

(51) International Patent Classification:
G10L 21/02 (2006.01)

(21) International Application Number:

PCT/IB2011/051391

(22) International Filing Date:

31 March 2011 (31.03.2011)

(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

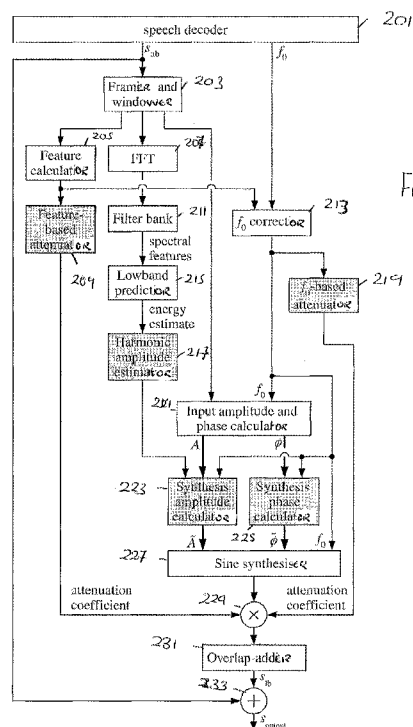
(75) Inventors/Applicants (for US only): **LAAKSONEN, Laura** [FI/FI]; Hyljeveentie 1 B 3, FIN-02260 Espoo (FI). **PULAKKA, Hannu, Juhani** [FI/FI]; Rumunlyöjankatu 2 A 8, FIN-02600 Espoo (FI). **REMES, Ulpu** [FI/FI]; Menninkäisentie 2 D 23, FIN-02110 Espoo (FI). **ALKU, Paavo, Ilmari** [FI/FI]; Eskelinsuo 16 B, FIN-00680 Helsinki (FI). **PALOMAKI, Kalle** [FI/FI]; Kielotie 44 A 5, FIN-01300 Vantaa (FI).(74) Agents: **SMITH, Gary, John** et al.; 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, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, 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, RS, SE, SI, SK, SM,

[Continued on next page]

(54) Title: A LOW BAND BANDWIDTH EXTENDER



(57) Abstract: Apparatus comprising: an input amplitude and phase calculator configured to determine at least one amplitude value and phase value dependent on a first audio signal; a synthesis amplitude calculator configured to synthesize a further amplitude value associated with each amplitude value dependent on a determined harmonic shaping function; a synthesis phase calculator configured to synthesize a further phase value associated with each phase value; and a signal synthesizer configured to generate a bandwidth extension signal dependent the further amplitude value and the further phase values.



TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, **Published:**
ML, MR, NE, SN, TD, TG).

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

A low band bandwidth extender

Field of the Application

- 5 The present invention relates to an apparatus and method for improving the quality of an audio signal. In particular, the present invention relates to an apparatus and method for extending the bandwidth of an audio signal.

Background of the Application

10

Audio signals, such as speech or music, can be encoded to enable efficient transmission or storage of the audio signals. The audio signal can be received or retrieved, decoded and presented to a user.

- 15 Audio signals can be limited to a bandwidth which is typically determined by the available capacity of the transmission system or storage medium. However, in some instances it may be desirable to perceive or present the decoded audio signal at a wider bandwidth than the bandwidth at which the audio signal was originally encoded. In these instances artificial bandwidth extension may be
20 deployed at the decoder, whereby the bandwidth of the decoded audio signal may be extended by using information solely determined from the decoded audio signal itself.

- 25 The audio bandwidth of 300 Hz to 3400 Hz which is used in today's fixed and mobile communication systems is comparable to that of conventional analogue telephony. This is because when digital standards were first established, a common audio bandwidth facilitated interoperability between the analogue and digital domains. This common narrowband signal is known as the telephone band.

- 30 These artificial bandwidth extensions can be higher or high frequency band (HB) extensions for example extending the output to 8 kHz and lower or low frequency band (LB) extensions for example extending the output to 50 Hz.

The capture and reproduction of frequencies below this range can often be limited by the characteristics of the terminal devices and by the filtering applied to the signal prior to encoding. However, human voice often contains frequency components below the telephone bandwidth. Consequently the quality of speech and the naturalness of the speech can be degraded by the limited frequency range. Artificial bandwidth extension (ABE) techniques have been proposed in which an extension band below the frequency range of the telephone band or narrowband signal called the low band which can, for example range from 50 Hz to 300 Hz is estimated from a received or recovered narrowband audio signal.

Current artificial bandwidth extension methods are known to apply, the low band or lower extension band (from 50 Hz to 300 Hz), irrelevant of whether or not it can improve the audio signal.

The embodiments of the application attempt to improve the perceived quality and intelligibility of the narrowband telephone speech by post-processing the speech signal received or recovered and by artificially widening the low frequency content below the telephone band, based solely on information extracted from the received speech signal when the sound reproduction system is capable of reproducing low frequencies. This can be employed in embodiments in a mobile terminal or in some other speech communication device or software, such as a teleconferencing system, or an ambient telephony system.

Summary of some embodiments

Embodiments aim to address the above problem.

There is provided according to a first aspect a method comprising: determining at least one amplitude value and phase value dependent on a first audio signal; synthesising a further amplitude value associated with each amplitude value dependent on a determined harmonic shaping function; synthesising a further phase value associated with each phase value; and generating a bandwidth extension signal dependent the further amplitude value and the further phase values.

The method may further comprise generating at least one attenuation factor, wherein the bandwidth extension signal may be further dependent on the attenuation factor.

5

The at least one attenuation factor may comprise an unvoiced signal attenuation factor, the unvoiced signal attenuation factor being dependent on an unvoiced component of the first audio signal.

- 10 The at least one attenuation factor may comprise a pause attenuation factor, the pause attenuation factor being dependent on determining a paused speech component of the first audio signal.

- 15 The at least one attenuation factor may comprise a fundamental frequency attenuation factor, the fundamental frequency attenuation factor being dependent on a fundamental frequency estimate associated with the first audio signal.

- 20 The at least one attenuation factor may comprise an octave error attenuation factor, the octave error attenuation factor being dependent on determining an error in a fundamental frequency estimate associated with the first audio signal.

The method may further comprise determining a harmonic shaping function dependent on the estimated bandwidth extension signal energy level.

- 25 The method may further comprise determining an estimated bandwidth extension signal energy level.

- 30 Determining an estimated bandwidth extension signal energy level may comprise: determining at least one feature value associated with the first signal; and applying the at least one feature to a trained modelling function to determine the estimated bandwidth extension signal energy level.

The modelling function may comprise at least one of: a Gaussian mixture model; a hidden Markov model; and a neural network model.

Synthesising the further amplitude value associated with each amplitude value may be further dependent on the first audio signal.

- 5 Synthesising the further amplitude value associated with each amplitude value to be further dependent on the first audio signal may comprise: determining the amplitudes of the first audio signal for the harmonic frequencies associated with the first amplitude; and synthesising the further amplitude dependent on the amplitudes of the first audio signal for the harmonic frequencies associated with
10 the first amplitude.

Synthesising a further phase value associated with each phase value may comprise: determining a condition associated with each phase value; and generating a further phase value dependent on the condition and the phase value.

- 15 Determining the condition associated with the phase value may comprise: determining the phase value is highly varying, wherein the further phase value is a reference phase value; determining the onset of the phase value, wherein the further phase value is the reference phase value; determining the phase value is sufficiently close to the reference phase value, wherein the further phase value is
20 the phase value; determining the phase value is different from the reference phase value and the phase value is consistent over a period of time, wherein the further phase value is phase value approaching the phase value from the reference phase value; and otherwise determining the phase value is inconsistent over a period of
25 time, wherein the further phase value is the reference phase value.

The reference phase value may be dependent on a previous period further phase value and the fundamental frequency estimates from the current and previous periods.

- 30 According to a second aspect there is provided an apparatus comprising at least one processor and at least one memory including computer code, the at least one memory and the computer code configured to with the at least one processor cause the apparatus to at least perform: determining at least one amplitude value

and phase value dependent on a first audio signal; synthesising a further amplitude value associated with each amplitude value dependent on a determined harmonic shaping function; synthesising a further phase value associated with each phase value; and generating a bandwidth extension signal dependent the further
5 amplitude value and the further phase values.

The apparatus may be further configured to perform generating at least one attenuation factor, wherein the bandwidth extension signal is further dependent on the attenuation factor.

10

The at least one attenuation factor may comprise an unvoiced signal attenuation factor, the unvoiced signal attenuation factor being dependent on an unvoiced component of the first audio signal.

15 The at least one attenuation factor may comprise a pause attenuation factor, the pause attenuation factor being dependent on determining a paused speech component of the first audio signal.

20 The at least one attenuation factor may comprise a fundamental frequency attenuation factor, the fundamental frequency attenuation factor being dependent on a fundamental frequency estimate associated with the first audio signal.

The at least one attenuation factor may comprise an octave error attenuation factor, the octave error attenuation factor being dependent on determining an error
25 in a fundamental frequency estimate associated with the first audio signal.

The apparatus may be further configured to perform determining a harmonic shaping function dependent on the estimated bandwidth extension signal energy level.

30

The apparatus may be further configured to perform determining an estimated bandwidth extension signal energy level.

Determining an estimated bandwidth extension signal energy level may cause the apparatus to further perform: determining at least one feature value associated with the first signal; and applying the at least one feature to a trained modelling function to determine the estimated bandwidth extension signal energy level.

5

The modelling function may comprise at least one of: a Gaussian mixture model; a hidden Markov model; and a neural network model.

10 Synthesising the further amplitude value associated with each amplitude value may be further dependent on the first audio signal.

15 Synthesising the further amplitude value associated with each amplitude value further dependent on the first audio signal further may cause the apparatus to perform: determining the amplitudes of the first audio signal for the harmonic frequencies associated with the first amplitude; and synthesising the further amplitude dependent on the amplitudes of the first audio signal for the harmonic frequencies associated with the first amplitude.

20 Synthesising a further phase value associated with each phase value may cause the apparatus to perform: determining a condition associated with each phase value; and generating a further phase value dependent on the condition and the phase value.

25 Determining the condition associated with the phase value may cause the apparatus to further perform: determining the phase value is highly varying, wherein the further phase value is a reference phase value; determining the onset of the phase value, wherein the further phase value is the reference phase value; determining the phase value is sufficiently close to the reference phase value, wherein the further phase value is the phase value; determining the phase value is different from the reference phase value and the phase value is consistent over a
30 period of time, wherein the further phase value is phase value approaching the phase value from the reference phase value; and otherwise determining the phase value is inconsistent over a period of time, wherein the further phase value is the reference phase value.

The reference phase value may be dependent on a previous period further phase value and the fundamental frequency estimates from the current and previous periods.

5

According to a third aspect there is provided apparatus comprising: means for determining at least one amplitude value and phase value dependent on a first audio signal; means for synthesising a further amplitude value associated with each amplitude value dependent on a determined harmonic shaping function;
10 means for synthesising a further phase value associated with each phase value; and means for generating a bandwidth extension signal dependent the further amplitude value and the further phase values.

The apparatus may further comprise means for generating at least one attenuation
15 factor, wherein the bandwidth extension signal is further dependent on the attenuation factor.

The at least one attenuation factor may comprise an unvoiced signal attenuation factor, the unvoiced signal attenuation factor being dependent on an unvoiced
20 component of the first audio signal.

The at least one attenuation factor may comprise a pause attenuation factor, the pause attenuation factor being dependent on determining a paused speech component of the first audio signal.

25

The at least one attenuation factor may comprise a fundamental frequency attenuation factor, the fundamental frequency attenuation factor being dependent on a fundamental frequency estimate associated with the first audio signal.

30 The at least one attenuation factor may comprise an octave error attenuation factor, the octave error attenuation factor being dependent on determining an error in a fundamental frequency estimate associated with the first audio signal.

The apparatus may further comprise means for determining a harmonic shaping function dependent on the estimated bandwidth extension signal energy level.

5 The apparatus may further comprise means for determining an estimated bandwidth extension signal energy level.

The means for determining an estimated bandwidth extension signal energy level may comprise: means for determining at least one feature value associated with the first signal; and means for applying the at least one feature to a trained
10 modelling function to determine the estimated bandwidth extension signal energy level.

The modelling function may comprise at least one of: a Gaussian mixture model; a hidden Markov model; and a neural network model.

15

The means for synthesising the further amplitude value associated with each amplitude value may be further dependent on the first audio signal.

