(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 3 252 763 A1

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
06.12.2017 Bulletin 2017/49

(51) Int Cl.:
G10L 19/18 *(2013.01)* *G10L 19/22* *(2013.01)*

(21) Application number: **16171853.1**

(22) Date of filing: **30.05.2016**

(84) Designated Contracting States:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR
Designated Extension States:
BA ME
Designated Validation States:
MA MD

(71) Applicant: **Nokia Technologies Oy**
02610 Espoo (FI)

(72) Inventors:
- **VASILACHE, Adriana**
  33580 Tampere (FI)
- **RÄMÖ, Anssi**
  33720 Tampere (FI)

(74) Representative: **Nokia EPO representatives**
Nokia Technologies Oy
Karaportti 3
02610 Espoo (FI)

(54) LOW-DELAY AUDIO CODING

(57) A technique for encoding a frame of an input audio signal that comprises a time series of input samples into a frame of an encoded audio signal is provided. In an example, the technique comprises encoding said frame of the input audio signal using at least two of a plurality of audio encoding modes, wherein each of said plurality of audio encoding modes is arranged to encode the frame of the input audio signal into a respective encoded signal, wherein said plurality of audio encoding modes include at least a first audio encoding mode that comprises linear predictive filtering of said time series of input samples using linear predictive filter coefficients computed using a backward prediction into a residual signal that comprises a respective time series of residual samples and quantizing the time series of residual samples, and a second audio encoding mode that comprises directly quantizing the time series of input samples, and selecting, in accordance with a mode selection rule, one of the respective encoded signals as the frame of the encoded audio signal.

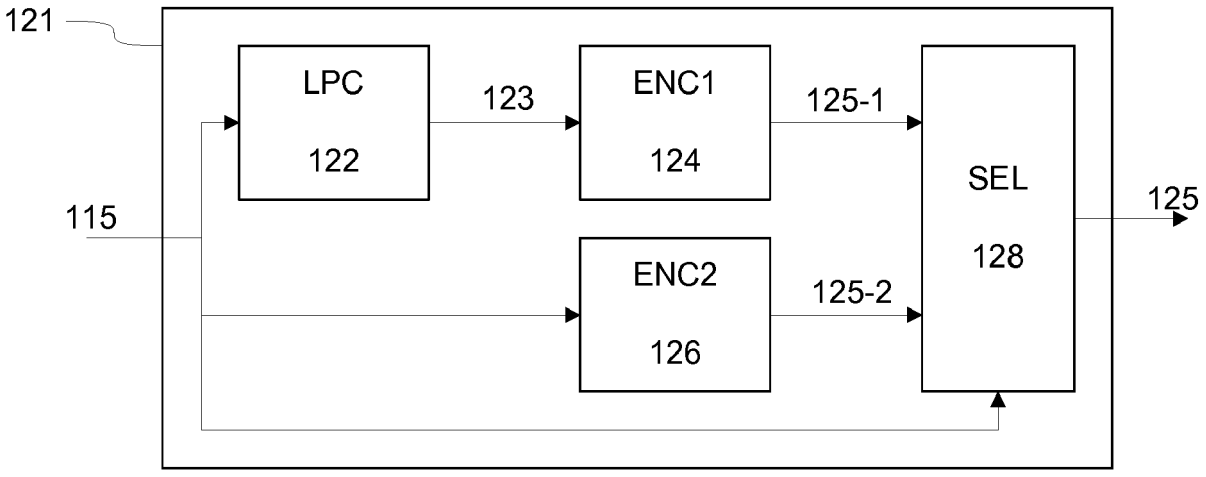

Figure 2

EP 3 252 763 A1

Description

TECHNICAL FIELD

[0001] The example and non-limiting embodiments of the present invention relate to very low-delay coding of audio signals at high sound quality.

BACKGROUND

[0002] Development of speech and audio coding techniques has evolved into solutions that enable high compression ratio at a good sound quality across input audio signals of various characteristics and across a wide-range of encoding bit-rates. Typically, achieving a high compression ratio in an audio coding technique that operates on a full-band audio signal (typically employing a sampling frequency of 48 kHz) requires usage of a relatively long analysis window in a range of 150 milliseconds (ms) or above to ensure sufficient sound quality. Consequently, a coding delay (or algorithmic delay) of such audio coding techniques is in the range of 150 ms or above. Examples of commonly employed audio coding techniques of this type include e.g. MPEG-1/MPEG-2 audio layer 3 (MP3) and MPEG-2/MPEG-4 advanced audio coding (AAC).

[0003] When such an audio coding technique is applied in an audio processing system that involves e.g. capturing and processing an audio signal and related processing, encoding the captured/processed audio signal, transmitting the encoded audio signal from one entity to another, decoding the received encoded audio signal and reproducing the decoded audio signal, the overall processing delay typically increases clearly beyond the mere coding delay, thereby rendering such audio coding techniques unsuitable for applications that cannot tolerate long latency such as telephony, wireless microphones or audio co-creation systems.

[0004] Speech coding techniques, such as adaptive multi-rate (AMR), adaptive multi-rate wideband (AMR-WB) and 3GPP enhanced voice services (EVS) employ coding delay in the range of 25 to 32 ms, which makes them somewhat better suited for some latency-critical applications. However, although enabling high compression ratio, these coding techniques are speech coding techniques that operate on bandwidth-limited audio signals at a relatively low-bitrates, thereby providing an audio quality that is not suited for applications that require high-quality full-band audio. There are also speech coding techniques such as. ITU-T G.726, G.728 and G.722 that enable very low coding delay even in a range below 1 ms, but also these coding techniques operate on voice band (e.g. at 8 or 16 kHz sampling frequency) and provide a rather modest compression ratio.

[0005] Some recently introduced audio coding techniques such as Opus (in a low-delay mode) and AAC-ULD enable relatively low coding delay in a range from 2.5 to 20 ms for full-band audio at a relatively good sound quality. As an example, assuming sampling frequency of 32 kHz, the AAC-ULD coding technique enables good sound quality using a coding delay of approximately 8 ms at bit-rates around 72 to 96 kilobits per second (kbps) or using a coding delay of approximately 2 ms at bit-rates around 128 to 192 kbps. While such coding delays make these audio coding techniques feasible candidates for many low-latency applications and usage scenarios, there is still a need for high-quality full-band audio coding technique that enables extremely low coding delay, e.g. one that is around 2.5 ms or below at bit rates at or close to 128 kbps and below.

SUMMARY

[0006] According to an example embodiment, a method for encoding a frame of an input audio signal that comprises a time series of input samples into a frame of an encoded audio signal is provided, the method comprising encoding said frame of the input audio signal using at least two of a plurality of audio encoding modes, wherein each of said plurality of audio encoding modes is arranged to encode the frame of the input audio signal into a respective encoded signal, wherein said plurality of audio encoding modes include at least a first audio encoding mode that comprises linear predictive filtering of said time series of input samples using linear predictive filter coefficients computed using a backward prediction into a residual signal that comprises a respective time series of residual samples and quantizing the time series of residual samples, and a second audio encoding mode that comprises directly quantizing the time series of input samples, and selecting, in accordance with a mode selection rule, one of the respective encoded signals as the frame of the encoded audio signal.

[0007] The selecting one of the respective encoded signals as the frame of the encoded audio signal may comprise: computing a respective distortion value for each of said respective encoded signals; and selecting the respective encoded signal that results in the smallest distortion value as the frame of the encoded audio signal.

[0008] The computing of a distortion value for a given respective encoded signal may comprise: creating a reconstructed audio signal on basis of the given respective encoded signal; and computing the distortion value as a value that is indicative of the difference between said frame of the input audio signal and the reconstructed audio signal.

[0009] The first audio encoding mode may comprise computing the linear predictive filter coefficients on basis of a reconstructed audio signal derived on basis of one or more frames of encoded audio signal that immediately precede said frame of the input audio signal.

[0010] The first audio encoding mode may comprise encoding said time series of the residual samples by using a first gain-shape encoder to generate a first gain and first relative sample values that represent said frame of the residual signal.

[0011] The first audio encoding mode may comprise quantizing the first gain and the first relative sample values that represent said frame of the residual signal by using a first pyramidally truncated lattice quantizer.

[0012] The second audio encoding mode may comprise encoding said time series of the input samples by using a second gain-shape encoder to generate a second gain and second relative sample values that represent said frame of the input audio signal.

[0013] The second audio encoding mode may comprise quantizing the second gain and the second relative sample values that represent said frame of the input audio signal by using a second pyramidally truncated lattice quantizer.

[0014] The second gain-shape encoder may comprise the first gain-shape encoder; and the second pyramidally truncated lattice quantizer may comprise the first pyramidally truncated lattice quantizer.

[0015] The method may further comprise providing an indication of the selected audio encoding mode in said frame of the encoded audio signal.

[0016] According to another example embodiment, a method for decoding a frame of an encoded audio signal into a frame of a reconstructed audio signal that comprises a time series of output samples, the method comprising decoding said frame of the encoded audio signal with one of a plurality of audio decoding modes, wherein said plurality of audio decoding modes include at least a first audio decoding mode and a second audio decoding mode, wherein the first audio decoding mode comprises dequantizing encoded residual parameters received in said frame of the encoded audio signal into a frame of reconstructed residual signal that comprises a time series of reconstructed residual samples and linear predictive filtering of said time series of reconstructed residual samples into said time series of output samples using linear predictive filter coefficients computed using a backward prediction, and wherein the second audio decoding mode comprises directly dequantizing encoded signal-domain parameters received in said frame of the encoded audio signal into said time series of output samples.

[0017] The method may further comprise: receiving an indication of one of the plurality of audio encoding modes; and decoding said frame of the encoded audio signal using one of the plurality of audio decoding modes in accordance with said received indication.

[0018] The first audio decoding mode may comprise computing the linear predictive filter coefficients on basis of a plurality of samples of reconstructed audio signal that immediately precede said frame of the reconstructed audio signal.

[0019] The encoded residual parameters may comprise a first gain and first relative sample values that represent said frame of the reconstructed residual signal; and the first audio decoding mode may comprise decoding said first gain and said first relative sample values using a first gain-shape decoder.

[0020] The first audio decoding mode may comprise dequantizing the first gain and the first relative sample values by using a first pyramidally truncated lattice quantizer.

[0021] The encoded signal-domain parameters may comprise a second gain and second relative sample values that represent said frame of the reconstructed audio signal; and the second audio decoding mode may comprise decoding said second gain and said second relative sample values using a second gain-shape decoder.

[0022] The second audio decoding mode may comprise dequantizing the second gain and the second relative sample values by using a second pyramidally truncated lattice quantizer.

[0023] The second gain-shape encoder may comprise the first gain-shape encoder; and the second pyramidally truncated lattice quantizer may comprise the first pyramidally truncated lattice quantizer.

