(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
22 September 2011 (22.09.2011)

PCT

(10) International Publication Number
WO 2011/114192 A1

(51) International Patent Classification:
G10L 19/02 (2006.01)

(21) International Application Number:
PCT/IB2010/051210

(22) International Filing Date:
19 March 2010 (19.03.2010)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant *(for all designated States except US)*: **NOKIA CORPORATION** [FI/FI]; Keilalahdentie 4, FIN-02150 Espoo (FI).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **LAAKSONEN, Lasse Juhani** [FI/FI]; Tiesenkuja 1 B 4, FIN-37120 Nokia (FI). **TAMMI, Mikko Tapio** [FI/FI]; Kohmankaari 1 C 13, FIN-33310 Tampere (FI). **VASILACHE, Adriana** [FI/FI]; Koiramäenkatu 27, FIN-33580 Tampere (FI). **RÄMÖ, Anssi Sakari** [FI/FI]; Velholehdenkatu 9, FIN-33720 Tampere (FI).

(74) Agent: **STYLE, Kelda Camilla Karen**; Page White & Farrer, Bedford House, John Street, London Greater London WC1N 2BF (GB).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

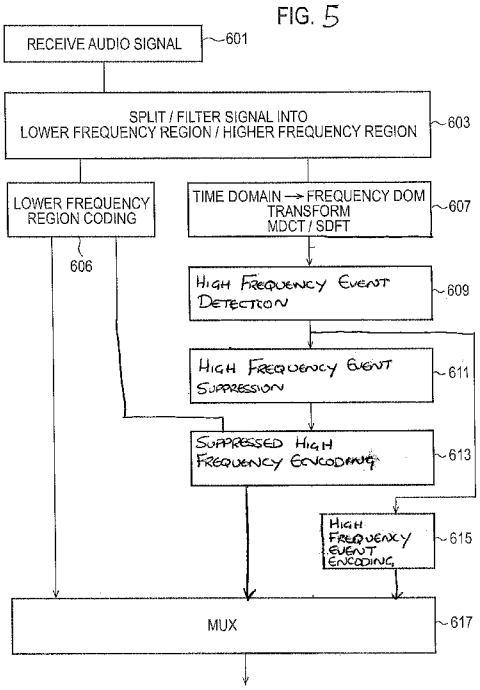
(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

(54) Title: METHOD AND APPARATUS FOR AUDIO CODING

FIG. 5



(57) Abstract: An apparatus comprising at least one processor and at least one memory including computer program code the at least one memory and the computer program code configured to, with the at least one processor, cause the apparatus at least to perform: determining at least one event from at least one audio signal, wherein the event comprises a region of frequency components of the at least one audio signal; generating a suppressed at least one audio signal by suppressing the at least one event from the at least one audio signal; and encoding at least one event from the at least one event.

WO 2011/114192 A1

Method and apparatus for audio coding

The present invention relates to coding, and in particular, but not exclusively to speech or audio coding.

Background

Audio signals, like speech or music, are encoded for example to enable efficient transmission or storage of the audio signals.

Audio encoders and decoders are used to represent audio based signals, such as music and ambient sounds (which in speech coding terms can be called background noise). These types of coders typically do not utilise a speech model for the coding process, rather they use processes for representing all types of audio signals, including speech.

Speech encoders and decoders (codecs) are usually optimised for speech signals, and can operate at either a fixed or variable bit rate.

An audio codec can also be configured to operate with varying bit rates. At lower bit rates, such an audio codec may work with speech signals at a coding rate equivalent to a pure speech codec. At higher bit rates, the audio codec may code any signal including music, background noise and speech, with higher quality and performance.

In some audio codecs the input signal is divided into a limited number of bands. Furthermore some codecs use the correlation between the low and high frequency bands or regions of an audio signal to improve the coding efficiency of the codecs.

As typically the higher frequency bands of the spectrum are generally quite similar to the lower frequency bands some codecs encode only the lower frequency bands and reproduce the upper frequency bands as a scaled lower frequency band copy. Thus by only using a small amount of additional control information considerable savings can be achieved in the total bit rate of the codec.

One such codec for coding the high frequency region is known as higher frequency region (HFR) coding. One form of higher frequency region coding is spectral-band-replication (SBR), which has been developed by Coding Technologies. In SBR, a known audio coder, such as Moving Pictures Expert Group MPEG-4 Advanced Audio Coding (AAC) or MPEG-1 Layer III (MP3) coder, codes the low frequency region. The higher frequency region is generated separately utilizing the coded low frequency region.

In SBR coding, the higher frequency region is obtained by transposing the lower frequency region to the higher frequencies. The transposition is based on a Quadrature Mirror Filters (QMF) filter bank with 32 bands and is performed such that it is predefined from which band samples each high frequency band sample is constructed. This is done independently of the characteristics of the input signal.

The higher frequency bands are modified based on additional information. The filtering is done to make particular features of the synthesized high frequency region more similar with the original one. Additional components, such as sinusoids or noise, are added to the high frequency region to increase the similarity with the original high frequency region. Finally, the envelope is adjusted to follow the envelope of the original high frequency spectrum.

Higher frequency region coding however does not produce an identical copy of the original high frequency region. Specifically, the known higher frequency region coding mechanisms perform relatively poorly where the input signal has narrow power spectral peaks, in other words does not have a spectrum with strongly noise-like characteristics.

Therefore based on the signal characteristics of a typical audio signal and the characteristics of human hearing, it has been found beneficial to exploit the dependencies between lower frequencies and higher frequencies of audio signals in audio coding. For example, after a division of a full band (20 kilohertz bandwidth) audio signal equally into two frequency regions it has been found that often the higher band is quite similar to the lower band. This is especially beneficial because the higher frequencies are not generally as sensitive to coding errors which are introduced during the compression process as the low frequency part of the signal. This suggests that a

lower bitrate (and a higher compression ratio) can be used for the high frequency content than the corresponding low frequency content and that the high frequency coding can be at least partially based on lower frequency coding. This has given rise to so-called bandwidth extension methods commonly employed in modern low rate audio coding. However, with decreasing bitrates noisy decoded signals are produced as the spectral shape of the low frequency signal does not provide as good a starting point for sufficiently good representation of the higher frequency coding.

For example an adaptive multi-rate wideband (AMR-WB) based low bitrate super wideband extension highlights limitations with regards to existing super wideband coding methods. The lower quality of the low band signal makes it harder to obtain good quality high band signals as it is more difficult to find spectral shapes with a good match for the high band when the low band is noisy. Furthermore using an adaptive multi-rate wideband based ACELP coding up to 6.4 kilohertz means that the super wideband extension requires to operate from fairly low starting frequencies which are still fairly sensitive to coding errors and as such may produce inaccurate encoding of the high band.

Summary of the Invention

This invention proceeds from the consideration that the currently proposed codecs lack flexibility with respect to being able to code efficient and accurate approximations to the signals.

Embodiments of the present invention aim to address the above problem.

There is provided according to a first aspect of the invention a method comprising: determining at least one event from at least one audio signal, wherein the event comprises a region of frequency components of the at least one audio signal; generating a suppressed at least one audio signal by suppressing the at least one event from the at least one audio signal; and encoding at least one event from the at least one event.

The method may further comprise: encoding the suppressed at least one audio signal; and multiplexing the encoded suppressed at least one audio signal with the encoded at least one event.

The method may further comprise: generating the at least one audio signal by filtering at least one full audio signal into the at least one audio signal representing a higher frequency part of the at least one full audio signal and a further at least one audio signal representing a lower frequency part of the at least one full audio signal, and wherein the encoding the suppressed at least one audio signal comprises: determining at least one part of the suppressed at least one audio signal being similar to at least one part of the further at least one audio signal; and generating an identifier for the at least one part of the further at least one audio signal.

Determining at least one event from the at least one audio signal may comprise: determining at least one activity value average for a first set of the at least one audio signal; and determining the event as a region with an activity value greater than the at least one activity value average by at least one predetermined value.

Determining the event as a region with an activity value greater than the at least one activity value average by at least one predetermined value may comprise: identifying the region with an activity value greater than a first of the activity value averages by a first predetermined value; and identifying at least part of the region with an activity value greater than a further of the activity value averages by a second predetermined value, wherein the first activity value average is the average over a first window of frequencies of the at least one audio signal and the second activity value is the average over a second window of the frequencies of the at least one audio signal.

Generating a suppressed at least one audio signal may comprise at least one of: setting at least one activity value for the region of frequency components to an average frequency components activity for frequency components encompassing and surrounding the region; setting at least one activity value for the region of frequency components a smoothed value of the activity value for the region of frequency components; and reducing the at the at least one activity value for the region of

frequency component by a value dependent on the average frequency components for frequency components encompassing the region.

Encoding at least one event from the at least one event may comprise: searching the at least one event to find the at least one event to encode; and encoding the at least one event selected.

Searching the at least one event may comprise: selecting a first event to be encoded; determining at least one search range dependent on a previously selected event; and selecting at least one further event to be encoded within the at least one search range dependent on a previously selected event.

Searching the at least one event may comprise: selecting a first event to be encoded; determining at least one masking function, each masking function surrounding and dependent on a previously selected event; and selecting at least one further event to be encoded dependent on the masking function.

Encoding at least event may comprise at least one of: identifying and encoding the amplitude, sign and location of each event region frequency component; and determining a shape and gain code book entry.

According to a second aspect of the invention there is provided an apparatus comprising at least one processor and at least one memory including computer program code the at least one memory and the computer program code configured to, with the at least one processor, cause the apparatus at least to perform: determining at least one event from at least one audio signal, wherein the event comprises a region of frequency components of the at least one audio signal; generating a suppressed at least one audio signal by suppressing the at least one event from the at least one audio signal; and encoding at least one event from the at least one event.

The apparatus may be further caused to perform: encoding the suppressed at least one audio signal; and multiplexing the encoded suppressed at least one audio signal with the encoded at least one event.

The apparatus may further be caused to perform: generating the at least one audio signal by filtering at least one full audio signal into the at least one audio signal representing a higher frequency part of the at least one full audio signal and a further at least one audio signal representing a lower frequency part of the at least one full audio signal, and wherein the encoding the suppressed at least one audio signal comprises: determining at least one part of the suppressed at least one audio signal being similar to at least one part of the further at least one audio signal; and generating an identifier for the at least one part of the further at least one audio signal.

The apparatus caused to perform determining at least one event from the at least one audio signal may be further caused to perform: determining at least one activity value average for a first set of the at least one audio signal; and determining the event as a region with an activity value greater than the at least one activity value average by at least one predetermined value.

The apparatus caused to perform determining the event as a region with an activity value greater than the at least one activity value average by at least one predetermined value, may be further caused to perform: identifying the region with an activity value greater than a first of the activity value averages by a first predetermined value; and identifying at least part of the region with an activity value greater than a further of the activity value averages by a second predetermined value, wherein the first activity value average is the average over a first window of frequencies of the at least one audio signal and the second activity value is the average over a second window of the frequencies of the at least one audio signal.

The apparatus caused to perform generating a suppressed at least one audio signal may be further caused to perform at least one of: setting at least one activity value for the region of frequency components to an average frequency components activity for frequency components encompassing and surrounding the region; setting at least one activity value for the region of frequency components a smoothed value of the activity value for the region of frequency components; and reducing the at the at least one activity value for the region of frequency component by a value dependent on the average frequency components for frequency components encompassing the region.

The apparatus caused to perform encoding at least one event from the at least one event further may be caused to perform: searching the at least one event to find the at least one event to encode; and encoding the at least one event selected.

The apparatus caused to perform searching the at least one event may be further caused to perform: selecting a first event to be encoded; determining at least one search range dependent on a previously selected event; and selecting at least one further event to be encoded within the at least one search range dependent on a previously selected event.

The apparatus caused to perform searching the at least one event may be further caused to perform: selecting a first event to be encoded; determining at least one masking function, each masking function surrounding and dependent on a previously selected event; and selecting at least one further event to be encoded dependent on the masking function.

The apparatus caused to perform encoding at least event may be further caused to perform at least one of: identifying and encoding the amplitude, sign and location of each event region frequency component; and determining a shape and gain code book entry.

According to a third aspect of the invention there is provided a method comprising: receiving at least one indicator representing at least one frequency component event from a region of frequency components; and modifying a decoded single frequency component dependent on the indicator.

The method may further comprise decoding the indicator to determine at least one of position, gain and sign of the at least one frequency component.

The method may further comprise: receiving an encoded event suppressed audio signal; and decoding the encoded event suppressed audio signal to generate decoded frequency components for the event suppressed audio signal.

According to a fourth aspect of the invention there is provided an apparatus comprising at least one processor and at least one memory including computer program code the at least one memory and the computer program code configured to, with the at least one processor, cause the apparatus at least to perform: receiving at least one indicator representing at least one frequency component event from a region of frequency components; and modifying at least one frequency component within the at least one event dependent on the indicator.

The apparatus may be further caused to decode the indicator to determine at least one of position, gain and sign of the at least one frequency component.

The apparatus may be further caused to perform: receiving an encoded event suppressed audio signal; decoding the encoded event suppressed audio signal to generate decoded frequency components for the event suppressed audio signal.

According to a fifth aspect of the invention there is provided an apparatus comprising: an event determiner configured to determine at least one event from at least one audio signal, wherein the event comprises a region of frequency components of the at least one audio signal; an event suppressor configured to generating a suppressed at least one audio signal by suppressing the at least one event from the at least one audio signal; and an event encoder configured to encode at least one event from the at least one event.

The apparatus may further comprise: a suppressed signal encoder configured to encode the suppressed at least one audio signal; and a multiplexer configured to multiplex the encoded suppressed at least one audio signal with the encoded at least one event.