20 The means for synthesising the further amplitude value associated with each amplitude value further dependent on the first audio signal may further comprise: means for determining the amplitudes of the first audio signal for the harmonic frequencies associated with the first amplitude; and means for synthesising the further amplitude dependent on the amplitudes of the first audio signal for the harmonic frequencies associated with the first amplitude.

25

The means for synthesising a further phase value associated with each phase value may comprise: means for determining a condition associated with each phase value; and means for generating a further phase value dependent on the condition and the phase value.

30

The means for determining the condition associated with the phase value may comprise: means for determining the phase value is highly varying, wherein the further phase value is a reference phase value; means for determining the onset of the phase value, wherein the further phase value is the reference phase value;

means for determining the phase value is sufficiently close to the reference phase value, wherein the further phase value is the phase value; means for determining the phase value is different from the reference phase value and the phase value is consistent over a period of time, wherein the further phase value is phase value
5 approaching the phase value from the reference phase value; and means for determining the phase value is inconsistent over a period of time, wherein the further phase value is the reference phase value.

The reference phase value may be dependent on a previous period further phase
10 value and the fundamental frequency estimates from the current and previous periods.

According to a fourth aspect there is provided apparatus comprising: an input amplitude and phase calculator configured to determine at least one amplitude
15 value and phase value dependent on a first audio signal; a synthesis amplitude calculator configured to synthesize a further amplitude value associated with each amplitude value dependent on a determined harmonic shaping function; a synthesis phase calculator configured to synthesize a further phase value associated with each phase value; and a signal synthesizer configured to generate
20 a bandwidth extension signal dependent the further amplitude value and the further phase values.

The apparatus may further comprising an attenuator gain determiner configured to generate at least one attenuation factor, wherein the bandwidth extension signal is
25 further dependent on the attenuation factor.

The at least one attenuation factor may comprise an unvoiced signal attenuation factor, the unvoiced signal attenuation factor being dependent on an unvoiced component of the first audio signal.
30

The at least one attenuation factor may comprise a pause attenuation factor, the pause attenuation factor being dependent on determining a paused speech component of the first audio signal.

The at least one attenuation factor may comprise a fundamental frequency attenuation factor, the fundamental frequency attenuation factor being dependent on a fundamental frequency estimate associated with the first audio signal.

- 5 The at least one attenuation factor may comprise an octave error attenuation factor, the octave error attenuation factor being dependent on determining an error in a fundamental frequency estimate associated with the first audio signal.

10 The apparatus may further comprise a harmonic amplitude estimator configured to perform determining a harmonic shaping function dependent on the estimated bandwidth extension signal energy level.

The apparatus may further comprise a lowband energy estimator configured to determine an estimated bandwidth extension signal energy level.

15 The lowband energy estimator may comprise: a feature determiner configured to determine at least one feature value associated with the first signal; and a trained modelling function configured to determine the estimated bandwidth extension signal energy level dependent on the at least one feature value.

20 The trained modelling function may comprise at least one of: a Gaussian mixture model; a hidden Markov model; and a neural network model.

25 The signal synthesizer configured to generate a bandwidth extension signal may be further dependent on the first audio signal.

30 The signal synthesizer configured to generate a bandwidth extension signal may further comprise: an amplitude determiner configured to determine the amplitudes of the first audio signal for the harmonic frequencies associated with the first amplitude; and wherein the synthesizer being configured to determine the further amplitude dependent on the amplitudes of the first audio signal for the harmonic frequencies associated with the first amplitude.

The synthesis phase calculator may comprise: a condition determiner configured to determine a condition associated with each phase value; a phase synthesizer configured to generate the further phase value dependent on the condition and the phase value.

5

The condition determiner may comprise: a first condition determiner configured to determine the phase value is highly varying, wherein the further phase value is a reference phase value; a second condition determiner configured to determine an onset of the phase value, wherein the further phase value is the reference phase value; a third condition determiner configured to determine the phase value is sufficiently close to the reference phase value, wherein the further phase value is the phase value; a fourth condition determiner configured to determine the phase value is different from the reference phase value and the phase value is consistent over a period of time, wherein the further phase value is phase value approaching the phase value from the reference phase value; and a fifth condition determiner configured to determine the phase value is inconsistent over a period of time, wherein the further phase value is the reference phase value.

The reference phase value may be dependent on a previous period further phase value and the fundamental frequency estimates from the current and previous periods.

An electronic device may comprise apparatus as described herein.

25 A chipset may comprise apparatus as described herein.

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:

30 Figure 1 shows schematically an electronic device employing embodiments of the invention;

Figure 2 shows schematically a decoder system employing embodiments of the invention;

Figure 3 shows schematically a decoder according to some embodiments of the application;

5 Figure 4 shows a flow diagram detailing the operation of the decoder shown in Figure 3;

Figure 5 shows relative performance for narrowband, adaptive multi rate-wide band, high band artificial bandwidth extension, and low band + high band artificial bandwidth extension for a voiced male speech short segment example;

10 Figure 6 shows relative performance for narrowband, adaptive multi rate-wide band, and low band extension + narrow band for a voiced male speech example; and

Figure 7 shows a further example of the relative performance characteristics in long-term average spectra shown by narrowband, adaptive multi-rate wideband
15 speech coding, and artificial bandwidth extension decoding.

Description of Some Embodiments of the Application

The following describes in more detail possible mechanisms for the provision of
20 artificially expanding the bandwidth of a decoded audio signal. In this regard reference is first made to Figure 1 which shows a schematic block diagram of an exemplary electronic device 10 or apparatus, which may incorporate an artificial bandwidth extension system according to some embodiments.

25 The electronic device or apparatus 10 can for example be as described herein a mobile terminal or user equipment of a wireless communication system. In some other embodiments the apparatus 10 can be any suitable audio or audio-subsystem component within an electronic device such as audio player (also known as MP3 players) or media players (also known as MP4 players). In some
30 other embodiments, the electronic device can be a teleconference terminal or ambient telephone terminal.

The electronic device 10 can comprise in some embodiments a microphone 11, which is linked via an analogue-to-digital converter (ADC) 14 to a processor 21.

The processor 21 is further linked in some embodiments via a digital-to-analogue converter (DAC) 32 to loudspeaker(s) 33. The processor 21 is in some embodiments further linked to a transceiver (RX/TX) 13, to a user interface (UI) 15 and to a memory 22.

5

The processor 21 can be in some embodiments configured to execute various program codes. The implemented program codes 23 can comprise an audio decoding code or speech decoding code implementing an artificial bandwidth extension code. The implemented program codes 23 can in some embodiments be
10 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 application.

The decoding code can in some embodiments be implemented in electronic based
15 hardware or firmware.

In some embodiments the device can comprise a user interface 15. 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 110,
20 for example via a display.

Furthermore in some embodiments the electronic device further comprises a transceiver 13. The transceiver 13 enables a communication with other electronic devices, for example via a wireless communication network.

25

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

The electronic device 10 can in some embodiments receive a bit stream with
30 suitably encoded data from another electronic device via its transceiver 13. Alternatively, coded data could be stored in the data section 24 of the memory 22, for instance for a later presentation by the same electronic device 10. In both cases, the processor 21 may execute the decoding program code stored in the memory 22.

The processor 21 can therefore in some embodiments decode the received data, for instance in the manner as described with reference to Figures 3 and 4, and provide the decoded data to the digital-to-analogue converter 32. The digital-to-analogue converter 32 can then in some embodiments convert the digital decoded data into analogue audio data and output the audio signal via the loudspeaker(s) 33. However it would be understood that the loudspeaker or loudspeakers 33 can in some embodiments be any suitable audio transducer converting electrical signals into presentable acoustic signals.

Execution of the decoding program code could in some embodiments be triggered by an application that has been called by the user via the user interface 15.

The received encoded data could also be stored instead of an immediate presentation via the loudspeaker(s) 33 in the data section 24 of the memory 22, for instance for enabling a later presentation or a forwarding to still further electronic device.

It would be appreciated that the schematic structures described in Figure 3 and the method steps in Figure 4 represent only a part of the operation of a decoder and artificial bandwidth extender as exemplarily shown implemented in the electronic device shown in Figure 1.

Embodiments of the application are now described in more detail with respect to Figures 2 to 7.

The general operation of speech and audio decoders as employed by embodiments of the application is shown in Figure 2. A general decoding system 102 is illustrated schematically in Figure 2. The system 102 may comprise a storage or media channel (also known as a communication channel) 106 and a 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.

5 The operation of speech and audio codecs are known from the art and features of such codecs which do not assist in the understanding of the operation of the embodiments of the application are not described in detail.

10 Figure 3 shows schematically a decoder 108 according to some embodiments of the application. Although the term decoder has been used with respect to the process of decoding the stored/received signal and generating the artificial bandwidth extension it would be understood that these functions could in some embodiments be divided into components which decode the signal and provide decoded values such as described hereafter as the speech decoder and components which receive the decoded values and generate an artificial bandwidth
15 extension to be combined with at least part of the decoded signal to form a wideband audio/speech signal. In other words in some embodiments an artificial bandwidth extension generator can comprise the decoder as described hereafter except the speech decoder. In such embodiments the artificial bandwidth extension generator can be configured to receive at least a narrowband signal as an input, and can furthermore optionally receive the fundamental frequency estimate. These
20 narrowband signal and optionally the fundamental frequency estimate can be received in some embodiments from a speech decoder or any other suitable source.

25 The decoder 108 in some embodiments comprises a speech decoder 201. The speech decoder in some embodiments receives the encoded bit stream via a receiver. In some other embodiments the speech decoder can retrieve or recover the encoded bit stream from the memory of the electronic apparatus 10.

30 The operation of receiving or recovering the encoded bit stream is shown in Figure 4 by step 301.

The speech decoder can be any suitable speech decoder, for example the adaptive multi rate (AMR) speech coding standard, details of which can be found in

the 3GPP TS 26.090 Technical Specification. However in other embodiments of the application, any suitable speech or audio codec decoding algorithm can be implemented to decode the encoded bit stream.

- 5 The decoder can in some embodiments generate the narrowband audio or speech signal s_{nb} from the encoded bit stream. Furthermore in some embodiments the decoder or speech decoder 201 can be configured to further generate or determine the fundamental frequency. The speech decoder 201 can furthermore generate or recover a fundamental frequency f_0 value or pitch estimate based on a pitch period
- 10 estimate performed in the associated encoder and passed along with the encoded narrowband signal. However in some embodiments the fundamental frequency can be estimated from the narrowband signal input to the bandwidth extension components as discussed herein.
- 15 The decoding of the bit stream, which can generate values for the fundamental frequency estimate f_0 and also output a narrowband audio or speech signal s_{nb} is shown in Figure 4 by step 303.

In some embodiments the decoder 108 can further comprise a framer and

20 windower 203. The framer and windower 203 can be configured in some embodiments to receive the narrowband audio or speech signal s_{nb} sample values and output a series of windowed time frame sampled data. In the following example the framer and windower 203 can be configured to output three differently windowed and framed audio signal outputs, however in some embodiments any

25 suitable number of frame formats can be output. For example in some embodiments the input or decoded narrowband (telephone) speech signal s_{nb} is sampled at 8 kHz and has frames of 5 ms. However any suitable input sample rate and frame length can be processed in some embodiments. The framer and windower 203 can in some embodiments process the input decoded narrowband

30 audio/speech signal using window functions and window lengths to generate various outputs for at least one analysis or component. The following frame formats are examples of the possible suitable framing and windowing operations.

For example in some embodiments the framer and windower 203 can perform a first framing and windowing to generate a time domain analysis frame format for a time domain feature calculator 205. In such embodiments the time domain frame format can implement a rectangular window of 20 ms onto the input signal and generate an output frame with a frame shift of 5ms. The framer and windower 203 can using the example input signal described above concatenate four input frames each of 5ms to generate a 20 ms frame. The framer and windower can be configured to output the time domain frame format data to the time domain feature calculator 205.

10

In some embodiments a second windowing and framing operation can be performed by the framer and windower 203 to generate a frequency domain analysis frame format for frequency domain analysis. For example in some embodiments the framer and windower 203 can be configured to output to a Fast Fourier Transformer 207, the narrowband signal with a 16 ms Hamming window computed for every 10 ms. However it would be possible to employ 5ms frame shifting in some embodiments.

Furthermore in some embodiments the framer and windower 203 can be configured to output a third framing and windowing operation for generate a low band analysis frame format for low band amplitude and phase analysis. For example the third framing and windowing operation can be to generate a 20 ms Hann window computed every 5 ms.

In such embodiments the maximum look ahead used in the framer and windower 203 is 5 ms.

The operation of framing and windowing is shown in Figure 4 by step 305.

