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

(11) EP 1 509 906 B1

(12) EUROPEAN PATENT SPECIFICATION

(45) Date of publication and mention of the grant of the patent:
25.06.2008 Bulletin 2008/26

(21) Application number: **03727092.3**

(22) Date of filing: **30.05.2003**

(51) Int Cl.:
G10L 21/02 (2006.01) *G10L 19/14* (2006.01)

(86) International application number:
PCT/CA2003/000828

(87) International publication number:
WO 2003/102923 (11.12.2003 Gazette 2003/50)

(54) METHOD AND DEVICE FOR PITCH ENHANCEMENT OF DECODED SPEECH

VERFAHREN UND ANORDNUNG ZUR GRUNDFREQUENZVERBESSERUNG EINES DECODIERTEN SPRACHSIGNALS

PROCEDE ET DISPOSITIF D'AMELIORATION DE LA HAUTEUR TONALE SELECTIVE EN FREQUENCE DE VOIX SYNTHETISEE

(84) Designated Contracting States:
AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IT LI LU MC NL PT RO SE SI SK TR

(30) Priority: **31.05.2002 CA 2388352**

(43) Date of publication of application:
02.03.2005 Bulletin 2005/09

(73) Proprietor: **Voiceage Corporation**
Ville Mont-Royal,
Quebec H3R 2H6 (CA)

(72) Inventors:
- **BESSETTE, Bruno**
  Rock Forest, Québec J1N 1L2 (CA)
- **LAFLAMME, Claude**
  Orford, Québec, J1X 6W1 (CA)
- **JELINEK, Milan**
  Sherbrooke, Quebec J1H 1K4 (CA)
- **LEFEBVRE, Roch**
  Canton de Magog, Québec J1X 5R9 (CA)

(74) Representative: **Schmit, Christian Norbert Marie**
SCHMIT-CHRETIEN-SCHIHIN
8, place du Ponceau
95000 Cergy (FR)

(56) References cited:
US-A- 5 806 025 **US-A- 5 864 798**

- **CHAN C-F ET AL: "Efficient frequency domain postfiltering for multiband excited linear predictive coding of speech" ELECTRONICS LETTERS, IEE STEVENAGE, GB, vol. 32, no. 12, 6 June 1996 (1996-06-06), pages 1061-1063, XP006005271 ISSN: 0013-5194**
- **"Wideband Coding of Speech at around 16 kbit/s using Adaptive Multi-Rate Wideband (AMR-WB)" ITU-T RECOMMENDATION G.722.2, XX, XX, January 2002 (2002-01), page complete, XP002274473 cited in the application**
- **CHEN H-H ET AL: "Adaptive postfiltering for quality enhancement of coded speech" IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING, IEEE INC. NEW YORK, US, vol. 3, no. 1, January 1995 (1995-01), pages 59-71, XP002225533 ISSN: 1063-6676**

Note: Within nine months of the publication of the mention of the grant of the European patent in the European Patent Bulletin, any person may give notice to the European Patent Office of opposition to that patent, in accordance with the Implementing Regulations. Notice of opposition shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

Printed by Jouve, 75001 PARIS (FR)

Description

BACKGROUND OF THE INVENTION

1. Field of the invention:

[0001] The present invention relates to a method and device for post-processing a decoded sound signal in view of enhancing a perceived quality of this decoded sound signal.

[0002] These post-processing method and device can be applied, in particular but not exclusively, to digital encoding of sound (including speech) signals. For example, these post-processing method and device can also be applied to the more general case of signal enhancement where the noise source can be from any medium or system, not necessarily related to encoding or quantization noise.

2. Brief description of the current technology:

2.1 *Speech encoders*

[0003] Speech encoders are widely used in digital communication systems to efficiently transmit and/or store speech signals. In digital systems, the analog input speech signal is first sampled at an appropriate sampling rate, and the successive speech samples are further processed in the digital domain. In particular, a speech encoder receives the speech samples as an input, and generates a compressed output bit stream to be transmitted through a channel or stored on an appropriate storage medium. At the receiver, a speech decoder receives the bit stream as an input, and produces an output reconstructed speech signal.

[0004] To be useful, a speech encoder must produce a compressed bit stream with a bit rate lower than the bit rate of the digital, sampled input speech signal. State-of-the-art speech encoders typically achieve a compression ratio of at least 16 to 1 and still enable the decoding of high quality speech. Many of these state-of-the-art speech encoders are based on the CELP (Code-Excited Linear Predictive) model, with different variants depending on the algorithm.