The apparatus may further comprise: a quadrature mirror filter configured to generate the at least one audio signal by filtering at least one full audio signal into the at least one audio signal representing a higher frequency part of the at least one full audio signal and a further at least one audio signal representing a lower frequency part of the at least one full audio signal.

The suppressed signal encoder may further comprise: a signal identifier configured to determine at least one part of the further at least one audio signal being similar to at least one part of the suppressed at least one audio signal; and; and a identifier generator configured to generate an identifier for the at least one part of the further at least one audio signal.

The event determiner may comprise: an activity value average configured to determine at least one activity value average for a first set of the at least one audio signal; and an event selector configured to determine the event as a region with an activity value greater than the at least one activity value average by at least one predetermined value.

The event selector may comprise: a first event region detector configured to identify the region with an activity value greater than a first of the activity value averages by a first predetermined value; and a second event region detector configured to identify at least part of the region with an activity value greater than a further of the activity value averages by a second predetermined value, wherein the first activity value average is the average over a first window of frequencies of the at least one audio signal and the second activity value is the average over a second window of the frequencies of the at least one audio signal.

The event suppressor may comprise an activity value suppressor configured to set at least one activity value for the region of frequency components to an average frequency components activity for frequency components encompassing and surrounding the region;

The event suppressor may comprise an activity value suppressor configured to set at least one activity value for the region of frequency components a smoothed value of the activity value for the region of frequency components.

The event suppressor may comprise an activity value suppressor configured to reduce the at least one activity value for the region of frequency component by a value dependent on the average frequency components for frequency components encompassing the region.

The event encoder may comprise: an event candidate identifier configured to search the at least one event to find the at least one event to encode; and a selected event encoder configured to encode the at least one event selected.

The event candidate identifier may comprise: an event selector configured to select a first event to be encoded; a search range determiner configured to determine at least one search range dependent on a previously selected event; wherein the event selector is further configured to select at least one further event to be encoded within the at least one search range dependent on the previously selected event.

The event candidate identifier may comprise: an event selector configured to select a first event to be encoded; a masking function generator configured to determine at least one masking function, each masking function surrounding and dependent on a previously selected event; wherein the event selector is further configured to select at least one further event to be encoded dependent on the masking function.

The event encoder may comprise: an event identifier configured to determine the sign and location of each event region frequency component; and event coder configure to code the amplitude, sign and location of each event region frequency component.

The event encoder may comprise an event codebook identifier configured to identify from a codebook a shape and gain of the event frequency components; and an event codebook coder configured to code the shape and gain using the codebook.

According to a sixth aspect of the invention there is provided apparatus comprising: an indicator receiver configured to receive at least one indicator representing at least one frequency component event from a region of frequency components; and a signal processor configured to modify at least one frequency component within the at least one event dependent on the indicator.

The apparatus may further comprise an indicator decoder configured to decode the indicator to determine at least one of position, gain and sign of the at least one frequency component.

The apparatus may further comprise: a suppressed signal buffer configured to receive an encoded event suppressed audio signal; and a suppressed signal decoder configured to decode the encoded event suppressed audio signal to generate at least one decoded frequency component for the event suppressed audio signal.

According to a seventh aspect of the invention there is provided a computer-readable medium encoded with instructions that, when executed by a computer perform: determining at least one event from at least one audio signal, wherein the event comprises a region of frequency components of the at least one audio signal; generating a suppressed at least one audio signal by suppressing the at least one event from the at least one audio signal; and encoding at least one event from the at least one event.

According to an eighth aspect of the invention there is provided a computer-readable medium encoded with instructions that, when executed by a computer perform: receiving at least one indicator representing at least one frequency component event from a region of frequency components; and modifying at least one frequency component within the at least one event dependent on the indicator.

According to a ninth aspect of the invention there is provided an apparatus comprising: an event determining means configured to determine at least one event from at least one audio signal, wherein the event comprises a region of frequency components of the at least one audio signal; an event suppressor means configured to generating a suppressed at least one audio signal by suppressing the at least one event from the at least one audio signal; and an event encoder means configured to encode at least one event from the at least one event.

According to a tenth aspect of the invention there is provided an apparatus comprising: receiver means configured to receive at least one indicator representing at least one frequency component event from a region of frequency components; and a signal processing means configured to modify at least one frequency component within the at least one event dependent on the indicator.

An electronic device may comprise apparatus as described above.

A chipset may comprise apparatus as described above.

Brief Description of Drawings

For better understanding of the present invention, reference will now be made by way of example to the accompanying drawings in which:

Figure 1 shows schematically an electronic device employing some embodiments of the application;

Figure 2 shows schematically an audio codec system employing some embodiments of the application;

Figure 3 shows schematically an encoder part of the audio codec system shown in figure 2 according to some embodiments of the application;

Figure 4 shows a schematic view of the higher frequency region encoder portion of the encoder as shown in figure 3 according to some embodiments of the application;

Figure 5 shows a flow diagram illustrating the operation the audio encoder as shown in figures 3 and 4 according to some embodiments of the application;

Figure 6 shows a schematic view of a high frequency event detector as shown in figure 4 according to some embodiments of the application;

Figure 7 shows a flow diagram illustrating the operation of the high frequency event detector as shown in figure 4 according to some embodiments of the application;

Figure 8 shows a schematic view of a high frequency event encoder as shown in figure 4 according to some embodiments of the application;

Figure 9 shows a flow diagram illustrating the operation of the high frequency event encoder as shown in figure 4 according to some embodiments of the application;

Figure 10 shows schematically a decoder part of the audio codec system as shown in Figure 2;

Figure 11 shows a flow diagram illustrating the operation the audio encoder as shown in figure 10 according to some embodiments of the application; and

Figure 12 shows a schematic view of an example audio signal and an example audio signal with suppressed high frequency events according to some embodiments of the application.

Description of Embodiments of the Application

The following describes in more detail possible codec mechanisms for the provision of layered or scalable variable rate audio codecs. In this regard reference is first made to Figure 1 which shows a schematic block diagram of an exemplary electronic device or apparatus 10, which may incorporate a codec according to an embodiment of the invention.

The apparatus 10 may for example be a mobile terminal or user equipment of a wireless communication system. In other embodiments the apparatus 10 may be an audio-video device such as video camera, a Television (TV) receiver, audio recorder or audio player such as a mp3 recorder/player, a media recorder (also known as a mp4 recorder/player), or any computer suitable for the processing of audio signals.

The apparatus 10 in some embodiments comprises a microphone 11, which is linked via an analogue-to-digital converter (ADC) 14 to a processor 21. The processor 21 is further linked via a digital-to-analogue (DAC) converter 32 to loudspeakers 33. The processor 21 is further linked to a transceiver (RX/TX) 13, to a user interface (UI) 15 and to a memory 22.

The processor 21 may be configured to execute various program codes. The implemented program codes in some embodiments comprise an audio encoding code for encoding a lower frequency band of an audio signal and a higher frequency band of an audio signal. The implemented program codes 23 in some embodiments further comprise an audio decoding code. The implemented program codes 23 can in some embodiments be stored for example in the memory 22 for retrieval by the processor 21 whenever needed. The memory 22 could further provide a section 24 for storing data, for example data that has been encoded in accordance with the invention.

The encoding and decoding code in embodiments can be implemented in hardware or firmware.

The user interface 15 enables a user to input commands to the electronic device 10, for example via a keypad, and/or to obtain information from the electronic device 10, for

example via a display. In some embodiments a touch screen may provide both input and output functions for the user interface. The apparatus 10 in some embodiments comprises a transceiver 13 suitable for enabling communication with other apparatus, for example via a wireless communication network.

It is to be understood again that the structure of the apparatus 10 could be supplemented and varied in many ways.

A user of the apparatus 10 for example can use the microphone 11 for inputting speech or other audio signals that are to be transmitted to some other apparatus or that are to be stored in the data section 24 of the memory 22. A corresponding application in some embodiments can be activated to this end by the user via the user interface 15. This application in these embodiments can be performed by the processor 21, causes the processor 21 to execute the encoding code stored in the memory 22.

The analogue-to-digital converter (ADC) 14 in some embodiments converts the input analogue audio signal into a digital audio signal and provides the digital audio signal to the processor 21. In some embodiments the microphone 11 can comprise an integrated microphone and ADC function and provide digital audio signals directly to the processor for processing.

The processor 21 in such embodiments then process the digital audio signal in the same way as described with reference to Figures 2 to 9.

The resulting bit stream can in some embodiments be provided to the transceiver 13 for transmission to another apparatus. Alternatively, the coded audio data in some embodiments can be stored in the data section 24 of the memory 22, for instance for a later transmission or for a later presentation by the same apparatus 10.

The apparatus 10 in some embodiments can also receive a bit stream with correspondingly encoded data from another apparatus via the transceiver 13. In this example, the processor 21 may execute the decoding program code stored in the memory 22. The processor 21 in such embodiments decodes the received data, and provides the decoded data to a digital-to-analogue converter 32. The digital-to-

analogue converter 32 converts the digital decoded data into analogue audio data and can in some embodiments output the analogue audio via the loudspeakers 33. Execution of the decoding program code in some embodiments can be triggered as well by an application called by the user via the user interface 15.

The received encoded data in some embodiment can also be stored instead of an immediate presentation via the loudspeakers 33 in the data section 24 of the memory 22, for instance for later decoding and presentation or decoding and forwarding to still another apparatus.

It would be appreciated that the schematic structures described in figures 2 to 4, 6, 8, and 10 and the method steps shown in figures 5, 7, 9 and 11 represent only a part of the operation of an audio codec as exemplarily shown implemented in the electronic device shown in figure 1.

The general operation of audio codecs as employed by embodiments of the application is shown in figure 2. General audio coding/decoding systems comprise both an encoder and a decoder, as illustrated schematically in figure 2. However, it would be understood that embodiments of the application may implement one of either the encoder or decoder, or both the encoder and decoder. Illustrated by Figure 2 is a system 102 with an encoder 104, a storage or media channel 106 and a decoder 108. It would be understood that as described above some embodiments of the apparatus 10 can comprise or implement one of the encoder 104 or decoder 108 or both the encoder 104 and decoder 108.

The encoder 104 compresses an input audio signal 110 producing a bit stream 112, which in some embodiments can be stored or transmitted through a media channel 106. The bit stream 112 can be received within the decoder 108. The decoder 108 decompresses the bit stream 112 and produces an output audio signal 114. The bit rate of the bit stream 112 and the quality of the output audio signal 114 in relation to the input signal 110 are the main features which define the performance of the coding system 102.

Figure 3 shows schematically an encoder 104 according to some embodiments of the application. The encoder 104 in such embodiments comprises an input 203 arranged to receive an audio signal. The input 203 is connected to a low pass filter 230 and high pass/band pass filter 235. The low pass filter 230 furthermore outputs a signal to the lower frequency region (LFR) coder (otherwise known as the core codec) 231. The lower frequency region coder 231 is configured to output signals to the higher frequency region (HFR) coder 232. The high pass/band pass filter 235 is connected to the HFR coder 232. The LFR coder 231, and the HFR coder 232 are configured to output signals to the bitstream formatter 234 (which in some embodiments of the invention is also known as the bitstream multiplexer). The bitstream formatter 234 is configured to output the output bitstream 112 via the output 205.

In some embodiments of the invention the high pass/band pass filter 235 may be optional, and the audio signal passed directly to the HFR coder 232. In some further embodiments the operation of the low pass filter 230 and high pass filter 235 can be implemented as a quadrature mirror filter (QMF) configuration which outputs a lower frequency component to the LFR coder 231 and a higher frequency component to the HFR coder 232.

The operation of these components is described in more detail with reference to the flow chart, figure 5, showing the operation of the coder 104.

The audio signal is received by the coder 104. In some embodiments the audio signal is a digitally sampled signal. In some other embodiments the audio input may be an analogue audio signal, for example from a microphone, which is analogue to digitally (A/D) converted in the coder 104. In some further embodiments the audio input is converted from a pulse code modulation digital signal to amplitude modulation digital signal.

The receiving of the audio signal is shown in figure 5 by step 601.

The low pass filter 230 and the high pass/band pass filter 235 receive the audio signal and define a cut-off frequency about which the input signal 110 is filtered. The received audio signal frequencies below the cut-off frequency are passed by the low pass filter 230 to the lower frequency region (LFR) coder 231. The received audio signal frequencies above the cut-off frequency are passed by the high pass filter 235 to the higher frequency region

(HFR) coder 232. In some embodiments of the invention the signal is optionally down sampled in order to further improve the coding efficiency of the lower frequency region coder 231.

The splitting or filtering of the signal into lower frequency regions and higher frequency regions is shown in figure 5 by step 603.

The LFR coder 231 receives the low frequency (and optionally down sampled) audio signal and applies a suitable low frequency coding upon the signal. In a first embodiment of the invention the low frequency coder 231 applies a quantization and Huffman coding with 32 low frequency sub-bands. The input signal 110 is divided into sub-bands using an analysis filter bank structure. Each sub-band may be quantized and coded utilizing the information provided by a psychoacoustic model. The quantization settings as well as the coding scheme may be dictated by the psychoacoustic model applied. The quantized, coded information is sent to the bit stream formatter 234 for creating a bit stream 112.

Furthermore the LFR coder 231 converts the low frequency content using a modified discrete cosine transform (MDCT) to produce frequency domain realizations of synthetic LFR signal. These frequency domain realizations are passed to the HFR coder 232.

This lower frequency region coding is shown in figure 5 by step 606.

In some other embodiments other low frequency codecs may be employed in order to generate the core coding output which is output to the bitstream formatter 234. Examples of these further embodiment low frequency codecs include but are not limited to advanced audio coding (AAC), MPEG layer 3 (MP3), the ITU-T Embedded variable rate (EV-VBR) speech codec which is also known as G.718, and ITU-T G.729.1.

