(19) DANMARK

(10) **DK/EP 3079153 T3**

Patent- og Varemærkestyrelsen

(12) Oversættelse af europæisk patentskrift

(51) Int.Cl.: ***G 10 L 19/26 (2013.01)*** *G 10 L 19/02 (2013.01)* *G 10 L 19/107 (2013.01)* *G 10 L 19/20 (2013.01)*

(45) Oversættelsen bekendtgjort den: **2018-11-05**

(80) Dato for Den Europæiske Patentmyndigheds bekendtgørelse om meddelelse af patentet: **2018-08-01**

(86) Europæisk ansøgning nr.: **16166357.0**

(86) Europæisk indleveringsdag: **2011-06-23**

(87) Den europæiske ansøgnings publiceringsdag: **2016-10-12**

(30) Prioritet: **2010-07-02 US 361237 P**

(62) Stamansøgningsnr: **14164773.5**

(84) Designerede stater: **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**

(73) Patenthaver: **Dolby International AB, Apollo Building, 3E , Herikerbergweg 1-35, 1101 CN Amsterdam Zuidoost, Holland**

(72) Opfinder: **RESCH, Barbara, c/o Dolby Sweden AB, Gävlegatan 12A, 113 30 Stockholm, Sverige**
KJOERLING, Kristofer, c/o Dolby Sweden AB, Gävlegatan 12 A, 113 30 Stockholm, Sverige
VILLEMOES, Lars, c/o Dolby Sweden AB, Gävlegatan 12A, 113 30 Stockholm, Sverige

(74) Fuldmægtig i Danmark: **Plougmann Vingtoft A/S, Rued Langgaards Vej 8, 2300 København S, Danmark**

(54) Benævnelse: **AUDIO-AFKODNING MED SELEKTIV EFTERFILTRERING**

(56) Fremdragne publikationer:
WO-A1-99/38155
MAX NEUENDORF (EDITOR): "WD7 of USAC", 92. MPEG MEETING; 19-04-2010 - 23-04-2010; DRESDEN; (MOTION PICTURE EXPERT GROUP OR ISO/IECJTC1/SC29/WG11), no. N11299, 26 April 2010 (2010-04-26), XP002658193,
BARBARA RESCH ET AL: "Finalization of CE on an improved basspost filter operation for the ACELP of USAC", 94. MPEG MEETING; 11-10-2010 - 15-10-2010; GUANGZHOU; (MOTION PICTURE EXPERT GROUP OR ISO/IEC JTC1/SC29/WG11),, no. M18379, 6 October 2010 (2010-10-06), XP030046969,
ANONYMOUS: "Study on ISO/IEC 23003-3:201x/CD of Unified Speech and Audio Coding", 94. MPEG MEETING;11-10-2010 - 15-10-2010; GUANGZHOU; (MOTION PICTUREEXPERT GROUP OR ISO/IEC JTC1/SC29/WG11),, no. N11659, 16 November 2010 (2010-11-16), XP030018154, ISSN: 0000-0009

DESCRIPTION

Technical field

[0001] The present invention generally relates to digital audio coding and more precisely to coding techniques for audio signals containing components of different characters.

Background

[0002] A widespread class of coding method for audio signals containing speech or singing includes *code excited linear prediction* (CELP) applied in time alternation with different coding methods, including frequency-domain coding methods especially adapted for music or methods of a general nature, to account for variations in character between successive time periods of the audio signal. For example, a simplified Moving Pictures Experts Group (MPEG) Unified Speech and Audio Coding (USAC; see standard ISO/IEC 23003-3) decoder is operable in at least three decoding modes, Advanced Audio Coding (AAC; see standard ISO/IEC 13818-7), algebraic CELP (ACELP) and transform-coded excitation (TCX), as shown in the upper portion of accompanying figure 2.

[0003] The various embodiments of CELP are adapted to the properties of the human organs of speech and, possibly, to the human auditory sense. As used in this application, CELP will refer to all possible embodiments and variants, including but not limited to ACELP, wide- and narrow-band CELP, SB-CELP (sub-band CELP), low- and high-rate CELP, RCELP (relaxed CELP), LD-CELP (low-delay CELP), CS-CELP (conjugate-structure CELP), CS-ACELP (conjugate-structure ACELP), PSI-CELP (pitch-synchronous innovation CELP) and VSELP (vector sum excited linear prediction). The principles of CELP are discussed by R. Schroeder and S. Atal in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), vol. 10, pp. 937-940, 1985, and some of its applications are described in references 25-29 cited in Chen and Gersho, IEEE Transactions on Speech and Audio Processing, vol. 3, no. 1, 1995. As further detailed in the former paper, a CELP decoder (or, analogously, a CELP speech synthesizer) may include a pitch predictor, which restores the periodic component of an encoded speech signal, and an pulse codebook, from which an innovation sequence is added. The pitch predictor may in turn include a long-delay predictor for restoring the pitch and a short-delay predictor for restoring formants by spectral envelope shaping. In this context, the pitch is generally understood as the fundamental frequency of the tonal sound component produced by the vocal chords and further coloured by resonating portions of the vocal tract. This frequency together with its harmonics will dominate speech or singing. Generally speaking, CELP methods are best suited for processing solo or one-part singing, for which the pitch frequency is well-defined and relatively easy to determine.

[0004] To improve the perceived quality of CELP-coded speech, it is common practice to combine it with *post filtering* (or *pitch enhancement* by another term). U.S. Patent No. 4 969 192 and section II of the paper by Chen and Gersho disclose desirable properties of such post filters, namely their ability to suppress noise components located between the harmonics of the detected voice pitch (long-term portion; see section IV). It is believed that an important portion of this noise stems from the spectral envelope shaping. The long-term portion of a simple post filter may be designed to have the following transfer function:

$$H_E(z) = 1 + \alpha\left(\frac{z^T + z^{-T}}{2} - 1\right),$$

where T is an estimated pitch period in terms of number of samples and α is a gain of the post filter, as shown in figures 1 and 2. In a manner similar to a comb filter, such a filter attenuates frequencies 1/(2T), 3/(2T), 5/(2T), ..., which are located midway between harmonics of the pitch frequency, and adjacent frequencies. The attenuation depends on the value of the gain α. Slightly more sophisticated post filters apply this attenuation only to low frequencies - hence the commonly used term *bass* post filter - where the noise is most perceptible. This can be expressed by cascading the transfer function H_E described above and a low-pass filter H_{LP}. Thus, the post-processed decoded S_E provided by the post filter will be given, in the transform domain, by

$$S_E(z) = S(z) - \alpha S(z) P_{LT}(z) H_{LP}(z),$$

where

$$P_{LT}(z) = 1 - \frac{z^T + z^{-T}}{2}$$

and S is the decoded signal which is supplied as input to the post filter. Figure 3 shows an embodiment of a post filter with these characteristics, which is further discussed in section 6.1.3 of the Technical Specification ETSI TS 126 290, version 6.3.0, release 6. As this figure suggests, the pitch information is encoded as a parameter in the bit stream signal and is retrieved by a pitch tracking module communicatively connected to the long-term prediction filter carrying out the operations expressed by P_{LT}.

[0005] The long-term portion described in the previous paragraph may be used alone. Alternatively, it is arranged in series with a noise-shaping filter that preserves components in frequency intervals corresponding to the formants and attenuates noise in other spectral regions (short-term portion; see section III), that is, in the 'spectral valleys' of the formant envelope. As another possible variation, this filter aggregate is further supplemented by a gradual high-pass-type filter to reduce a perceived deterioration due to spectral tilt of the short-term portion. Document M. Neuendorf (ed.): WD7 of USAC, 92th MPEG Meeting, Dresden, N11299, and the International Application published as WO 99/38144 A1 disclose audio processing systems with selective application of post-filtering.

[0006] Audio signals containing a mixture of components of different origins - e.g., tonal, non-tonal, vocal, instrumental, non-musical - are not always reproduced by available digital coding technologies in a satisfactory manner. It has more precisely been noted that available technologies are deficient in handling such non-homogeneous audio material, generally favouring one of the components to the detriment of the other. In particular, music containing singing accompanied by one or more instruments or choir parts which has been encoded by methods of the nature described above, will often be decoded with perceptible artefacts spoiling part of the listening experience.