[0005] In CELP encoding, the digital speech signal is processed in successive blocks of speech samples called *frames*. For each frame, the encoder extracts from the digital speech samples a number of parameters that are digitally encoded, and then transmitted and/or stored. The decoder is designed to process the received parameters to reconstruct, or synthesize the given frame of speech signal. Typically, the following parameters are extracted from the digital speech samples by a CELP encoder:

- Linear Prediction Coefficients (LP coefficients), transmitted in a transformed domain such as the Line Spectral Frequencies (LSF) or
  Immitance Spectral Frequencies, (ISF);
- Pitch parameters, including a pitch delay (or lag) and a pitch gain; and
- Innovative excitation parameters (fixed codebook index and gain). The pitch parameters and the innovative excitation parameters together describe what is called the excitation signal. This excitation signal is supplied as an input to a Linear Prediction (LP) filter described by the LP coefficients. The LP filter can be viewed as a mode1 of the vocal tract, whereas the excitation signal can be viewed as the output of the glottis. The LP or LSF coefficients are typically calculated and transmitted every frame, whereas the pitch and innovative excitation parameters are calculated and transmitted several times per frame. More specifically, each frame is divided into several signal blocks called *sub-frames,* and the pitch parameters and the innovative excitation parameters are calculated and transmitted every subframe. A frame typically has a duration of 10 to 30 milliseconds, whereas a subframe typically has a duration of 5 milliseconds.

[0006] Several speech encoding standards are based on the Algebraic CELP (ACELP) model, and more precisely on the ACELP algorithm. One of the main features of ACELP is the use of algebraic codebooks to encode the innovative excitation at each subframe. An algebraic codebook divides a subframe in a set of *tracks* of interleaved pulse positions. Only a few non-zero-amplitude pulses per track are' allowed, and, each non-zero-amplitude pulse is restricted to the positions of the corresponding track. The encoder uses fast search algorithms to find the optimal pulse positions and amplitudes for the pulses of each subframe. A description of the ACELP algorithm can be found in the article of R. SALAMI et al., "Design and description of CS-ACELP: a toll quality 8 kb/s speech code", IEEE Trans. on Speech and Audio Proc., Vol. 6, No. 2, pp. 116-130, March 1998, herein incorporated be reference, and which describes the ITU-TG.729CS-ACELP narrowband speech encoding algorithm at 8 kbits/second. It should be noted that there are several variations of the ACELP innovation codebook search, depending on the standard of concern. The present invention is not dependent on these variations, since it only applies to post-processing of the decoded (synthesized) speech signal.

[0007] A recent standard based on the ACELP algorithm is the ETSV3GPP AMR-WB speech encoding algorithm, which was also adopted by the ITU-T (Telecommunication Standardization Sector of ITU (International Telecommunication Union)) as recommendation *G. 722.2 [ITU-T Recommendation* G.722.2 *"Wideband coding of speech at around 16 kbit/s using Adaptive Multi-Rate Wideband (AMR-WB)", Geneva, ZOOZ], [3GPP TS 26.190, 'AMR Wideband Speech Codec: Transcoding Functions, " 3GPP Technical Specification]*. The AMR-WB is a multi-rate algorithm designed to operate at nine different bit rates between 6.6 and 23.85 kbits/second. Those of ordinary skill in the art know that the quality of the decoded speech generally increases with the bit rate. The AMR-WB has been designed to allow cellular communication systems to reduce the bit rate of the speech encoder in the case of bad channel conditions; the bits are converted to channel encoding bits to increase the protection of the transmitted bits. In this manner, the overall quality of the transmitted bits can be kept higher than in the case where the speech encoder operates at a single fixed bit rate.

[0008] Figure 7 is a schematic block diagram showing the principle of the AMR-WB decoder. More specifically, Figure 7 is a high-level representation of the decoder, emphasizing the fact that the received bitstream encodes the speech signal only up to 6.4 kHz (12.8 kHz sampling frequency), and the frequencies higher than 6.4 kHz are synthesized at the decoder from the lower-band parameters. This implies that, in the encoder, the original wideband, 16 kHz-sampled speech signal was first down-sampled to the 12.8 kHz sampling frequency, using multi-rate conversion techniques well known to those of ordinary skill in the art. The parameter decoder 701 and the speech decoder 702 of Figure 7 are analogous to the parameter decoder 106 and the source decoder 107 of Figure 1. The received bitstream 709 is first decoded by the parameter decoder 701 to recover parameters 710 supplied to the speech decoder 702 to resynthesize the speech signal. In the specific case of the AMR-WB decoder, these parameters are:

- ISF coefficients for every frame of 20 milliseconds;
- An integer pitch delay *T0,* a fractional pitch value *T0-frac* around *T0,* and a pitch gain for every 5 millisecond subframe; and
- An algebraic codebook shape (pulse positions and signs) and gain for every 5 millisecond subframe.