Where the lower frequency region coder 231 does not effectively output a frequency domain synthetic output as part of the coding process the low frequency region (LFR) coder 231 may furthermore comprise a low frequency decoder and frequency domain converter (not shown in figure 3) to generate a synthetic reproduction of the low frequency signal. These may then in embodiments of the invention be converted into frequency domain representations and, if needed, partitioned into a series of low frequency sub-

bands which are sent to the HFR coder 232.

This allows in some embodiments the choice of the lower frequency region coder 231 to be made from a wide range of possible coder/decoders and as such the embodiments are not limited to a specific low frequency or core code algorithm which produces frequency domain information as part of the output.

The higher frequency region (HFR) coder 232 is schematically shown in further detail in figure 4.

The higher frequency region coder 232 receives the signal from the high pass/band pass filter 235. In some embodiments the HFR coder 232 comprises a modified discrete cosine transform (MDCT)/shifted discrete Fourier transform (SDFT) processor 301 configured to receive the signal from the high pass/band pass filter 235 and transform a time domain signal into a frequency domain signal. It would be understood that any suitable time domain to frequency domain converter may be employed.

In some embodiments the MDCT/SDFT processor furthermore divides the higher frequency audio signal into short frequency sub-bands. These frequency sub-bands in some embodiments can be of the order of 500-800Hz wide. In some embodiments of the invention the frequency sub-bands have non-equal band-widths.

In some embodiments, the frequency sub-band bandwidth is constant, in other words does not change from frame to frame. In some other embodiments, the frequency sub-band bandwidth is not constant and a frequency sub-band may have bandwidth which changes over time.

In some embodiments, this variable frequency sub-band bandwidth allocation may be determined based on a psycho-acoustic modelling of the audio signal. These frequency sub-bands may furthermore be in various embodiments successive (in other words, one after another and producing a continuous spectral realisation) or partially overlapping.

The frequency domain output from the MDCT/SDFT transformer 301 is passed to the high frequency event detector 303, and the higher frequency event suppressor 305.

The time domain to frequency domain transformation and sub-band organisation step is shown in figure 5 by step 607.

The high frequency region coder 232 in some embodiments comprises a high frequency event detector 303. The high frequency event detector 303 in these embodiments receives the frequency components from the MDCT/SDFT 301 and outputs a series of detected high frequency events.

The operation of the high frequency event detection is shown in Figure 5 by step 609.

A high frequency event is one which within the high frequency band differs the most from otherwise noisy type content and is therefore typically audible and perceptually important.

The high frequency event detector 303 in some embodiments receives the high frequency part of the current frame of the signal in the frequency or transformed domain. This may be represented mathematically by $X_H(k)$.

The transformed domain signal used in the coding may also additionally include the lower frequencies which are represented mathematically $X_L(k)$.

Furthermore the synthesised lower frequency components may be represented mathematically by $\hat{X}_L(k)$.

For example where the input signal is sampled at 32 kHz with a 14 kHz signal bandwidth, the low frequency synthesis components $\hat{X}_L(k)$ may have a frequency range of 6.4 kilohertz (where k is divided into 256 frequency components, i.e. k = 0, 1, 2..., 255), while the high frequency part $X_H(k)$ covers frequencies between 6.4 kHz and 14 kHz (where k = 256, 257..., 559).

With respect to Figure 6, a high frequency event detector 303 is shown in further detail according to some embodiments of the application. With respect to Figure 7, the operation of the high frequency event detector according to some embodiments is shown.

The high frequency event detector 303 in some embodiments comprises a sub-band generator 311 which receives the high frequency components and from the high frequency components generates a range of sub-bands covering the complete spectral range of the high frequency band. In some embodiments there is only one sub-band, in other words the sub-band is the complete spectral range. In some other embodiments there are between 1 and N sub-bands generated by the sub-band generator 311. Each sub-band may be contiguous, in other words non-overlapping or in some embodiments the sub-bands may overlap. Furthermore each sub-band in some embodiments can be equal in frequency range or in some embodiments the range or bandwidth for each sub-and can differ. In some embodiments the sub-band generator 311 generates sub-bands according to a psychoacoustical process.

The sub-band generator outputs the sub-band components to the average amplitude generator 313 and to the metric generator 315.

In some embodiments the high frequency event detector further comprises an average amplitude generator 313. The average amplitude generator 313 in such embodiments is configured to generate an average amplitude for the analysis window (which in some embodiments comprises the length of the sub-band). The mathematical description of which can be represented as:

$$aver_i = \frac{\sum_{m=i_{i,start}}^{i_{i,start}+Len_i-1} |X_H(m)|}{Len_i}$$

where $aver_i$ is the average amplitude over the analysis window for the i^{th} sub-band $i_{i,\,start}$ is the index of the start of the i^{th} sub-band and L_{en} is the length of the i^{th} sub-band in frequency index components..

The average amplitude generator 313 in such embodiments outputs the average amplitude values $aver_i$ to the metric generator 313.

The generation of the amplitude average for each sub-band is shown in Figure 7 by step 651.

In some embodiments the high frequency event detector further comprises a metric generator 315. The metric generator 315 can be configured to perform an index search comparing the index amplitude against the amplitude average for the sub-band.

The setting up of the index search is shown in Figure 7 by step 653. Furthermore the metric generator 315 can in some embodiments generate an error metric based on the comparison. The generation of the error metric is shown in Figure 7 by step 655. The generated error metric can in some embodiments be passed to the event detector 317.

In some embodiments the high frequency event detector 303 further comprises an event detector 317. The event detector 317 compares the error metric against a predefined value to determine whether or not the index value is an event candidate.

In some embodiments the high frequency event is defined as a segment of a relatively short duration that differs considerably from the overall high frequency region, a sub-band of it, or a sliding window that is considerably longer than the event itself. The high frequency event may differ from the overall signal, for example due to it having a different local spectral shape or significantly different energy or amplitude. The high frequency event in some embodiments can be assumed to be audible and thus perceptually important. Thus the high frequency event has to be replicated quite accurately to maintain good quality. Where the bitrate does not permit accurate modelling of the high frequency event it should at least be made sure that it is not coded incorrectly in a way that introduces perceptual degradation.

In the following examples, the high frequency event is defined as a short energy peak which may be represented mathematically as a high frequency event located at sub-band index n as a single sample.

$$X_j^{EV}(n) = X_H(n)$$

Furthermore the high frequency event may be more than a single sample, for example mathematically by

$$X_j^{EV}(n) = \begin{cases} w_{-1} X_H(n-1) \\ w_0 X_H(n) \\ w_1 X_H(n+1) \end{cases}$$

where W_p are weighting factors where typically at least W_0 would be a value of 1. Furthermore in some embodiments, the high frequency event is not defined by three sample sub-bands but more than three sample sub-bands.

The operation of detecting whether or not the error metric is significant is shown in step 657.

The event detector may then store the index as the event candidate if it is significant.

The event detector may then determine whether or not the index search can be terminated. The event detector 317 when detecting that there is further searches to occur may then progress the search to the next index and request from the metric generator the next index metric value. The operation of performing a check step to determine whether or not the search is at the end, is shown in Figure 7 by step 659.

Furthermore the progression operation to determine the next index is shown in Figure 7 by step 658.

Where the event detector has determined that it is the end of the search, the event detector then compiles a list of event candidates and outputs these event candidates as detected events to the high frequency event suppressor 305 and to the high frequency event encoder 305.

The compilation of the list of event candidates is shown in Figure 7 by step 661.

In the above example, the whole high frequency band is divided into at least one sub-bands, however it would be understood that using the single sub-band, in other words the whole of the high frequency band as the analysis window the values of i = 0, $i_{i,\ start}$ = 256 and len_i = 304. In some other embodiments the sub-bands generated by the sub-band generator 311 are similar to the sub-bands used later during the suppressed high frequency encoder 307 to encode the high frequency region when the selected one or more events have been suppressed.

Furthermore in some embodiments, rather than the high frequency band X_H (k) a smoothed version can be used $\tilde{X}_H(k)$ to determine event location, size, length, etc. The use of the smoothed version is in some embodiments useful as this value can be used in some embodiments in the high frequency event suppression operations described later.

In some embodiments the metric generator 315 may generate more than a single metric for any single index or frequency component. For example in some embodiments not only is the average amplitude generator 313 configured to generate a global (with regards to the sub-band) average amplitude generator 313 but can also generate a "local" average amplitude value. For example in some embodiments global average amplitude generation may be calculated for the sub-band or the whole of the HF bandwidth whereas the local average considers a much smaller group or set of frequency components surrounding the index value being tested. This global and local metric generation has a further advantage in that it reduces the possibility of detecting local peak plateau regions as being events whereas they can be conveniently encoded due to their plateau noise behaviour.

In some embodiments the event detector 317 is configured to select a predefined minimum number of events in each sub-band. For example the event detector 317 in some embodiments does not compare the metric or metrics against a predefined value but instead orders and selects the index value with the biggest error values up to the number of events required to be detected within that sub-band.

The detection of high frequency events has a first aim to determine where such events are in order to suppress them, and a second aim to find events to then encode separately in order for them to be more accurately represented.

In some embodiments the high frequency event detector 303 can first perform a first search of the high frequency band to find a minimum number of high frequency events for encoding purposes and then perform a second search to locate a second number of high frequency events (and in some embodiments "all" high frequency events) that fulfil a predefined criteria for suppression. For each determined event, the event detector 317 in some embodiments can save a set of parameters for subsequent processing (i.e. suppression or coding) for each determined high frequency event. In some embodiments the event detector 317 stores the absolute amplitude and sign of all of the samples of the event and the positions of the event. In some other embodiments, the event detector 317 stores only the position of the event (especially where it has been predetermined that this event is not to be encoded, i.e. found on a second or "less important" search).

In some embodiments the higher frequency region coder 232 further comprises a high frequency event suppressor 305. The high frequency event suppressor 305 is configured in some embodiments to receive the high frequency domain components X_H and also receive indications for the determined high frequency events X_j^{EV}. The high frequency event suppressor 305 is thus configured to suppress or remove the high frequency event occurrences in order that the output high frequency components from the suppressor 305 are more easily encoded using conventional codecs.

In some embodiments the high frequency suppression can be performed by setting the sample index value of an event to the average value of the corresponding analysis window or sub-band $aver_i$. In some other embodiments the samples may be suppressed by using the smoothed version of the HF band signal X_H. In some other embodiments the high frequency event suppressor 305 suppress the HF events from the HF signal by reducing the corresponding amplitude of each index value determined to be a high frequency event candidate towards the average value. In some further embodiments the corresponding index high frequency event candidate amplitude may

be reduced by an amount of the average value $aver_i$. In some further embodiments the weighting factor w_{-1}, w_0, w_1 of a multiple sample form of event candidate can be used to window the suppression effect.

In some embodiments the operation of suppression can differ from event candidate to event candidate dependent on the signal characteristics. For example in a single sample event suppression the average value may be used whereas in a multiple sample event suppression, the weighting factors can be used (as well as in some embodiments the smoothed HF band signal). It would be appreciated that any suitable smoothing or suppression may be used in order to obtain a more "noise-like" sample signal. An example of the suppression of events can be seen for example with respect to Figure 12 which show an example frequency domain representation and the suppressed frequency domain representation in Figure 12 .

The suppression of the high frequency event is shown in Figure 5 by step 611. Furthermore the high frequency event suppressor outputs the suppressed high frequency samples to the suppressed high frequency encoder 307.

In some embodiments the high frequency event suppressor may perform more than one suppression operation upon the same sample group. For example in some embodiments where more than one cycle of search and suppression is carried out in a first operation a short analysis window or single sample peak search is carried out to find a suppress single sample peaks and then a further operation to find and suppress is performed where the longer segment has longer segments a milder suppression of the values then applied to the single sample values. In such samples the second operation or search may be seen to be a localised smoothing of the signal.

The suppressed high frequency encoder 307 receives the suppressed high frequency component and the low frequency region coded signal and may perform any suitable bandwidth extension encoding which is suitable for noise-like signals. For example the suppressed high frequency encoder 307 can in some embodiments use the quantised and locally synthesised low frequency signal $\hat{X}_L(k)$ as the basis for selecting and encoding the high frequency suppressed components. In such embodiments it is not

necessary to have this signal available for event suppression/coding and it may in some embodiments be possible to encode the noise part without the low frequency synthesis. In these non-synthesised embodiments the computational complexity may be lower.

Furthermore in some embodiments, for example where tonal signals are found to be common and thus where the output of the high frequency event suppression 305 is not necessarily very noise-like non-noise like encoding may be carried out such as multi-mode or tonal encoding. In some embodiments where multi-mode encoding is used, at least one of the modes is suitable for encoding noise-like characteristics and another mode suitable for non-noise type or tonal characteristics of the signal. For example tonal encoding operations are described in detail in Nokia published patent application WO2009/059633.

The output of the suppressed high frequency encoder signal is passed to the multiplexer 311.

The high frequency event encoder 309 receives the high frequency event detection candidates from the high frequency event detector 303. The high frequency event encoder 309 can in some embodiments select a sub-set of the event candidates and encodes these sub-set events for transmission or storage.

With respect to Figures 8 and 9, the structure of suitable high frequency event encoder 309 and operation of the high frequency event encoder according to some embodiments is described.

In some embodiments the high frequency event encoder 309 comprises an event selector 321 configured to receive the candidate high frequency event parameters and is configured to attempt to select event candidates for encoding. In some embodiments the event selector 321 performs a recursive search which aims to maximise at each stage the perceptual importance of the combination of the selected high frequency

events. This may be expressed mathematically as $X_k^{EV,SEL} = \max\left(\sum_{j=j_{first}}^{j_{last}} w_{j,MASK} X_j^{EV} \right)$ where k is the current high frequency event index and j is greater than k.