In some embodiments the decoder 108 further comprises a time domain feature calculator 205 or feature calculator. The feature calculator 205 can, as described previously, be configured to receive frames or segments of 20 ms narrowband speech or audio signals s_{nb} with frame shifting of 5 ms. The time domain feature calculator 205 can then determine or generate from each windowed frame at least

one of the following feature or characteristic values of the narrowband audio or speech signal.

Frame Energy

5

The frame energy E_{dB} of the input signal s_{nb} for each frame can be computed and converted to a decibel scale using the following equation:

$$E_{dB} = 10 \log_{10} \left(\sum_{k=0}^{N_k-1} s_{nb}(k)^2 \right),$$

10

where k is the time index within the frame, N_k is the frame length, and s_{nb} is the narrowband input signal.

Noise Floor Estimate

15

The noise floor estimate $N_{dB}(n)$ for the frame n can be determined such that it approximates the lowest frame energy value. For example in some embodiments the noise floor estimate can be computed from the frame energy value by filtering the frame energy value $E_{dB}(n)$ with a first order recursive filter such as:

20

$$N_{dB}(n) = \alpha_n E_{dB}(n) + (1 - \alpha_n) N_{dB}(n-1),$$

where $\alpha_n = 0.0015$ for an upward change and $\alpha_n = 0.2$ for a downward change.

The noise floor estimate thus rises slowly during speech but quickly approaches energy minima. In such embodiments the value of the noise floor estimate $N_{dB}(n)$ can in some embodiments be configured to be not allowed to go below a fixed low limit.

25

Active Speech Level

30

The active speech level estimate $S_{dB}(n)$ furthermore approximates a typical maximum value of the frame energy in the input signal. In a manner similar to the

noise floor estimate, the active speech level estimate can be determined in some embodiments by a first order recursive filter arrangement such as:

$$S_{dB}(n) = \alpha_s E_{dB}(n) + (1 - \alpha_s) S_{dB}(n-1),$$

5

where $\alpha_s = 0.2$ for upward change and $\alpha_s = 0.0005$ for downward change. The speech level estimate thus decays slowly during pauses but quickly approaches the energy maxima during active speech. Furthermore in such embodiments the value of the active speech level $S_{dB}(n)$ can be configured not to be allowed to go below the noise floor estimate $N_{dB}(n)$.

10

Gradient Index

The gradient index x_{gi} is defined as the sum of the signal gradient magnitude at each change of signal direction normalised by the frame energy and can be determined using the following equation:

15

$$x_{gi} = \frac{\sum_{k=2}^{N_k-1} \Psi(k) |s_{nb}(k) - s_{nb}(k-1)|}{\sum_{k=0}^{N_k-1} s_{nb}(k)^2},$$

where N_k is the frame index length and s_{nb} is the narrowband input signal, $\Psi(k)$ is equal to 1 when the gradient $s_{nb}(k) - s_{nb}(k-1)$ changes sign and 0 otherwise. This value feature provides low values for voiced speech and high values for unvoiced speech. In other words generating a low value when the signal contains components when the vocal folds are vibrating (voiced).

25

In some embodiments other feature values suitable for predicting voiced or unvoiced characteristics of speech could furthermore be determined by the time domain feature calculator either in combination or to replace the gradient index value. Furthermore although in the above examples the determination of voiced or unvoiced speech is based on the time domain features, it would be understood that

30

in some embodiments the determination could be performed based on at least one frequency domain feature.

5 The performance of calculating time domain features is shown in Figure 4 by step 307.

10 In some embodiments the decoder 108 comprises a feature based attenuator 209. The feature based attenuator 209 can be configured to detect or determine, for example when the audio signal comprises voiced segments and generate an attenuation or amplification factor to be applied to the generated low band whenever the audio signal is lacking in voiced components. This operation is particularly useful as low band extension is useful only for voiced speech and adding energy to the low band during unvoiced or non-speech segments can be perceived as low frequency noise. The feature based attenuator 209 in some
15 embodiments could be implemented as any suitable means for generating attenuator factors or attenuator gain determiner and could for example generate the attenuation factors or gains for the fundamental frequency determined or based factors or gains as well.

20 The feature based attenuator 209 can therefore be configured to receive feature values from the time domain feature calculator 205 to determine whether or not the current frame is voiced, non voiced speech or non-speech.

25 The feature based attenuator 209 can in some embodiments determine at least one attenuation factor for a frame based on the time domain feature values to control applications of the generated low band. The output of the low band synthesis process can then be modified by the at least one attenuation factor before generating the final output. In some embodiments, two attenuation factors can be generated by the feature based attenuator 209.

30

In some embodiments an 'voiced' attenuation factor g_{gi} can be determined based on the value of the gradient index feature x_{gi} by using fixed or determined threshold values. For example in some embodiments the attenuation factor g_{gi} can be set to be a value of 0 when the gradient index feature x_{gi} is greater than 5.0 and set to a

value of 1 when the gradient index feature x_{gi} is less than 3.0 with a linear transition from 0 to 1 between these threshold values. However it would be understood that any suitable transition function can be implemented between such threshold values and similarly the threshold values themselves can in some embodiments be values other than those described above.

In some embodiments a pause attenuation factor g_p can also be generated by the feature based attenuator 209. Where the current frame energy $E_{dB}(n)$ does not exceed the noise floor estimate $N_{dB}(n)$ by a determined value or amount, the generated pause attenuation factor can be configured to enable the low synthesis signal to be attenuated. In some embodiments, for example, the attenuation factor g_p can be set to -40dB where the frame energy and the noise floor estimate differ by less than 4dB and the attenuation factor g_p is set to 0dB where the difference between the current frame energy and the noise floor estimate is greater than 10dB with a linear transition on the decibel scale between these thresholds. It would be understood that the threshold values of 4dB and 10dB and also the linear transition between these thresholds can be any suitable value and function in some other embodiments.

In some embodiments the feature based attenuator 209 could alternatively implement the 'pause' attenuation factor by using a received external VAD (voice activity detector) signal. In some embodiments the VAD signal could be received from the speech decoder, that predicts whether the current frame contains speech or not.

These attenuation factors can then be passed to an attenuation amplifier 229.

The generation of at least one attenuation factor dependent on the time domain features of the narrowband signal is shown in Figure 4 by step 311.

In some embodiments the decoder 108 can further comprise a Fast Fourier Transformer 207. The Fast Fourier Transformer 207 receives from the framer and windower 203 frequency domain analysis frame sample data and converts the time domain samples in each frame into suitable frequency domain values. For

example in some embodiments the input signal to the Fast Fourier Transformer 207 is a series of frames, each 16 ms long with a frame shift of 10 ms having been windowed, for example using a Hamming window. The FFT 207 is then configured to transform the input signals into the frequency domain using, for example, a 128 point Fast Fourier Transform. The output frequency domain characteristics of the narrowband audio signal can then be passed in some embodiments to a filterbank 211. It would be understood that any suitable time to frequency domain transformer could be used in some embodiments of the application.

10 The operation of performing a Fast Fourier Transform is shown in Figure 4 by step 309.

In some embodiments the decoder 108 further comprises a filterbank 211. The filterbank 211 can be configured to divide the frequency domain representation of the narrowband signal frame into sub-bands with linear spacing on a perceptually motivated mel-scale. The filterbank 211 can in some embodiments comprise a bank of 7 trapezoidal filters with the centre frequencies of each of the sub-bands located at 448 Hz, 729 Hz, 1079 Hz, 1515 Hz, 2058 Hz, 2733 Hz, and 3574 Hz.

20 It would be understood that in some other embodiments the filterbank can be any suitable filterbank with any suitable filter characteristics being performed on the frequency domain signal values.

In some embodiments the sub-band energies can then be calculated by squaring the magnitude of each of the sub-band frequency components amplitudes within each frequency bin generated by the filterbank.

In some embodiments the sub-band energies can be determined by squaring the magnitude of each FFT output to get the power spectrum, then for each sub-band, weight the squared frequency components by the corresponding filter window, before summing the weighted frequency components to get the sub-band energy.

In some embodiments the sub-band energy values can be log compressed using the mapping $\log(x+1)$.

The output of the spectral feature values can be passed to a low band predictor 215.

- 5 The operation of filtering and generation of spectral features is shown in Figure 4 by step 313.

In some embodiments the decoder 108 comprises a fundamental frequency estimate corrector 213. The fundamental frequency estimator corrector 213 can be
10 configured to receive the initial fundamental frequency estimate from the speech decoder 201 f_0 and produce a more accurate estimate of the fundamental frequency.

The fundamental frequency f_0 estimate from the audio signal can in some
15 embodiments be determined for each input frame. For example as previously described, in some embodiments the speech decoder 201 can obtain as part of the adaptive multi-rate (AMR) speech codec decoder a pitch period estimate for f_0 that the speech decoder receives from the encoder. In some other embodiments the decoder can also determine the pitch period of the audio signal by any suitable
20 pitch estimator of sufficient accuracy. In some embodiments, for example in speech or audio codecs which do not provide a fundamental frequency pitch period estimate, the fundamental frequency f_0 can be estimated from the narrowband input signal.

25 In some embodiments the fundamental frequency corrector 213 can be configured to perform an initial determination or decision on the consistency of the fundamental frequency estimate f_0 . For example in some embodiments the f_0 corrector 213 can be configured to compare the current fundamental frequency estimate to a previous fundamental frequency estimate and furthermore evaluate
30 the range of variation of fundamental frequency values within a determined number of previous frames. In some embodiments the fundamental frequency corrector 213 can be configured to generate an initial smoothed long term estimate or long term average of the fundamental frequency.

In some embodiments the long term average can be determined by using a first order recursive filter where the smoothed estimate can be updated in some embodiments dependent on whether or not the frame has been classified as being voiced or non-voiced. For example, in some embodiments the fundamental frequency corrector 213 can be configured to receive from the feature calculator 205 a value of the active speech level for the current frame to assist in determining whether or not the current frame is voiced or non-voiced. The fundamental frequency corrector 213 can thus using the feature based attenuation factors, the consistency of the fundamental frequency estimate and the comparison of the frame energy with the noise floor and the active speech level estimate perform a classification of the frame.

Short term octave errors can then be detected and corrected based on the assumption that the fundamental frequency contour is continuous.

In other words the fundamental frequency corrector 213 can be configured to double the fundamental frequency estimate, in other words the estimated f_0 is corrected to be $2f_0$ when the current frame is classified as voiced speech, the corrected estimate is close to the previous frame fundamental frequency estimate and the corrected fundamental frequency estimate is closer to the long term fundamental frequency estimate.

Similarly the fundamental frequency corrector 213 can be configured to halve the fundamental frequency estimate, in other words the current estimate f_0 is corrected to $0.5 f_0$ when the current frame is classified as voiced speech, the current frame corrected estimate $0.5 f_0$ is closer to the previous frame fundamental frequency estimate and the corrected fundamental frequency estimate is close to the long term fundamental frequency estimate.

Furthermore other short term deviations in the fundamental frequency estimate can be allowed for in the fundamental frequency corrector 213 by replacing the estimated fundamental frequency $f_0(n)$ by a corrected estimate from a previous frame $f_0(n-1)$ when the current frame is classified as voiced, the current fundamental frequency deviates greatly from the previous frame fundamental

frequency estimate, and the previous frame fundamental frequency estimate is closer to the long term estimate.

5 In some embodiments the fundamental frequency corrector 213 can be configured to perform such modifications to the fundamental frequency to a small number of successive frames only. In other words, should the fundamental frequency estimator corrector 213 determine that the correction has to be applied to a number of frames greater than a determined threshold then the fundamental frequency corrector performs a further change to re-correct or edit the fundamental frequency
10 estimate values back to the original estimated value.

The fundamental frequency corrector 213 can furthermore output the corrected fundamental frequency values in some embodiments to a fundamental frequency estimate attenuation factor generator 219 and to the amplitude and phase
15 calculator 221.

The decoder 108 in some embodiments can comprise a fundamental frequency attenuator generator 219. The fundamental frequency attenuator generator 219 is configured to generate at least one attenuation or gain factor that can be used to
20 attenuate the artificial bandwidth extension low band output depending on the reliability of the fundamental frequency estimate. In other words where a highly variable fundamental frequency estimate is determined and considered not to be reliable (and therefore unlikely to be correct) the artificial bandwidth extension low band synthesis for such frames should be attenuated in order to prevent incorrect
25 low band energy being heard by the user. The consistency or reliability of the fundamental frequency estimate can be determined by comparing the fundamental frequency estimate for the current frame against the estimate of at least one previous frame and evaluating the range of variation of fundamental frequency estimates. Where a small variation of fundamental frequency estimate is
30 determined there is a high likelihood of consistent estimates.