Summary of the invention

[0007] In order to mitigate at least some of the drawbacks outlined in the previous section, a method of decoding, a computer-program product and a decoding system in accordance with the invention are provided in claims 1, 4 and 5, respectively. Dependent claims 2 and 3 define further embodiments. The inventors have realized that some artefacts perceived in decoded audio signals of non-homogeneous origin derive from an inappropriate switching between several coding modes of which at least one includes post filtering at the decoder and at least one does not. More precisely, available post filters remove not only interharmonic noise (and, where applicable, noise in spectral valleys) but also signal components representing instrumental or vocal accompaniment and other material of a 'desirable' nature. The fact that the *just noticeable difference* in spectral valleys may be as large as 10 dB (as noted by Ghitza and Goldstein, IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-4, pp. 697-708, 1986) may have been taken as a justification by many designers to filter these frequency bands severely. The quality degradation by the interharmonic (and spectral-valley) attenuation itself may however be less important than that of the switching occasions. When the post filter is switched on, the background of a singing voice sounds suddenly muffled, and when the filter is deactivated, the background instantly becomes more sonorous. If the switching takes place frequently, due to the nature of the audio signal or to the configuration of the coding device, there will be a switching artefact. As one example, a USAC decoder may be operable either in an ACELP mode combined with post filtering or in a TCX mode without post filtering. The ACELP mode is used in episodes where a dominant vocal component is present. Thus, the switching into the ACELP mode may be triggered by the onset of singing, such as at the beginning of a new musical phrase, at the beginning of a new verse, or simply after an episode where the accompaniment is deemed to drown the singing voice in the sense that the vocal component is no longer prominent. Experiments have confirmed that an alternative solution, or rather circumvention of the problem, by which TCX coding is used throughout (and the ACELP mode is disabled) does not remedy the problem, as reverb-like artefacts appear.

[0008] Accordingly, in a first and a second aspects, there is provided an audio encoding method (and an audio encoding system with the corresponding features) characterized by a decision being made as to whether the device which will decode the bit stream, which is output by the encoding method, should apply post filtering including attenuation of interharmonic noise. The outcome of the decision is encoded in the bit stream and is accessible to the decoding device. By the invention, the decision whether to use the post filter is taken separately from the decision as to the most suitable coding mode. This makes it possible to maintain one post filtering status throughout a period of such length that the switching will not annoy the listener. Thus, the encoding method may prescribe that the post filter will be kept inactive even though it switches into a coding mode where the filter is conventionally active.

[0009] It is noted that the decision whether to apply post filtering is normally taken frame-wise. Thus, firstly, post filtering is not applied for less than one frame at a time. Secondly, the decision whether to disable post filtering is only valid for the duration of a current frame and may be either maintained or reassessed for the subsequent frame. In a coding format enabling a main frame format and a reduced format, which is a fraction of the normal format, e.g., 1/8 of its length, it may not be necessary to take post-filtering decisions for individual reduced frames. Instead, a number of reduced frames summing up to a normal frame may be considered, and the parameters relevant for the filtering decision may be obtained by computing the mean or median of the reduced frames comprised therein.

[0010] A decoding method in accordance with the invention is well suited for coding of mixed-origin audio signals by virtue of its capability to deactivate the post filter in dependence of the post filtering information only, hence independently of factors such as the current coding mode. When applied to coding techniques wherein post filter activity is conventionally associated with particular coding modes, the post-filtering disabling capability enables a new operative mode, namely the unfiltered application of a conventionally filtered decoding mode.

[0011] It is noted that the methods and apparatus disclosed in this section may be applied, after appropriate modifications within the skilled person's abilities including routine experimentation, to coding of signals having several components, possibly corresponding to different channels, such as stereo channels. Throughout the present application, pitch enhancement and post filtering are used as synonyms. It is further noted that AAC is discussed as a representative example of frequency-domain coding methods. Indeed, applying the invention to a decoder or encoder operable in a frequency-domain coding mode other than AAC will only require small modifications, if any, within the skilled person's abilities. Similarly, TCX is mentioned as an example of weighted linear prediction transform coding and of transform coding in general.

[0012] Features from two or more embodiments described hereinabove can be combined, unless they are clearly complementary, in further embodiments. The fact that two features are recited in different claims does not preclude that they can be combined to advantage. Likewise, further embodiments can also be provided by the omission of certain features that are not necessary or not essential for the desired purpose.

Brief description of the drawings

[0013] Embodiments of the present invention will now be described with reference to the accompanying drawings, on which:

figures 1 is a block diagram showing a conventional decoder with post filter;

figure 2 is a schematic block diagram of a conventional decoder operable in AAC, ACELP and TCX mode and including a post filter permanently connected downstream of the ACELP module;

figure 3 is a block diagram illustrating the structure of a post filter;

figures 4 and 5 are block diagrams of two decoders;

figure 6 and 7 are block diagrams illustrating differences between a conventional decoder (figure 6) and a decoder (figure 7) according to the invention;

figure 8 is a block diagram of an encoder;

figures 9 and 10 are a block diagrams illustrating differences between a conventional decoder (figure 9) and a decoder (figure 10) according to the invention; and

figure 11 is a block diagram of an autonomous post filter which can be selectively activated and deactivated.

Detailed description of embodiments

[0014] Figure 4 is a schematic drawing of a decoder system 400, having as its input a bit stream signal and as its output an audio signal. As in the conventional decoders shown in figure 1, a post filter 440 is arranged downstream of a decoding module 410 but can be switched into or out of the decoding path by operating a switch 442. The post filter is enabled in the switch position shown in the figure. It would be disabled if the switch was set in the opposite position, whereby the signal from the decoding module 410 would instead be conducted over the bypass line 444. The switch 442 is controllable by post filtering information contained in the bit stream signal, so that post filtering may be applied and removed irrespectively of the current status of the decoding module 410. Because a post filter 440 operates at some delay - for example, the post filter shown in figure 3 will introduce a delay amounting to at least the pitch period T - a compensation delay module 443 is arranged on the bypass line 444 to maintain the modules in a synchronized condition at switching. The delay module 443 delays the signal by the same period as the post filter 440 would, but does not otherwise process the signal. To minimize the change-over time, the compensation delay module 443 receives the same signal as the post filter 440 at all times. In an alternative embodiment where the post filter 440 is replaced by a zero-delay post filter (e.g., a causal filter, such as a filter with two taps, independent of future signal values), the compensation delay module 443 can be omitted.

[0015] Figure 5 illustrates a further development according to the teachings of the invention of the triple-mode decoder system 500 of figure 2. An ACELP decoding module 511 is arranged in parallel with a TCX decoding module 512 and an AAC decoding module 513. In series with the ACELP decoding module 511 is arranged a post filter 540 for attenuating noise, particularly noise located between harmonics of a pitch frequency directly or indirectly derivable from the bit stream signal for which the decoder system 500 is adapted. The bit stream signal also encodes post filtering information governing the positions of an upper switch 541 operable to switch the post filter 540 out of the processing path and replace it with a compensation delay 543 like in figure 4. A lower switch 542 is used for switching between different decoding modes. With this structure, the position of the upper switch 541 is immaterial when one of the TCX or AAC modules 512, 513 is used; hence, the post filtering information does not necessary indicate this position except in the ACELP mode. Whatever decoding mode is currently used, the signal is supplied from the downstream connection point of the lower switch 542 to a spectral band replication (SBR) module 550, which outputs an audio signal. The skilled person will realize that the drawing is of a conceptual nature, as is clear notably from the switches which are shown schematically as separate physical entities with movable contacting means. In a possible realistic implementation of the decoder system, the switches as well as the other modules will be embodied by computer-readable instructions.