The perceptual importance can in some embodiments be estimated by considering at least one of the amplitudes of the events, the amplitude of the events in relation to the local noise level, and/or the amplitude of the events when taking into account also their position. Thus in some embodiments the perceptual importance selection can be implemented as a type of psychoacoustical selection where the position of a single high frequency event creates a mask around the index making close-by events less important. In some embodiments the effect of this mask may be denoted mathematically by $w_{j, MASK}$. In these embodiments the relative importance of an events amplitude is decreased as a function of the closeness to the pre-selected events (j, j-1, ..., 0).

In some other embodiments the HF events themselves may be provided with a positive bias which are further away from each other such that as wide a frequency range as possible is covered by the selected HF events. In these embodiments the bias may be implemented as a weight factor larger than 1 which increases the further apart two events are. In these embodiments the weighting factor may be used in common with or instead of the masking weighting factor described above.

In such embodiments the selection may limit the recursive searching to only a few high frequency event candidates for encoding at any recursive step to limit the computational complexity. Furthermore in some embodiments where high frequency events have been found and suppressed very close to each other (for example in adjacent sample index values), the encoding event selector 321 skips such samples when selecting the search range for the encoded events.

In some embodiments the encoding event selector 321 divides the high frequency bands into sub-bands for the search and suppression step as well as the high frequency event encoding step.

As can be seen in Figure 8 for example, the encoding event selector 321 selects a predefined minimum number of suppressed HF events for each sub-band (or performs a recursive search such as described above for each sub-band). By using sub-bands rather than the whole frequency range to encode events there is a greater probability that the whole of the high frequency band is covered.

The output of the encoding event value selected is shown in Figure 9 by step 673.

The selected events are passed to the event classifier 323. The event classifier 323 can in some embodiments determine whether or not the event is a single sample or multiple sample event and thus provide the event encoder quantizer 325 with information enabling a more optimal encoding of the event value.

The event encoder/quantizer 325 having received the selected event candidates and the classification values of each may then represent each event in an encoded form.

For example where the HF event has been determined to be a single sample, in some embodiments the event encoder may represent the event as an amplitude, sign and location. The amplitude values in some embodiments for example can be scalar quantized or in some other embodiments vector quantized efficiently. The sign of single samples can be transmitted using a single bit for each event, wherein the location information may be similarly encoded using any suitable encoding mechanism.

Where the event encoder has received a selected event which is classified as a multiple or group event, a similar encoding of each sample can be employed.

However, in some embodiments the event encoder/quantizer 325 when indicated that a multiple sample event is to be encoded can use a shape and gain codebook to select an entry from the codebook with a representation of sign and relative amplitude of the samples to shape the event and a overall gain factor to scale the whole event. In some embodiments the sign of the candidate can be considered in the selection process for coding the high frequency event by giving more weight in selection to those events that can be efficiently represented by a single sign. The encoding of selected events dependent on classification is shown in Figure 9 by step 677. These encoded events

are passed to be comprised with the suppressed high frequency encoded bits to form an encoded higher frequency encoded data stream and passed to the bitstream formatter 232.

The bitstream formatter 234 receives the lower frequency coder 231 output, the higher frequency region processor 232 output and formats the bitstream to produce the bitstream output. The bitstream formatter 234 in some embodiments of the invention may interleave the received inputs and may generate error detecting and error correcting codes to be inserted into the bitstream output 112.

The step of multiplexing the HFR coder 232 and LFR coder 231 information into the output bitstream is shown in figure 5 by step 617.

To further assist the understanding of the invention the operation of the decoder 108 with respect to the some embodiments is shown with respect to the decoder schematically shown in figure 10 and the flow chart showing the operation of the decoder in figure 11.

The decoder comprises an input 413 from which the encoded bitstream 112 may be received. The decoder furthermore in some embodiments comprises a bitstream unpacker 401 configured to receive the input 413.

The bitstream unpacker 401 in such embodiments demultiplexes, partitions, or unpacks the encoded bitstream 112 into three separate bitstreams. The low frequency encoded bitstream is in these embodiments passed to a lower frequency region decoder 403, the high frequency event encoded bitstream is passed to a high frequency event decoder 405 and the encoded suppressed high frequency bitstream is passed to a high frequency region decoder 407.

This unpacking process is shown in figure 11 by step 701.

The lower frequency region decoder 403 receives the low frequency encoded data and constructs a synthesized low frequency signal by performing the inverse process to

that performed in the lower frequency region coder 231. This synthesized low frequency signal is passed to the higher frequency region decoder 407.

This lower frequency region decoding process is shown in figure 11 by step 707.

The decoder in some embodiments comprises a high frequency event decoder 405 which receives encoded event information from the bitstream unpacker 401 and generates event reconstruction data which may be passed to the higher frequency region decoder 407. The high frequency event decoder 405 in these embodiments can carry out the inverse operation to the high frequency event encoder. For example in some embodiments the high frequency event decoder 405 generate a position, orientation, amplitude information for single index events.

The decoding of the data is shown in figure 11 by step 703.

In some embodiments the high frequency event decoder controller 405 can be implemented as part of the higher frequency region decoder 407.

The HFR decoder 407 in some embodiments receives the suppressed event high frequency encoded data, the lower frequency region decoded data and the event indications and carries out a replicant HFR reconstruction operation.

The HFR decoder 407 in these embodiments performs the inverse to the suppressed high frequency encoder 307. For example the HFR decoder in some embodiments replicates and scales the low frequency components from the synthesized low frequency signal as indicated by the high frequency reconstruction bitstream in terms of the bands indicated by the band selection information.

This high frequency suppressed replica construction is shown in figure 11 by step 705.

The HFR decoder 407 furthermore in some embodiments carries out an event decoding and injection operation. Thus in other words the HFR decoder performs the inverse operation to the high frequency event suppressor 305.

The injection of high frequency events is shown in figure 11 by step 709.

The reconstructed high frequency component bitstream is passed to the reconstruction decoder 409.

The reconstruction decoder 409 receives the decoded low frequency bitstream and the reconstructed high frequency bitstream to form a bitstream representing the original signal and outputs the output audio signal 114 on the decoder output 415.

This reconstruction of the signal is shown in figure 11 by step 711.

The embodiments of the invention described above describe the codec in terms of separate encoders 104 and decoders 108 apparatus in order to assist the understanding of the processes involved. However, it would be appreciated that the apparatus, structures and operations may be implemented as a single encoder-decoder apparatus/structure/operation. Furthermore in some embodiments of the invention the coder and decoder may share some/or all common elements.

Although the above examples describe embodiments of the invention operating within a codec within an apparatus 10, it would be appreciated that the invention as described below may be implemented as part of any audio (or speech) codec, including any variable rate/adaptive rate audio (or speech) codec. Thus, for example, embodiments of the invention may be implemented in an audio codec which may implement audio coding over fixed or wired communication paths.

Thus user equipment may comprise an audio codec such as those described in embodiments of the invention above.

It shall be appreciated that the term user equipment is intended to cover any suitable type of wireless user equipment, such as mobile telephones, portable data processing devices or portable web browsers.

Furthermore elements of a public land mobile network (PLMN) may also comprise audio codecs as described above.

In general, the various embodiments of the invention may be implemented in hardware or special purpose circuits, software, logic or any combination thereof. For example, some aspects may be implemented in hardware, while other aspects may be implemented in firmware or software which may be executed by a controller, microprocessor or other computing device, although the invention is not limited thereto. While various aspects of the invention may be illustrated and described as block diagrams, flow charts, or using some other pictorial representation, it is well understood that these blocks, apparatus, systems, techniques or methods described herein may be implemented in, as non-limiting examples, hardware, software, firmware, special purpose circuits or logic, general purpose hardware or controller or other computing devices, or some combination thereof.

Thus at least some embodiments of the encoder may be an apparatus comprising at least one processor and at least one memory including computer program code the at least one memory and the computer program code configured to, with the at least one processor, cause the apparatus at least to perform: determining at least one event from at least one audio signal, wherein the event comprises a region of frequency components of the at least one audio signal; generating a suppressed at least one audio signal by suppressing the at least one event from the at least one audio signal; and encoding at least one event from the at least one event.

In some embodiments of the decoder there may be an apparatus comprising at least one processor and at least one memory including computer program code the at least one memory and the computer program code configured to, with the at least one processor, cause the apparatus at least to perform: receiving at least one indicator representing at least one frequency component event from a region of frequency components; and modifying at least one frequency component within the at least one event dependent on the indicator.

The embodiments of this invention may be implemented by computer software executable by a data processor of the mobile device, such as in the processor entity, or by hardware, or by a combination of software and hardware. Further in this regard it should be noted that any blocks of the logic flow as in the Figures may represent

program steps, or interconnected logic circuits, blocks and functions, or a combination of program steps and logic circuits, blocks and functions.

Thus at least some embodiments of the encoder may be a computer-readable medium encoded with instructions that, when executed by a computer perform: determining at least one event from at least one audio signal, wherein the event comprises a region of frequency components of the at least one audio signal; generating a suppressed at least one audio signal by suppressing the at least one event from the at least one audio signal; and encoding at least one event from the at least one event.

Furthermore at least some of the embodiments of the decoder may be provided a computer-readable medium encoded with instructions that, when executed by a computer perform: receiving at least one indicator representing at least one frequency component event from a region of frequency components; and modifying at least one frequency component within the at least one event dependent on the indicator.

The memory may be of any type suitable to the local technical environment and may be implemented using any suitable data storage technology, such as semiconductor-based memory devices, magnetic memory devices and systems, optical memory devices and systems, fixed memory and removable memory. The data processors may be of any type suitable to the local technical environment, and may include one or more of general purpose computers, special purpose computers, microprocessors, digital signal processors (DSPs), application specific integrated circuits (ASIC), gate level circuits and processors based on multi-core processor architecture, as non-limiting examples.

Embodiments of the inventions may be practiced in various components such as integrated circuit modules. The design of integrated circuits is by and large a highly automated process. Complex and powerful software tools are available for converting a logic level design into a semiconductor circuit design ready to be etched and formed on a semiconductor substrate.

Programs, such as those provided by Synopsys, Inc. of Mountain View, California and Cadence Design, of San Jose, California automatically route conductors and locate

components on a semiconductor chip using well established rules of design as well as libraries of pre-stored design modules. Once the design for a semiconductor circuit has been completed, the resultant design, in a standardized electronic format (e.g., Opus, GDSII, or the like) may be transmitted to a semiconductor fabrication facility or "fab" for fabrication.

As used in this application, the term 'circuitry' refers to all of the following:

(a) hardware-only circuit implementations (such as implementations in only analog and/or digital circuitry) and

(b) to combinations of circuits and software (and/or firmware), such as: (i) to a combination of processor(s) or (ii) to portions of processor(s)/software (including digital signal processor(s)), software, and memory(ies) that work together to cause an apparatus, such as a mobile phone or server, to perform various functions and

(c) to circuits, such as a microprocessor(s) or a portion of a microprocessor(s), that require software or firmware for operation, even if the software or firmware is not physically present.

This definition of 'circuitry' applies to all uses of this term in this application, including any claims. As a further example, as used in this application, the term 'circuitry' would also cover an implementation of merely a processor (or multiple processors) or portion of a processor and its (or their) accompanying software and/or firmware. The term 'circuitry' would also cover, for example and if applicable to the particular claim element, a baseband integrated circuit or applications processor integrated circuit for a mobile phone or similar integrated circuit in server, a cellular network device, or other network device.

The foregoing description has provided by way of exemplary and non-limiting examples a full and informative description of the exemplary embodiment of this invention. However, various modifications and adaptations may become apparent to those skilled in the relevant arts in view of the foregoing description, when read in conjunction with the accompanying drawings and the appended claims. However, all such and similar modifications of the teachings of this invention will still fall within the scope of this invention as defined in the appended claims.

Claims:

1. A method comprising:

determining at least one event from at least one audio signal, wherein the event comprises a region of frequency components of the at least one audio signal;

generating a suppressed at least one audio signal by suppressing the at least one event from the at least one audio signal; and

encoding at least one event from the at least one event.

2. The method as claimed in claim 1, further comprising:

encoding the suppressed at least one audio signal; and

multiplexing the encoded suppressed at least one audio signal with the encoded at least one event.

3. The method as claimed in claims 1 and 2, further comprising:

generating the at least one audio signal by filtering at least one full audio signal into the at least one audio signal representing a higher frequency part of the at least one full audio signal and a further at least one audio signal representing a lower frequency part of the at least one full audio signal, and wherein the encoding the suppressed at least one audio signal comprises:

determining at least one part of the suppressed at least one audio signal being similar to at least one part of the further at least one audio signal; and

generating an identifier for the at least one part of the further at least one audio signal.

4. The method as claimed in claims 1 to 3, wherein determining at least one event from the at least one audio signal comprises:

determining at least one activity value average for a first set of the at least one audio signal;

determining the event as a region with an activity value greater than the at least one activity value average by at least one predetermined value.

5. The method as claimed in claim 4, wherein determining the event as a region with an activity value greater than the at least one activity value average by at least one predetermined value comprises:

identifying the region with an activity value greater than a first of the activity value averages by a first predetermined value; and

identifying at least part of the region with an activity value greater than a further of the activity value averages by a second predetermined value, wherein the first activity value average is the average over a first window of frequencies of the at least one audio signal and the second activity value is the average over a second window of the frequencies of the at least one audio signal.

6. The method as claimed in claims 1 to 5, wherein generating a suppressed at least one audio signal comprises at least one of:

setting at least one activity value for the region of frequency components to an average frequency components activity for frequency components encompassing and surrounding the region;

setting at least one activity value for the region of frequency components a smoothed value of the activity value for the region of frequency components; and

reducing the at the at least one activity value for the region of frequency component by a value dependent on the average frequency components for frequency components encompassing the region.