The fundamental frequency attenuator generator can in some embodiments thus generate a binary attenuation factor g_{f0} to silence or mute the low band output when the fundamental frequency estimate f_0 is considered to be unreliable.

Furthermore “downward” octave errors in the fundamental frequency estimate have occasionally been observed especially when female speech, in particular where the voice is determined to be “creaky”. In order to reduce artefacts generated by these fundamental frequency error estimates the artificial bandwidth extension low band can be muted where the fundamental frequency estimate is lower than an adaptive threshold value. For example in some embodiments an updated long term estimate of fundamental frequency values f_0 can be calculated or determined from the corrected fundamental frequency f_0 values in frames classified as voiced speech. Furthermore a lower limit for an acceptable fundamental frequency is set at, for example 70% of the long term estimate and the fundamental frequency attenuator can generate an attenuation factor g_l so that the low band output is muted when the current frame fundamental frequency estimate is below this limit.

In order that transitions are smooth, in some embodiments a transition range of a few Hz can be defined around the threshold from complete muting to no attenuation.

These attenuation factors can then be passed to an attenuation amplifier 229.

The determination of at least one fundamental frequency based attenuation factor is shown in Figure 4 by step 321.

In some embodiments the decoder 108 comprises an artificial bandwidth extension low band energy predictor 215 or any suitable means for determining an estimated bandwidth extension energy level. The low band energy predictor 215 can be configured to produce an estimate of the low band energy required in order to synthesis the low band signal. In some embodiments the low band energy estimate can be determined or produced by using statistical techniques using training data derived from wideband speech recordings. In some embodiments the seven spectral feature values calculated from the narrowband input speech and output by the filterbank 211 can be used as an input to the low band energy estimator.

For example the training data can be any suitable speech database or part of speech database. An example database of speech is "SPEECON - speech database for consumer devices: database specification and validation" published in Proceedings of the Third International Conference on Language Resources and Evaluation (LREC), 2002, pages 329 to 333 by D. Iskra et al. Similarly any suitable training method can be implemented in some embodiments.

In some embodiments the speech database can be used to train the low band energy estimator by high pass filtering the database signals to simulate the input response of a mobile terminal and to generate a suitable narrowband training signal and scale the filtered values to a level of -26dBov. The filtered and scaled samples can then in some embodiments be coded and decoded using a suitable adaptive multi-rate (AMR) narrowband speech codec. The signals can then be split into frames and the associated spectral features as described earlier generated. For example the database signals a series of seven log compressed sub-band energy feature values as described earlier can be extracted and from these sub-band energy values the associated low band energy values stored also for later use. The lowband energy values are calculated from the same original signals but without highpass filtering in such embodiments as filtering would remove the lowband information. The lowband is not included in the 7 sub-bands that are used as input features.

In some other embodiments, other training samples can be processed in order to permit the low-band energy levels to be calculated. For example in some embodiments the speech samples can be scaled with an equivalent scaling factor as the samples for input feature calculation but without the use of a high pass filtering or adaptive multi-rate coding.

The associated low band energy values in some embodiments can be calculated through applying a 128 point Fast Fourier Transform (FFT) and using a trapezoidal filter window applied to the power spectrum to extract the low band energy from the database signals. The filter window in such embodiments can for example have a flat unit gain from 81 Hz to 272 Hz, with the trapezoid tail extending from 0 Hz to 81 Hz and 272 Hz to 385 Hz and the upper -3dB point at 330 Hz. In such

embodiments a logarithmic mapping of the form $\log(x+1)$ can be used to log compress the low band energy values.

In such embodiments a Gaussian mixture model (GMM) with ten components can
5 be trained using the data from the database to model the joint probability distribution of the log compressed low band energy of a current frame, the log compressed sub-band energy features of the current frame and two proceeding frames. However it would be understood more than or fewer than ten components can be used in some embodiments. In other words: denoting the input spectral
10 features (log-compressed sub-band energies) of the current frame and two preceding frames by x and the log-compressed lowband energy of the current frame by y the GMM models the joint distribution of x and y . The model can be used to estimate the log compressed low band energy from the input features using the minimum mean square error (MMSE) estimate. In other words the model
15 can be used to get the MMSE estimate of y when having observed x .

In such embodiments the GMM predictor utilised in this example can be similar to those described for high band artificial bandwidth extension. Although this example describes the implementation of the low band prediction energy estimate
20 being formed using a Gaussian mixture model, any suitable pattern recognition model or modelling function or means could be implemented, for example a neural network or a Hidden Markov Model (HMM). Furthermore although the features used as an input feature set are those of the spectral features generated from the filterbank 211 any suitable input feature set could be used in addition to or to
25 replace the spectral features used in this example.

In some embodiments the energy estimates are calculated for every 10 ms frame and a linear interpolation between two successive estimates can be used to generate an estimate for every 5 ms sub-frame. In such embodiments where the
30 spectral features and thus lowband energy estimates are determined every 5ms no interpolation operation is required.

The output of the Gaussian mixture model predictor $y(n)$ for frame n can then be converted to the energy estimate $E_{lb}(n)$ by reversing the log compression

$$E_{lb}(n) = e^{y(n)} - 1.$$

- 5 The output of the low band energy predictor 215 can be passed to a harmonic amplitude estimator 217.

The operation of determining the low band energy estimate is shown in Figure 4 by step 315.

10

In some embodiments the decoder 108 comprises a harmonic amplitude estimator 217 or means for determining a harmonic shaping function. The harmonic amplitude estimator 217 is configured to determine or generate estimates of the amplitudes of the artificial bandwidth extension low band harmonics dependent on
15 the low band energy estimate. As conventional low band determinations have occasionally produced short peaks of high or too high values which can be heard as momentary artefacts, the harmonic amplitude estimator 217 can perform an adaptive compression of the low band energy estimates.

- 20 Furthermore in some embodiments the harmonic amplitude estimator 217 can apply a logarithmic compression curve to the energy estimates that exceed the smoothed contour by greater than a determined amount. For example in some embodiments the logarithmic compression can be applied to energy estimates which exceed the smoothed contour by a factor greater than 150%.

25

The sinusoidal components or single frequency components in the low band are generated in some embodiments up to a frequency of 400 Hz. In some embodiments the harmonic amplitude estimator generates an indicator or range of harmonic indicators whereby initially all of the harmonics to be generated have
30 equal amplitudes. The amplitude of the sine waves generated in the synthesis generator is set such that the energy estimate of the low band is approximately realised. For example the harmonic amplitude estimator can generate an amplitude using the following equation:

$$A_e = \sqrt{kE_{lb}(n) \left(\frac{f_0}{330 \text{ Hz}} - 0.5 \right)},$$

where A_e is the amplitude, the constant k represents the effects of windowing, Fast Fourier Transform and filtering in the computation of the low band energy such that

- 5 a single sine wave with the amplitude $\sqrt{kE_{lb}(n)}$ can yield the low band energy of $E_{lb}(n)$. The term in brackets can adjust the amplitude such that the total energy of the harmonics generated in the low frequency extension band approximately matches the estimated low band energy.

- 10 However in some embodiments the harmonic amplitude estimator 217 can then be configured to apply a frequency dependent attenuation or generate an attenuation profile or function so to provide a smooth transition from the low frequency extension band to the telephone band. The profile or function can be passed in some embodiments to the synthesis amplitude calculator.

15

The generation of the harmonic amplitude profile is shown in Figure 4 by step 317.

In some embodiments the decoder comprises an input amplitude and phase calculator 221.

20

The input amplitude and phase calculator 221 or means for determining at least one amplitude value and phase value dependent on a first audio signal in some embodiments determines an amplitude for the artificial bandwidth extension low band which is dependent on the fundamental frequency estimate and the low band analysis framed narrowband audio signal (the first audio signal). This is because the number of harmonic components within the low band can vary dependent on the fundamental frequency.

25

- 30 The input amplitude and phase calculator in some embodiments analyses the input narrowband signal in 5 ms steps using a segment length of 20 milliseconds and a look ahead of 5 ms, where each segment has been windowed with a Hann window. The amplitude and phase at the frequency of each multiple of the

estimated fundamental frequency can then be analysed, according to the following equation:

$$S(n, l) = \sum_{m=0}^{N-1} s_{n,w}(m) e^{-2\pi i \frac{lf_0}{f_s} m},$$

5

where N is the length of the segment to be analysed, $S_{n,w}$ is the windowed signal segment for frame n and f_s is the sampling frequency. In other words, this analysis can be considered to be a Discrete Fourier Transform of the input signal computed for only a few specific frequencies $lf_0(n)$ below 400 Hz. In some other
 10 embodiments a computation of Fast Fourier Transform of sufficient lengths that the frequency bins corresponding to the harmonic frequencies can be extracted can be implemented.

In such embodiments the input amplitude and phase calculator 221 can then
 15 generate an amplitude for the l'th harmonic in the input signal as:

$$A(n, l) = c_A |S(n, l)|,$$

where C_A is a constant which compensates for the effects of the segment length
 20 and windowing such that $A(n, l)$ represents the amplitude of the partial.

The input amplitude and phase calculator can then pass the "A" value to the synthesis amplitude calculator 223 for further processing.

25 The input amplitude and phase calculator 213 can furthermore generate an observed phase of the l'th harmonic for-frame n of the input signal using the following equation:

$$\varphi(n, l) = \arg(S(n, l)),$$

30

where arg is the argument of the value. The observed phase values can then be passed to the synthesis phase calculator 225.

The operation of generating an initial or observed amplitude and phase value is shown in Figure 4 by step 323.

5 In some embodiments the decoder 108 comprises a synthesis amplitude calculator 223, or means for synthesizing a further amplitude value. The synthesis amplitude calculator is configured to receive the input amplitude estimate, harmonic amplitude estimate and corrected f_0 estimate and determine at least one single frequency component or sinusoid amplitude value.

10 Thus in some embodiments the synthesis amplitude calculator 217 uses a first order recursive filter to smooth the fundamental frequency estimates for consecutive frames and thus reduce a rapid variation of sine wave amplitudes.

15 As described previously the output of the low band predictor 215 generates a single low band energy estimate produced from the predictor (such as the Gaussian mixed model predictor). Furthermore all of the low band harmonic partials can be determined or generated with equal amplitudes such that the energy estimate is approximately realised. This approach has been evaluated by
20 replacing the low band harmonics of a wideband speech signal by sinusoidal or single frequency components with correct frequencies but using the amplitude of the first partial for all low band harmonics. In such embodiments during informal listening evaluations, only a slight difference was noticed in comparison to a signal with correct frequencies and amplitudes of low band harmonics.

25

In some embodiments, frequency dependent attenuation can be applied to the amplitudes A to provide a smooth transition from the extension band to the telephone band. In principle, the synthesis low band signal can smoothly extend the spectrum of the telephone band signal. However in practice, the detailed low
30 cut characteristics of the telephone connection are generally unknown and can vary largely from case-to-case. In such embodiments the low band synthesis should ideally be adjusted to the frequency characteristics of the narrowband signal but can in some embodiments and for simplicity use a fixed transition.

In such embodiments the upper end of the extension band can apply a gradual transition from the extension band to the telephone band by limiting the synthesis amplitudes relative to the observed amplitudes of the harmonics. Thus in some embodiments the amplification of observed harmonics is limited between 250 Hz and 400 Hz using a smooth curve that approaches infinity at 250 Hz and approximately 10dB at 300 Hz and 0dB at 400 Hz. However it would be appreciated that any suitable filtering approach could be implemented.

In some embodiments the synthesis amplitude calculator can further take into account the observed low band harmonics of the input signal when synthesising the low band such that the sum of the input signal and synthesised signal approximately produces the estimated amplitude for the harmonic partials. The amplitude for the synthesis of each harmonic is computed, for example by subtracting the observed harmonic amplitude from the limited target amplitude if the target amplitude exceeds the observed amplitude. For example where the observed amplitude is larger, no synthetic signal is generated.

Furthermore in some embodiments the input amplitude and phase calculator 221 can apply a smoothing filter to the harmonic amplitudes to reduce the rapid variation in the extension band signal.

In some embodiments the decoder 108 comprises a synthesis phase calculator or means for synthesising a further phase value. The synthesis phase calculator 225 can be configured to receive an initial phase observation from the input amplitude and phase calculator and further receive a fundamental frequency estimate from the fundamental frequency corrector 213. The synthesis phase calculator 225 can use the observed phase from the input signal when it is considered to be reliable and consistent.