[0016] Figures 6 and 7 are also block diagrams of two triple-mode decoder systems operable in an ACELP, TCX or frequency-domain decoding mode. With reference to the latter figure, which shows an embodiment of the invention, a bit stream signal is supplied to an input point 701, which is in turn permanently connected via respective branches to the three decoding modules 711, 712, 713. The input point 701 also has a connecting branch 702 (not present in the conventional decoding system of figure 6) to a pitch enhancement module 740, which acts as a post filter of the general type described above. As is common practice in the art, a first transition windowing module 703 is arranged downstream of the ACELP and TCX modules 711, 712, to carry out transitions between the decoding modules. A second transition module 704 is arranged downstream of the frequency-domain decoding module 713 and the first transition windowing module 703, to carry out transition between the two super-modes. Further a SBR module 750 is provided immediately upstream of the output point 705. Clearly, the bit stream signal is supplied directly (or after demultiplexing, as appropriate) to all three decoding modules 711, 712, 713 and to the pitch enhancement module 740. Information contained in the bit stream controls what decoding module is to be active. By the invention however, the pitch enhancement module 740 performs an analogous self actuation , which responsive to post filtering information in the bit stream may act as a post filter or simply as a pass-through. This may for instance be realized through the provision of a control section (not shown) in the pitch enhancement module 740, by means of which the post filtering action can be turned on or off. The pitch enhancement module 740 is always in its pass-through mode when the decoder system operates in the frequency-domain or TCX decoding mode, wherein strictly speaking no post filtering information is necessary. It is understood that modules not forming part of the inventive contribution and whose presence is obvious to the skilled person, e.g., a demultiplexer, have been omitted from figure 7 and other similar drawings to increase clarity.

[0017] As a variation, the decoder system of figure 7 may be equipped with a control module (not shown) for deciding whether post filtering is to be applied using an analysis-by-synthesis approach. Such control module is communicatively connected to the pitch enhancement module 740 and to the ACELP module 711, from which it extracts an intermediate decoded signal $s_{i\_DEC}(n)$ representing an intermediate stage in the decoding process, preferably one

corresponding to the excitation of the signal. The detection module has the necessary information to simulate the action of the pitch enhancement module 740, as defined by the transfer functions $P_{LT}(z)$ and $H_{LP}(z)$ (cf. *Background* section and figure 3), or equivalently their filter impulse responses $p_{LT}(z)$ and $h_{LP}(n)$. As follows by the discussion in the *Background* section, the component to be subtracted at post filtering can be estimated by an approximate difference signal $s_{AD}(n)$ which is proportional to $[(S_{i\_DEC} * p_{LT}) * h_{LP}](n)$, where * denotes discrete convolution. This is an approximation of the true difference between the original audio signal and the post-filtered decoded signal, namely
$s_{ORIG}(n) - s_E(n) = s_{ORIG}(n) - (s_{DEC}(n) - \alpha[s_{DEC} * p_{LT} * h_{LP}](n))$,
where α is the post filter gain. By studying the total energy, low-band energy, tonality, actual magnitude spectrum or past magnitude spectra of this signal, as disclosed in the *Summary* section and the claims, the control section may find a basis for the decision whether to activate or deactivate the pitch enhancement module 740.

[0018] Figure 8 shows an encoder system 800. The encoder system 800 is adapted to process digital audio signals, which are generally obtained by capturing a sound wave by a microphone and transducing the wave into an analog electric signal. The electric signal is then sampled into a digital signal susceptible to be provided, in a suitable format, to the encoder system 800. The system generally consists of an encoding module 810, a decision module 820 and a multiplexer 830. By virtue of switches 814, 815 (symbolically represented), the encoding module 810 is operable in either a CELP, a TCX or an AAC mode, by selectively activating modules 811, 812, 813. The decision module 820 applies one or more predefined criteria to decide whether to disable post filtering during decoding of a bit stream signal produced by the encoder system 800 to encode an audio signal. For this purpose, the decision module 820 may examine the audio signal directly or may receive data from the encoding module 810 via a connection line 816. A signal indicative of the decision taken by the decision module 820 is provided, together with the encoded audio signal from the encoding module 810, to a multiplexer 830, which concatenates the signals into a bit stream constituting the output of the encoder system 800.

[0019] Preferably, the decision module 820 bases its decision on an approximate difference signal computed from an intermediate decoded signal s_{i_DEC}, which can be subtracted from the encoding module 810. The intermediate decoded signal represents an intermediate stage in the *decoding* process, as discussed in preceding paragraphs, but may be extracted from a corresponding stage of the *encoding* process. However, in the encoder system 800 the original audio signal s_{ORIG} is available so that, advantageously, the approximate difference signal is formed as:
$s_{ORIG}(n) - (s_{i\_DEC}(n) - \alpha\,[(s_{i\_DEC} * p_{LT}) * h_{LP}](n))$.
The approximation resides in the fact that the intermediate decoded signal is used in lieu of the final decoded signal. This enables an appraisal of the nature of the component that a post filter would remove at decoding, and by applying one of the criteria discussed in the *Summary* section, the decision module 820 will be able to take a decision whether to disable post filtering.

[0020] As a variation to this, the decision module 820 may use the original signal in place of an intermediate decoded signal, so that the approximate difference signal will be $[(s_{i\_DEC} * p_{LT}) * h_{LP}](n)$. This is likely to be a less faithful approximation but on the other hand makes the presence of a connection line 816 between the decision module 820 and the encoding module 810 optional.

[0021] In such other variations of this embodiment where the decision module 820 studies the audio signal directly, one or more of the following criteria may be applied:

- Does the audio signal contain both a component with dominant fundamental frequency and a component located below the fundamental frequency? (The fundamental frequency may be supplied as a by-product of the encoding module 810.)
- Does the audio signal contain both a component with dominant fundamental frequency and a component located between the harmonics of the fundamental frequency?
- Does the audio signal contain significant signal energy below the fundamental frequency?
- Is post-filtered decoding (likely to be) preferable to unfiltered decoding with respect to rate-distortion optimality?

[0022] In all the described variations of the encoder structure shown in figure 8 - that is, irrespectively of the basis of the detection criterion - the decision section 820 may be enabled to decide on a gradual onset or gradual removal of post filtering, so as to achieve smooth transitions. The gradual onset and removal may be controlled by adjusting the post filter gain.

[0023] Figure 9 shows a conventional decoder operable in a frequency-decoding mode and a CELP decoding mode depending on the bit stream signal supplied to the decoder. Post filtering is applied whenever the CELP decoding mode is selected. An improvement of this decoder is illustrated in figure 10, which shows an decoder 1000 according to an embodiment of the invention. This decoder is operable not only in a frequency-domain-based decoding mode, wherein the frequency-domain decoding module 1013 is active, and a filtered CELP decoding mode, wherein the CELP decoding module 1011 and the post filter 1040 are active, but also in an unfiltered CELP mode, in which the CELP module 1011 supplies its signal to a compensation delay module 1043 via a bypass line 1044. A switch 1042 controls what decoding mode is currently used responsive to post filtering information contained in the bit stream signal provided to the decoder 1000. In this decoder and that of figure 9, the last processing step is effected by an SBR module 1050, from which the final audio signal is output.

[0024] Figure 11 shows a post filter 1100 suitable to be arranged downstream of a decoder 1199. The filter 1100 includes a post filtering module 1140, which is enabled or disabled by a control module (not shown), notably a binary or non-binary gain controller, in response to a post filtering signal received from a decision module 1120 within the post filter 1100. The decision module performs one or more tests on the signal obtained from the decoder to arrive at a decision whether the post filtering module 1140 is to be active or inactive. The decision may be taken along the lines of the functionality of the decision module 820 in figure 8, which uses the original signal and/or an intermediate decoded signal to predict the action of the post filter. The decision of the decision module 1120 may also be based on similar information as the decision modules uses in those embodiments where an intermediate decoded signal is formed. As one example, the decision module 1120 may estimate a pitch frequency (unless this is readily extractable from the bit stream signal) and compute the energy content in the signal below the pitch frequency and between its harmonics. If this energy content is significant, it probably represents a relevant signal component rather than noise, which motivates a decision to disable the post filtering module 1140.