[0024] According to another example embodiment, an apparatus for encoding a frame of an input audio signal that comprises a time series of input samples into a frame of an encoded audio signal is provided, the apparatus configured to: encode said frame of the input audio signal using at least two of a plurality of audio encoding modes, wherein each of said plurality of audio encoding modes is arranged to encode the frame of the input audio signal into a respective encoded signal, wherein said plurality of audio encoding modes include at least a first audio encoding mode configured to linear predictive filter said time series of input samples using linear predictive filter coefficients computed using a backward prediction into a residual signal that comprises a respective time series of residual samples and quantize the time series of residual samples, and a second audio encoding mode configured to directly quantize the time series of input samples, and select, in accordance with a mode selection rule, one of the respective encoded signals as the frame of the encoded audio signal.

[0025] According to another example embodiment, an apparatus for decoding a frame of an encoded audio signal into a frame of a reconstructed audio signal that comprises a time series of output samples is provided, the apparatus configured to decode said frame of the encoded audio signal with one of a plurality of audio decoding modes, wherein said plurality of audio decoding modes include at least a first audio decoding mode and a second audio decoding mode, wherein the first audio decoding mode is configured to dequantize encoded residual parameters received in said frame

of the encoded audio signal into a frame of reconstructed residual signal that comprises a time series of reconstructed residual samples and linear predictive filter said time series of reconstructed residual samples into said time series of output samples using linear predictive filter coefficients computed using a backward prediction, and wherein the second audio decoding mode is configured to directly dequantize encoded signal-domain parameters received in said frame of the encoded audio signal into said time series of output samples.

[0026] According to another example embodiment, an apparatus for encoding a frame of an input audio signal that comprises a time series of input samples into a frame of an encoded audio signal is provided, the apparatus comprising audio encoding means for encoding said frame of the input audio signal using at least two of a plurality of audio encoding modes, wherein each of said plurality of audio encoding modes is arranged to encode the frame of the input audio signal into a respective encoded signal, wherein said plurality of audio encoding modes include at least a first audio encoding mode that comprises linear predictive filtering of said time series of input samples using linear predictive filter coefficients computed using a backward prediction into a residual signal that comprises a respective time series of residual samples and quantizing the time series of residual samples, and a second audio encoding mode that comprises directly quantizing the time series of input samples, and selection means for selecting, in accordance with a mode selection rule, one of the respective encoded signals as the frame of the encoded audio signal.

[0027] According to another example embodiment, an apparatus for decoding a frame of an encoded audio signal into a frame of a reconstructed audio signal that comprises a time series of output samples is provided, the apparatus comprising audio decoding means for decoding said frame of the encoded audio signal with one of a plurality of audio decoding modes, wherein said plurality of audio decoding modes include at least a first audio decoding mode and a second audio decoding mode, wherein the first audio decoding mode comprises dequantizing encoded residual parameters received in said frame of the encoded audio signal into a frame of reconstructed residual signal that comprises a time series of reconstructed residual samples and linear predictive filtering of said time series of reconstructed residual samples into said time series of output samples using linear predictive filter coefficients computed using a backward prediction, and wherein the second audio decoding mode comprises directly dequantizing encoded signal-domain parameters received in said frame of the encoded audio signal into said time series of output samples.

[0028] According to another example embodiment, an apparatus for encoding a frame of an input audio signal that comprises a time series of input samples into a frame of an encoded audio signal is provided, wherein the apparatus comprises at least one processor; and at least one memory including computer program code, which when executed by the at least one processor, causes the apparatus to: encode said frame of the input audio signal using at least two of a plurality of audio encoding modes, wherein each of said plurality of audio encoding modes is arranged to encode the frame of the input audio signal into a respective encoded signal, wherein said plurality of audio encoding modes include at least a first audio encoding mode configured to linear predictive filter said time series of input samples using linear predictive filter coefficients computed using a backward prediction into a residual signal that comprises a respective time series of residual samples and quantize the time series of residual samples, and a second audio encoding mode configured to directly quantizing the time series of input samples, and select, in accordance with a mode selection rule, one of the respective encoded signals as the frame of the encoded audio signal.

[0029] According to another example embodiment, an apparatus for decoding a frame of an encoded audio signal into a frame of a reconstructed audio signal that comprises a time series of output samples is provided, wherein the apparatus comprise at least one processor; and at least one memory including computer program code, which when executed by the at least one processor, causes the apparatus to: decode said frame of the encoded audio signal with one of a plurality of audio decoding modes, wherein said plurality of audio decoding modes include at least a first audio decoding mode and a second audio decoding mode, wherein the first audio decoding mode is configured to dequantize encoded residual parameters received in said frame of the encoded audio signal into a frame of reconstructed residual signal that comprises a time series of reconstructed residual samples and linear predictive filter said time series of reconstructed residual samples into said time series of output samples using linear predictive filter coefficients computed using a backward prediction, and wherein the second audio decoding mode is configured to directly dequantize encoded signal-domain parameters received in said frame of the encoded audio signal into said time series of output samples.

[0030] According to another example embodiment, a computer program is provided, the computer program comprising computer readable program code configured to cause performing at least a method according to the example embodiment described in the foregoing when said program code is executed on a computing apparatus:

The computer program according to an example embodiment may be embodied on a volatile or a non-volatile computer-readable record medium, for example as a computer program product comprising at least one computer readable non-transitory medium having program code stored thereon, the program which when executed by an apparatus cause the apparatus at least to perform the operations described hereinbefore for the computer program according to an example embodiment of the invention.

The exemplifying embodiments of the invention presented in this patent application are not to be interpreted to pose

limitations to the applicability of the appended claims. The verb "to comprise" and its derivatives are used in this patent application as an open limitation that does not exclude the existence of also unrecited features. The features described hereinafter are mutually freely combinable unless explicitly stated otherwise.

[0031] Some features of the invention are set forth in the appended claims. Aspects of the invention, however, both as to its construction and its method of operation, together with additional objects and advantages thereof, will be best understood from the following description of some example embodiments when read in connection with the accompanying drawings.

BRIEF DESCRIPTION OF FIGURES

[0032] The embodiments of the invention are illustrated by way of example, and not by way of limitation, in the figures of the accompanying drawings, where

Figure 1 illustrates a block diagram of some components and/or entities of an audio processing system within which one or more example embodiments may be implemented.

Figure 2 illustrates a block diagram of some components and/or entities of an audio encoder according to an example embodiment;

Figure 3 illustrates a block diagram of some components and/or entities of an audio decoder according to an example embodiment;

Figure 4 illustrates a method according to an example embodiment;

Figure 5 illustrates a method according to an example embodiment; and

Figure 6 illustrates a block diagram of some components and/or entities of an apparatus for implementing an audio encoder and/or an audio decoder according to an example embodiment.

DESCRIPTION OF SOME EMBODIMENTS

[0033] Figure 1 schematically illustrates a block diagram of some components and/or entities of an audio processing system 100. The audio processing system comprises an audio capturing entity 110 for capturing an input audio signal 115 that represents at least one sound, an audio encoding entity 120 for encoding the input audio signal 115 into an encoded audio signal 125, an audio decoding entity 130 for decoding the encoded audio signal 125 obtained from the audio encoding entity into a reconstructed audio signal 135, and an audio reproduction entity 140 for playing back the reconstructed audio signal 135.

[0034] The audio capturing entity 110 may comprise e.g. a microphone, an arrangement of two or more microphones or a microphone array, each operable for capturing a respective sound signal. The audio capturing entity 110 serves to process one or more sound signals that each represent an aspect of the captured sound into the (single-channel) input audio signal 115 for provision to the audio encoding entity 120 and/or for storage in a storage means for subsequent use.

[0035] The audio encoding entity 120 employs an audio coding algorithm, referred herein to as an audio encoder, to process the input audio signal 115 into the encoded audio signal 125. In this regard, the audio encoder may be considered to implement a transform from a signal domain (the input audio signal 115) to the compressed domain (the encoded audio signal 125). The audio encoding entity 120 may further include a pre-processing entity for processing the input audio signal 115 from a format in which it is received from the audio capturing entity 110 into a format suited for the audio encoder. This pre-processing may involve, for example, level control of the input audio signal 115 and/or modification of frequency characteristics of the input audio signal 115 (e.g. low-pass, high-pass or bandpass filtering). The pre-processing may be provided as a pre-processing entity that is separate from the audio encoder, as a sub-entity of the audio encoder or as a processing entity whose functionality is shared between a separate pre-processing and the audio encoder.

[0036] The audio decoding entity 130 employs an audio decoding algorithm, referred herein to as an audio decoder, to process the encoded audio signal 125 into the reconstructed audio signal 135. The audio encoder may be considered to implement a transform from an encoded domain (the encoded audio signal 125) back to the signal domain (the reconstructed audio signal 135). The audio decoding entity 130 may further include a post-processing entity for processing the reconstructed audio signal 115 from a format in which it is received from the audio decoder into a format suited for the audio reproduction entity 140. This post-processing may involve, for example, level control of the reconstructed

audio signal 135 and/or modification of frequency characteristics of the reconstructed audio signal 135 (e.g. low-pass, high-pass or bandpass filtering). The post-processing may be provided as a post-processing entity that is separate from the audio decoder, as a sub-entity of the audio decoder or as a processing entity whose functionality is shared between a separate post-processing and the audio decoder.

[0037] The audio reproduction entity 140 may comprise, for example, headphones, a headset, a loudspeaker or an arrangement of one or more loudspeakers.

[0038] Instead of using the audio capturing entity 110, the audio processing system 100 may include a storage means for storing pre-captured or pre-created audio signals, among which the audio input signal for provision to the audio encoding entity 120 can be selected.

[0039] Instead of using the audio reproduction entity 140, the audio processing system 100 may comprise a storage means for storing the reconstructed audio signal 135 for subsequent analysis, processing, playback and/or transmission to a further entity.

[0040] The dotted vertical line in Figure 1 serves to denote that, typically, the audio encoding entity 120 and the audio decoding entity 130 are provided in separate devices that may be connected to each other via a network or via a transmission channel. The network/channel may enable a wireless connection, a wired connection or a combination of the two between the audio encoding entity 120 and the audio decoding entity 130. As an example in this regard, the audio encoding entity 120 may further comprise a (first) network interface for encapsulating the encoded audio signal 125 into a sequence of protocol data units (PDUs) for transfer to the decoding entity 130 over a network/channel, whereas the audio decoding entity 130 may further comprise a (second) network interface for decapsulating the encoded audio signal 125 from the sequence of PDUs received from the audio encoding entity 120 over the network/channel.

[0041] Figure 2 illustrates a block diagram of some components and/or entities of an audio encoder 121 that may be provided as part of the audio encoding entity 120 according to an example. The audio encoder 121 combines encoding in a signal domain and in an excitation domain to enable high sound quality in combination with a low delay, as will be described in more detail in examples in the following. The audio encoding entity 120 may include further components or entities in addition to the audio encoder 121, e.g. the pre-processing entity referred to in the foregoing, which pre-processing entity may be arranged to process the input audio signal 115 before passing it for the audio encoder 121.

