configured for quantizing the first gain parameter for obtaining a first quantized gain parameter and for quantizing the second gain parameter for obtaining a second quantized gain parameter and wherein the bitstream former is configured for forming the output signal based on the first quantized gain parameter and the second quantized gain parameter.

3. Encoder according to claim 1 or 2, further comprising a formant information calculator configured for calculating a speech related spectral shaping information from the prediction coefficients and wherein the gain parameter calculator is configured to calculate the first gain parameter information and the second gain parameter information based on the speech related spectral shaping information.

4. Encoder according to one of previous claims, wherein the gain parameter calculator comprises:

a first amplifier configured for amplifying the first excitation signal by applying the first gain parameter to obtain a first amplified excitation signal;

a second amplifier configured for amplifying the second excitation signal different from the first excitation signal by applying the second gain parameter to obtain a second amplified excitation signal;

a combiner configured for combining the first amplified excitation signal and the second amplified excitation signal to obtain a combined excitation signal;

a controller configured for filtering the combined excitation signal with a synthesis filter to obtain a synthesized signal, for comparing the synthesized signal and the audio signal frame to obtain a comparison result, to adapt the first gain parameter or the second gain parameter based on the comparison result; and

wherein the bitstream former is configured for forming the output signal based on an information related to the first gain parameter and the second gain parameter.

5. Encoder according to one of previous claims, wherein the gain parameter controller further comprises at least one shaper configured for spectrally shaping the first excitation signal or a signal derived thereof or the second excitation signal or a signal derived thereof based on a spectral shaping information.

6. Encoder according to one of previous claims, wherein the encoder is configured for encoding the audio signal framewise in a sequence of frames and wherein the gain parameter calculator is configured for determining the first gain parameter and the second gain parameter for each of a plurality of subframes of a processed frame and wherein the gain parameter controller is configured for determining an average energy value associated to the processed frame.

7. Encoder according to one of previous claims, further comprising:

a formant information calculator configured for calculating at least a first a speech related spectral shaping information from the prediction coefficients;

a decider configured for determining if the residual signal was determined from an unvoiced signal audio frame.

8. Encoder according to one of previous claims, wherein the gain parameter controller comprises a controller configured for determining the first gain parameter based on:

$$g_c = \frac{\sum_{n=0}^{Lsf-1} xw(n) \cdot cw(n)}{\sum_{n=0}^{Lsf-1} cw(n) \cdot cw(n)}$$

wherein cw(n) is a filtered excitation signal of an innovative codebook and xw(n) is a perceptual target excitation computed in CELP encoder;

wherein the controller is configured to determine the quantized noise gain based on quantized value of the first gain parameter and the root square energy ratio between the first excitation and the second excitation:

$$\frac{\sum_{n=0}^{Lsf-1} \sqrt{c(n) \cdot c(n)}}{\sum_{n=0}^{Lsf-1} \sqrt{n(n) \cdot n(n)}}$$

wherein Lsf is the size in samples of a subframe.

9. Encoder according to one of previous claims, further comprising a quantizer configured for quantizing the first gain parameter to obtain a quantized first gain parameter, wherein the gain parameter controller is configured for determining the first gain parameter as a based on:

$$g_c = \frac{\sum_{n=0}^{Lsf-1} xw(n) \cdot cw(n)}{\sum_{n=0}^{Lsf-1} cw(n) \cdot cw(n)}$$

wherein gc is the first gain parameter, Lsf is the size of the subframe in samples, cw(n) denotes the first shaped excitation signal, xw(n) denotes a Code Excited Linear Prediction encoding signal

wherein the gain parameter controller or the quantizer is further configured for normalizing the first gain parameter to obtain a normalized first gain parameter based on:

$$g_{nc} = g_c \cdot \frac{\sum_{n=0}^{Lsf-1} \sqrt{c(n) \cdot c(n)}}{Lsf \cdot 10^{\widehat{nrg}/20}}$$

wherein g_{nc} denotes the normalized first gain parameter and $\widehat{nrg}$ is a measure for an average energy of the unvoiced residual signal over the whole frame; and

wherein the quantizer is configured for quantizing the normalized first gain parameter to obtain the quantized first gain parameter.