[0025] A 6-person listening test has been carried out, during which music samples encoded and decoded according to the invention were compared with reference samples containing the same music coded while applying post filtering in the conventional fashion but maintaining all other parameters unchanged. The results confirm a perceived quality improvement.

[0026] Further embodiments of the present invention will become apparent to a person skilled in the art after reading the description above. Even though the present description and drawings disclose embodiments and examples, the invention is not restricted to these specific examples. Numerous modifications and variations can be made without departing from the scope of the present invention, which is defined by the accompanying claims.

[0027] The systems and methods disclosed hereinabove may be implemented as software, firmware, hardware or a combination thereof. Certain components or all components may be implemented as software executed by a digital signal processor or microprocessor, or be implemented as hardware or as an application-specific integrated circuit. Such software may be distributed on computer readable media, which may comprise computer storage media (or non-transitory media) and communication media (or transitory media). As is well known to a person skilled in the art, computer storage media includes both volatile and nonvolatile, removable and non-removable media implemented in any method or technology for storage of information such as computer readable instructions, data structures, program modules or other data. Computer storage media includes, but is not limited to, RAM, ROM, EEPROM, flash memory or other memory technology, CD-ROM, digital versatile disks (DVD) or other optical disk storage, magnetic cassettes, magnetic tape, magnetic disk storage or other magnetic storage devices, or any other medium which can be used to store the desired information and which can be accessed by a computer. Further, it is well known to the skilled person that communication media typically embodies computer readable instructions, data structures, program modules or other data in a modulated data signal such as a carrier wave or other transport mechanism and includes any information delivery media.

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.

Patent documents cited in the description

- US4969192A [0004]
- WO9938144A1 [0005]

Non-patent literature cited in the description

- **R. SCHROEDERS. ATAL**Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 1985, vol. 10, 937-940 [0003]
- **CHENGERSHO**IEEE Transactions on Speech and Audio Processing, 1995, vol. 3, 1 [0003]
- WD7 of USAC, 92th MPEG Meeting [0005]
- **GHITZAGOLDSTEIN**IEEE Trans. Acoust., Speech, Signal Processing, 1986, vol. ASSP-4, 697-708 [0007]

Patentkrav

1. Fremgangsmåde til afkodning af et bitstrømsignal som et audio-tidssignal, hvilken inkluderer trinnene:

at afkode et bitstrømsignal som et foreløbigt audio-tidssignal ifølge en kodningsmetode valgt fra en flerhed af kodningsmetoder, hvor flerheden af kodningsmetoder inkluderer mindst en første kodningsmetode, der inkluderer et efterfiltreringstrin og mindst en anden kodningsmetode, der ikke inkluderer efterfiltreringstrinnet,

hvor efterfiltreringstrinnet anvender et toneforstærkningsfilter til det foreløbige audio-tidssignal, hvorved der opnås et audio-tidssignal,

hvor efterfiltreringstrinnet selektivt udelades som reaktion på efterfiltreringsinformation indkodet i bitstrømsignalet, hvor efterfiltreringsinformationen indikerer en indkoder-side-beslutning om hvorvidt at udelade efterfiltreringstrinnet eller ej, hvorved efterfiltreringstrinnet selektivt udelades i den første kodningsmetode.

2. Fremgangsmåden ifølge krav 1, hvor afkodningstrinnet inkluderer at anvendekode-exciteret lineær forudsigelse, CELP, afkodning.

3. Fremgangsmåden ifølge krav 1, hvor bitstrømsignalet er segmenteret i tidsrammer og efterfiltreringstrinnet udelades for en hel tidsramme eller en sekvens af hele tidsrammer.

4. Computerprogramprodukt, der inkluderer en databærer med lagrede instruktioner, som, når eksekveret af en digital signalprocessor, får den digitale signalprocessor til at udføre fremgangsmåden ifølge krav 1, 2 eller 3.

5. Afkodningssystem konfigureret til at udføre fremgangsmåden ifølge krav 1, 2 eller 3.

DRAWINGS

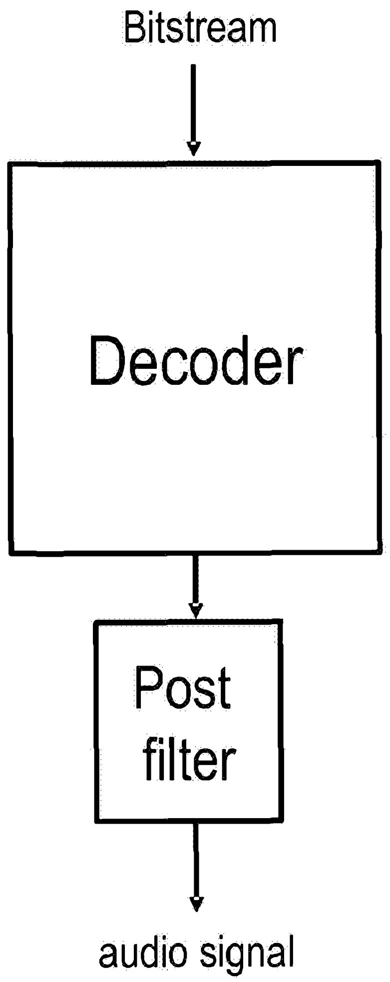

Fig. 1
(prior art)

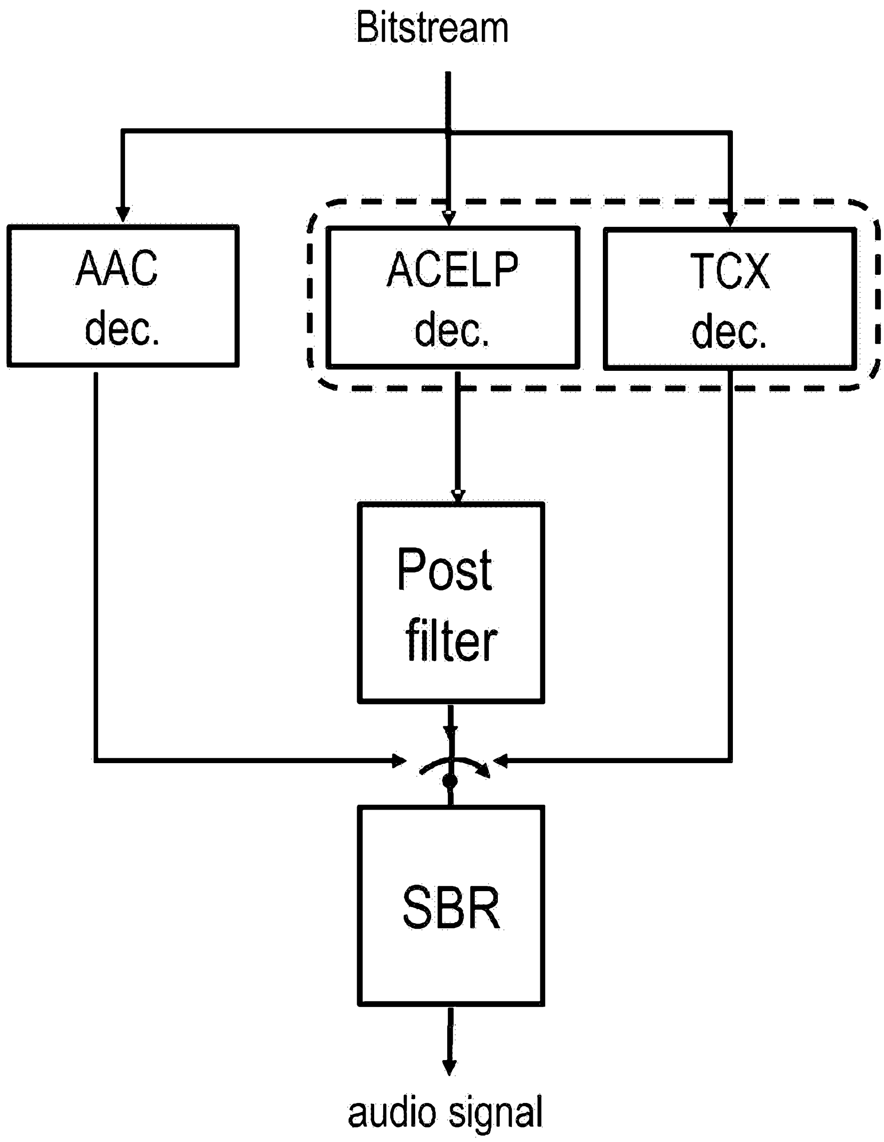

Fig. 2
(prior art)