7. The method as claimed in claims 1 to 6, wherein encoding at least one event from the at least one event comprises:

searching the at least one event to find the at least one event to encode; and

encoding the at least one event selected.

8. The method as claimed in claim 7, wherein searching the at least one event comprises:

selecting a first event to be encoded;

determining at least one search range dependent on a previously selected event; and

selecting at least one further event to be encoded within the at least one search range dependent on a previously selected event.

9. The method as claimed in claim 7, wherein searching the at least one event comprises:

selecting a first event to be encoded;

determining at least one masking function, each masking function surrounding and dependent on a previously selected event; and

selecting at least one further event to be encoded dependent on the masking function.

10. The method as claimed in claims 1 to 9, wherein encoding at least event comprises at least one of:

identifying and encoding the amplitude, sign and location of each event region frequency component; and

determining a shape and gain code book entry.

11. An apparatus comprising at least one processor and at least one memory including computer program code the at least one memory and the computer program code configured to, with the at least one processor, cause the apparatus at least to perform:

determining at least one event from at least one audio signal, wherein the event comprises a region of frequency components of the at least one audio signal;

generating a suppressed at least one audio signal by suppressing the at least one event from the at least one audio signal; and

encoding at least one event from the at least one event.

12. The apparatus as claimed in claim 11, further caused to perform:

encoding the suppressed at least one audio signal; and

multiplexing the encoded suppressed at least one audio signal with the encoded at least one event.

13. The apparatus as claimed in claims 11 and 12, further caused to perform:

generating the at least one audio signal by filtering at least one full audio signal into the at least one audio signal representing a higher frequency part of the at least one full audio signal and a further at least one audio signal representing a lower

frequency part of the at least one full audio signal, and wherein the encoding the suppressed at least one audio signal comprises:

determining at least one part of the suppressed at least one audio signal being similar to at least one part of the further at least one audio signal; and

generating an identifier for the at least one part of the further at least one audio signal.

14. The apparatus as claimed in claims 11 to 13, caused to perform determining at least one event from the at least one audio signal, further caused to perform:

determining at least one activity value average for a first set of the at least one audio signal;

determining the event as a region with an activity value greater than the at least one activity value average by at least one predetermined value.

15. The apparatus as claimed in claim 14, caused to perform determining the event as a region with an activity value greater than the at least one activity value average by at least one predetermined value, further caused to perform:

identifying the region with an activity value greater than a first of the activity value averages by a first predetermined value; and

identifying at least part of the region with an activity value greater than a further of the activity value averages by a second predetermined value, wherein the first activity value average is the average over a first window of frequencies of the at least one audio signal and the second activity value is the average over a second window of the frequencies of the at least one audio signal.

16. The apparatus as claimed in claims 11 to 15, caused to perform generating a suppressed at least one audio signal further caused to perform at least one of:

setting at least one activity value for the region of frequency components to an average frequency components activity for frequency components encompassing and surrounding the region;

setting at least one activity value for the region of frequency components a smoothed value of the activity value for the region of frequency components; and

reducing the at the at least one activity value for the region of frequency component by a value dependent on the average frequency components for frequency components encompassing the region.

17. The apparatus as claimed in claims 11 to 16, caused to perform encoding at least one event from the at least one event further caused to perform:

searching the at least one event to find the at least one event to encode; and

encoding the at least one event selected.

18. The apparatus as claimed in claim 17, caused to perform searching the at least one event further caused to perform:

selecting a first event to be encoded;

determining at least one search range dependent on a previously selected event; and

selecting at least one further event to be encoded within the at least one search range dependent on a previously selected event.

19. The apparatus as claimed in claim 17, caused to perform searching the at least one event further caused to perform:

selecting a first event to be encoded;

determining at least one masking function, each masking function surrounding and dependent on a previously selected event; and

selecting at least one further event to be encoded dependent on the masking function.

20. The apparatus as claimed in claims 11 to 19, caused to perform encoding at least event further caused to perform at least one of:

identifying and encoding the amplitude, sign and location of each event region frequency component; and

determining a shape and gain code book entry.

21. A method comprising:

receiving at least one indicator representing at least one frequency component event from a region of frequency components; and

modifying at least one frequency component within the at least one event dependent on the indicator.

22. The method as claimed in claim 21, further comprising decoding the indicator to determine at least one of position, gain and sign of the at least one frequency component.

23. The method as claimed in claims 21 and 22, further comprising:

receiving an encoded event suppressed audio signal;

decoding the encoded event suppressed audio signal to generate decoded frequency components for the event suppressed audio signal.

24. An apparatus comprising at least one processor and at least one memory including computer program code the at least one memory and the computer program code configured to, with the at least one processor, cause the apparatus at least to perform:

receiving at least one indicator representing at least one frequency component event from a region of frequency components; and

modifying at least one frequency component within the at least one event dependent on the indicator.

25. The apparatus as claimed in claim 24, further caused to decode the indicator to determine at least one of position, gain and sign of the at least one frequency component.

26. The apparatus as claimed in claims 24 and 25, further caused to perform:

receiving an encoded event suppressed audio signal;

decoding the encoded event suppressed audio signal to generate decoded frequency components for the event suppressed audio signal.

FIG. 1

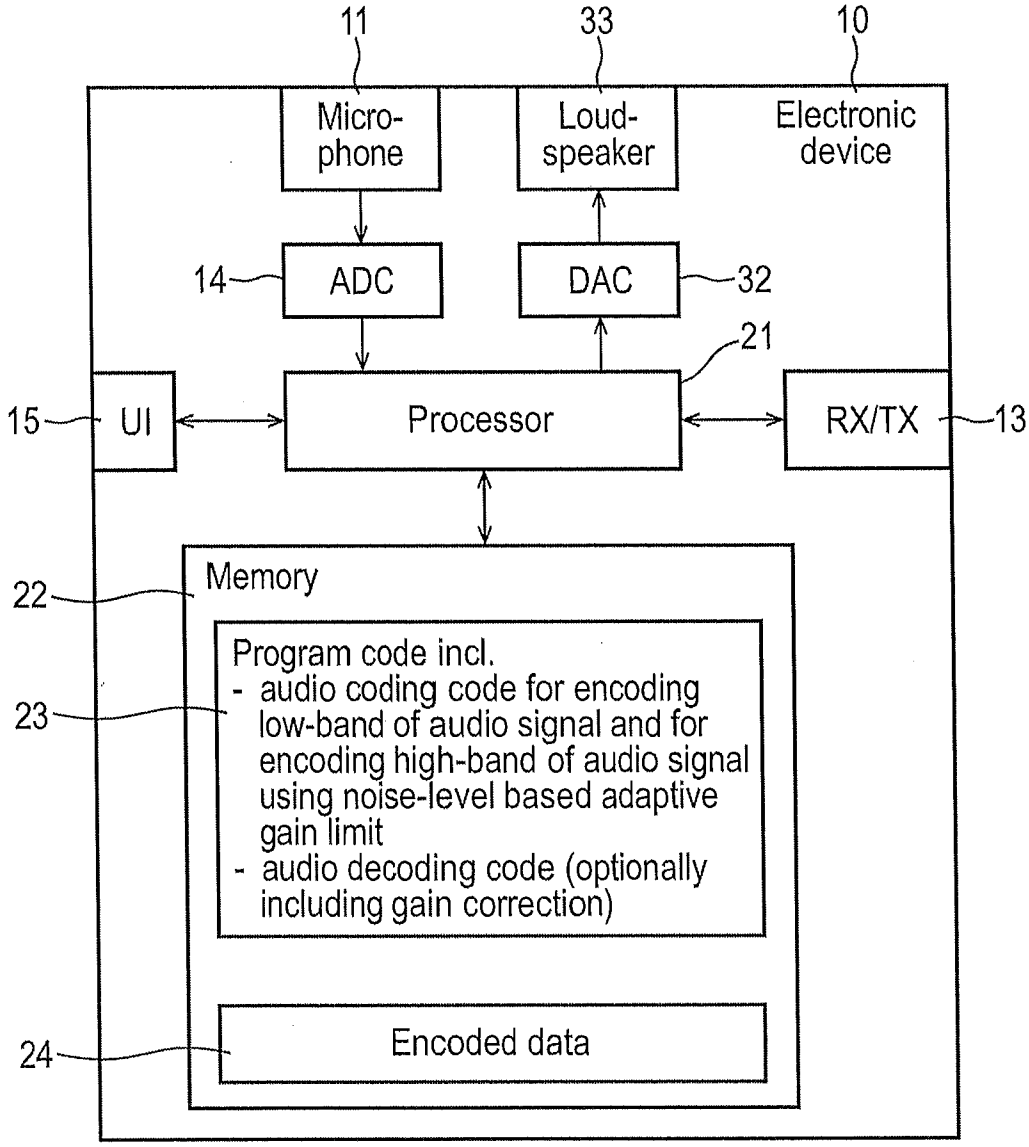

11
33
10
Micro-phone
Loud-speaker
Electronic device
14
ADC
DAC
32
21
15
UI
Processor
RX/TX
13
22
Memory
Program code incl.
- audio coding code for encoding low-band of audio signal and for encoding high-band of audio signal using noise-level based adaptive gain limit
- audio decoding code (optionally including gain correction)
23
Encoded data
24

FIG. 2

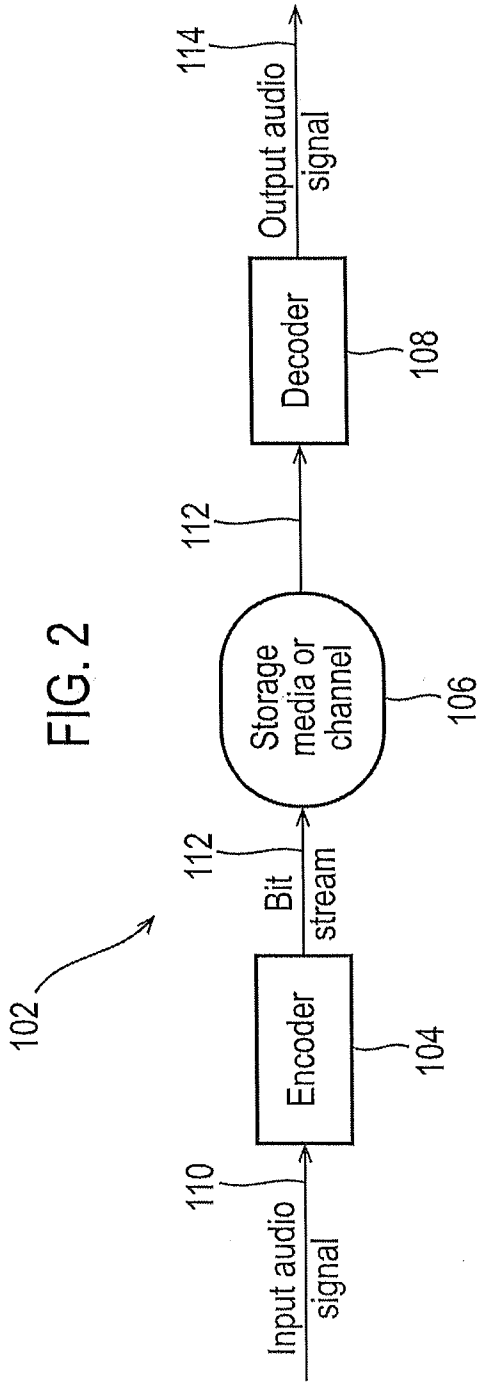

102
Input audio signal
110
Encoder
104
Bit stream
112
Storage media or channel
106
112
Decoder
108
Output audio signal
114

FIG. 3

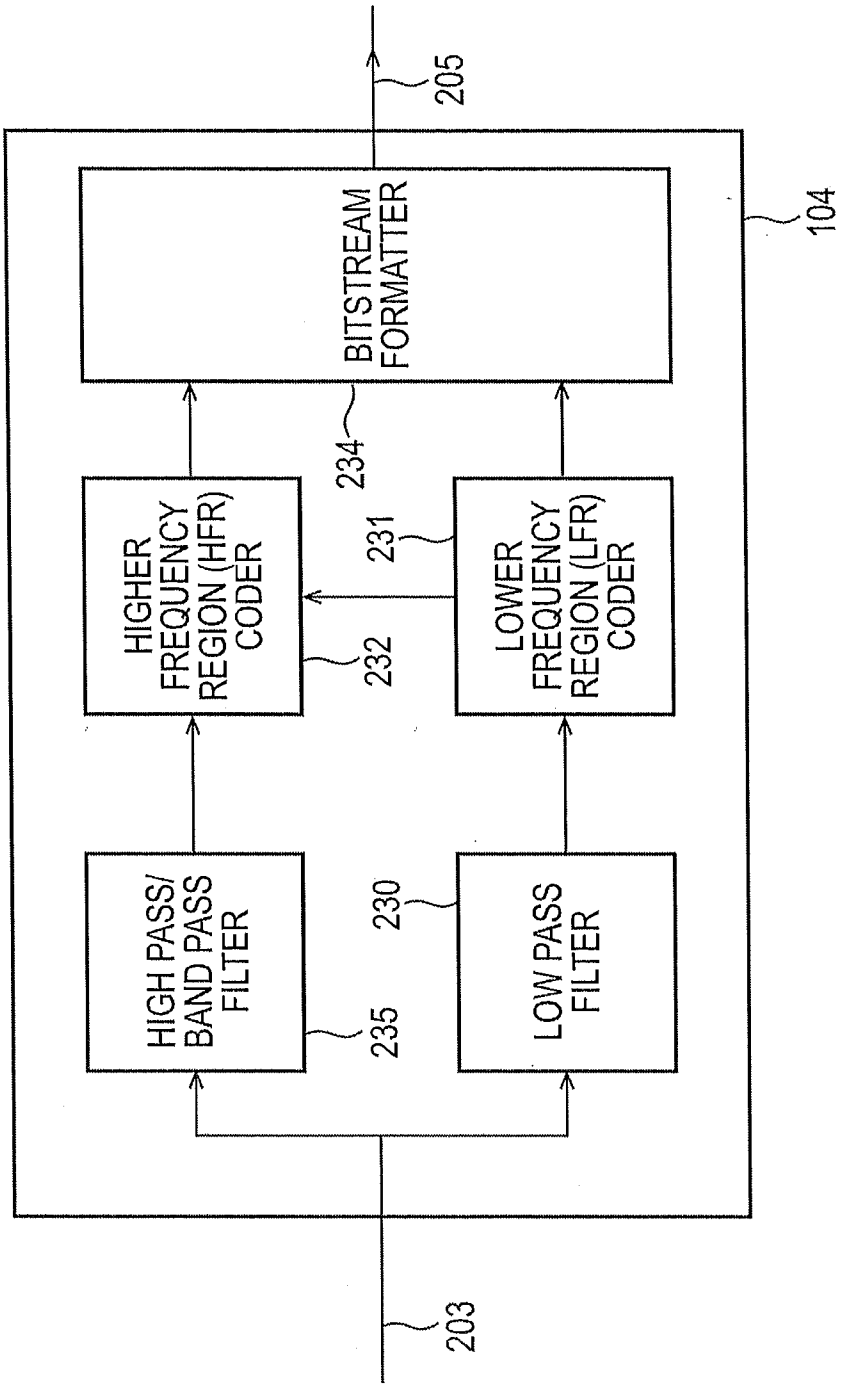

205
104
BITSTREAM FORMATTER
234
HIGHER FREQUENCY REGION (HFR) CODER
231
LOWER FREQUENCY REGION (LFR) CODER
232
HIGH PASS/ BAND PASS FILTER
230
LOW PASS FILTER
235
203