[0042] The audio encoder 121 carries out encoding of the input audio signal 115 into the encoded audio signal 125, i.e. the audio encoder 121 implements a transform from the signal domain to the encoded domain. The audio encoder 121 may be arranged to process the input audio signal 115 arranged into a sequence of input frames, each input frame including digital audio signal at a predefined sampling frequency and comprising a time series of input samples. Typically, the audio encoder 121 employs a fixed predefined frame length. In other examples, the frame length may be a selectable frame length that may be selected from a plurality of predefined frame lengths, or the frame length may be an adjustable frame length that may be selected from a predefined range of frame lengths. A frame length may be defined as number samples L included in the frame, which at the predefined sampling frequency maps to a corresponding duration in time.

[0043] As an example in this regard, the audio encoder 121 may employ a fixed frame length of 1 ms and sampling frequency of 48 kHz, resulting in frames of L=48 samples. These values, however, serve as non-limiting examples and different frame length and/or sampling frequency may be employed instead, depending e.g. on the desired audio bandwidth, on desired framing delay and/or on available processing capacity.

[0044] The audio encoder 121 includes two signal paths: a first signal path that involves a linear predictive coding (LPC) encoder 122 followed by a residual encoder 124 and a second signal path that involves a signal-domain encoder or can be referred to as a time sample domain encoder 126. LPC encoding is a coding technique well known in the art and it makes use of short-term redundancies in the input audio signal 125. In the first signal path, the LPC encoder 122 carries out an LPC encoding procedure to process the input audio signal 115 into a residual signal 123, which is provided as input to the residual encoder 124. The residual encoder 124 carries out residual encoding procedure to process the residual signal 123 into a first encoded signal 125-1 for provision to the selection entity 128. In the second signal path, the signal-domain encoder 126 carries out input signal encoding procedure to process the input audio signal 115 into a second encoded signal 125-2 for provision to the selection entity 128. The selection entity further receives the input audio signal 115 and carries out selection of one of the first and second encoded signals 125-1, 125-2 as the encoded audio signal 125.

[0045] In each of the first and second signal paths, the input audio signal 115 is processed into the respective encoded signal 125-1, 125-2 frame by frame. In other words, in the first signal path the LPC encoder 122 carries out the LPC encoding for a frame of input audio signal 115 and produces a corresponding frame of the residual signal 123, which in turn is processed by the residual encoder 124 into a corresponding frame of the first encoded signal 125-1. In the second signal path, the signal-domain encoder 126 processes the frame of input audio signal 115 into a corresponding frame of the second encoded signal 125-2. The first signal path constitutes a first audio encoding mode and the second signal path constitutes a second audio encoding mode.

[0046] The first and second signal paths (i.e. the first and second audio encoding modes, respectively) outlined above and described in more detail in the following serve as non-limiting examples and hence one or both of the first and

second signal paths may include additional processing components or entities. As an example in this regard, the first signal path may further comprise a long-term prediction (LTP) encoder that encodes the residual signal 123 provided by the LPC encoder 122 into a second residual signal for provision instead of the residual signal 123 to the residual encoder 124 for residual encoding therein. LTP encoding is a coding technique well known in the art and makes use of long(er) term redundancies (e.g. in a range above approximately 2 ms) in the input audio signal 125: while the LPC encoder 122 is typically successful in modeling any short-term redundancies, possible long-term redundancies are still there in the residual signal 123 and hence the LPC encoder may provide an improvement for encoding of audio input signals 125 that include a periodic or a quasi-periodic signal component whose periodicity falls into the range of long(er) term redundancies (e.g. a voice of a human subject).

[0047] In the audio encoding mode, the LPC encoder 122 carries out an LPC analysis based on past values of the reconstructed audio signal 135 using a backward prediction technique known in the art. A 'local' copy of the reconstructed audio signal 135 may be stored in a past audio buffer, which may be provided e.g. in a memory in the audio encoder 121 or in the LPC encoder 122, thereby making the reconstructed audio signal 135 available for the LPC analysis in the LPC encoder 122. Hence, the references to the reconstructed audio signal 135 in context of the audio encoder 121 refer to the local copy available therein. This aspect will be described in more detail later below.

[0048] In the LPC analysis, the LPC encoder 122 may find the LPC filter coefficients e.g. by minimizing the error term

$$e(t) = \left\| \sum_{i=0}^{K_{LPC}} a_i x(t-i) \right\|, \quad t = t+1 : t+N_{lpc}$$

where a_i, $i = 0: {}_{KLPC}$, $\alpha_0 = 1$ denote the LPC filter coefficients, N_{lpc} denotes the analysis window length (in number of samples), $x(t)$, $t = t - N_{LPC}$: t denotes a signal reconstructed on basis of one or more past frames of the encoded audio signal, i.e. the most recent samples of the reconstructed audio signal 135, and the symbol $\|\cdot\|$ denotes an applied norm, e.g. the Euclidean norm.

[0049] The backward prediction computes LPC filter coefficients on basis of past samples of the reconstructed audio signal and carries out LPC analysis filtering for a frame of the input audio signal 115 using the computed LPC filter coefficients to produce a corresponding frame of the residual signal 123. In other words, the LPC analysis filtering involves processing a time series of input samples into a corresponding time series of residual samples. The LPC analysis filtering to compute the residual signal 123 on basis of the input audio signal 115 may be carried out e.g. by using the following

$$r(t) = \sum_{i=0}^{K_{LPC}} a_i x(t-i), \qquad t = t+1 : t+L$$

where $a_i i = 0: K_{LPC}$, $a_0 = 1$ denote the LPC filter coefficients, L denotes the frame length (in number of samples), $x(t)$, $t = t + 1$: t + L denotes a frame of the input audio signal 115 (i.e. the time series of input samples), and $r(t)$, $t = t + 1$: $t + L$ denotes a corresponding frame of the residual signal 123 (i.e. the time series of residual samples).

[0050] The LPC encoder 122 passes the residual signal 123 to the residual encoder 124 for computation of the first encoded signal 125-1 therein. The LPC encoder 122 may further pass the LPC filter coefficients computed therein to the residual encoder 124 for subsequent forwarding to the selection entity 128 or the LPC encoder 122 may pass the computed LPC filter coefficients directly to the selection entity 128.

[0051] The backward prediction in the LPC encoder 122 employs a predefined window length, denoted as N_{lpc}, implying that the backward prediction bases the LPC analysis on N_{lpc} most recent samples of the reconstructed audio signal 135. In an example, the analysis window covers 608 most recent samples of the reconstructed audio signal 135, which at the sampling frequency of 48 kHz corresponds to approx. 12.7 ms. This, however, is a non-limiting example and a shorter or longer window may be employed instead, e.g. a window having a duration of 16 ms or a duration selected from the range 12 to 30 ms. A suitable length of the analysis window, in ms, depends also on the existence and/or characteristics of other encoding components employed in the first audio encoding mode. As an example, the first audio encoding mode may, additionally, involve LTP referred to in the foregoing, and the range of delays considered by the LTP encoder may have an effect on the most appropriate choice for the temporal length of the analysis window for the backward predictive LPC analysis. The analysis window has a predefined shape, which may be selected in view of desired LPC analysis characteristics. Several analysis windows for the LPC analysis applicable for the LPC encoder 122 are known in the art,

e.g. a (modified) Hamming window and a (modified) Hanning window, as well as hybrid windows such as one specified in the ITU-T Recommendation G.728 (section 3.3).

[0052] The LPC encoder 122 employs a predefined LPC model order, denoted as K_{lpc}, resulting in a set of K_{lpc} LPC filter coefficients. Since the LPC analysis in the LPC encoder 122 relies on past values of the reconstructed audio signal 135, there is no need to transmit parameters that are descriptive of the computed LPC filter coefficients to the decoding entity 130, but the decoding entity 130 is able to compute an identical set of LPC filter coefficients for LPC synthesis filtering therein on basis of the reconstructed audio signal 135 available in the audio decoding entity 130. Consequently, a relatively high LPC model order K_{lpc} may be employed since it does not have an effect on the resulting bit-rate of the encoded audio signal 125, thereby enabling accurate modeling of spectral envelope of the input audio signal 115 especially for input audio signals 115 that include a periodic or a quasi-periodic signal component. On the other hand, required computing capacity increases with increasing LPC model order K_{lpc}, and hence selection of the most appropriate LPC model order K_{lpc} for a given use case may involve a trade-off between the desired accuracy of modeling the spectral envelope of the input audio signal 115 and the available computational resources. As a non-limiting example, the LPC model order K_{lpc} may be selected as a value between 30 and 60.

[0053] The residual encoder 124 carries out a residual encoding procedure that involves computing the first encoded signal 125-1 on basis of the residual signal 123 received from the LPC encoder 122. The residual encoding may employ, for example, a gain-shape coding technique (e.g. a gain-shape encoder) known in the art, where the relative amplitudes of samples in a frame of the residual signal 123 are encoded separately from the gain of the frame of the residual signal 123. Therein, the encoded residual parameters for a frame of the residual signal 123 hence include a vector $\mathbf{v}_r$ (or two or more sub-vectors $\mathbf{v}_{r,i}$) of amplitude values and a gain value g_r, where a reconstructed frame of the residual signal 123 can be formed by multiplying each amplitude value of the vector $\mathbf{v}_r$ (or the two or more sub-vectors $\mathbf{v}_{r,i}$) by the gain value g_r. In an example, the gain-shape coding technique makes use of pyramidally truncated lattice quantization in generating quantized values of the vector $\mathbf{v}_r$ (or the sub-vectors $\mathbf{v}_{r,i}$), whereas quantized value of the gain g_r may be generated separately e.g. by using a suitable scalar quantizer. In the example case of the frame length of L=48 samples (i.e. 1 ms at 48 kHz sampling frequency) may include pyramidally truncated Z_{48} lattice, e.g. one described in the article by Thomas R. Fisher titled "A pyramid Vector Quantizer", IEEE Transactions on Information Theory, Vol. 32, Issue 4, pp. 568-583, July 1986, ISSN 0018-9448.

[0054] In other examples, a coding technique different from the gain-shape coding and/or quantization technique different from the lattice quantization may be employed instead. However, the lattice quantization has an advantage that it enables computationally feasible approach for encoding relatively long vectors (e.g. 48 samples or even higher) at a good quantization accuracy without the need to store large codebooks for the residual encoder 124.

[0055] The residual encoder 124 passes the encoded parameters that are descriptive of the residual signal 123 as the first encoded signal 125-1 to the selection entity 128. In a scenario where the residual encoder 124 has received the LPC filter coefficients from the LPC encoder 122, it may further pass the LPC filter coefficients to the selection entity 128 together with the first encoded signal 125-1.

[0056] In an example, the zero-input response of the LPC analysis filter derived in the LPC encoder 122 can be removed from the residual signal 123 before encoding the residual signal 123 in the residual encoder 124. The zero-input response removal may be provided, for example, as part of the LPC encoder 122 (before passing the residual signal 123 obtained by the LPC analysis filtering to the residual encoder 124) or in the residual encoder 124 (before carrying out the encoding procedure therein).