From the parameters 710, the speech decoder 702 is designed to synthesize a given frame of speech signal for the frequencies equal to and lower than 6.4 kHz, and thereby produce a low-band synthesized speech signal 712 at the 12.8 kHz sampling frequency. TO recover the full-band signal corresponding to the 16 kHz sampling frequency, the AMR-WB decoder comprises a high-band resynthesis processor 707 responsive to the decoded parameters 710 from the parameter decoder 701 to resynthesize a high-band signal 711 at the sampling frequency of 16 kHz. The details of the high-band signal resynthesis processor 707 can be found in the following publications which are herein incorporated by reference:

- *ITU- T Recommendation* G. 722.2 *"Wideband coding of speech at around 16 kbit/s using Adaptive Multi-Rate Wideband (AMR- INB) ", Geneva, 2002;* and

- *3GPP TS 26.190, "AMR Wideband Speech Codec: Transcoding Functions, " 3GPP Technical Specification.*

The output of the high-band resynthesis processor 707, referred to as the high-band signal 711 of Figure 7, is a signal at the 16 kHz sampling frequency, having an energy concentrated above 6.4 kHz. The processor 708 sums the high-band signal 711 to a 16-kHz up-sampled low-band speech signal 713 to form the complete decoded speech signal 714 of the AMR-WB decoder at the 16 kHz sampling frequency.

[0009] The US Patent 5,806,025 discloses a method for adaptively filtering a speech signal for noise suppression.

2.2 Need for post-processing

[0010] Whenever a speech encoder is used in a communication system, the synthesized or decoded speech signal is never identical to the original speech signal even in the absence of transmission errors. The higher the compression ratio, the higher the distortion introduced by the encoder. This distortion can be made subjectively small using different approaches. A first approach is to condition the signal at the encoder to better describe, or encode, subjectively relevant information in the speech signal. The use of a formant weighting filter, often represented as W(z), is a widely used example of this first approach [B. Kleijn and K. Paliwal editors, «Speech Coding and Synthesis, » Elsevier, 1995]. This filter W(z) is typically made adaptive, and is computed in such a way that it reduces the signal energy near the spectral formants, thereby increasing the relative energy of lower energy bands. The encoder can then better quantize lower energy bands, which would otherwise be masked by encoding noise, increasing the perceived distortion. Another example of signal conditioning at the encoder is the so-called *pitch sharpening* filter which 1 0 enhances the harmonic structure of the excitation signal at the encoder. Pitch sharpening aims at ensuring that the inter-harmonic noise level is kept low enough in the perceptual sense.

[0011] A second approach to minimize the perceived distortion introduced by a speech encoder is to apply a so-called *post-processing* algorithm. Post-processing is applied at the decoder, as shown in Figure 1. In Figure 1, the speech encoder 101 and the speech decoder 105 are broken down in two modules. In the case of the speech encoder 101, a source encoder 102 produces a series of speech encoding parameters 109 to be transmitted or 20 stored. These parameters 109 are then binary encoded by the parameter encoder 103 using a specific encoding method, depending on the speech encoding algorithm and on the parameters to encode. The encoded speech signal (binary encoded parameters) 110 is then transmitted to the decoder through a communication channel 104. At the decoder, the received bit stream II 1 is first analysed by a parameter decoder 106 to decode the received, encoded sound signal encoding parameters, which are then used by the source decoder 107 to generate the synthesized speech signal 112. The aim of post-processing (see post-processor 108 of Figure 1) is to enhance the perceptually relevant information in the synthesized speech signal, or equivalently to reduce or remove the perceptually annoying information. Two commonly used forms of post-processing are formant post-processing and pitch post-processing. In the first case, the formant structure of the synthesized speech signal is amplified by the use of an adaptive filter with a frequency response correlated to the speech formants. The spectral peaks of the synthesized speech signal are then accentuated at the expense of spectral valleys whose relative energy becomes smaller. In the case of pitch post-processing, an adaptive filter is also applied to the synthesized speech signal. However in this case, the filters frequency response is correlated to the fine spectral structure, namely the harmonics. A pitch post-filter then accentuates the harmonics at the expense of inter-harmonic energy which becomes relatively smaller. Note that the frequency response of a pitch post-filter typically covers the whole frequency range. The impact is that a harmonic structure is imposed on the post-processed speech even in frequency bands that did not exhibit a harmonic structure in the decoded speech. This is not a perceptually optimal approach for wideband speech (speech sampled at 16 kHz), which rarely exhibits a periodic structure on the whole frequency range.