FIG. 4

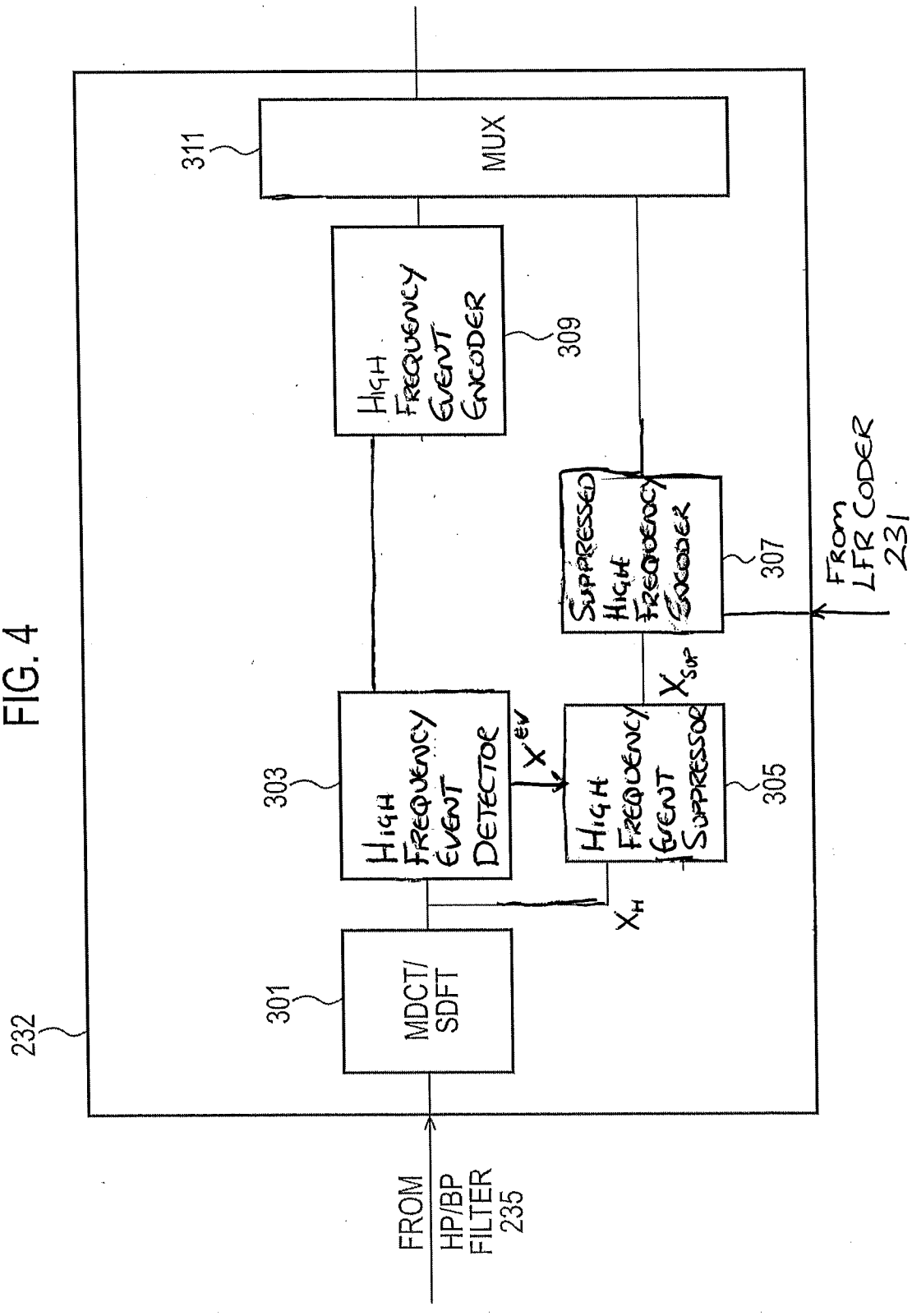
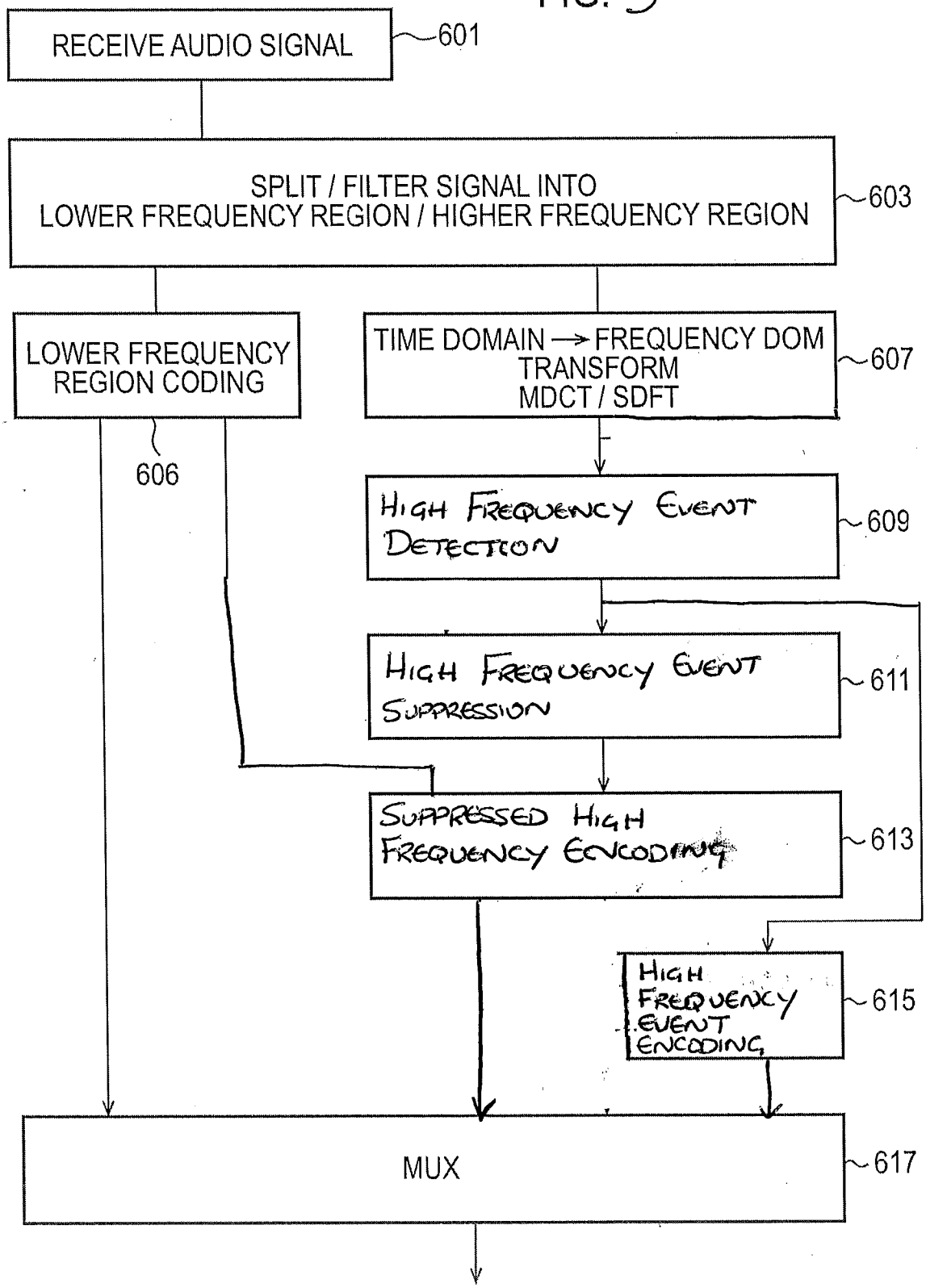

232
301
MDCT/ SDFT
303
HIGH FREQUENCY EVENT DETECTOR
X^{ev}
HIGH FREQUENCY EVENT SUPPRESSOR
305
X_H
X_{sup}
SUPPRESSED HIGH FREQUENCY ENCODER
307
FROM LFR CODER 231
HIGH FREQUENCY EVENT ENCODER
309
311
MUX
FROM HP/BP FILTER 235

FIG. 5


RECEIVE AUDIO SIGNAL
601
SPLIT / FILTER SIGNAL INTO LOWER FREQUENCY REGION / HIGHER FREQUENCY REGION
603
LOWER FREQUENCY REGION CODING
606
TIME DOMAIN → FREQUENCY DOM TRANSFORM MDCT / SDFT
607
HIGH FREQUENCY EVENT DETECTION
609
HIGH FREQUENCY EVENT SUPPRESSION
611
SUPPRESSED HIGH FREQUENCY ENCODING
613
HIGH FREQUENCY EVENT ENCODING
615
MUX
617