[0057] The zero input response may be calculated as

$$x(t) = -\sum_{i=1}^{K_{LPC}} a_i x(t-i), \qquad t = t+1 : t+L$$

where a_i, $i = 1: K_{LPC}$ denote the LPC filter coefficients, L denotes the frame length (in number of samples), and $x(t)$, $t = t - K_{LPC} + 1: t$ denotes a signal reconstructed on basis of one or more past frames of the encoded audio signal, i.e. the most recent samples of the reconstructed audio signal 135. The computation of the zero input response is a recursive process: for the first sample of the zero input response all $x(t)$ refer to past samples of the reconstructed audio signal 135, whereas the following samples of the zero input response are computed at least in part using signal samples computed for the zero input response.

[0058] After encoding a frame of the residual signal 123 in the audio encoder 121, the calculated zero input response is added back to the reconstructed audio signal 135. Consequently, also in the audio decoder, after reconstructing the residual signal therein and filtering it through the LPC synthesis filter, the zero input response is added to the reconstructed audio signal 135, as described in the following.

[0059] In the second audio encoding mode, the signal-domain encoder 126, also referred to as the time sample encoder 126 (as described in the foregoing), carries out an encoding procedure that involves computing the second encoded signal 125-2 directly on basis of the input audio signal 115. In this regard, the signal-domain encoder 126 may directly encode and/or quantize the time series of input samples, i.e. the input samples that constitute a frame of the input audio signal 115, into encoded signal-domain parameters that are descriptive of the frame of the input audio signal 115. The signal-domain encoder 126 further passes the encoded signal-domain parameters as the second encoded signal 125-2 to the selection entity 128.

[0060] In an example, the signal-domain encoder 126 employs the same or similar coding technique as applied in the residual encoder 124. Such an approach enables efficient re-use of components within the audio encoder 121 while enabling high quality of the reconstructed audio. Hence, the signal-domain encoder 126 may employ a gain-shape coding technique (e.g. a gain-shape encoder) known in the art (as outlined in the foregoing), wherein the vector of amplitude values is denoted as $\mathbf{v}_s$ (or two or more sub-vectors denoted as $\mathbf{v}_{s,i}$) and the gain value is denoted as g_s, and use the pyramidally truncated lattice quantization (e.g. the Z_{48} lattice) in generating quantized values of the vector $\mathbf{v}_s$ (or the sub-vectors $\mathbf{v}_{s,i}$) together with a suitable separate scalar quantizer for generating the quantized value of the gain g_s.

[0061] In other examples, the signal-domain encoder 126 employs a coding technique and/or quantization technique different from those employed in the residual encoder 124. While this approach would fall short of providing the benefit that arises from sharing the respective component(s) with the residual encoder 124, on the other hand it may enable tailoring the respective coding techniques and/or quantization techniques employed in the residual encoder 124 and the signal-domain encoder 126 in accordance with characteristics of the respective input signals these coding entities are arranged to process.

[0062] The selection entity 128 receives, for each frame, the first and second encoded signals 125-1, 125-2 together with the input audio signal 115 and the LPC filter coefficients computed in the LPC encoder 122. Based at least in part on this information, the selection entity 128 selects one of the first and second encoded signals 125-1, 125-2 for provision in the encoded audio signal 125.

[0063] In an example, the selection entity 128 computes a first distortion value D_1 on basis of the first encoded signal 125-1 and the input audio signal 115, which first distortion value D_1 is descriptive of the difference between the input audio signal 115 and a first reconstructed audio signal that is derivable on basis of the first encoded signal 125-1. To enable computation of the first distortion value D_1, the selection entity 128 derives the first reconstructed audio signal by carrying out LPC synthesis filtering of a reconstructed residual signal by using the LPC filter coefficients derived for the current frame in the LPC encoder 122. The reconstructed residual signal, in turn, may be received as side information from the residual encoder 124 or the selection entity 128 may apply the encoded parameters carried in the first encoded signal 125-1 to derive the reconstructed residual signal therein. The selection entity 128 may compute first distortion value D_1 e.g. as a mean squared deviation (MSD) between the first reconstructed audio signal and the input audio signal 115 or as a mean absolute deviation (MAD) between the first reconstructed audio signal and the input audio signal 115.

[0064] Moreover, in this example, the selection entity 128 further computes a second distortion value D_2 on basis of the second encoded signal 125-2 and the input audio signal 115, which second distortion value D_2 is descriptive of the difference between the input audio signal 115 and a second reconstructed audio signal that is derivable on basis of the second encoded signal 125-2. The second reconstructed audio signal may be received as side information from the signal-domain encoder 126 or the selection entity 128 may apply the encoded parameters carried in the second encoded signal 125-2 to derive the second reconstructed audio signal therein. As in case of the first distortion value D_1, the selection entity 128 may derive the second distortion value D_2, for example, as the MSD or the MAE between the second reconstructed audio signal and the input audio signal 115.

[0065] Consequently, the selection entity 128 may select one of the first and second encoded signals 125-1, 125-2 for the encoded audio signal 125 on basis of comparison of the first and second distortion values D_1 and D_2. In an example, the selection entity may select the first encoded signal 125-1 for the current frame in response to the first distortion value D_1 being smaller than the second distortion value D_1 (e.g. in case $D_1 < D_2$ holds true) and, conversely, select the second encoded signal 125-1 for the current frame in response to the first distortion value D_1 being larger than or equal to the second distortion value D_2 (e.g. in case $D_1 \geq D_2$ holds true).

[0066] In another example, the selection entity 128 may select the second encoded signal 125-2 for the current frame in case the first distortion value D_1 exceeds the second distortion value D_2 by at least a predefined margin. Application of the margin serves to avoid unnecessarily switching between selecting first and second encoded signal 125-1, 125-2 for the encoded audio signal 125 from frame to frame by favoring the first audio encoding mode that involves also the LPC encoding. This enhances sound quality in the reconstructed audio signal 135 by avoiding the switching that is likely to result in distortion especially at high frequencies. The margin may be defined as a relative value or as an absolute value:

- As an example of a relative margin, the selection entity 128 may select the second encoded signal 125-2 in response to the ratio of the first distortion value D_1 and the second distortion value D_2 exceeding a predefined threshold T_r, where the threshold T_r has a value that is larger than unity, e.g. in case the condition $D_1 / D_2 > T_r$ with $T_r > 1$ holds

true (and conversely, select the first encoded signal 125-1 in response to the ratio of the first distortion value D_1 and the second distortion value D_2 failing to exceed the threshold T_r, e.g. in case the above-mentioned condition in this regard does not hold). Herein, the value of threshold T_r may be set to a value selected, for example, from the range 1.25 to 3, e.g. 2.

- As an example of an absolute margin, the selection entity 128 may select the second encoded signal 125-2 in response to the first distortion value D_1 exceeding the second distortion value D_2 at least by a predefined margin M_a, where the margin M_a has a positive value, e.g. in case the condition $D_1 > D_2 + M_a$ with $M_a > 0$ holds true (and conversely, select the first encoded signal 125-1 in response to the first distortion value D_1 failing to exceed the second distortion value D_2 by at least the margin M_a, e.g. in case the above-mentioned condition in this regard does not hold).

[0067] The selection entity 128 appends the selected one of the first and second encoded signals 125-1, 125-2 with an indication of the selected one of the first and second encoded signals 125-1, 125-2 to provide the encoded audio signal 125 for the current frame. Such indication may be referred to as a coding mode indication that serves to identify which one of the first and second audio encoding modes has been selected by the selection entity 128 to represent the current frame. The coding mode indication enables the decoding entity 130 to correctly reconstruct the audio signal therein.

[0068] The audio encoder 121 stores at least a predefined number of most recent samples of the reconstructed audio signal 135 to enable the backward prediction in the LPC encoder 122. As described in the foregoing, this may be implemented by generating a local copy of the reconstructed audio signal 135 in the audio encoder 121 (e.g. in the selection entity 128) and storing the local copy of the reconstructed audio signal 135 in the past audio buffer in the LPC encoder 122 or otherwise within the audio encoder 121. In this regard, the past audio buffer stores at least the N_{lpc} most recent samples of the reconstructed audio signal 135 to cover the analysis window applied by the LPC encoder 122.

[0069] After having selected one of the first and second encoded signals 125-1, 125-2 for the current frame, the selection entity 128 updates the past audio buffer by discarding the L oldest samples in the past audio buffer and, depending on the selection of the first or the second encoded signal 125-1 to represent the current frame, inserting corresponding one of the first and second reconstructed audio signals in the past audio buffer to facilitate LPC analysis in the next frame.

[0070] Figure 3 illustrates a block diagram of some components and/or entities of an audio decoder 131 that may be provided as part of the audio decoding entity 130 according to an example. The audio decoder 131 carries out decoding of the encoded audio signal 125 into the reconstructed audio signal 135, thereby serving to implement a transform from the encoded domain (back) to the signal domain and, in a way, reversing the encoding operation carried out in the audio encoder 121. The audio decoder 131 process the encoded audio signal 125 frame by frame.

[0071] The audio decoder 131 can also have two signal paths: a first signal path that involves a residual decoder 134 followed by a LPC decoder 132 and a second signal path that involves a signal-domain decoder 136. A frame of the encoded audio signal 125 received at the audio decoder 131 is processed through one of the first and second signal paths in accordance with the coding mode indication received in the encoded audio signal 125. The first and second signal paths in the audio decoder 132 constitute first and second audio decoding modes, respectively. In this regard, a selection entity 138 receives the frame of encoded audio signal 125, reads the coding mode indication for the current frame, extracts the encoded signal from the frame of encoded audio signal 125, and passes the extracted encoded signal to one of the first and second signal paths in the audio decoder 131 accordingly. In other words, if the coding mode indication indicates that the encoded signal from first signal path was selected for the current frame in the audio encoder 121, the encoded signal in the encoded audio signal 125 comprises the first encoded signal 125-1 and the selection entity 138 passes this signal to the first signal path in the audio decoder 131 for decoding according to the first audio decoding mode. On the other hand, in case the coding mode indication indicates that the encoded signal from second signal path was selected for the current frame in the audio encoder 121, the encoded signal in the encoded audio signal 125 comprises the second encoded signal 125-2 and the selection entity 138 passes this signal to the second signal path in the audio decoder 131 for decoding according to the second audio decoding mode.

[0072] If the first audio decoding mode is invoked, the residual decoder 134 processes the first encoded signal 125-1 into a reconstructed residual signal 133, which is provided as input to the LPC decoder 132, which in turn carries out LPC synthesis on basis of the reconstructed residual signal 133 to output a reconstructed audio signal 135-1, which will serve as the reconstructed audio signal 135. If the second audio decoding mode is invoked, the signal-domain decoder 136 processes the second encoded signal 125-2 into a reconstructed audio signal 135-2, which will serve as the reconstructed audio signal 135.