10. Encoder according to claim 9, wherein the quantizer is configured for quantizing the second gain parameter to obtain a quantized second gain parameter wherein the gain parameter controller is configured to determine the second gain parameter by determining an error value based on:

$$\frac{1}{Lsf}\left|\sum_{n=0}^{Lsf-1} k \cdot xw^2(n) - \sum_{n=0}^{Lsf-1} (\widehat{g_c} \cdot cw(n) + g_n nw(n))^2\right|$$

wherein is a variable attenuation factor in a range between 0.5 and 1, Lsf corresponds to the size of a subframe of a processed audio frame, , cw(n) denotes the first shaped excitation signal, xw(n) denotes a Code Excited Linear Prediction encoding signal, gn denotes the second gain parameter and $\widehat{g_c}$ denotes a quantized first gain parameter;

wherein the gain parameter controller is configured for determining the error for the current subframe and wherein the quantizer is configured for determining the quantized second gain which minimizes the error and for obtaining the quantized second gain based on:

$$\widehat{g_n} = Q(index_n) \cdot \widehat{g_c} \cdot \frac{\sum_{n=0}^{Lsf-1} \sqrt{c(n) \cdot c(n)}}{\sum_{n=0}^{Lsf-1} \sqrt{n(n) \cdot n(n)}}$$

where $Q(index_n)$ denotes a scalar value from a finite set a possible values.

11. Encoder according to claim 10, wherein the combiner is configured for combining the first gain parameter and the second gain parameter to obtain a combines excitation signal based on:

$$e(n) = \widehat{g_c} \cdot c(n) + \widehat{g_n} \cdot n(n)$$

12. Decoder for decoding a received audio signal comprising an information related to prediction coefficients, the decoder comprising:

a first signal generator configured for generating a first excitation signal from a deterministic codebook for a portion of a synthesized signal;

a second signal generator configured for generating a second excitation signal from a noise-like signal for the portion of the synthesized signal;

a combiner configured for combining the first excitation signal and the second excitation signal for generating a combined excitation signal for the portion of the synthesized signal; and

a synthesizer configured for synthesizing the portion of the synthesized signal from the combined excitation signal and the prediction coefficients.

13. Decoder according to claim 12, wherein the received audio signal comprises an information related to a first gain parameter and to a second gain parameter, wherein the decoder further comprises:

a first amplifier configured for amplifying the first excitation signal or a signal derived thereof by applying the first gain parameter to obtain a first amplified excitation signal;

a second amplifier configured for amplifying the second excitation signal or a signal derived by applying the second gain parameter to obtain a second amplified excitation signal;

14. Decoder according to claim 12 or 13, further comprising:

a formant information calculator configured for calculating a first spectral shaping information and a second spectral shaping information from the prediction coefficients;

a first shaper for spectrally shaping a spectrum of the first excitation signal or a signal derived thereof using the first spectral shaping information; and

a second shaper for spectrally shaping a spectrum of the second excitation signal or a signal derived thereof using the second shaping information;

15. Encoded audio signal comprising an information related to prediction coefficients, an information related to a deterministic codebook, an information related to a first gain parameter and a second gain parameter and an information related to a voiced and an unvoiced signal frame.

16. Method for encoding an audio signal, the method comprising:

deriving prediction coefficients and a residual signal from an unvoiced frame of the audio signal;

calculating a first gain parameter information for defining a first excitation signal related to a deterministic codebook and for calculating a second gain parameter information for defining a second excitation signal related to a noise-like signal for the unvoiced frame; and

forming an output signal based on an information related to a voiced signal frame, the first gain parameter information and the second gain parameter information.

17. Method for decoding a received audio signal comprising an information related to prediction coefficients, the decoder comprising:

generating a first excitation signal from a deterministic codebook for a portion of a synthesized signal;

generating a second excitation signal from a noise-like signal for the portion of the synthesized signal;

combining the first excitation signal and the second excitation signal for generating a combined excitation signal for the portion of the synthesized signal; and

synthesizing the portion of the synthesized signal from the combined excitation signal and the prediction coefficients.

18. Computer program having a program code for executing a method according to claim 16 or 17 when running on a computer.