The harmonics may be attenuated in the input signal (due to the transmission chain or the transmitting device, for example) but the phase information can be detected reliably. In such embodiments it can be beneficial to use the observed phase to maximise the quality of the output signal. However in these embodiments if or when the phase of the frequency of the l 'th harmonic is lost due to the speech

transmission chain generating a continuous phase from frame-to-frame can be implemented.

A reference phase value $\varphi_{\text{ref}}(n, l)$ can thus be generated by the synthesis phase calculator 225 using harmonic values for each frame n and harmonic l from the previous synthesis phase values $\tilde{\varphi}(n-1, l)$ using the estimates of the fundamental frequency for the previous and current frames $f_0(n-1)$ and $f_0(n)$ and assuming phase continuity at the frame boundary in the middle of the overlapping region. The difference $\delta(n, l)$ between the observed phase and reference phase can be calculated according to the following equation:

$$\delta(n, l) = \varphi(n, l) - \varphi_{\text{ref}}(n, l)$$

and wrapped to the range $-\pi$ to $+\pi$. Furthermore in some embodiments the synthesis phase calculator 225 can determine the difference between successive values of the difference $\delta(n, l)$ according to the following equation:

$$\Delta\delta(n, l) = \delta(n, l) - \delta(n-1, l)$$

which can also be wrapped within the range $-\pi$ to $+\pi$.

The synthesis phase calculator 225 can then apply a series of following rules within which the synthesis phase of the l 'th harmonic in each frame n can be determined by the first matching condition of the following list. In other words the synthesis phase calculator or means for synthesising a further phase value associated with each phase value can therefore be considered to comprise in at least one embodiment a condition determiner or means for determining a condition associated with each phase value; and also a further phase generator or means for generating a further phase value dependent on the condition and the phase value.

Therefore the synthesis phase calculator 225 is configured to perform the following operation in order 1-5 and set the phase on finding a first matching operation.

1. When the observed phase of the l 'th harmonic is highly varying, the observed phase information in the frequency range of this harmonic is considered unreliable and a continuous phase contour is generated for synthesis. For example in some embodiments the phase variability can be assessed by
 5 generating an expected phase angle $\varphi_e(n, l)$ which can be determined from the observed phase $\varphi(n-2, l)$ and the estimated fundamental frequency values $f_0(n-2)$, $f_0(n-1)$, and $f_0(n)$. A phase error between the expected and observed phase $\varphi_e(n, l) - \varphi(n, l)$ can then be determined wrapped within the range $-\pi$ to $+\pi$ and smoothed in time using a recursive filter. In such embodiments the current value of
 10 the smoothed phase error is compared with a fixed threshold value. When the threshold is exceeded, the phase is considered to fluctuate too wildly and the continuous phase contour is used. In other words in some embodiments the output synthesis phase is determined as:

$$15 \quad \tilde{\varphi}(n, l) = \varphi_{\text{ref}}(n, l)$$

2. In some embodiments on signal onset the observed phase can be implemented or used. In such embodiments the low band energy estimate is compared against its smoothed copy from the previous frame or frames other than
 20 the current frame. For example in some embodiments the synthesis phase calculator determines when the previous energy estimate has a low relative value and the current value has a sufficiently high relative value to use the observed phase value. In other words:

$$25 \quad \tilde{\varphi}(n, l) = \varphi(n, l)$$

3. In some embodiments the synthesis phase calculator can be configured such that when the phase mismatch between the observed phase and the continuous reference phase is small then the observed phase is used. In some
 30 embodiments this difference within which the observed phase is used can be $\pi/8$. In which the synthesis phase calculator outputs the following:

$$\tilde{\varphi}(n, l) = \varphi(n, l)$$

4. In some embodiments the synthesis phase calculator can determine when there is a mismatch between the observed phase and reference values but the observed phase is consistent in successive frames then the observed phase is approached gradually. For example the synthesis phase calculator 225 can determine when:

$$\max_{j=0 \dots 3} |\Delta \delta(n-j, l)| < \frac{\pi}{8},$$

- 10 and then generate an output phase of:

$$\tilde{\varphi}(n, l) = \varphi_{\text{ref}}(n, l) \pm \frac{\pi}{8},$$

- where the sign is chosen such that the output phase $\tilde{\varphi}(n, l)$ is closer to $\varphi(n, l)$ than $\varphi_{\text{ref}}(n, l)$.

5. In some embodiments the synthesis phase calculator 225 can be configured to output the reference phase when determining that the observed phase of the harmonic partial in question is inconsistent from frame-to-frame. In other words outputting a low band synthesis value based only on the criteria of the phase continuity at the frame boundary. Thus the synthesis phase calculator 225 in such embodiments can output the following phase value:

$$\tilde{\varphi}(n, l) = \varphi_{\text{ref}}(n, l)$$

25

The operation of generating at least one synthesised phase value, each associated with a synthesised harmonic, is shown in Figure 4 by step 327.

- In some embodiments the decoder comprises a sine synthesiser 227. The sine synthesiser can receive the outputs of the synthesis amplitude calculator 223, the synthesis phase calculator 225 and also the corrected fundamental frequency estimate from the fundamental frequency corrector 213 and generate the artificial

30

bandwidth extension from the harmonics formed from sinusoidal signal (or as seen from the frequency domain single frequency component). In some embodiments this can be represented by the following equation:

$$s_s(n, k) = \sum_l \tilde{A}(n, l) \cos\left(2\pi n \frac{l f_0(n)}{f_s} + \tilde{\varphi}(n, l)\right),$$

where k is the time index within the frame n , l is the index of the harmonic from f_0 up to 400Hz where the number of harmonics being determined by the value of f_0 .

10 The output signal can then be passed to an attenuator amplifier 229.

The generation of the synthesized artificial bandwidth signal is shown in Figure 4 by step 329.

15 The attenuation amplifier 229 can receive the output from the sinusoidal synthesiser 227 and the attenuation factors from the time domain attenuator 209 and the fundamental frequency based attenuator 219 to generate an attenuated or amplified, in other words synthesised frames are then multiplied by the attenuation factors g_{gi} , g_p , g_{f0} , and g_l . The output of the attenuation amplifier 229 can then be
20 passed to the overlap adder 231.

The operation of performing the attenuation amplification is shown in Figure 4 by step 331.

25 In some embodiments the decoder 108 comprises an overlap adder 231 configured to window the output artificial bandwidth extension low band signal with a 10 ms Hann window and add overlaps to get a continuous low band signal with smooth transitions between adjacent frames. The output s_{lb} can then be passed to the full band summer configured to receive both the narrowband s_{nb} and band
30 extension s_{lb} and output a full band signal s_{output} . The full band addition is shown in Figure 4 by step 335.

In such embodiments the low band extension can be determined by using existing signals at narrowband frequencies and adapting to different passband characteristics closer to the lower end of the telephone band. The algorithmic delay of such an embodiment is relatively low (a few ms in addition to the framing
5 delay) and furthermore by combining the low band bandwidth extension with artificial bandwidth extension to frequencies above the telephone band, a more balanced and natural speech spectrum can be developed than using the narrowband signal. In other words by using both low band and high band artificial bandwidth extension, a total bandwidth which is close to the bandwidth of
10 wideband telephone speech transmitted by an adaptive multi-rate wideband codec (AMR-WB) can be achieved.

For example as shown in Figures 5, 6 and 7, a series of simulated bandwidths are shown for the adaptive multi-rate wideband codec, a narrowband codec, a
15 narrowband plus high band artificial bandwidth extension and a narrowband with both low band and high band artificial bandwidth extension are shown.

Figure 5 for example shows the relative performance for narrowband, adaptive multi rate-wide band, high band artificial bandwidth extension, and low band + high
20 band artificial bandwidth extension for a voiced male speech short segment example wherein the lowband artificial bandwidth extension simulated signal performance is significantly improved over the narrowband signal.

Figure 6 furthermore shows the relative performance for narrowband, Adaptive
25 Multi-Rate Wideband, and low band extension + narrowband for the voiced male speech example shown in Figure 5, further demonstrating that lowband extension performs only slightly worse than the AMR-WB codec.

Figure 7 shows a further example of the relative performance characteristics in
30 long-term average spectra shown by narrowband, adaptive multi-rate wideband speech coding, and artificial bandwidth extension decoding where once again the lowband artificial bandwidth extension performance is significantly better than narrowband and only slightly worse than AMR-WB.

Although the above examples describe embodiments of the invention operating within a codec within an electronic device 10 or apparatus, it would be appreciated that the invention as described below may be implemented as part of any audio decoding process. Thus, for example, embodiments of the application may be
5 implemented in an audio decoder which may implement lowband artificial bandwidth extension.

Thus user equipment may comprise a bandwidth extender such as those described in embodiments of the invention above.

10

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.

15 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.
20 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
25 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.

30

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.

- 5 Therefore in summary at least one embodiment of the invention comprises an apparatus configured to: determine at least one amplitude value and phase value dependent on a first audio signal; synthesize a further amplitude value associated with each amplitude value dependent on a determined harmonic shaping function; synthesize a further phase value associated with each phase value; and generate
10 a bandwidth extension signal dependent the further amplitude value and the further phase values.

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
15 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) and processors
20 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
25 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
30 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.

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 amplitude value and phase value dependent on a
5 first audio signal;
synthesising a further amplitude value associated with each amplitude value
dependent on a determined harmonic shaping function;
synthesising a further phase value associated with each phase value; and
generating a bandwidth extension signal dependent the further amplitude
10 value and the further phase values.
2. The method as claimed in claim 1, further comprising generating at least
one attenuation factor, wherein the bandwidth extension signal is further
dependent on the attenuation factor.
15
3. The method as claimed in claim 2, wherein the at least one attenuation
factor comprises an unvoiced signal attenuation factor, the unvoiced signal
attenuation factor being dependent on an unvoiced component of the first audio
signal.
20
- 4 The method as claimed in claims 2 to 3, wherein the at least one attenuation
factor comprises a pause attenuation factor, the pause attenuation factor being
dependent on determining a paused speech component of the first audio signal.
- 25 5. The method as claimed in claims 2 to 4, wherein the at least one attenuation
factor comprises a fundamental frequency attenuation factor, the fundamental
frequency attenuation factor being dependent on a fundamental frequency
estimate associated with the first audio signal.
- 30 6. The method as claimed in claims 2 to 5, wherein the at least one attenuation
factor comprises an octave error attenuation factor, the octave error attenuation
factor being dependent on determining an error in a fundamental frequency
estimate associated with the first audio signal.

7. The method as claimed in claims 1 to 6, further comprising determining a harmonic shaping function dependent on the estimated bandwidth extension signal energy level.
- 5 8. The method as claimed in claim 7, further comprising determining an estimated bandwidth extension signal energy level.
9. The method as claimed in claim 8, wherein determining an estimated bandwidth extension signal energy level comprises:
- 10 determining at least one feature value associated with the first signal; and
applying the at least one feature to a trained modelling function to determine the estimated bandwidth extension signal energy level.
10. The method as claimed in claim 9, wherein the modelling function comprises
15 at least one of:
a Gaussian mixture model;
a hidden Markov model; and
a neural network model.
- 20 11. The method as claimed in claims 1 to 10, wherein synthesising the further amplitude value associated with each amplitude value is further dependent on the first audio signal.
12. The method as claimed in claim 11, wherein synthesising the further
25 amplitude value associated with each amplitude value further dependent on the first audio signal comprises:
determining the amplitudes of the first audio signal for the harmonic frequencies associated with the first amplitude; and
synthesising the further amplitude dependent on the amplitudes of the first
30 audio signal for the harmonic frequencies associated with the first amplitude.
13. The method as claimed in claim 1 to 12, wherein synthesising a further phase value associated with each phase value comprises performing:
determining a condition associated with each phase value;

generating a further phase value dependent on the condition and the phase value.

14. The method as claimed in claim 13, wherein determining the condition
5 associated with the phase value comprises:

determining the phase value is highly varying, wherein the further phase value is a reference phase value;

determining the onset of the phase value, wherein the further phase value is the reference phase value;

10 determining the phase value is sufficiently close to the reference phase value, wherein the further phase value is the phase value;

determining the phase value is different from the reference phase value and the phase value is consistent over a period of time, wherein the further phase value is phase value approaching the phase value from the reference phase value; and

15 otherwise determining the phase value is inconsistent over a period of time, wherein the further phase value is the reference phase value.

15. The method as claimed in claim 14, wherein the reference phase value is dependent on a previous period further phase value and the fundamental
20 frequency estimates from the current and previous periods.