[0073] In the first signal path of the audio decoder 131, the residual decoder 134 carries out a residual decoding procedure that involves computing the reconstructed residual signal 133 on basis of the first encoded signal 125-1 received from the selection entity 138. A frame of reconstructed residual signal 133 is provided as respective time series of reconstructed residual samples. The reconstructed residual signal 133 is passed to the LPC decoder 132 for LPC

synthesis therein. In order to enable meaningful reconstruction of the residual signal, the residual decoder 134 must employ the same or otherwise matching residual coding technique as employed in the residual encoder 124. In an example, the residual decoding procedure involves dequantizing the encoded residual parameters received as part of the encoded audio signal 125 and using the dequantized residual parameters to create a frame of the reconstructed residual signal 133, i.e. the time series of reconstructed residual samples. As an example, the gain-shape coding technique (e.g. a gain-shape decoder) may be employed, where the dequantization may comprise using the received encoded residual parameter to find the vector $\mathbf{v}_r$ (or the two or more sub-vectors $\mathbf{v}_{r,i}$) of amplitude values and the gain value g_r and creation of the frame of the reconstructed residual signal 133 may comprise multiplying each amplitude value of the vector $\mathbf{v}_r$ (or the two or more sub-vectors $\mathbf{v}_{r,i}$) by the gain value g_r.

[0074] Further in the first signal path of the audio decoder 131, the LPC decoder 132 carries out the LPC analysis based on past values of the reconstructed audio signal 135 using the same backward prediction technique as applied in the LPC encoder 122. Hence, the backward prediction computes LPC filter coefficients on basis of past samples of the reconstructed audio signal 135. The LPC decoder further carries out LPC synthesis filtering of the reconstructed residual signal 133 by using the LPC filter coefficients derived for the current frame in the LPC decoder 132, thereby generating the reconstructed audio signal 135-1.

[0075] The LPC synthesis filtering in the LPC decoder 132 involves processing a time series of reconstructed residual samples into a corresponding time series of output samples that hence constitute a corresponding frame of the reconstructed audio signal 135. The LPC decoder 132 may find the LPC filter coefficients for the LPC synthesis therein, for example, using the procedure outlined in the foregoing for the LPC encoder 122. The LPC synthesis may be carried out e.g. by using the following equation:

$$x(t) = r(t) - \sum_{i=1}^{K_{LPC}} a_i x(t-i), \qquad t = t+1:t+L$$

where a_i, $i = 1: K_{LPC}$ denote the LPC filter coefficients, L denotes the frame length (in number of samples), $x(t)$, $t = t + 1: t + L$ denotes a frame of the reconstructed audio signal 135-1 (i.e. the time series of output samples), and $r(t)$, $t = t + 1: t + L$ denotes a corresponding frame of the reconstructed residual signal 133 (i.e. the time series of reconstructed residual samples).

[0076] Since the LPC analyses in the LPC encoder and the LPC decoder 132 are carried out using the same approach and they are further performed on the same or similar audio signals, the resulting LPC filter coefficients are also the same or similar. The past values of the reconstructed audio signal 135 required for the LPC analysis in the LPC decoder 131 are stored in a past audio buffer, which may be provided e.g. in a memory in the audio decoder 131 or in the LPC decoder 132.

[0077] After having derived the reconstructed audio signal 135-1, the LPC decoder 132 further adds the zero input response of the LPC synthesis filter to the reconstructed audio signal 135-1 before using the reconstructed audio signal 135-1 from the LPC decoder 132 as the reconstructed audio signal 135 provided as output from the audio decoder 131 and before using this signal to update the past audio buffer of the audio decoder 131 (as will be described later in this text). The zero input response may be calculated on basis of the reconstructed audio signal 135-1, for example, as described in the foregoing for computation of the zero input response in the audio encoder 121.

[0078] In the second signal path of the audio decoder 131, the signal-domain decoder 136, which may be alternatively referred to as a time sample decoder or as a time sample domain decoder, carries out a decoding procedure that involves computing the reconstructed audio signal 135-1 directly on basis of the encoded signal-domain parameters received as part of the second encoded signal 125-2 received from the selection entity 138. Consequently, a frame of reconstructed audio signal 133 is provided as respective time series of output samples. In order to enable meaningful reconstruction of the audio signal, the signal-domain decoder 136 must employ the same or otherwise matching coding technique as employed in the signal-domain encoder 126. In an example, the decoding procedure involves dequantizing the encoded signal-domain parameters and using the dequantized signal-domain parameters to create a frame of the reconstructed audio signal 135-1. As an example, the gain-shape coding technique (e.g. a gain-shape decoder) may be employed, where the dequantization may comprise using the received encoded signal-domain parameter to find the vector $\mathbf{v}_s$ (or the two or more sub-vectors $\mathbf{v}_{s,i}$) of amplitude values and the gain value g_s and creation of the frame of the reconstructed audio signal 135-2 may comprise multiplying each amplitude value of the vector $\mathbf{v}_s$ (or the two or more sub-vectors $\mathbf{v}_{s,i}$) by the gain value g_s.

[0079] Along the lines described in the foregoing for the audio encoder 121, also the audio decoder 131 stores at least N_{lpc} most recent samples of the reconstructed audio signal 135 to enable the backward prediction in the LPC decoder 132. This may be implemented by storing sufficient number of most recent samples in the past audio buffer of the audio

decoder 131. After having carried out decoding using one of the first and second decoding modes, the audio decoder 131 updates the past audio buffer therein by discarding the *L* oldest samples in the past audio buffer and inserting the samples of the reconstructed audio signal 135 in the past audio buffer to facilitate the LPC analysis in the next frame.

[0080] In order to ensure keeping the memory of the LPC synthesis filter in the LPC decoder 132 up to date, the audio decoder carries out the LPC analysis to derive the LPC filter coefficients therein also for those frames of audio signal that are encoded by the audio encoder 121 by using the second encoding mode. The LPC synthesis for such frames may be carried out by the LPC decoder 132. Further in this regard, the audio encoder 131 further carries out the LPC analysis filtering (e.g. by the LPC decoder 132) of the current frame of the reconstructed audio signal 135 to derive the respective residual signal also in the audio decoder 131. The residual signal derived in the audio decoder 131 is employed as part of the memory of the LPC synthesis filter in decoding of the following frame of the encoded audio signal 125.

[0081] Instead of carrying out the LPC synthesis in the audio decoder 131 (e.g. by the LPC decoder 132) in order to update LPC synthesis filter memory therein, the memory update may be provided by using the following equation:

$$\begin{pmatrix} y(t+1) \\ \vdots \\ y(t+L) \end{pmatrix} = \begin{pmatrix} h_1 & 0 & 0 \\ h_2 & h_1 & 0 \\ \vdots & \vdots & \vdots \\ h_n & \dots & h_1 \end{pmatrix} \begin{pmatrix} r(t+1) \\ \vdots \\ r(t+L) \end{pmatrix}$$

where $n = K_{LPC}$, $y(t)$, $t = t + 1: t + L$ is the zero input response removed reconstructed audio signal 135 (i.e. the reconstructed audio signal 135 without the zero input response), $r(t)$ denotes the residual signal obtained (by the LPC analysis filtering) in the audio decoder 131 and ($h_1 h_2 \dots h_n$) denotes the LPC synthesis filter impulse response. Also the reciprocal equation can be used for the analysis part (i.e. $r=H^{-1}y$). The components of the inverse matrix H^{-1} , h_i' can be obtained as follows:

$$h'(1) = \frac{1}{h(1)}$$

$$h'(i) = -h'(1) \sum_{k=2}^{i} h(k) h'(i-k+2), i = 2:n$$

[0082] In an example, the residual encoder 124 and the signal-domain encoder 126 of the audio encoder 121 employ the same or substantially the same bit-rate of the encoded audio signal to ensure constant or substantially constant bit-rate regardless of the currently employed audio encoding mode. Such an approach results in a constant or substantially constant transmission bandwidth requirement throughout the audio coding session. The bit-rate of the encoded audio signal may be selected, for example, from the range from 80 to 150 kilobits per second (kbps), e.g. as approximately 100 kbps, 119 kbps or 133 kbps, depending on the desired tradeoff between the required transmission bandwidth and sound quality in the reconstructed audio signal 135. If assuming any of the exemplifying bit-rates 100, 119 or 133 kbps, assuming the frame length of 1 ms (e.g. 48-sample frames at 48 kHz sampling frequency), the encoded audio signal 135 is provided as frames of 100, 119 or 133 bits, respectively.

[0083] Tables 1, 2 and 3 in the following provide examples of performance gain enables by an audio coding arrangement that makes use of the audio encoder 121 and the audio decoder 131 according to respective examples.

[0084] Each of Tables 1, 2 and 3 provides respective signal to noise ratio (SNR) values computed for 12 test signals that comprise audio of different characteristics (identified in the first column of a table). For each test signal, the second column of the table provides the SNR obtained by using a reference audio coding arrangement that enables only the first audio encoding mode operated at a certain bit-rate while the third column of the table provides the SNR obtained by using an audio coding arrangement that makes use of the audio encoder 121 and the audio decoder 131 arranged to operate at the same bit-rate as the reference audio coding arrangement. The fourth column of the table indicates the relative increase in the SNR obtained by using the audio coding arrangement that makes use of the audio encoder 121 and the audio decoder 131 instead of the reference audio coding arrangement at the same bit-rate, and the fifth column of the table indicates the percentage of frames for which the second encoding mode has been selected by the audio encoder 121. Tables 1, 2 and 3 provide this information for the two audio coding arrangements operated at 133 kbps, 119 kbps and 100 kbps, respectively.

Table 1

Test signal	Reference SNR [dB]	Obtained SNR [dB]	Improvement in SNR [%]	Usage of the 2nd audio encoding mode [%]
Vocal	16.8940	21.2785	25.9530	29.2%
German male speech	17.0272	22.9015	34.4995	24.8%
English female speech	15.5659	23.4642	50.7410	24.5%
Trumpet solo and orch.	22.9232	24.8984	8.6166	16.2%
Classical orch. music	18.8988	20.0848	6.2755	24.6%
Contemp. pop music	15.8702	17.7997	12.1580	16.2%
Harpsichord	15.3343	19.9265	29.9472	24.6%
Castanets	6.8766	17.1686	149.6670	27.4%
Pitch pipe	19.5439	23.2357	18.8898	33.1%
Bagpipes	18.6216	21.9669	17.9646	25.4%
Glockenspiel	16.1310	27.6679	71.5201	31.4%
Plucked strings	15.9745	20.2925	27.0306	19.8%

Table 2

Test signal	Reference SNR [dB]	Obtained SNR [dB]	Improvement in SNR [%]	Usage of the 2nd audio encoding mode [%]
Vocal (S. Vega)	14.7567	18.8965	28.0537	27.8%
German male speech	15.0560	20.6880	37.4070	22.9%
English female speech	11.1678	20.7141	85.4806	23.8%
Trumpet solo and orch.	20.9197	22.5552	7.8180	15.2%
Classical orch. music	16.4628	17.8675	8.5326	13.6%
Contemp. pop music	13.6088	15.5205	14.0475	13.4%
Harpsichord	13.6955	17.4038	27.0768	23.7%
Castanets	6.5807	14.8308	125.3681	24.8%
Pitch pipe	17.1496	20.8216	21.4116	30.3%
Bagpipes	16.4810	19.4764	18.1749	22.8%
Glockenspiel	15.4877	24.9040	60.7986	29.0%
Plucked strings	13.9776	17.9217	28.2173	17.0%

Table 3

Test signal	Reference SNR [dB]	Obtained SNR [dB]	Improvement in SNR [%]	Usage of the 2nd audio encoding mode [%]
Vocal (S. Vega)	12.3742	16.2469	31.2966	25.2%
German male speech	13.0146	17.4418	34.0172	22.0%
English female speech	10.9116	18.6103	70.5552	21.2%
Trumpet solo and orch.	18.0884	19.3952	7.2245	13.2%
Classical orch. music	13.8288	15.2516	10.2887	11.7%
Contemp. pop music	11.1108	13.0383	17.3480	11.6%
Harpsichord	11.2863	14.8947	31.9715	22.7%
Castanets	4.8771	11.9507	145.0370	23.4%
Pitch pipe	14.4011	18.0850	25.5807	27.7%
Bagpipes	13.8669	16.9362	22.1340	20.4%
Glockenspiel	13.7021	22.4115	63.5625	27.9%
Plucked strings	12.0257	15.4607	28.5638	15.2%

[0085] Comparison of the performance figures in Tables 1 and 3 suggests that the sound quality enabled by the audio coding arrangement that makes use of the audio encoder 121 and the audio decoder 131 as outlined in the foregoing at 100 kbps can be reached at 133 kbps if using the reference audio coding arrangement that only provides the first audio encoding mode. While an improvement in the SNR values does not typically directly translate into a corresponding perceived sound quality, the SNR values nevertheless suggest that the audio coding arrangement using of the audio encoder 121 and the audio decoder 131 enables a significant improvement, which has also been valeted by informal listening tests.