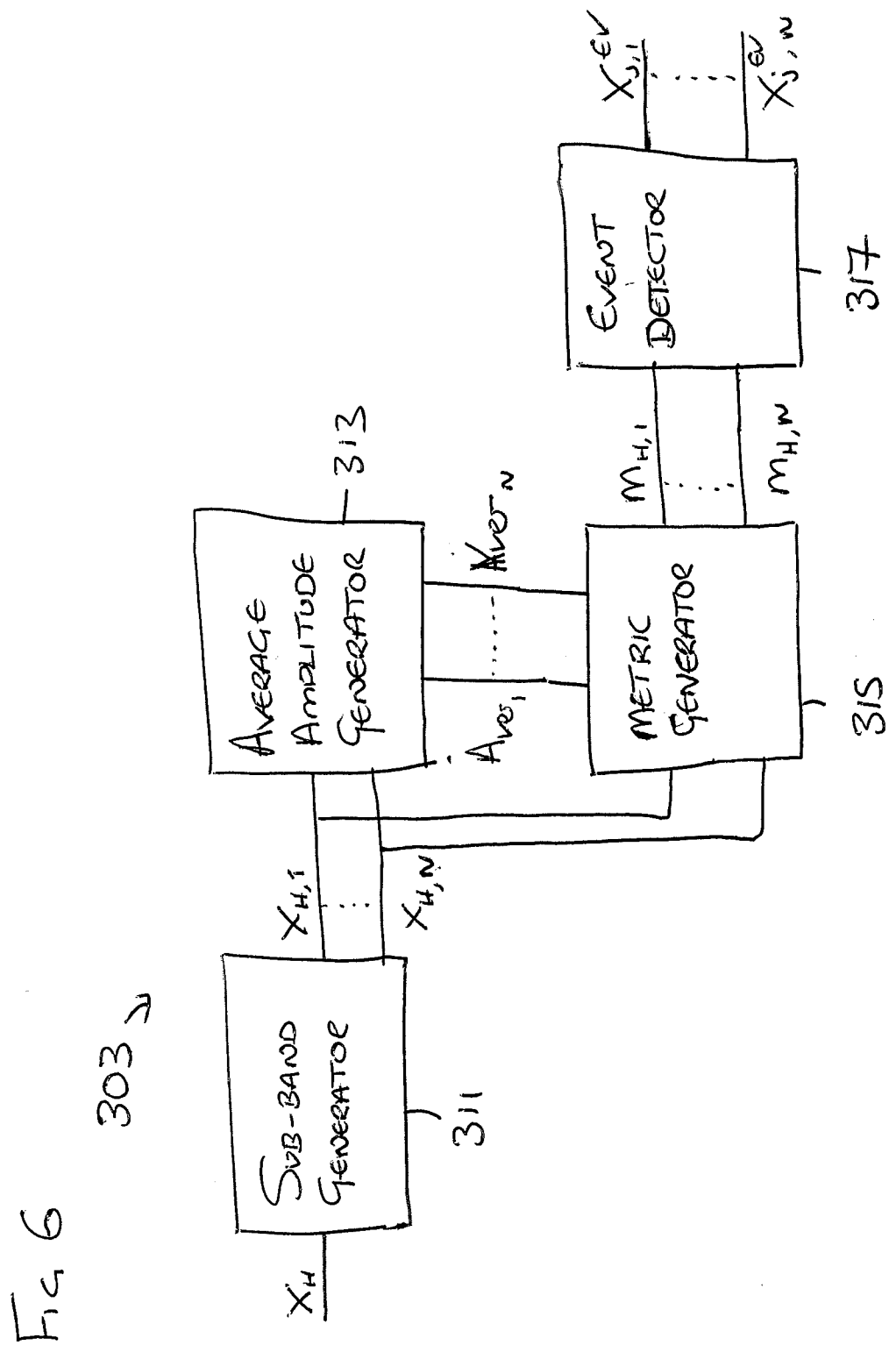
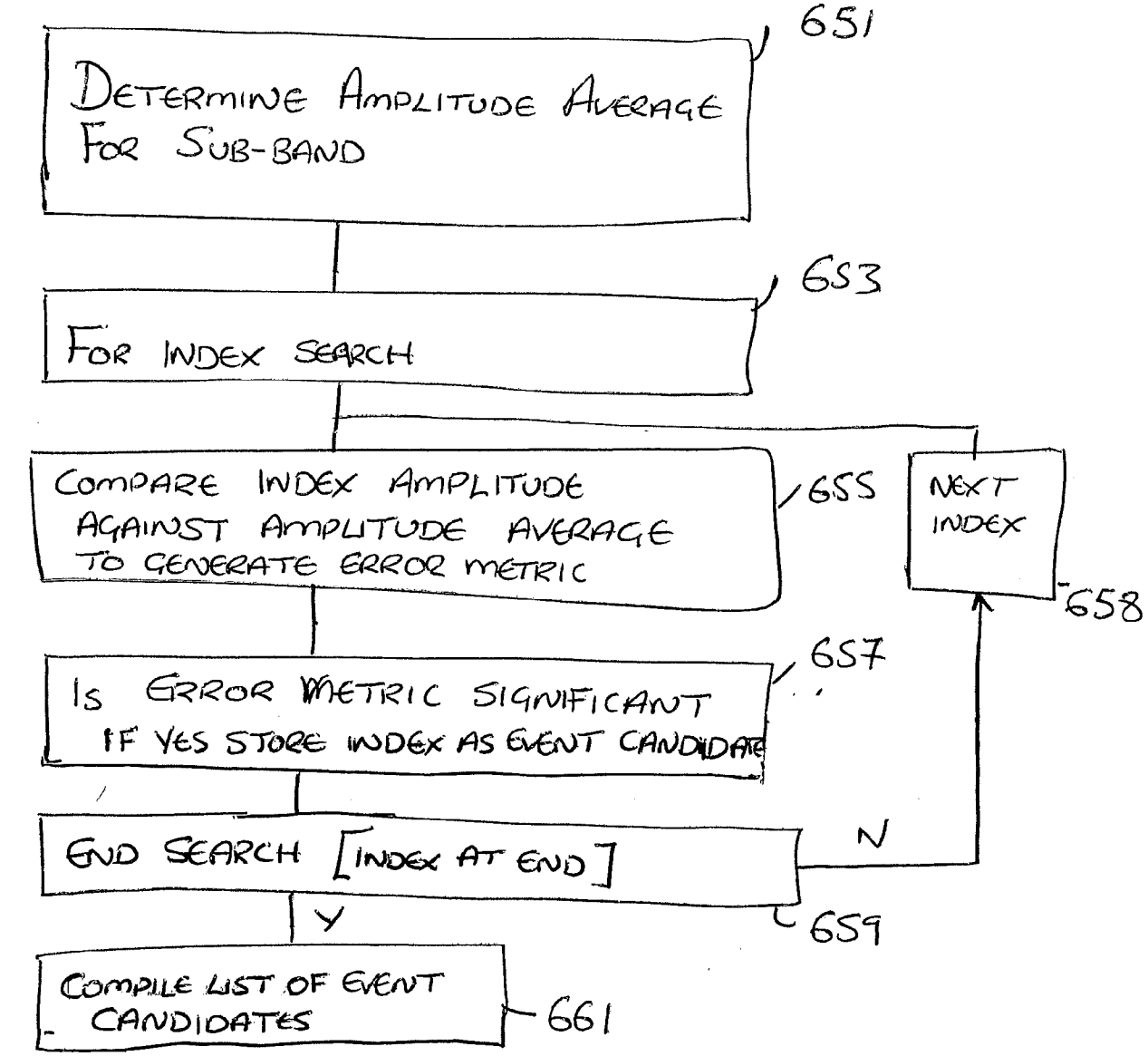

303
SUB-BAND GENERATOR
311
AVERAGE AMPLITUDE GENERATOR
313
METRIC GENERATOR
315
EVENT DETECTOR
317
X_H
$X_{H,1}$
$X_{H,N}$
A_{VE_1}
A_{VE_N}
$M_{H,1}$
$M_{H,N}$
$X_{j,1}^{EV}$
$X_{j,N}^{EV}$

Fig 6

Fig 7


651
DETERMINE AMPLITUDE AVERAGE FOR SUB-BAND
653
FOR INDEX SEARCH
655
COMPARE INDEX AMPLITUDE AGAINST AMPLITUDE AVERAGE TO GENERATE ERROR METRIC
NEXT INDEX
658
657
IS ERROR METRIC SIGNIFICANT IF YES STORE INDEX AS EVENT CANDIDATE
END SEARCH [INDEX AT END]
N
659
Y
COMPILE LIST OF EVENT CANDIDATES
661

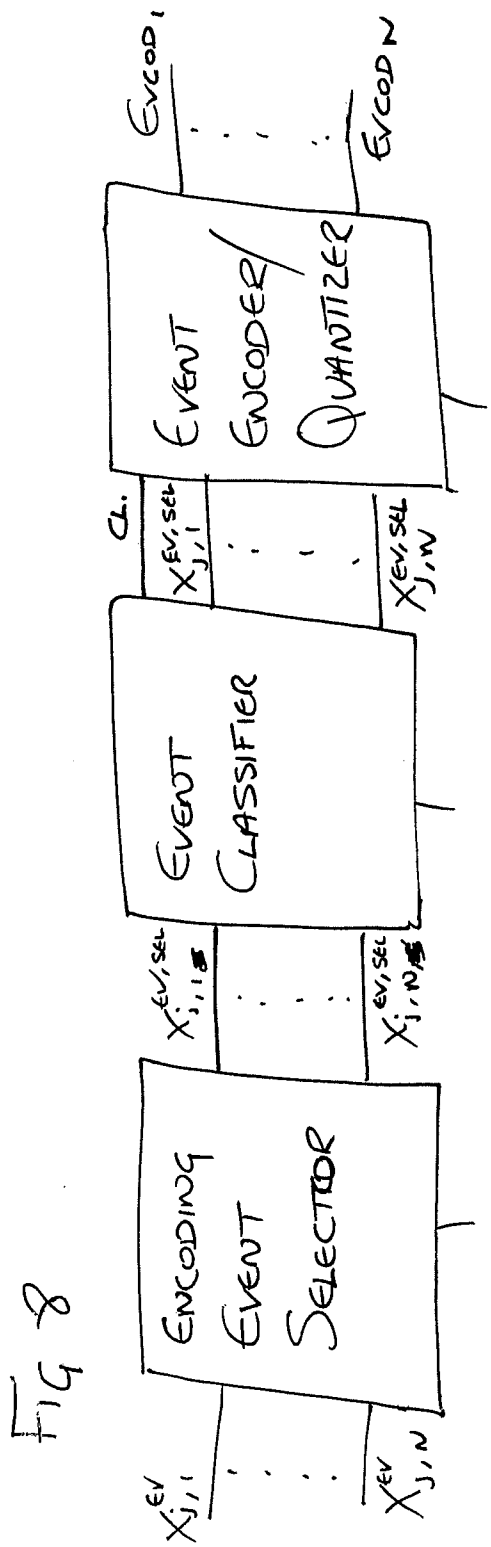

ENCODING EVENT SELECTOR
EVENT CLASSIFIER
EVENT ENCODER/ QUANTIZER
321
323
325

FIG 8

FIG 9

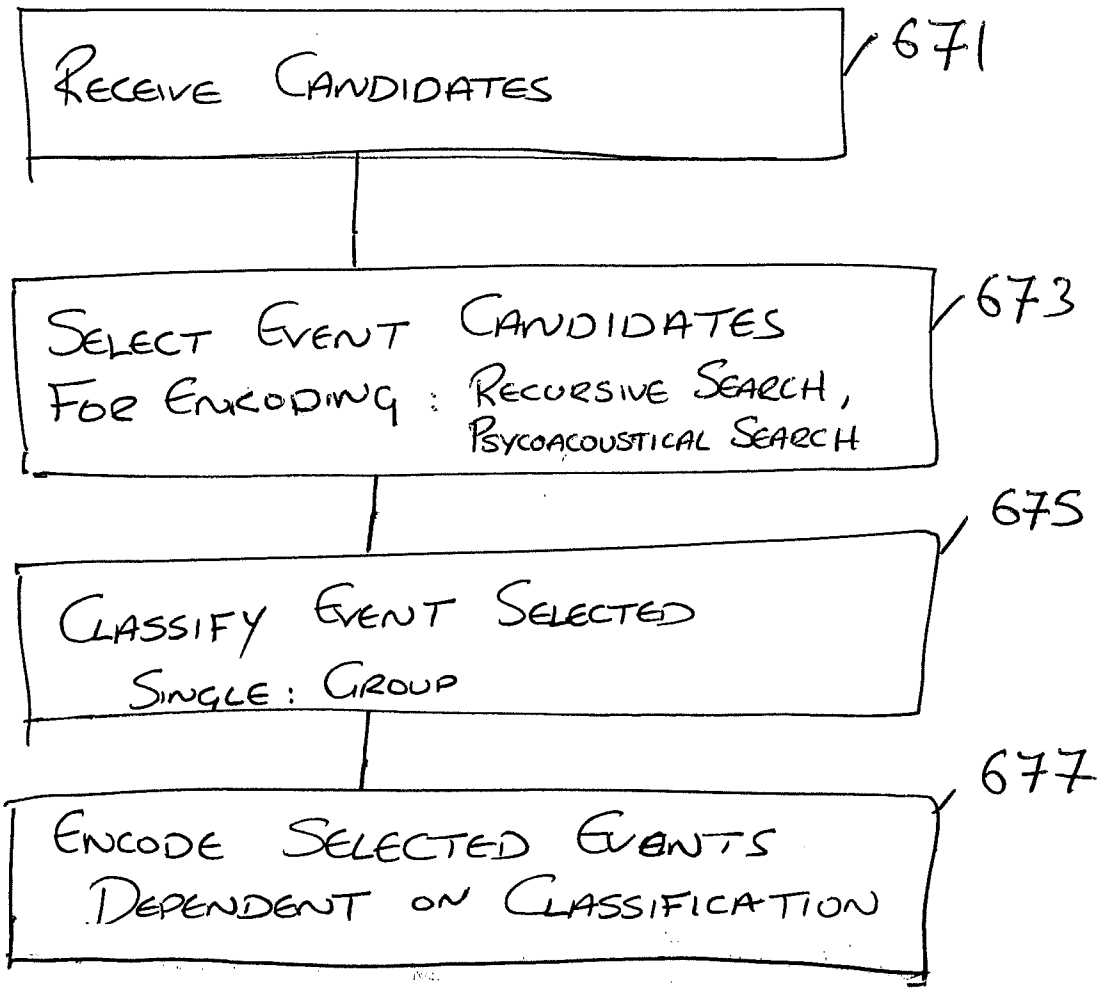

RECEIVE CANDIDATES
671
SELECT EVENT CANDIDATES FOR ENCODING : RECURSIVE SEARCH, PSYCOACOUSTICAL SEARCH
673
CLASSIFY EVENT SELECTED SINGLE : GROUP
675
ENCODE SELECTED EVENTS DEPENDENT ON CLASSIFICATION
677