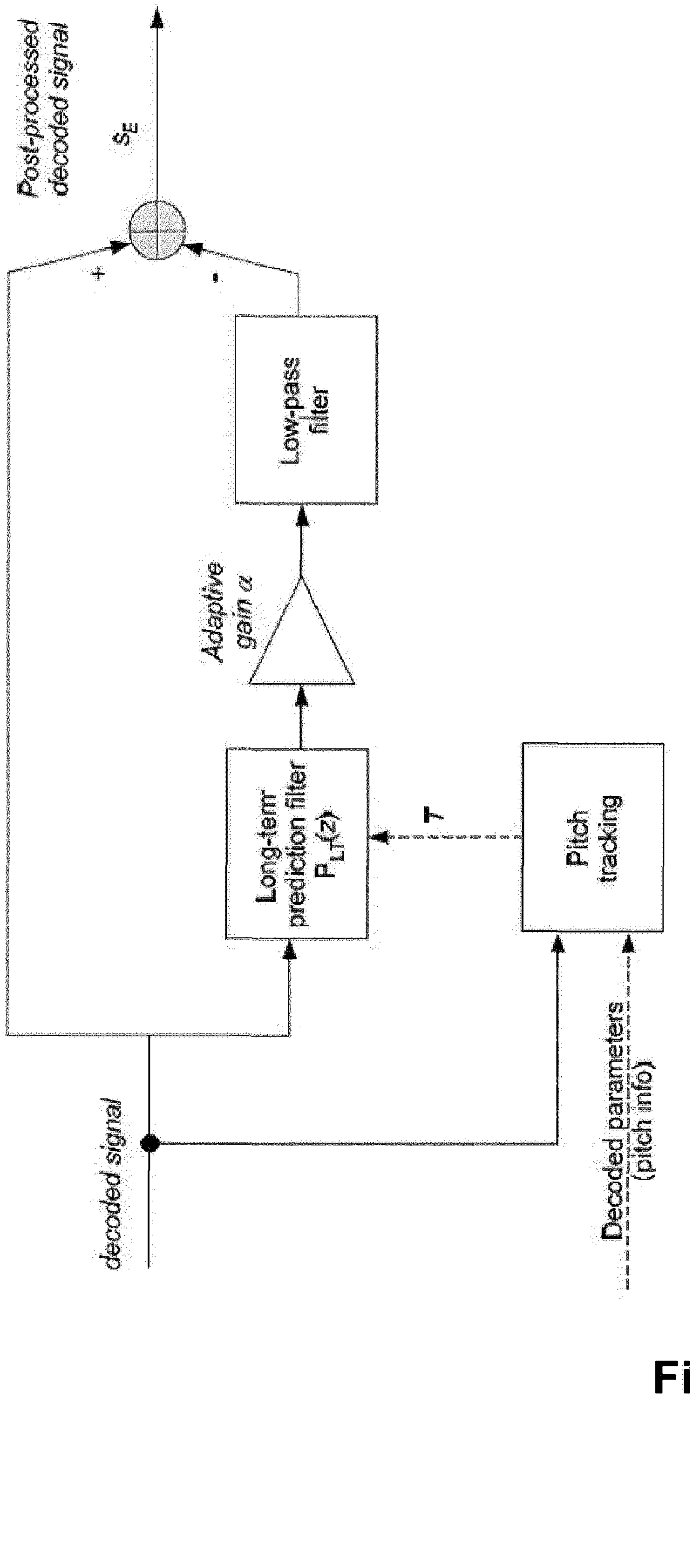
Post-processed decoded signal
s_E
+
-
Low-pass filter
Adaptive gain α
Long-term prediction filter $P_{LT}(z)$
T
Pitch tracking
decoded signal
Decoded parameters (pitch info)

Fig. 3

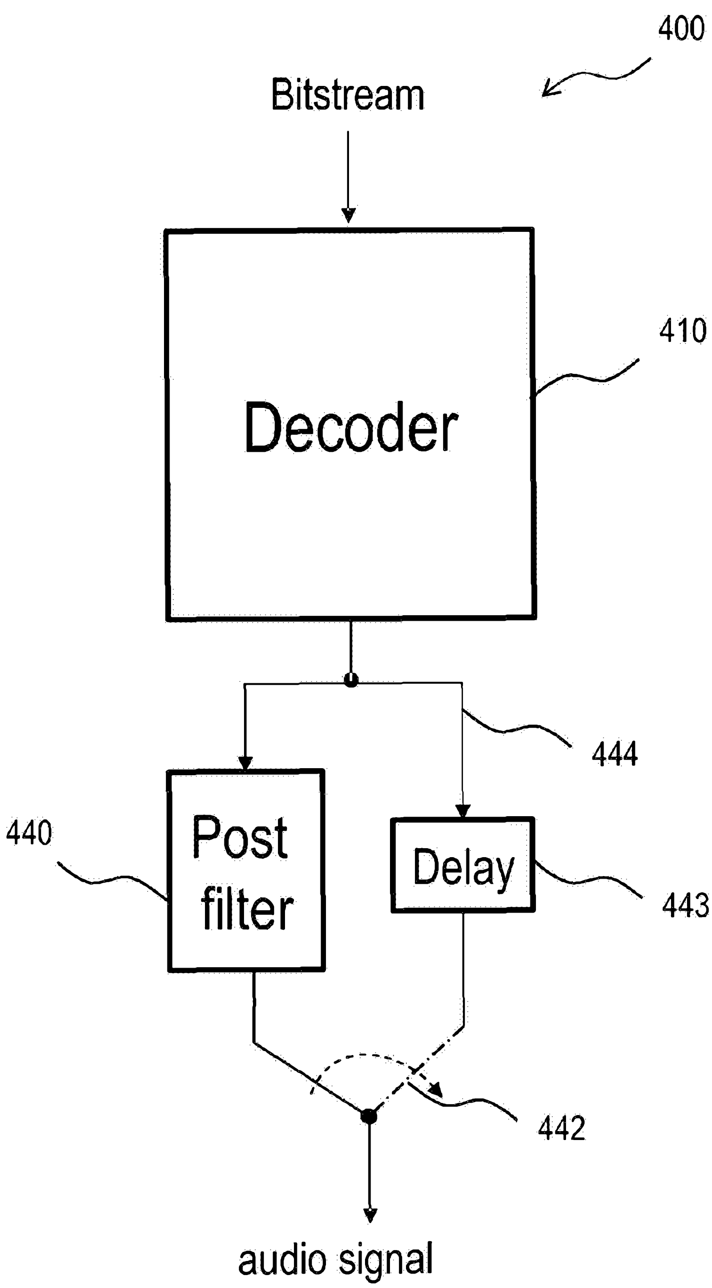
400
Bitstream
Decoder
410
444
440
Post filter
Delay
443
442
audio signal