[0086] In the foregoing, the operation of the audio encoder 121 and the audio decoder 131 is described using an example that involves two audio encoding modes in the audio encoder 121 and respective two audio decoding modes in the audio decoder 131. This, however, is a non-limiting example and in other examples an arrangement where the audio encoder 121 comprises two or more audio encoding modes and the audio decoder 131 comprises respective two or more audio decoding modes may be employed instead. As a non-limiting example in this regard, the audio encoder 121 may include three audio encoding modes, including the first and second audio encoding modes described in the foregoing together with a third audio encoding mode that is otherwise similar to the first audio encoding mode but further includes the LTP encoder envisaged in the foregoing as an exemplifying variation of the first signal path.

[0087] In an example of such an arrangement, the audio encoder 121 carries out the encoding procedure via two or more signal paths that each correspond to a respective audio encoding mode. Moreover, the selection entity 128 receives the encoded signals 125-k from each of the signal paths and makes, derives respective reconstructed audio signals, derives for each reconstructed audio signal a respective distortion value D_k that is descriptive of the difference between the input audio signal 115 and the reconstructed audio signal that is derivable on basis of the respective encoded signal 125-k. Each of the distortion values D_k may be computed, for example, as MSD or MAE as described in the foregoing. Yet further, the selection entity 128 may select the encoding mode that yields the lowest distortion value D_k or the encoding mode that yields the lowest distortion value $D_{k,w} = w_k * D_k$, where w_k denotes a predefined weighting factor assigned for the encoding mode *k*. In the audio decoder 131, the selection entity 138 extracts the coding mode indication and the encoded signal from a frame of the encoded audio signal 135 and carries out audio decoding on basis of the extracted encoded signal using the indicated audio decoding mode.

[0088] Figure 4 depicts an outline of a method 200, which serves as an exemplifying method for encoding a frame of the input audio signal 115 that comprises a time series of input samples into a corresponding frame of the encoded audio signal 125 according to an example. The method 200 commences from encoding the frame of the input audio

signal 115 using at least one of a plurality of audio encoding modes that include at least the first audio encoding mode and the second audio encoding mode.

[0089] The method 200 comprises encoding the frame of the input audio signal 115 using the first audio encoding mode that comprises linear predictive filtering of the time series of input samples using a linear predictive filter coefficients computed using a backward prediction into a residual signal 123 that comprises a respective time series of residual samples and quantizing the time series of residual samples, as indicated in block 210. The method 200 further comprises encoding the frame of input audio signal 115 using the second audio encoding mode that comprises directly quantizing the time series of input samples, as indicated in block 220. The method 200 further comprises selecting one of the input audio signal 115 encoded using the first audio encoding mode and the input audio signal 115 encoded using the second audio encoding mode for provision as the encoded audio signal 125, as indicated in block 230. Although described herein with explicit references to the first and second audio encoding modes, the method 200 generalizes into encoding the input audio signal 115 using a desired number of audio encoding modes (e.g. two or more) and selecting the input audio signal 115 encoded using one of the audio encoding modes for provision as the encoded audio signal 125.

[0090] Figure 5 depicts an outline of a method 300, which serves as an exemplifying method for decoding a frame of the encoded audio signal 125 into a corresponding frame of the reconstructed audio signal 135 that comprises a time series of output samples according to an example. The method 300 commences from receiving an indication of the employed audio encoding mode, as indicated in block 310, and decoding the encoded audio signal 125 using one of a plurality of audio decoding modes in accordance with the received indication of the employed audio encoding mode.

[0091] The method 300 further comprises decoding the frame of encoded audio signal 125 using the first audio decoding mode in response the received indication indicating the first audio encoding mode, wherein the first audio decoding mode comprises dequantizing encoded residual parameters received in the frame of the encoded audio signal 215 into a frame of reconstructed residual signal 133 that comprises a time series of reconstructed residual samples and linear predictive filtering of the time series of reconstructed residual samples into the time series of output samples using a linear predictive filter coefficients computed using a backward prediction, as indicated in block 320.

[0092] The method 300 further comprises decoding the frame of encoded audio signal 125 using the second audio decoding mode in response to the received indication indicating the second audio encoding mode, wherein the second audio decoding mode comprises directly dequantizing encoded signal-domain parameters received in the frame of encoded audio signal 125 into the time series of output samples.

[0093] Although described herein with explicit references to the first and second audio decoding modes, the method 300 generalizes into decoding the frame of encoded audio signal 125 using one of a plurality of audio decoding modes (including two or more audio decoding modes) in accordance with the received indication of the audio encoding mode employed by the audio encoder 121.

[0094] The method 200 may be provided, for example, in the audio encoding entity 120 or in a device that operates as or implements the audio encoding entity 120. Along similar lines, the method 300 may be provided, for example, in the audio decoding entity 130 or in a device that operates as or implements the audio decoding entity 130. The method 200 and/or the method 300 may be varied in a number of ways, e.g. in accordance with the examples provided in context of description of the audio encoder 121 and the audio decoder 131 in the foregoing.

[0095] Figure 6 illustrates a block diagram of some components of an exemplifying apparatus 400. The apparatus 400 may comprise further components, elements or portions that are not depicted in Figure 6. The apparatus 400 may be employed in implementing e.g. the audio encoder 121 or the audio decoder 131.

[0096] The apparatus 400 further comprises a processor 416 and a memory 415 for storing data and computer program code 417. The memory 415 and a portion of the computer program code 417 stored therein may be further arranged to, with the processor 416, to implement the function(s) described in the foregoing in context of the audio encoder 121 or the audio decoder 131.

[0097] The apparatus 400 comprises a communication portion 412 for communication with other devices. The communication portion 412 comprises at least one communication apparatus that enables wired or wireless communication with other apparatuses. A communication apparatus of the communication portion 412 may also be referred to as a respective communication means.

[0098] The apparatus 400 may further comprise user I/O (input/output) components 418 that may be arranged, possibly together with the processor 416 and a portion of the computer program code 417, to provide a user interface for receiving input from a user of the apparatus 400 and/or providing output to the user of the apparatus 400 to control at least some aspects of operation of the audio encoder 121 or the audio decoder 131 implemented by the apparatus 400. The user I/O components 418 may comprise hardware components such as a display, a touchscreen, a touchpad, a mouse, a keyboard, and/or an arrangement of one or more keys or buttons, etc. The user I/O components 418 may be also referred to as peripherals. The processor 416 may be arranged to control operation of the apparatus 400 e.g. in accordance with a portion of the computer program code 417 and possibly further in accordance with the user input received via the user I/O components 418 and/or in accordance with information received via the communication portion 412.

[0099] Although the processor 416 is depicted as a single component, it may be implemented as one or more separate

processing components. Similarly, although the memory 415 is depicted as a single component, it may be implemented as one or more separate components, some or all of which may be integrated/removable and/or may provide permanent / semi-permanent/ dynamic/cached storage.

[0100] The computer program code 417 stored in the memory 415, may comprise computer-executable instructions that control one or more aspects of operation of the apparatus 400 when loaded into the processor 416. As an example, the computer-executable instructions may be provided as one or more sequences of one or more instructions. The processor 416 is able to load and execute the computer program code 417 by reading the one or more sequences of one or more instructions included therein from the memory 415. The one or more sequences of one or more instructions may be configured to, when executed by the processor 416, cause the apparatus 400 to carry out operations, procedures and/or functions described in the foregoing in context of the audio encoder 121 or the audio decoder 131.

[0101] Hence, the apparatus 400 may comprise at least one processor 416 and at least one memory 415 including the computer program code 417 for one or more programs, the at least one memory 415 and the computer program code 417 configured to, with the at least one processor 416, cause the apparatus 400 to perform operations, procedures and/or functions described in the foregoing in context of the audio encoder 121 or the audio decoder 131.

[0102] The computer programs stored in the memory 415 may be provided e.g. as a respective computer program product comprising at least one computer-readable non-transitory medium having the computer program code 417 stored thereon, the computer program code, when executed by the apparatus 400, causes the apparatus 400 at least to perform operations, procedures and/or functions described in the foregoing in context of the audio encoder 121 or the audio decoder 131. The computer-readable non-transitory medium may comprise a memory device or a record medium such as a CD-ROM, a DVD, a Blu-ray disc or another article of manufacture that tangibly embodies the computer program. As another example, the computer program may be provided as a signal configured to reliably transfer the computer program.

[0103] Reference(s) to a processor should not be understood to encompass only programmable processors, but also dedicated circuits such as field-programmable gate arrays (FPGA), application specific circuits (ASIC), signal processors, etc. Features described in the preceding description may be used in combinations other than the combinations explicitly described.

[0104] Although functions have been described with reference to certain features, those functions may be performable by other features whether described or not. Although features have been described with reference to certain embodiments, those features may also be present in other embodiments whether described or not.

Claims

1. An apparatus for encoding a frame of an input audio signal that comprises a time series of input samples into a frame of an encoded audio signal, the apparatus configured to:

   encode said frame of the input audio signal using at least two of a plurality of audio encoding modes, wherein each of said plurality of audio encoding modes is arranged to encode the frame of the input audio signal into a respective encoded signal, wherein said plurality of audio encoding modes include at least;

   a first audio encoding mode configured to linear predictive filter said time series of input samples using linear predictive filter coefficients computed using a backward prediction into a residual signal that comprises a respective time series of residual samples and quantize the time series of residual samples, and
   a second audio encoding mode configured to directly quantize the time series of input samples; and

   select, in accordance with a mode selection rule, one of the respective encoded signals as the frame of the encoded audio signal.

2. An apparatus according to claim 1, wherein the apparatus configured to select one of the respective encoded signals as the frame of the encoded audio signal is further configured to:

   compute a respective distortion value for each of said respective encoded signals; and
   select the respective encoded signal that results in the smallest distortion value as the frame of the encoded audio signal.