FIG. 10

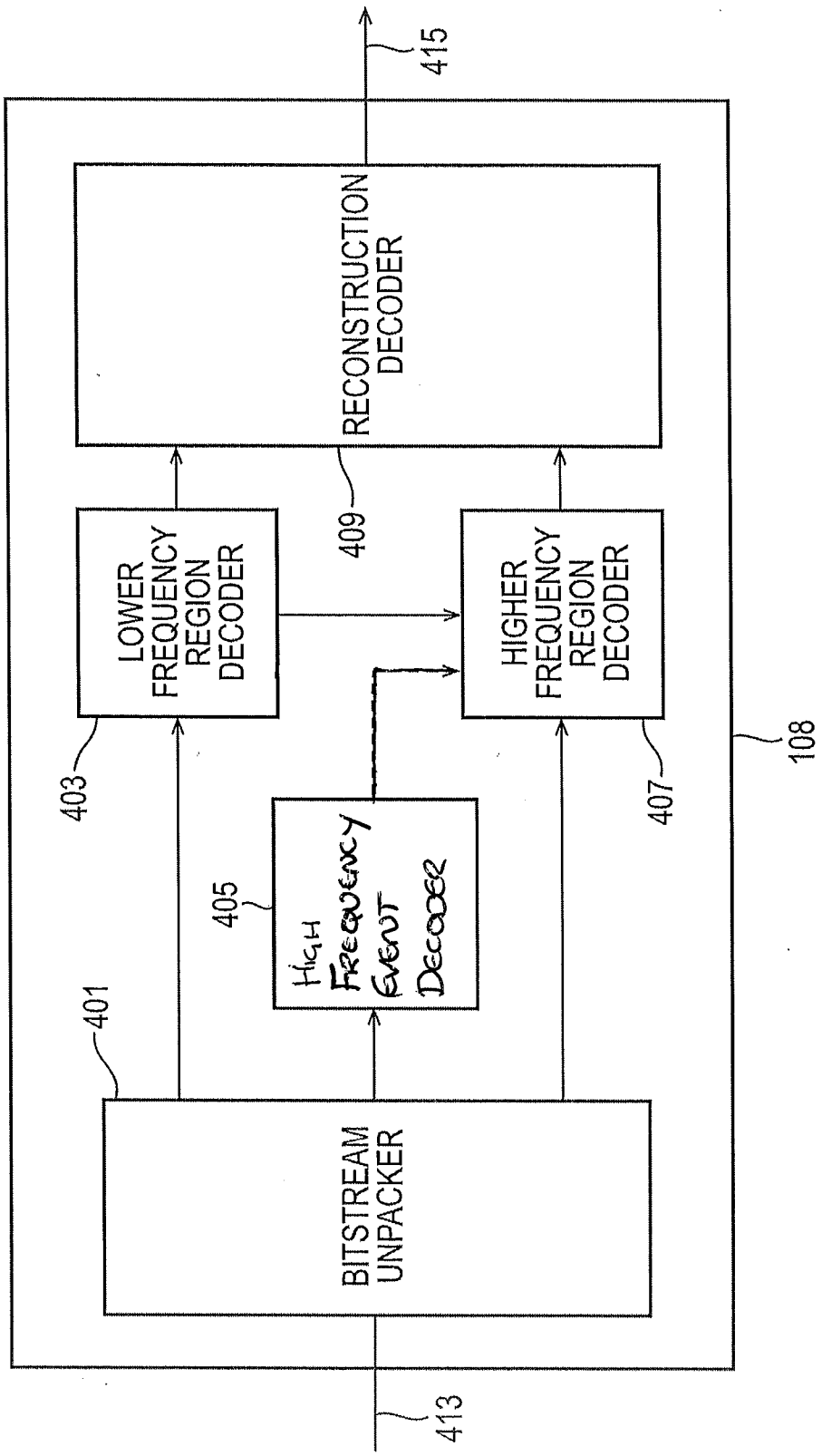
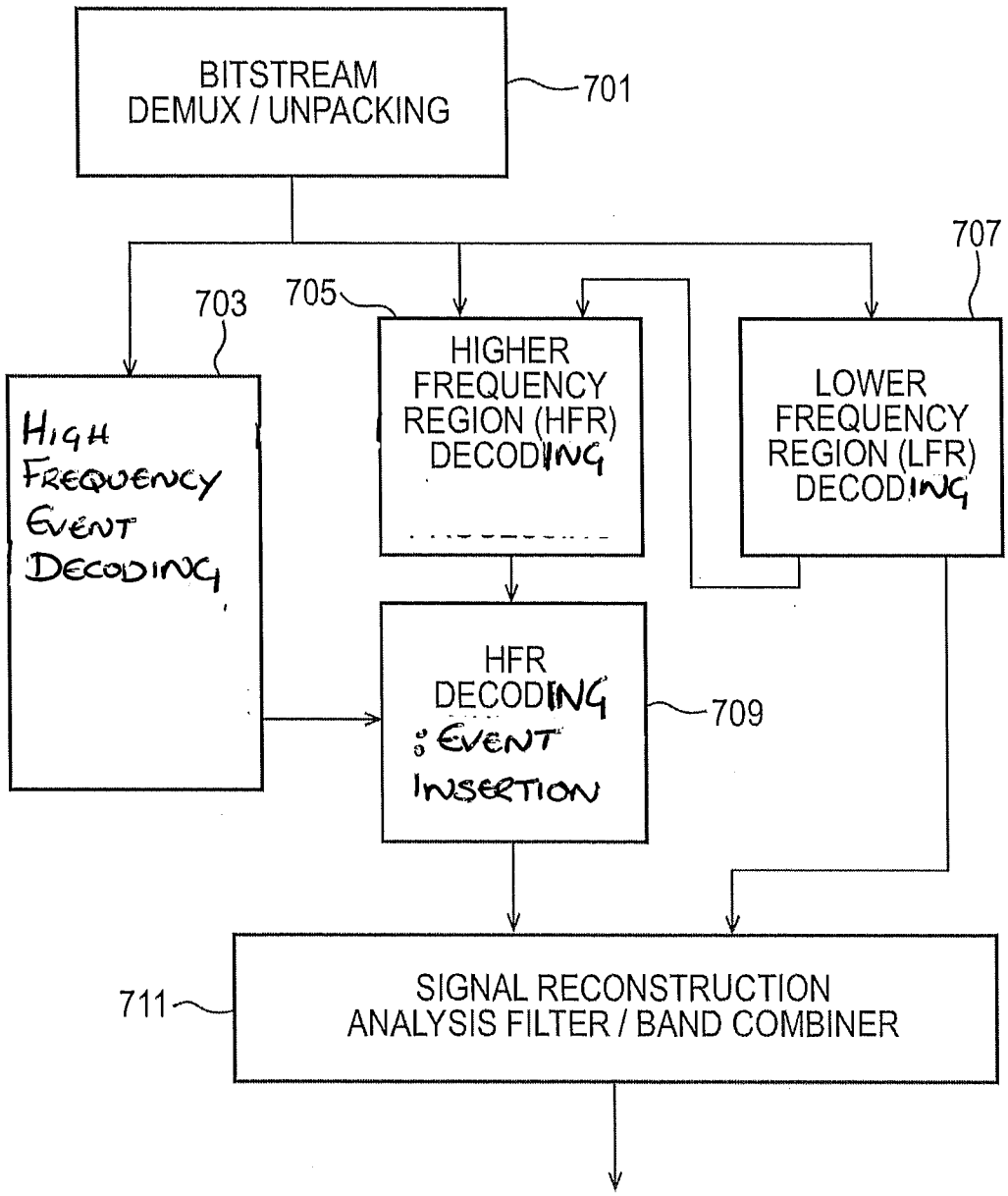
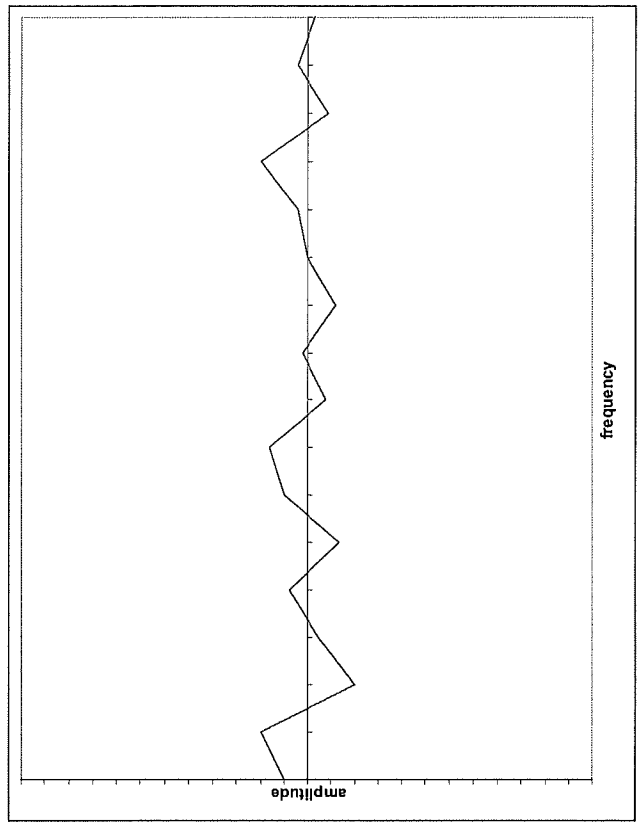
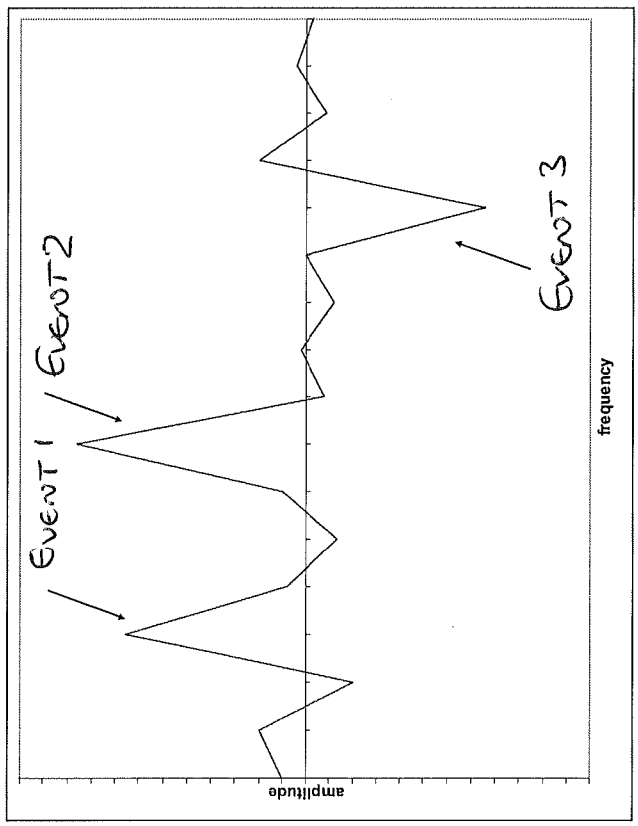

415
RECONSTRUCTION DECODER
409
LOWER FREQUENCY REGION DECODER
HIGHER FREQUENCY REGION DECODER
403
405
HIGH FREQUENCY EVENT DECODER
407
401
BITSTREAM UNPACKER
413
108

FIG 11


BITSTREAM DEMUX / UNPACKING
701
707
703
705
HIGHER FREQUENCY REGION (HFR) DECODING
LOWER FREQUENCY REGION (LFR) DECODING
HIGH FREQUENCY EVENT DECODING
HFR DECODING & EVENT INSERTION
709
711
SIGNAL RECONSTRUCTION ANALYSIS FILTER / BAND COMBINER



Suppressed Frequency Domain Representation



Example Frequency Domain Representation (with events)

Fig 12

INTERNATIONAL SEARCH REPORT

International application No.
PCT/IB2010/051210

A. CLASSIFICATION OF SUBJECT MATTER

IPC: see extra sheet
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC: G10L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

SE,DK,FI,NO classes as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-INTERNAL, WPI DATA, PAJ, INSPEC, COMPENDEX

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2009059633 A1 (NOKIA CORPORATION), 14 May 2009 (14.05.2009), page 3, line 13 - line 15; page 19, line 18 - line 24; page 19, line 29 - line 31, page 21, line 17-line 18, page 22, line 17-line 19, page 23, line 17-line 19, page 26, line 27 - page 27, line 7, page 36, line 13-line 15, page 37, line 6-line 9, line 26-line 29	1-3,11-13, 21-26
X	--	
A	WO 03046891 A1 (CODING TECHNOLOGIES SWEDEN AB), 5 June 2003 (05.06.2003), page 21, line 8 - page 22, line 2, figures 7-9, claim 7	1-26
	--	

[X] Further documents are listed in the continuation of Box C. [X] See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
25 November 2010	26-11-2010
Name and mailing address of the ISA/ Swedish Patent Office Box 5055, S-102 42 STOCKHOLM Facsimile No. +46 8 666 02 86	Authorized officer Jesper Bergstrand / EÖ Telephone No. +46 8 782 25 00

Form PCT/ISA/210 (second sheet) (July 2009)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/IB2010/051210

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 20050149339 A1 (TANAKA, NAOYA ET AL), 7 July 2005 (07.07.2005), figures 5-7, paragraph [0037], figures 5-7	1-26
A	WO 2007052088 A1 (NOKIA CORPORATION), 10 May 2007 (10.05.2007), paragraph [0027]	1-26
A	US 20070156397 A1 (CHONG, KOK SENG ET AL), 5 July 2007 (05.07.2007), paragraphs [0003], [0096], [0104]	1-26
A	US 20050080621 A1 (TSUSHIMA, MINEO ET AL), 14 April 2005 (14.04.2005), claim 1	1-26

INTERNATIONAL SEARCH REPORT

International application No.
PCT/IB2010/051210

International patent classification (IPC)
G10L 19/02 (2006.01)

Download your patent documents at www.prv.se
The cited patent documents can be downloaded:

- From "Cited documents" found under our online services at www.prv.se (English version)
- From "Anförda dokument" found under "e-tjänster" at www.prv.se (Swedish version)

Use the application number as username. The password is **PDGTOBMHFU**.

Paper copies can be ordered at a cost of 50 SEK per copy from PRV InterPat (telephone number 08-782 28 85).

Cited literature, if any, will be enclosed in paper form.

WO	2009059633	A1	14/05/2009	CA	2704812	A	14/05/2009
				EP	2212884	A	04/08/2010
WO	03046891	A1	05/06/2003	AT	288617	T	15/02/2005
				AU	2002352182	A	10/06/2003
				CN	1279512	C	11/10/2006
				CN	1571993	A	26/01/2005
				DE	60202881	D,T	19/01/2006
				EP	1423847	A,B	02/02/2005
				SE	1423847	T3	
				ES	2237706	T	01/08/2005
				HK	1062350	A	20/05/2005
				JP	3870193	B	17/01/2007
				JP	2005510772	T	21/04/2005
				PT	1423847	E	31/05/2005
				US	7469206	B	23/12/2008
				US	20050096917	A	05/05/2005
				US	20090132261	A	21/05/2009
				US	20090326929	A	31/12/2009
US	20050149339	A1	07/07/2005	NONE			
WO	2007052088	A1	10/05/2007	AU	2005337961	A	10/05/2007
				BR	PI0520729	A	26/05/2009
				CN	101297356	A	29/10/2008
				EP	1943643	A	16/07/2008
				JP	2009515212	T	09/04/2009
				US	20090271204	A	29/10/2009
US	20070156397	A1	05/07/2007	NONE			
US	20050080621	A1	14/04/2005	NONE			