Fig. 4

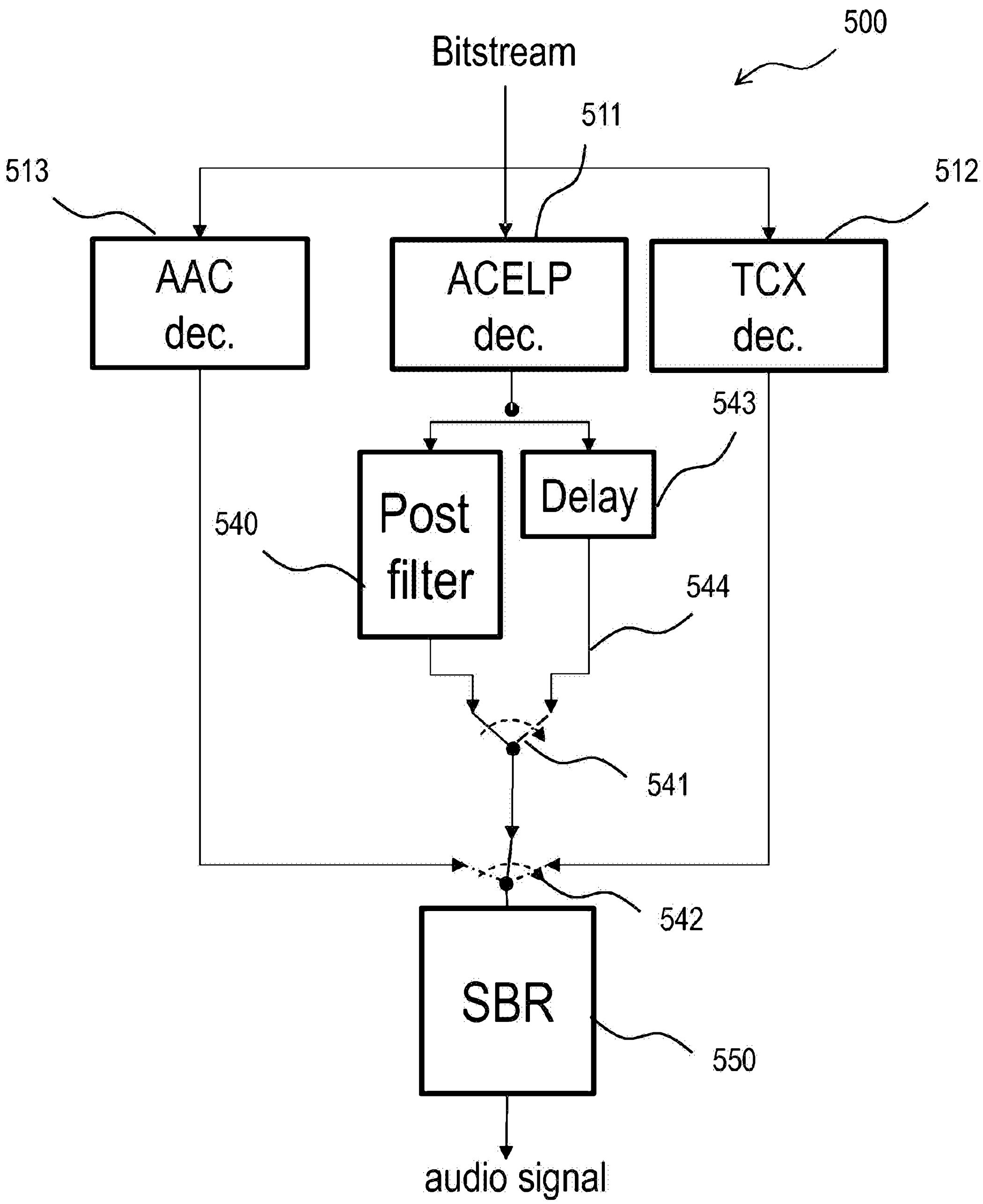
500
Bitstream
511
513
512
AAC
dec.
ACELP
dec.
TCX
dec.
543
Post
filter
Delay
540
544
541
542
SBR
550
audio signal

Fig. 5

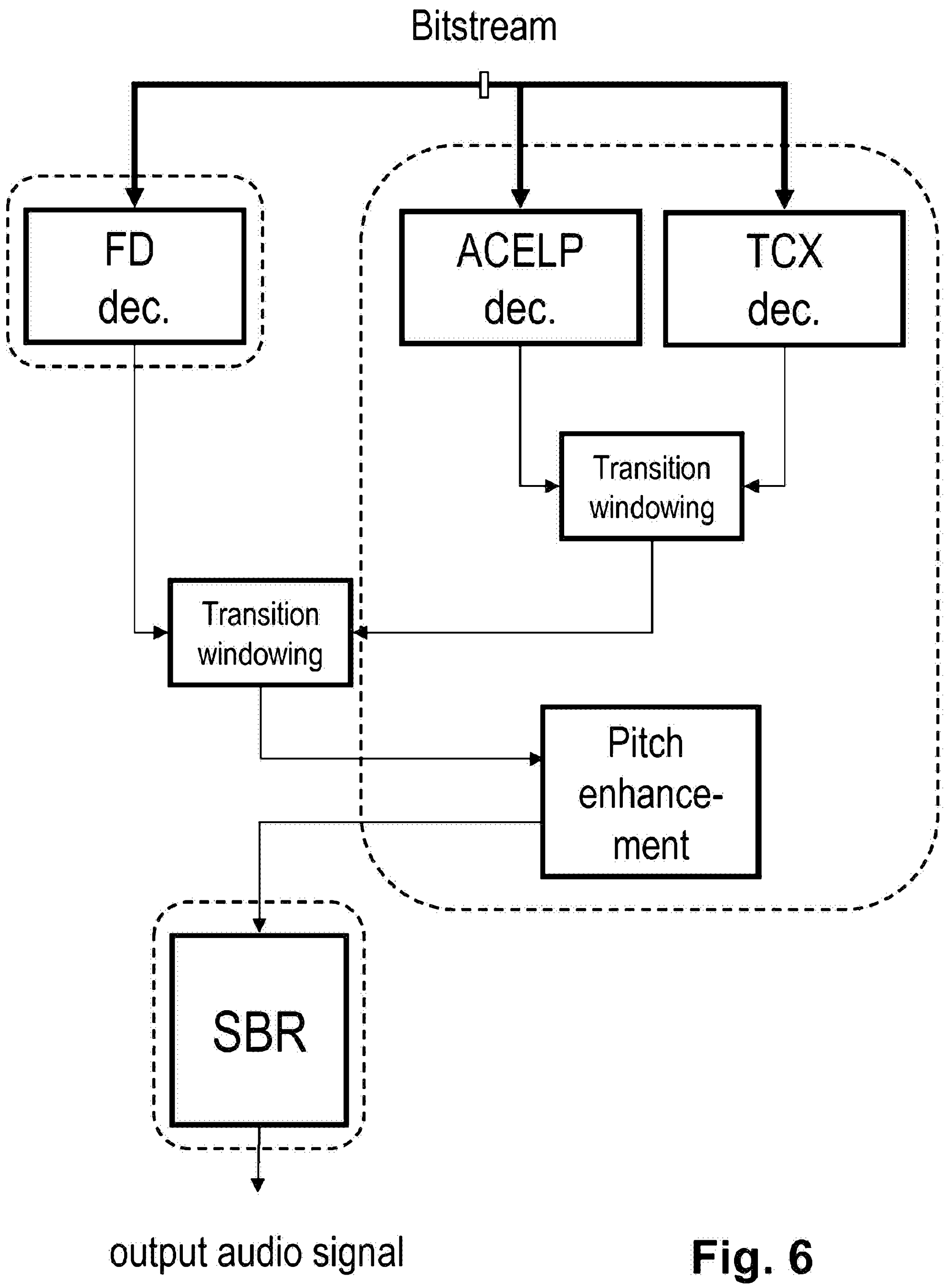

Fig. 6
(prior art)