3. An apparatus according to claim 2, wherein the apparatus configured to compute a distortion value for a given respective encoded signal is further configured to:

create a reconstructed audio signal on basis of the given respective encoded signal; and
compute the distortion value as a value that is indicative of the difference between said frame of the input audio signal and the reconstructed audio signal.

4. An apparatus according to any of claims 1 to 3, wherein said first audio encoding mode is configured to compute the linear predictive filter coefficients on basis of a reconstructed audio signal derived on basis of one or more frames of encoded audio signal that immediately precede said frame of the input audio signal.

5. An apparatus according to any of claims 1 to 4, wherein said first audio encoding mode is configured to encode said time series of the residual samples by using a first gain-shape encoder to generate a first gain and first relative sample values that represent said frame of the residual signal.

6. An apparatus according to claim 5, wherein said first audio encoding mode is configured to quantize the first gain and the first relative sample values that represent said frame of the residual signal by using a first pyramidally truncated lattice quantizer.

7. An apparatus according to any of claims 1 to 6, wherein said second audio encoding mode is further configured to encode said time series of the input samples by using a second gain-shape encoder to generate a second gain and second relative sample values that represent said frame of the input audio signal.

8. An apparatus according to claim 7, wherein said second audio encoding mode is further configured to quantize the second gain and the second relative sample values that represent said frame of the input audio signal by using a second pyramidally truncated lattice quantizer.

9. An apparatus according to any of claims 5 to 8,
wherein the second gain-shape encoder comprises the first gain-shape encoder; and/or
wherein the second pyramidally truncated lattice quantizer comprises the first pyramidally truncated lattice quantizer.

10. An apparatus according any of claims 1 to 9, wherein the apparatus configured to select one of the respective encoded signals as the frame of the encoded audio signal is further configured to provide an indication of the selected audio encoding mode in said frame of the encoded audio signal.

11. An apparatus for decoding a frame of an encoded audio signal into a frame of a reconstructed audio signal that comprises a time series of output samples, the apparatus configured to:

decode said frame of the encoded audio signal with one of a plurality of audio decoding modes, wherein said plurality of audio decoding modes include at least a first audio decoding mode and a second audio decoding mode;

wherein the first audio decoding mode is configured to dequantize encoded residual parameters received in said frame of the encoded audio signal into a frame of reconstructed residual signal that comprises a time series of reconstructed residual samples and linear predictive filter said time series of reconstructed residual samples into said time series of output samples using linear predictive filter coefficients computed using a backward prediction, and
wherein the second audio decoding mode is configured to directly dequantize encoded signal-domain parameters received in said frame of the encoded audio signal into said time series of output samples.

12. An apparatus according to claim 11, wherein the apparatus is further configured to:

receive an indication of one of the plurality of audio encoding modes; and
decode said frame of the encoded audio signal using one of the plurality of audio decoding modes in accordance with said received indication.

13. An apparatus according to claim 11 or 12, wherein said first audio decoding mode is configured to compute the linear predictive filter coefficients on basis of a plurality of samples of reconstructed audio signal that immediately precede said frame of the reconstructed audio signal.

14. An apparatus according to any of claims 11 or 13,

wherein said encoded residual parameters comprise a first gain and first relative sample values that represent said frame of the reconstructed residual signal; and
wherein said first audio decoding mode comprises decoding said first gain and said first relative sample values using a first gain-shape decoder.

15. An apparatus according to claim 14, wherein said first audio decoding mode is configured to dequantize the first gain and the first relative sample values by using a first pyramidally truncated lattice quantizer.

16. An apparatus according to any of claims 11 to 15,
wherein said encoded signal-domain parameters comprise a second gain and second relative sample values that represent said frame of the reconstructed audio signal; and
wherein said second audio decoding mode is configured to decode said second gain and said second relative sample values using a second gain-shape decoder.

17. An apparatus according to claim 16, wherein said second audio decoding mode is configured to dequantizie the second gain and the second relative sample values by using a second pyramidally truncated lattice quantizer.

18. An apparatus according to any of claims 14 to 17,
wherein the second gain-shape encoder comprises the first gain-shape encoder; and/or
wherein the second pyramidally truncated lattice quantizer comprises the first pyramidally truncated lattice quantizer.

19. A method for encoding a frame of an input audio signal that comprises a time series of input samples into a frame of an encoded audio signal, the method comprising,
encoding said frame of the input audio signal using at least two of a plurality of audio encoding modes, wherein each of said plurality of audio encoding modes is arranged to encode the frame of the input audio signal into a respective encoded signal, wherein said plurality of audio encoding modes include at least

a first audio encoding mode that comprises linear predictive filtering of said time series of input samples using linear predictive filter coefficients computed using a backward prediction into a residual signal that comprises a respective time series of residual samples and quantizing the time series of residual samples, and
a second audio encoding mode that comprises directly quantizing the time series of input samples; and

selecting, in accordance with a mode selection rule, one of the respective encoded signals as the frame of the encoded audio signal.

20. A method for decoding a frame of an encoded audio signal into a frame of a reconstructed audio signal that comprises a time series of output samples, the comprising:

decoding said frame of the encoded audio signal with one of a plurality of audio decoding modes, wherein said plurality of audio decoding modes include at least a first audio decoding mode and a second audio decoding mode;

wherein the first audio decoding mode is comprises to dequantizing encoded residual parameters received in said frame of the encoded audio signal into a frame of reconstructed residual signal that comprises a time series of reconstructed residual samples and linear predictive filter said time series of reconstructed residual samples into said time series of output samples using linear predictive filter coefficients computed using a backward prediction, and
wherein the second audio decoding mode comprises directly dequantizing encoded signal-domain parameters received in said frame of the encoded audio signal into said time series of output samples.