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

determining at least one amplitude value and phase value dependent on a first audio signal;

synthesising a further amplitude value associated with each amplitude value dependent on a determined harmonic shaping function;

30 synthesising a further phase value associated with each phase value; and

generating a bandwidth extension signal dependent the further amplitude value and the further phase values.

17. The apparatus as claimed in claim 16, further configured to perform generating at least one attenuation factor, wherein the bandwidth extension signal is further dependent on the attenuation factor.
- 5 18. The apparatus as claimed in claim 17, wherein the at least one attenuation factor comprises an unvoiced signal attenuation factor, the unvoiced signal attenuation factor being dependent on an unvoiced component of the first audio signal.
- 10 19. The apparatus as claimed in claims 17 to 18, wherein the at least one attenuation factor comprises a pause attenuation factor, the pause attenuation factor being dependent on determining a paused speech component of the first audio signal.
- 15 20. The apparatus as claimed in claims 17 to 19, wherein the at least one attenuation factor comprises a fundamental frequency attenuation factor, the fundamental frequency attenuation factor being dependent on a fundamental frequency estimate associated with the first audio signal.
- 20 21. The apparatus as claimed in claims 17 to 20, wherein the at least one attenuation factor comprises an octave error attenuation factor, the octave error attenuation factor being dependent on determining an error in a fundamental frequency estimate associated with the first audio signal.
- 25 22. The apparatus as claimed in claims 16 to 21, further configured to perform determining a harmonic shaping function dependent on the estimated bandwidth extension signal energy level.
23. The apparatus as claimed in claim 22, further configured to perform
30 determining an estimated bandwidth extension signal energy level.
24. The apparatus as claimed in claim 23, wherein determining an estimated bandwidth extension signal energy level causes the apparatus to further perform:
determining at least one feature value associated with the first signal; and

applying the at least one feature to a trained modelling function to determine the estimated bandwidth extension signal energy level.

25. The apparatus as claimed in claim 24, wherein the modelling function
5 comprises at least one of:

- a Gaussian mixture model;
- a hidden Markov model; and
- a neural network model.

10 26. The apparatus as claimed in claims 16 to 25, wherein synthesising the further amplitude value associated with each amplitude value is further dependent on the first audio signal.

27. The apparatus as claimed in claim 26, wherein synthesising the further
15 amplitude value associated with each amplitude value further dependent on the first audio signal further causes the apparatus to perform:

determining the amplitudes of the first audio signal for the harmonic frequencies associated with the first amplitude; and

20 synthesising the further amplitude dependent on the amplitudes of the first audio signal for the harmonic frequencies associated with the first amplitude.

28. The apparatus as claimed in claim 16 to 27, wherein synthesising a further phase value associated with each phase value causes the apparatus to perform:

25 determining a condition associated with each phase value;
generating a further phase value dependent on the condition and the phase value.

29. The apparatus as claimed in claim 28, wherein determining the condition associated with the phase value causes the apparatus to further perform:

30 determining the phase value is highly varying, wherein the further phase value is a reference phase value;

determining the onset of the phase value, wherein the further phase value is the reference phase value;

determining the phase value is sufficiently close to the reference phase value, wherein the further phase value is the phase value;

- 5 determining the phase value is different from the reference phase value and the phase value is consistent over a period of time, wherein the further phase value is phase value approaching the phase value from the reference phase value; and
otherwise determining the phase value is inconsistent over a period of time, wherein the further phase value is the reference phase value.

- 10 30. The apparatus as claimed in claim 29, wherein the reference phase value is dependent on a previous period further phase value and the fundamental frequency estimates from the current and previous periods.

31. Apparatus comprising:
means for determining at least one amplitude value and phase value
15 dependent on a first audio signal;
means for synthesising a further amplitude value associated with each amplitude value dependent on a determined harmonic shaping function;
means for synthesising a further phase value associated with each phase value; and
20 means for generating a bandwidth extension signal dependent the further amplitude value and the further phase values.

32. The apparatus as claimed in claim 31, further comprising means for generating at least one attenuation factor, wherein the bandwidth extension signal
25 is further dependent on the attenuation factor.

33. The apparatus as claimed in claim 32, wherein the at least one attenuation factor comprises an unvoiced signal attenuation factor, the unvoiced signal attenuation factor being dependent on an unvoiced component of the first audio
30 signal.

34. The apparatus as claimed in claims 32 to 33, wherein the at least one attenuation factor comprises a pause attenuation factor, the pause attenuation

factor being dependent on determining a paused speech component of the first audio signal.

35. The apparatus as claimed in claims 32 to 34, wherein the at least one
5 attenuation factor comprises a fundamental frequency attenuation factor, the
fundamental frequency attenuation factor being dependent on a fundamental
frequency estimate associated with the first audio signal.

36. The apparatus as claimed in claims 32 to 35, wherein the at least one
10 attenuation factor comprises an octave error attenuation factor, the octave error
attenuation factor being dependent on determining an error in a fundamental
frequency estimate associated with the first audio signal.

37. The apparatus as claimed in claims 31 to 36, further comprising means for
15 determining a harmonic shaping function dependent on the estimated bandwidth
extension signal energy level.

38. The apparatus as claimed in claim 37, further comprising means for
determining an estimated bandwidth extension signal energy level.

20 39. The apparatus as claimed in claim 38, wherein the means for determining
an estimated bandwidth extension signal energy level comprises:

means for determining at least one feature value associated with the first
signal; and

25 means for applying the at least one feature to a trained modelling function to
determine the estimated bandwidth extension signal level.

40. The apparatus as claimed in claim 39, wherein the modelling function
comprises at least one of:

30 a Gaussian mixture model;
a hidden Markov model; and
a neural network model.

41. The apparatus as claimed in claims 31 to 40, wherein the means for synthesising the further amplitude value associated with each amplitude value is further dependent on the first audio signal.

5 42. The apparatus as claimed in claim 41, wherein the means for synthesising the further amplitude value associated with each amplitude value further dependent on the first audio signal further comprises:

means for determining the amplitudes of the first audio signal for the harmonic frequencies associated with the first amplitude; and

10 means for synthesising the further amplitude dependent on the amplitudes of the first audio signal for the harmonic frequencies associated with the first amplitude.

43. The apparatus as claimed in claims 31 to 42, wherein the means for synthesising a further phase value associated with each phase value comprises:

means for determining a condition associated with each phase value; and

means for generating a further phase value dependent on the condition and the phase value.

20 44. The apparatus as claimed in claim 43, wherein the means for determining the condition associated with the phase value comprises:

means for determining the phase value is highly varying, wherein the further phase value is a reference phase value;

25 means for determining the onset of the phase value, wherein the further phase value is the reference phase value;

means for determining the phase value is sufficiently close to the reference phase value, wherein the further phase value is the phase value;

30 means for determining the phase value is different from the reference phase value and the phase value is consistent over a period of time, wherein the further phase value is phase value approaching the phase value from the reference phase value; and

means for determining the phase value is inconsistent over a period of time, wherein the further phase value is the reference phase value.

45. The apparatus as claimed in claim 44, wherein the reference phase value is dependent on a previous period further phase value and the fundamental frequency estimates from the current and previous periods.

5 46. Apparatus comprising:

an input amplitude and phase calculator configured to determine at least one amplitude value and phase value dependent on a first audio signal;

10 a synthesis amplitude calculator configured to synthesize a further amplitude value associated with each amplitude value dependent on a determined harmonic shaping function;

a synthesis phase calculator configured to synthesize a further phase value associated with each phase value; and

a signal synthesizer configured to generate a bandwidth extension signal dependent the further amplitude value and the further phase values.

15

47. The apparatus as claimed in claim 46, further comprising an attenuator gain determiner configured to generate at least one attenuation factor, wherein the bandwidth extension signal is further dependent on the attenuation factor.

20 48. The apparatus as claimed in claim 47, wherein the at least one attenuation factor comprises an unvoiced signal attenuation factor, the unvoiced signal attenuation factor being dependent on an unvoiced component of the first audio signal.

25 49. The apparatus as claimed in claims 47 to 48, wherein the at least one attenuation factor comprises a pause attenuation factor, the pause attenuation factor being dependent on determining a paused speech component of the first audio signal.

30 50. The apparatus as claimed in claims 47 to 49, wherein the at least one attenuation factor comprises a fundamental frequency attenuation factor, the fundamental frequency attenuation factor being dependent on a fundamental frequency estimate associated with the first audio signal.

51. The apparatus as claimed in claims 47 to 50, wherein the at least one attenuation factor comprises an octave error attenuation factor, the octave error attenuation factor being dependent on determining an error in a fundamental frequency estimate associated with the first audio signal.
- 5 52. The apparatus as claimed in claims 46 to 51, further comprising a harmonic amplitude estimator configured to perform determining a harmonic shaping function dependent on the estimated bandwidth extension signal energy level.
- 10 53. The apparatus as claimed in claim 52, further comprising a lowband energy estimator configured to determine an estimated bandwidth extension signal energy level.
54. The apparatus as claimed in claim 53, wherein the lowband energy
15 estimator comprises:
a feature determiner configured to determine at least one feature value associated with the first signal; and
a trained modelling function configured to determine the estimated
bandwidth extension signal energy level dependent on the at least one feature
20 value.
55. The apparatus as claimed in claim 54, wherein the trained modelling function comprises at least one of:
a Gaussian mixture model;
25 a hidden Markov model; and
a neural network model.
56. The apparatus as claimed in claims 46 to 55, wherein the signal synthesizer configured to generate a bandwidth extension signal is further dependent on the
30 first audio signal.
57. The apparatus as claimed in claim 56, wherein the signal synthesizer configured to generate a bandwidth extension signal further comprises:

an amplitude determiner configured to determine the amplitudes of the first audio signal for the harmonic frequencies associated with the first amplitude; and wherein the synthesizer being configured to determine the further amplitude dependent on the amplitudes of the first audio signal for the harmonic frequencies associated with the first amplitude.

58. The apparatus as claimed in claim 46 to 57, wherein the synthesis phase calculator comprises:

a condition determiner configured to determine a condition associated with each phase value;

a phase synthesizer configured to generate the further phase value dependent on the condition and the phase value.

59. The apparatus as claimed in claim 58, wherein the condition determiner comprises:

a first condition determiner configured to determine the phase value is highly varying, wherein the further phase value is a reference phase value;

a second condition determiner configured to determine an onset of the phase value, wherein the further phase value is the reference phase value;

a third condition determiner configured to determine the phase value is sufficiently close to the reference phase value, wherein the further phase value is the phase value;

a fourth condition determiner configured to determine the phase value is different from the reference phase value and the phase value is consistent over a period of time, wherein the further phase value is phase value approaching the phase value from the reference phase value; and

a fifth condition determiner configured to determine the phase value is inconsistent over a period of time, wherein the further phase value is the reference phase value.

30

60. The apparatus as claimed in claim 59, wherein the reference phase value is dependent on a previous period further phase value and the fundamental frequency estimates from the current and previous periods.

61. An electronic device comprising apparatus as claimed in claims 16 to 60.
62. A chipset comprising apparatus as claimed in claims 16 to 60.