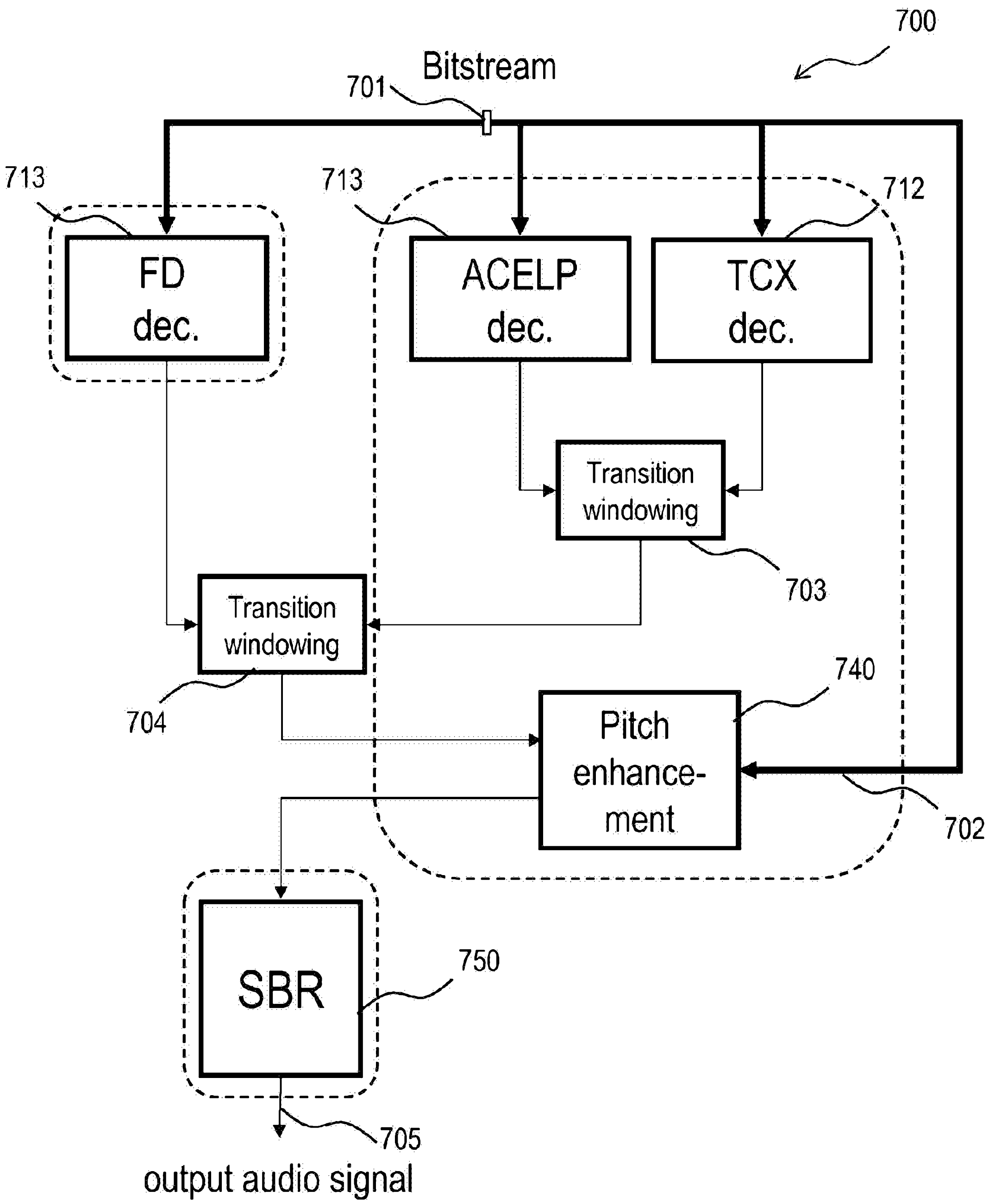
700
Bitstream
701
713
713
712
FD
dec.
ACELP
dec.
TCX
dec.
Transition
windowing
703
Transition
windowing
704
740
Pitch
enhance-
ment
702
SBR
750
705
output audio signal

Fig. 7

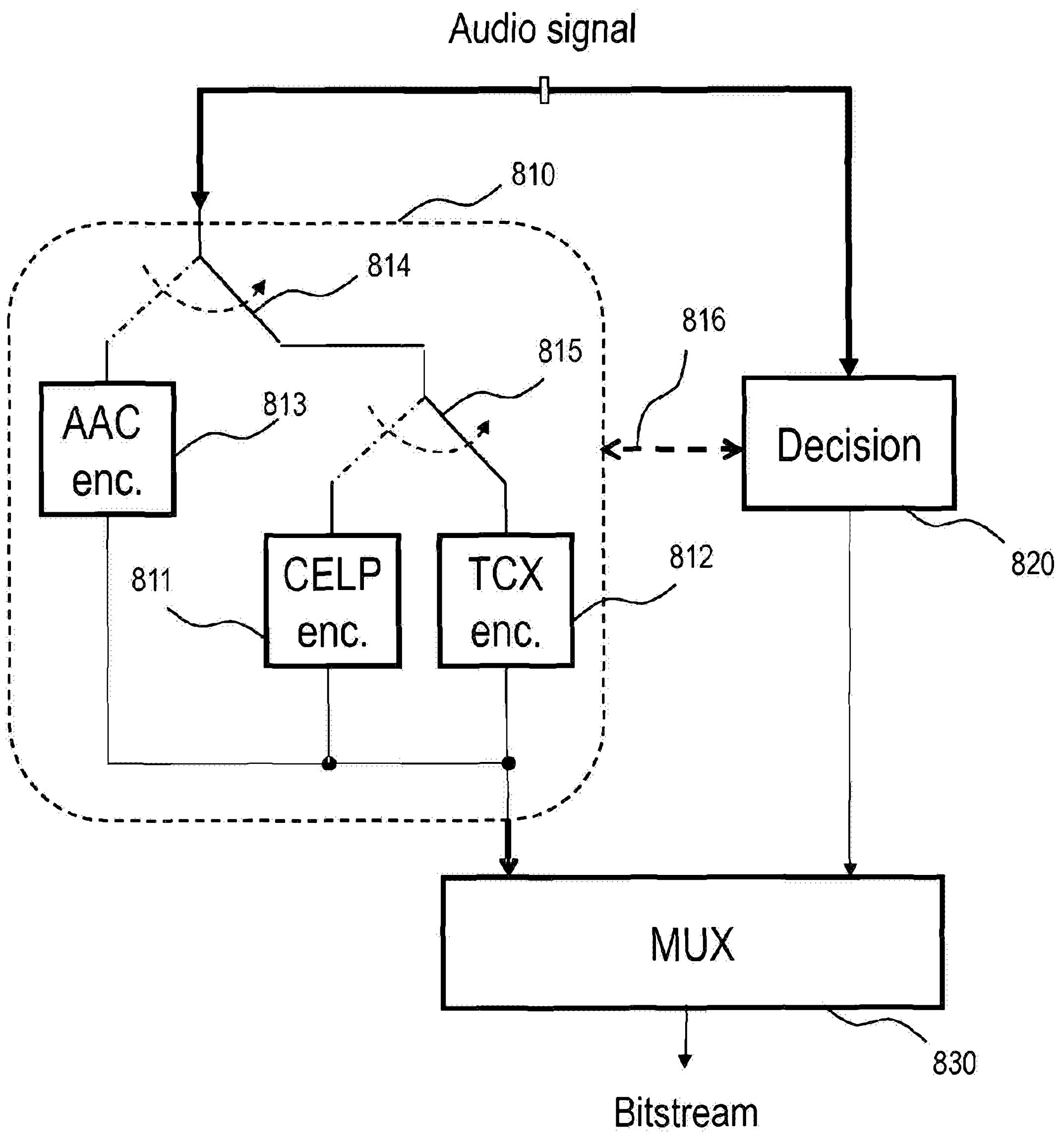
Audio signal
810
814
816
815
813
AAC
enc.
Decision
820
811
CELP
enc.
TCX
enc.
812
MUX
830
Bitstream