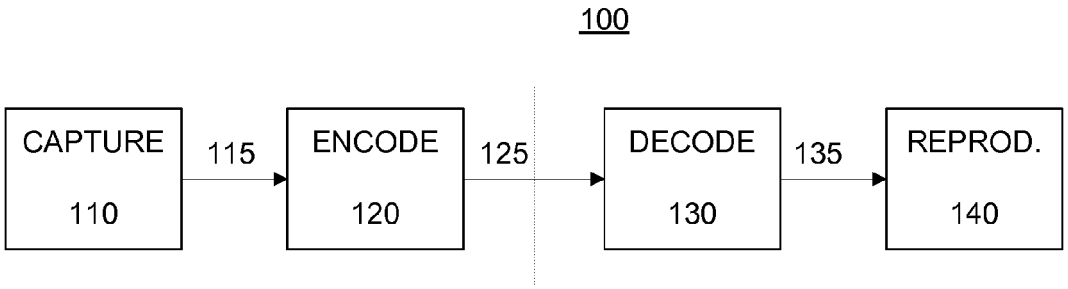

Figure 1

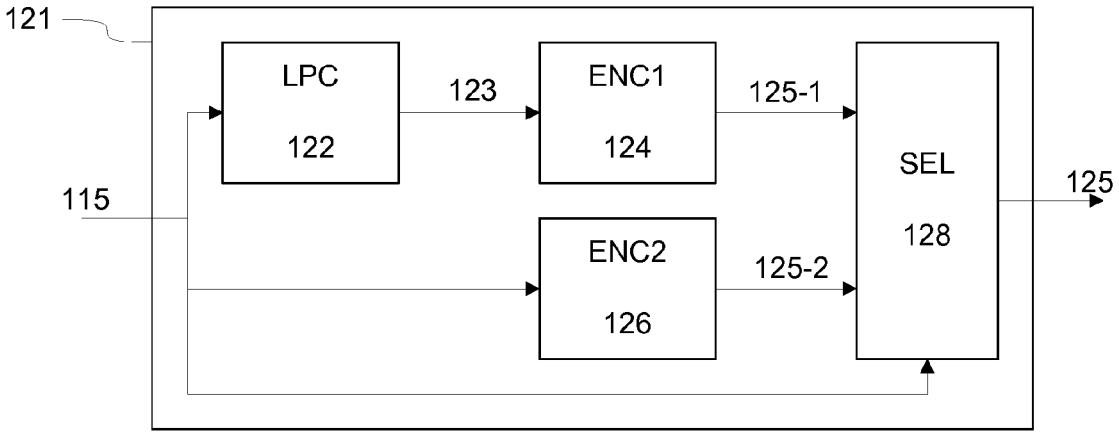

Figure 2

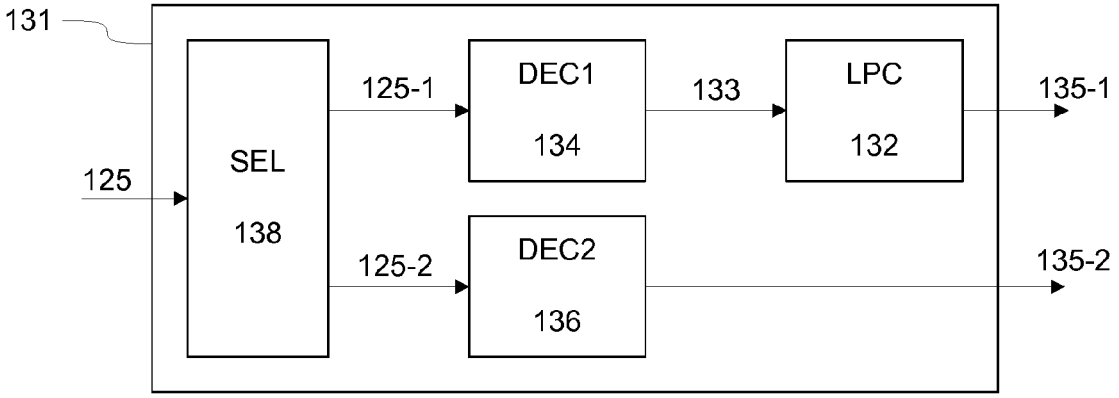

Figure 3

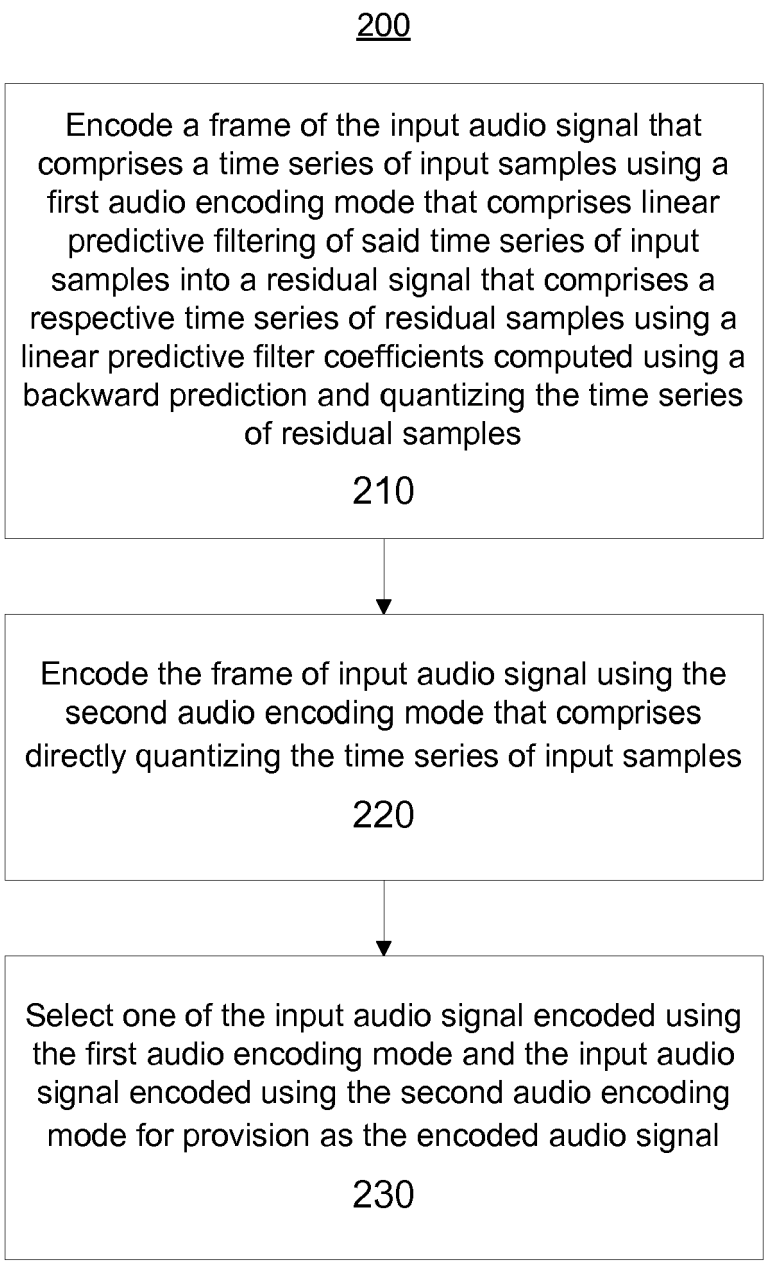
200
Encode a frame of the input audio signal that comprises a time series of input samples using a first audio encoding mode that comprises linear predictive filtering of said time series of input samples into a residual signal that comprises a respective time series of residual samples using a linear predictive filter coefficients computed using a backward prediction and quantizing the time series of residual samples
210
Encode the frame of input audio signal using the second audio encoding mode that comprises directly quantizing the time series of input samples
220
Select one of the input audio signal encoded using the first audio encoding mode and the input audio signal encoded using the second audio encoding mode for provision as the encoded audio signal
230

Figure 4

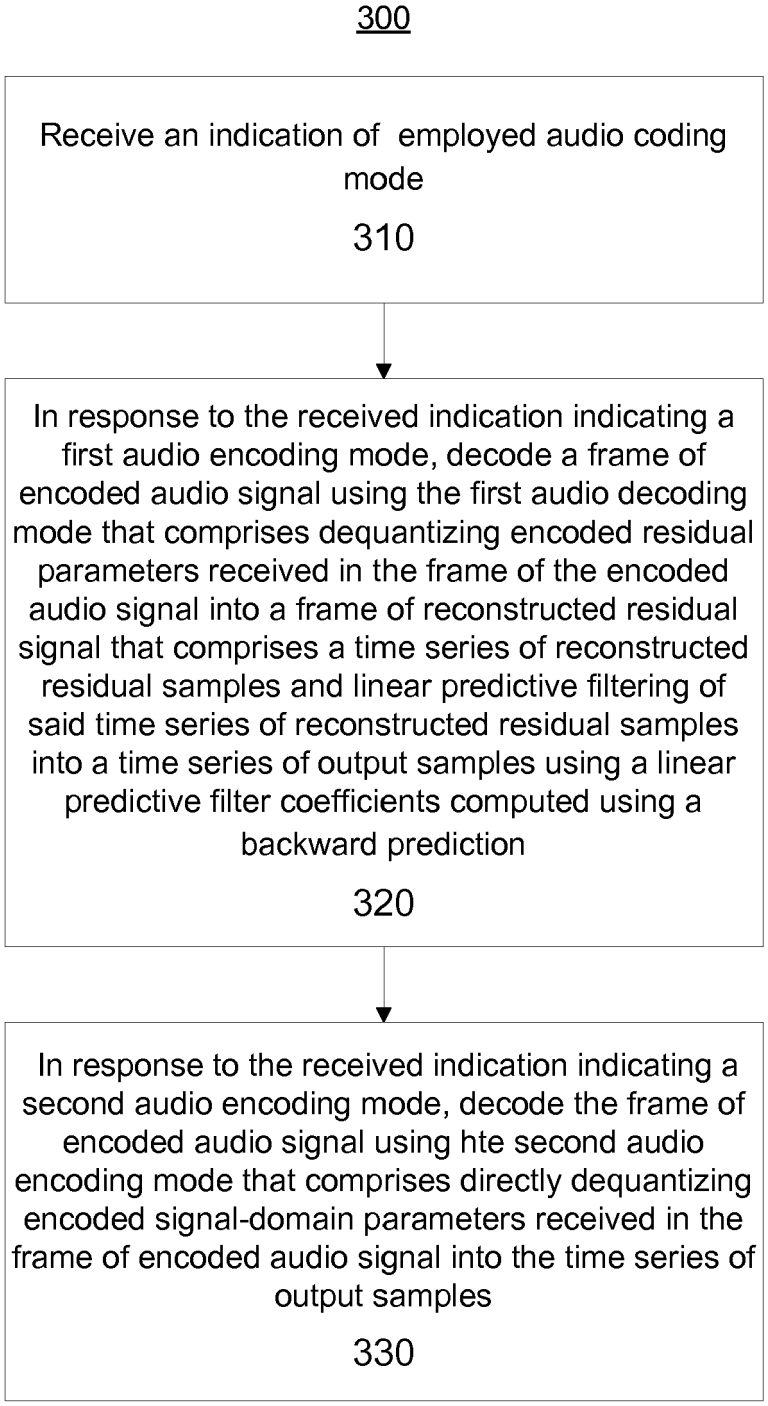
300
Receive an indication of employed audio coding mode
310
In response to the received indication indicating a first audio encoding mode, decode a frame of encoded audio signal using the first audio decoding mode that comprises dequantizing encoded residual parameters received in the frame of the encoded audio signal into a frame of reconstructed residual signal that comprises a time series of reconstructed residual samples and linear predictive filtering of said time series of reconstructed residual samples into a time series of output samples using a linear predictive filter coefficients computed using a backward prediction
320
In response to the received indication indicating a second audio encoding mode, decode the frame of encoded audio signal using hte second audio encoding mode that comprises directly dequantizing encoded signal-domain parameters received in the frame of encoded audio signal into the time series of output samples
330

Figure 5

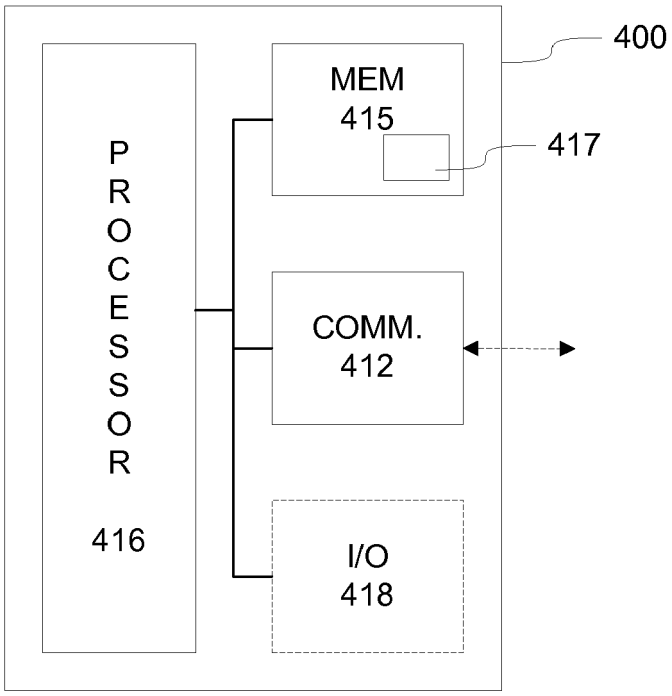

Figure 6

Europäisches Patentamt
European Patent Office
Office européen des brevets

EUROPEAN SEARCH REPORT

Application Number
EP 16 17 1853

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
X	US 2002/069075 A1 (MIET GILLES [FR]) 6 June 2002 (2002-06-06)	1-4, 10-13, 19,20	INV. G10L19/18
Y	* paragraph [0037] * * paragraph [0045] * * claim 1 * * paragraph [0041] - paragraph [0042] *	5-9, 14-18	ADD. G10L19/22
A	US 2012/093213 A1 (MORIYA TAKEHIRO [JP] ET AL) 19 April 2012 (2012-04-19) * paragraph [0007] * * paragraph [0036] * * paragraph [0036] - paragraph [0038] * * paragraph [0043] - paragraph [0048] * * paragraph [0039] - paragraph [0042] * * paragraph [0055] *	1-20	
Y	EP 0 417 739 A2 (FUJITSU LTD [JP]) 20 March 1991 (1991-03-20) * page 3, line 1 - line 20 * * page 6, line 16 - line 26 * * page 16, line 12 - page 17, line 24 *	5-9, 14-18	
Y	US 2012/232913 A1 (TERRIBERRY TIMOTHY B [US] ET AL) 13 September 2012 (2012-09-13) * paragraph [0018] * * paragraph [0023] *	5-9, 14-18	

TECHNICAL FIELDS SEARCHED (IPC)
G10L

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
The Hague	10 November 2016	Burchett, Stefanie

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

ANNEX TO THE EUROPEAN SEARCH REPORT ON EUROPEAN PATENT APPLICATION NO. EP 16 17 1853

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report. The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

10-11-2016

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2002069075 A1	06-06-2002	CN 1237073 A	01-12-1999
		DE 69922582 D1	20-01-2005
		DE 69922582 T2	06-10-2005
		EP 0961264 A1	01-12-1999
		ES 2235431 T3	01-07-2005
		JP 2000091922 A	31-03-2000
		KR 19990088502 A	27-12-1999
		US 2002069075 A1	06-06-2002
US 2012093213 A1	19-04-2012	CN 102449689 A	09-05-2012
		JP 5486597 B2	07-05-2014
		US 2012093213 A1	19-04-2012
		WO 2010140546 A1	09-12-2010
EP 0417739 A2	20-03-1991	CA 2024742 A1	12-03-1991
		DE 69020269 D1	27-07-1995
		DE 69020269 T2	25-01-1996
		EP 0417739 A2	20-03-1991
		JP H0398318 A	23-04-1991
		US 5224167 A	29-06-1993
US 2012232913 A1	13-09-2012	US 2012232913 A1	13-09-2012
		WO 2012122299 A1	13-09-2012

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Non-patent literature cited in the description

- **THOMAS R.** A pyramid Vector Quantizer. *IEEE Transactions on Information Theory,* July 1986, vol. 32 (4), ISSN 0018-9448, 568-583 **[0053]**