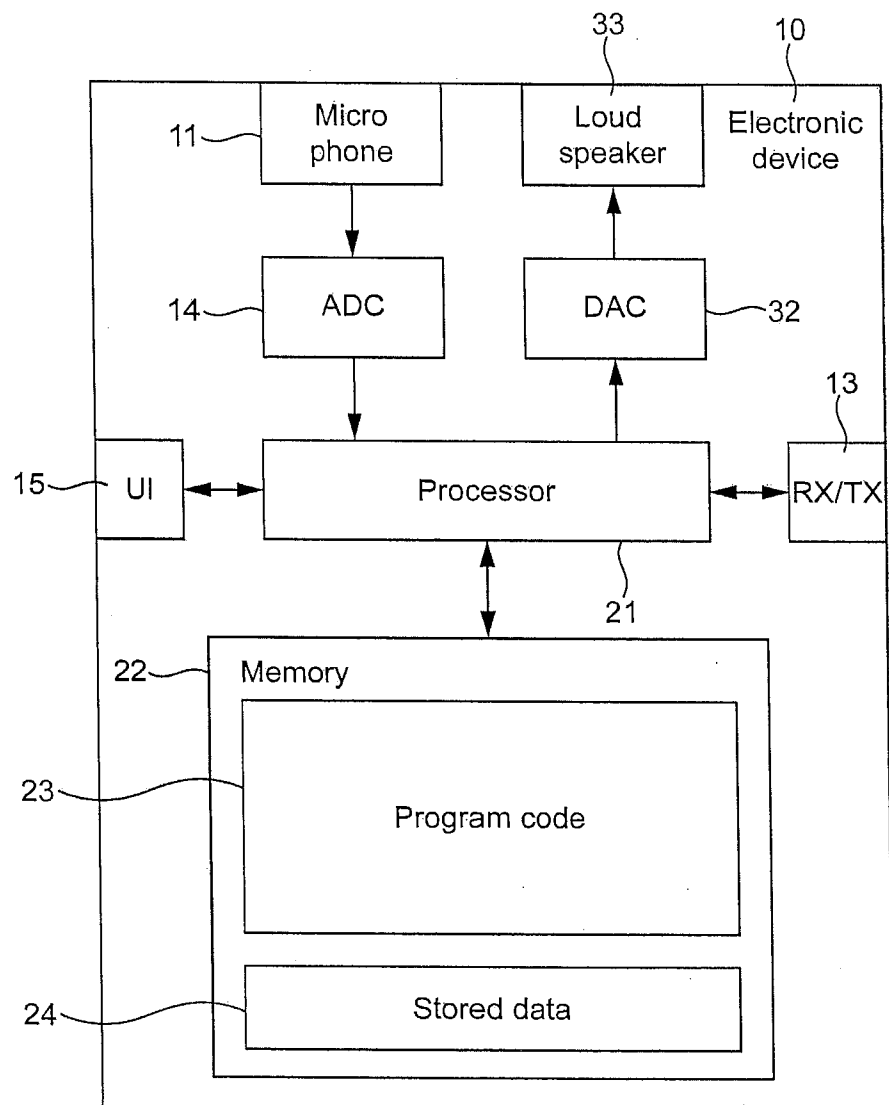


FIG. 1

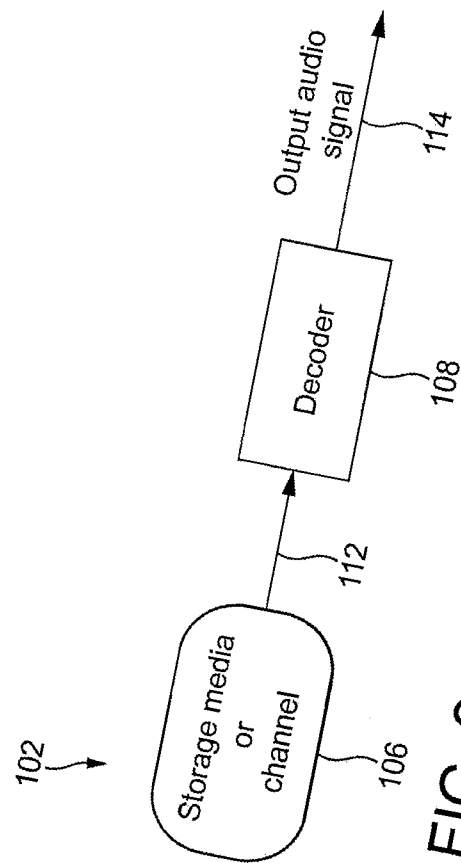


FIG. 2

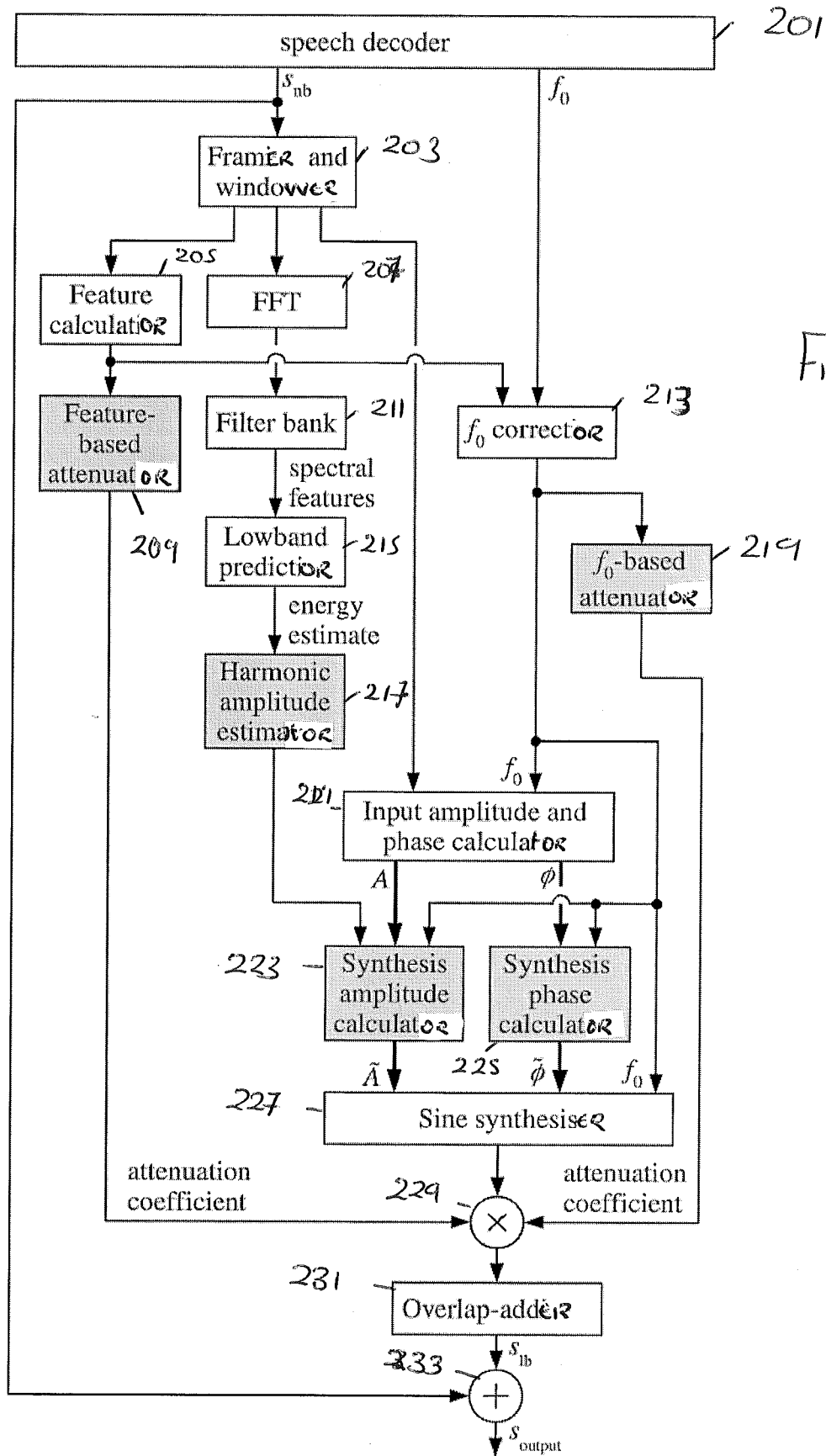


Fig 4

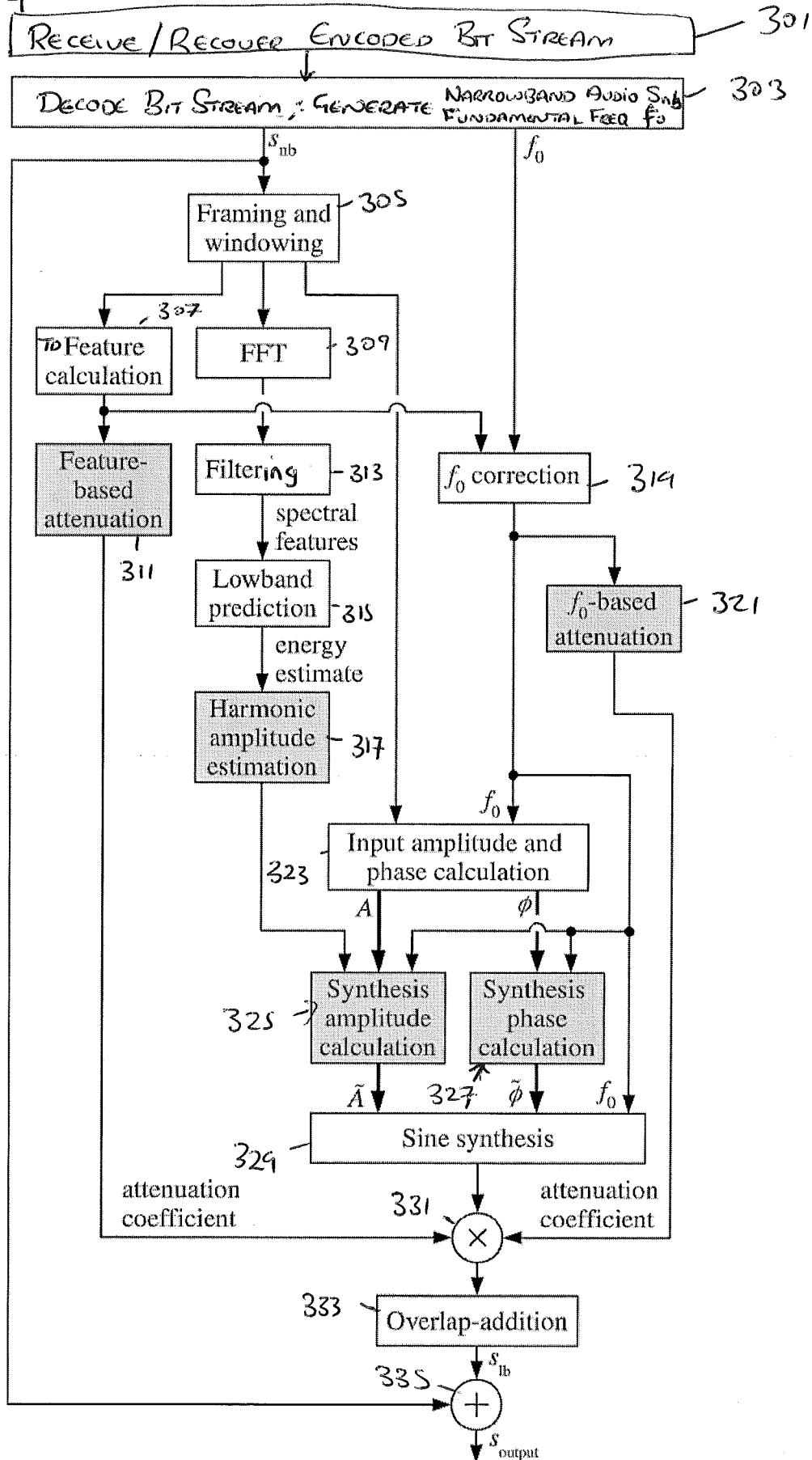


Fig 5

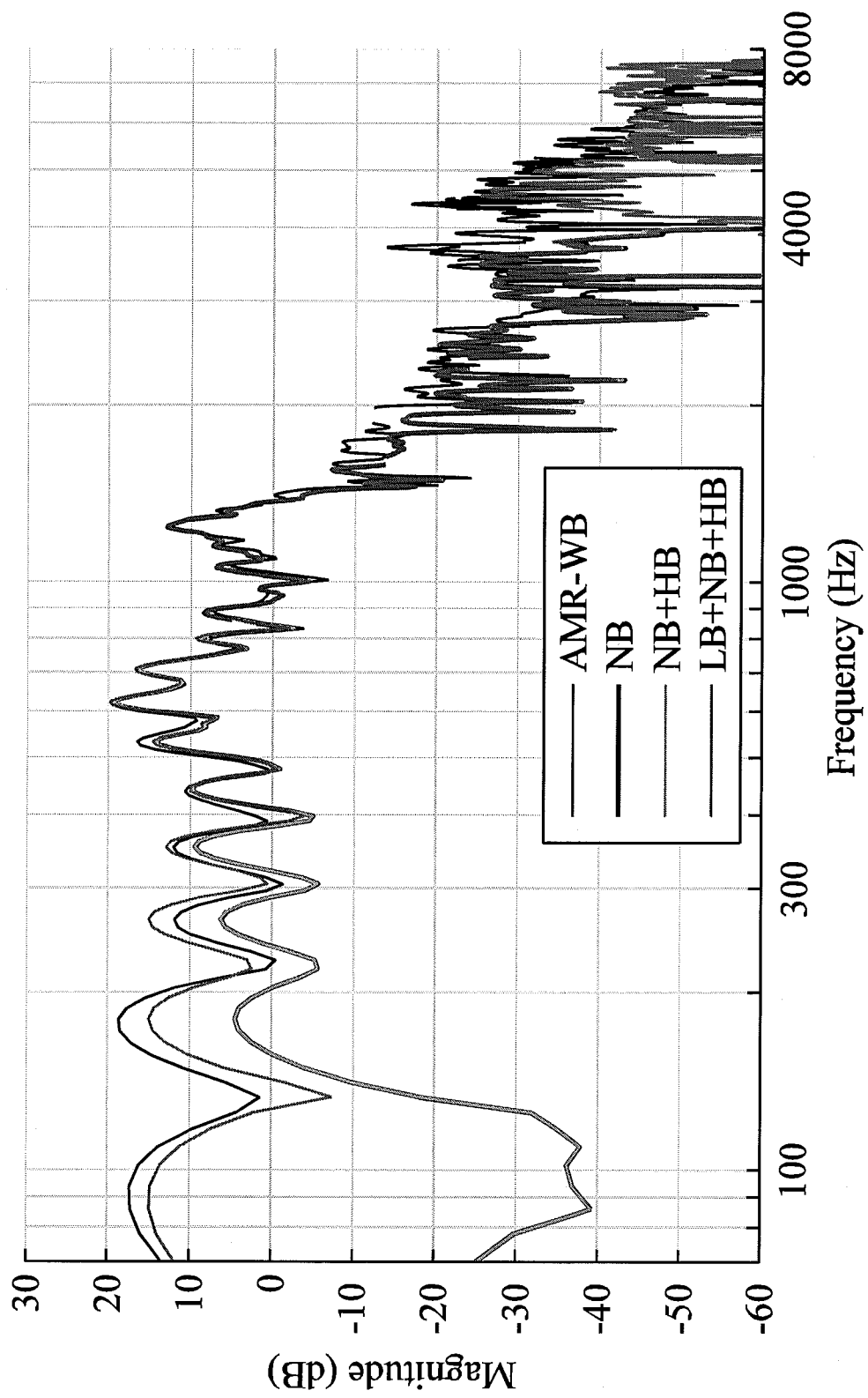


Fig 6

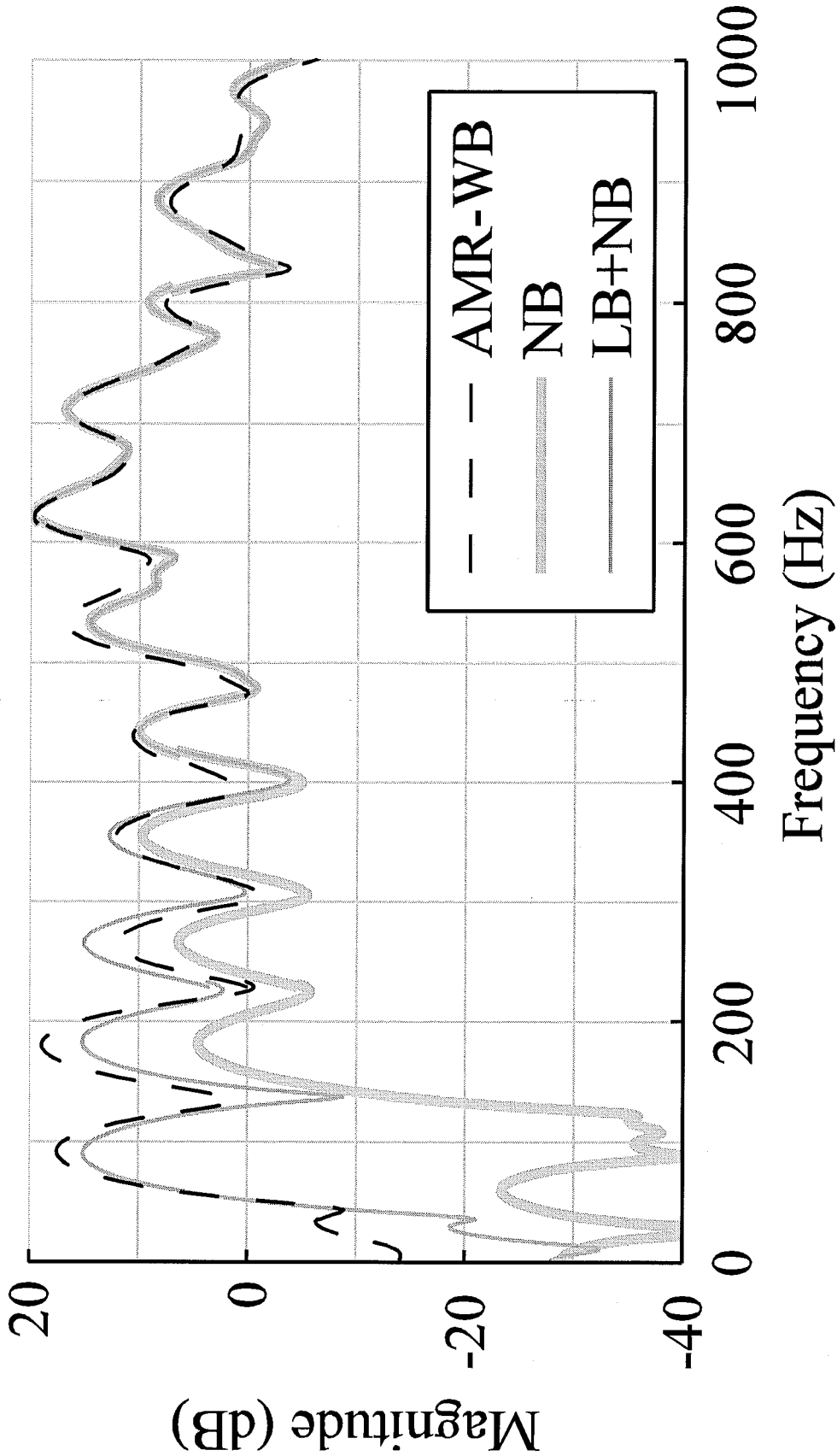
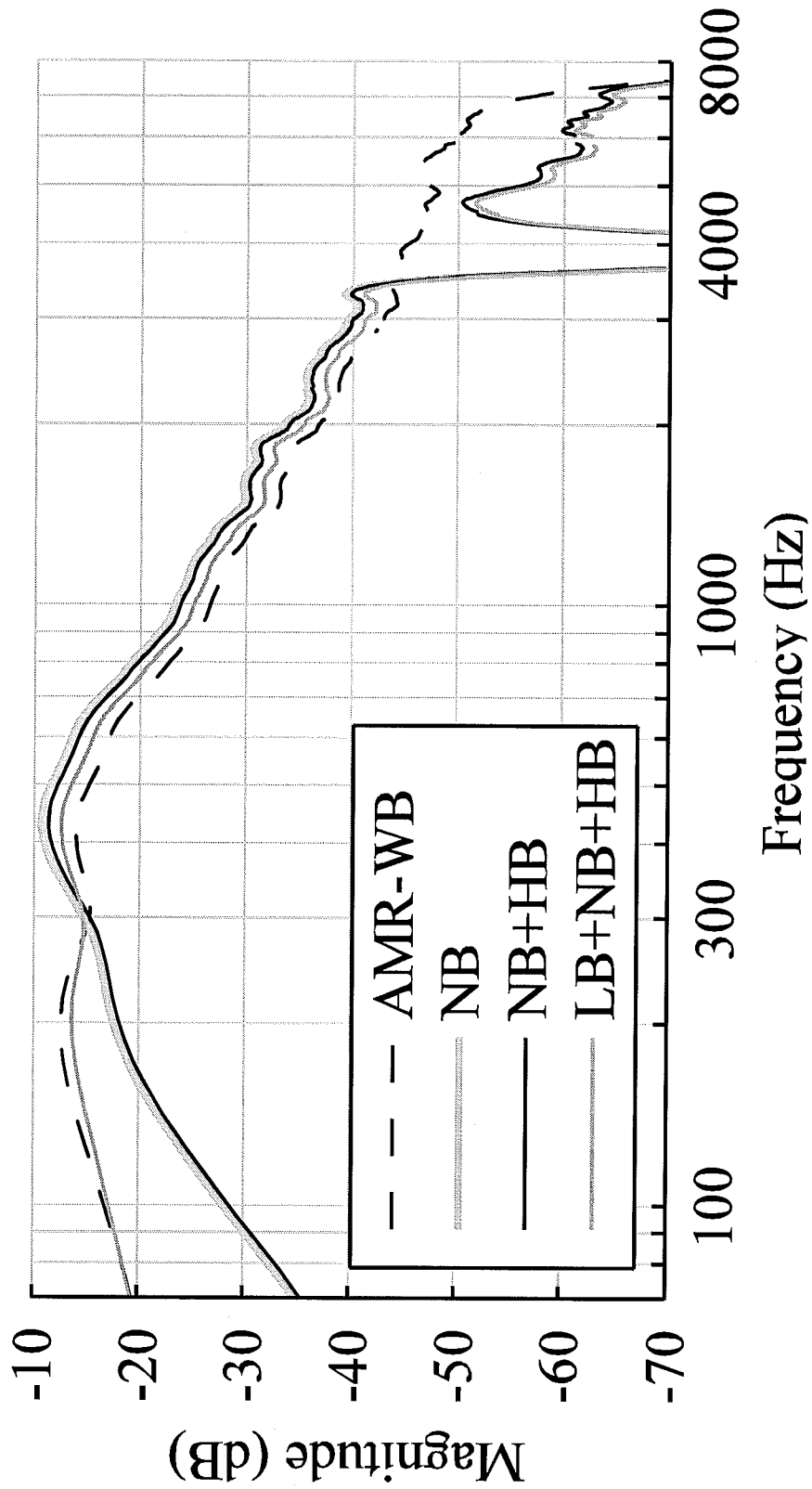


Fig 7



INTERNATIONAL SEARCH REPORT

International application No.
PCT/IB2011/051391

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, PAJ, WPI data, COMPENDEX, INSPEC, IBM-TDB

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 20020138268 A1 (GUSTAFSSON HARALD), 26 September 2002 (2002-09-26); paragraphs [0014]-[0016], [0029]-[0030], [0032], [0056]-[0060]; figures 1,3,4 --	1, 16, 31, 46
A	VALIN J et al. "Bandwidth extension of narrowband speech for low bit-rate wideband coding", Proceedings. 2000 IEEE Workshop on September 17-20, 2000, p. 130-132, ISBN 978-0-7803-6416-5 ; ISBN 0-7803-6416-3; abstract; figure 2; Section 3 --	1-62



Further documents are listed in the continuation of Box C.



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

31-01-2012

Date of mailing of the international search report

31-01-2012

Name and mailing address of the ISA/SE

Patent- och registreringsverket
Box 5055
S-102 42 STOCKHOLM
Facsimile No. + 46 8 666 02 86

Authorized officer

Alexander Lakic

Telephone No. + 46 8 782 25 00

INTERNATIONAL SEARCH REPORT

International application No.
PCT/IB2011/051391