Fig. 8

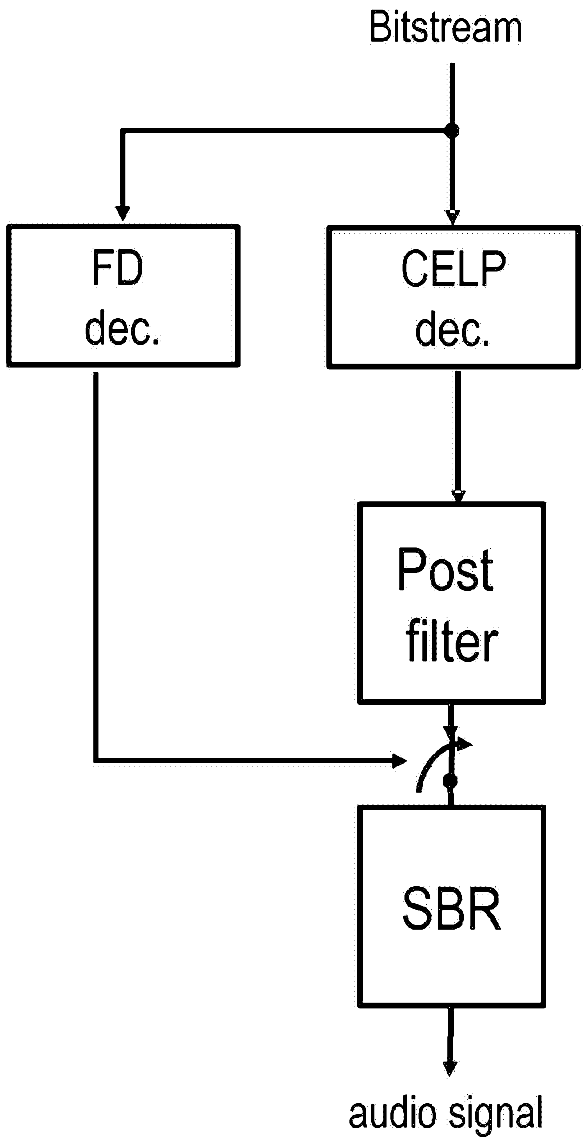

Fig. 9
(prior art)

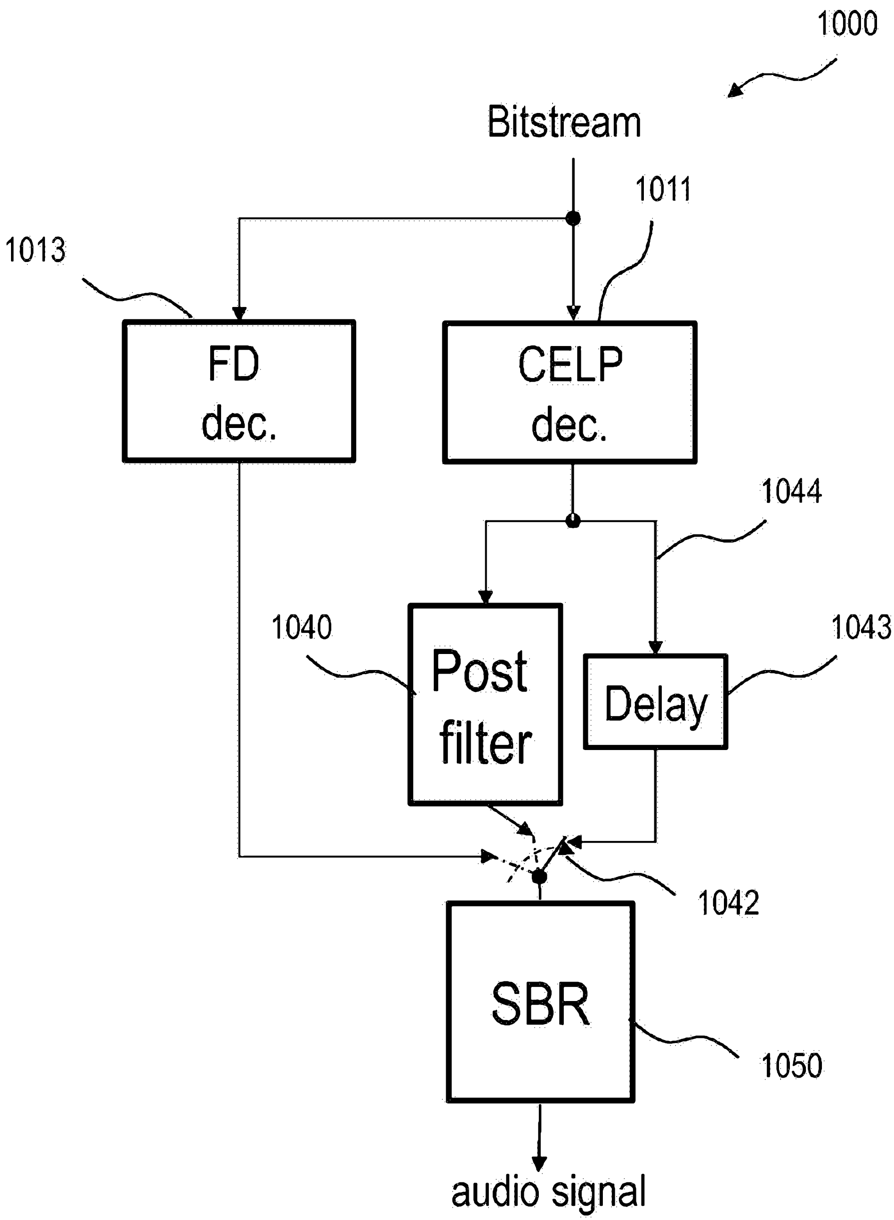
1000
Bitstream
1011
1013
FD
dec.
CELP
dec.
1044
1040
Post
filter
Delay
1043
1042
SBR
1050
audio signal

Fig. 10

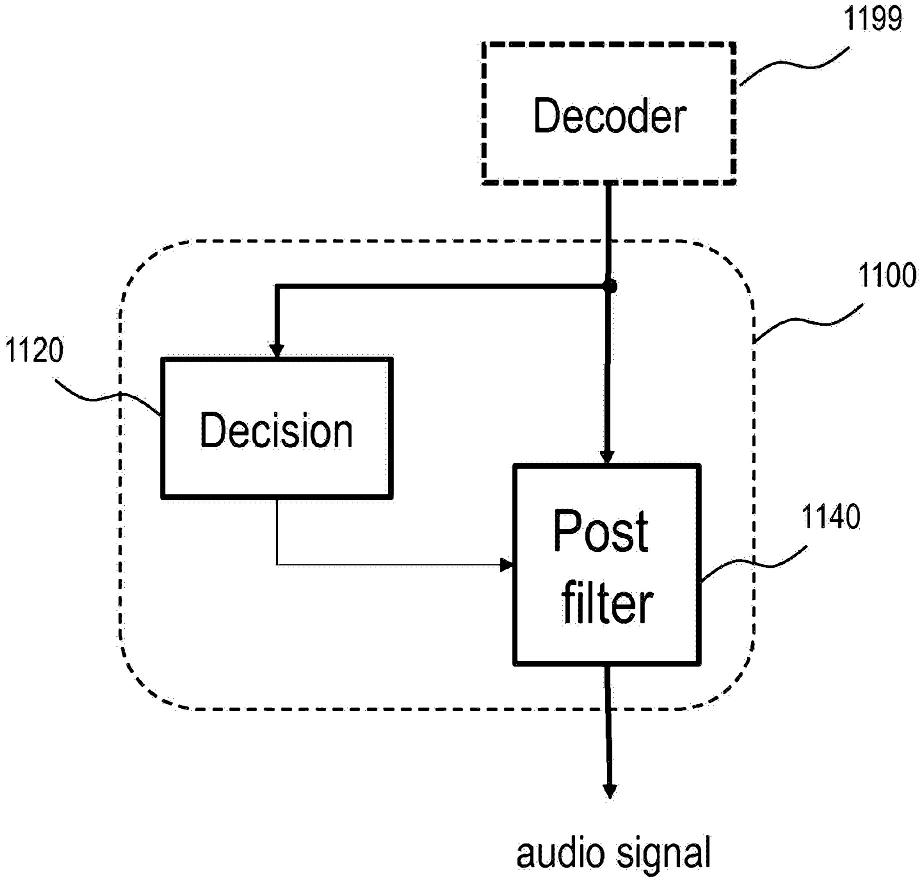
1199
Decoder
1100
1120
Decision
Post filter
1140
audio signal

Fig. 11