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 20090201983 A1 (JASIUK MARK A ET AL), 13 August 2009 (2009-08-13); abstract; paragraphs [0002]-[0005], [0017]-[0020], [0026]-[0030], [0033]-[0034], [0036], [0057], [0062]-[0065], [0097]-[0100]; figure 1 --	1-62
A	US 20040166820 A1 (SLUIJTER ROBERT JOHANNES ET AL), 26 August 2004 (2004-08-26); paragraphs [0001], [0021]-[0023], [0026], [0032]; figure 2 --	1-62
A	MIET G et al. "Low-band extension of telephone-band speech", Acoustics, Speech, and Signal Processing, 2000. ICASSP '00. Proceedings. 2000 IEEE International Conference on 5-9 June 2000, Vol. 3, p. 1851-1854, ISBN 978-0-7803-6293-2 ; ISBN 0-7803-6293-4; whole document -- -----	1-62

Continuation of: second sheet

International Patent Classification (IPC)

G10L 21/02 (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 **LCHCTIRQJE**.

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.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/IB2011/051391

US	20020138268 A1	26/09/2002	AU	2002237264 A1	24/07/2002
			EP	1350243 A2	08/10/2003
			JP	2004517368 A	10/06/2004
			US	6889182 B2	03/05/2005
			WO	02056295 A3	28/11/2002
US	20090201983 A1	13/08/2009	CN	101939783 A	05/01/2011
			EP	2238593 A1	13/10/2010
			KR	20100123712 A	24/11/2010
			MX	2010008288 A	31/08/2010
			US	20110112844 A1	12/05/2011
			US	20110112845 A1	12/05/2011
US	20040166820 A1	26/08/2004	WO	2009100182 A1	13/08/2009
			CN	1235192 C	04/01/2006
			CN	1520590 A	11/08/2004
			EP	1405303 A1	07/04/2004
			JP	2004521394 A	15/07/2004
			US	7174135 B2	06/02/2007
			WO	03003350 A1	09/01/2003