SUMMARY OF THE INVENTION

[0012] The present invention relates to a method, as claimed in claim 1, for post-processing a decoded sound signal in view of enhancing a perceived quality of this decoded sound signal, comprising dividing the decoded sound signal into a plurality of frequency sub-band signals, and applying post-processing to at least one of the frequency sub-band signals, but not all the frequency sub-band signals, characterized in that, for pitch enhancement, post-processing is applied to only a lower sub-band of the frequency sub-band signals.

[0013] The present invention is also concerned with a device, as claimed in claim 32, for post-processing a decoded sound signal in view of enhancing a perceived quality of this decoded sound signal, comprising means for dividing the decoded sound signal into a plurality of frequency sub-band signals, and means for post-processing only the lower sub-band of the frequency sub-band signals.

[0014] According to an illustrative embodiment, after post-processing of the above mentioned lower frequency sub-band signal, the frequency sub-band signals are summed to produce an output post-processed decoded sound signal.

[0015] Accordingly, the post-processing method and device make it possible to localize the post-processing in the desired sub-band and to leave other subbands virtually unaltered.

[0016] The present invention further relates to a sound signal decoder, as claimed in claim 63, comprising an input for receiving an encoded sound signal, a parameter decoder supplied with the encoded sound signal for decoding sound signal encoding parameters, a sound signal decoder supplied with the decoded sound signal encoding parameters for producing a decoded sound signal, and a post processing device as described above for post-processing the decoded sound signal in view of enhancing a perceived quality of this decoded sound signal.

[0017] The foregoing and other objects, advantages and features of the present invention will become more apparent upon reading of the following, non restrictive description of illustrative embodiments thereof, given by way of example only with reference to the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

[0018] In the appended drawings:

Figure 1 is a schematic block diagram of the high-level structure of an example of speech encoder/decoder system using post-processing at the decoder,

Figure 2 is a schematic block diagram showing the general principle of an illustrative embodiment of the present invention using a bank of adaptive filters and sub-band filters, in which the input of the adaptive filters is the decoded (synthesized) speech signal (solid line) and the decoded parameters (dotted line);

Figure 3 is a schematic block diagram of a two-band pitch enhancer, which constitutes a special case of the illustrative embodiment of Figure 2;

Figure 4 is a schematic block diagram of an illustrative embodiment of the present invention, as applied to the special case of the AMR-WB wideband speech decoder;

Figure 5 is a schematic block diagram of an alternative implementation of the illustrative embodiment of Figure 4;

Figure 6a is a graph illustrating an example of spectrum of a preprocessed signal;

Figure 6b is a graph illustrating an example of spectrum of the post-processed signal obtained when using the method described in Figure 3;

Figure 7 is a schematic block diagram showing the principle of operation of the 3GPP AMR-WB decoder;

Figures 8a and 8b are graphs showing an example of the frequency response of a pitch enhancer filter as described by Equation (I), with the special case of a pitch ,period T=10 samples;

Figure 9a is a graph showing an example of frequency response for the low-pass filter 404 of Figure 4;

Figure 9b is a graph showing an example of frequency response for the band-pass filter 407 of Figure 4;

Figure 9c is a graph showing an example of combined frequency response for the low-pass filter 404 and band-pass filters 407 of Figure 4; and

Figure 10 is a graph showing an example of the frequency response of an inter-harmonic filter as described by Equation (2), and used in the inter-harmonic filter 503 of Figure 5, for the specific case of T= 10 samples.

DETAILED DESCRIPTION OF THE ILLUSTRATIVE EMBODIMENTS

[0019] Figure 2 is a schematic block diagram illustrating the general principle of an illustrative embodiment of the present invention.

[0020] In Figure 1, the input signal (signal on which post-processing is applied) is the decoded (synthesized) speech signal 112 produced by the speech decoder 105 (Figure 1) at the receiver of a communications system (output of the source decoder 107 of Figure 1). The aim is to produce a post-processed decoded speech signal at the output 113 of the post-processor 108 of Figure 1 (which is also the output of processor 203 of Figure 2) with enhanced perceived quality. This is achieved by first applying at least one, and possibly more than one, adaptive filtering operation to the input signal. 112 (see adaptive filters 201 a, 201 b, ... , 201 N). These adaptive filters will be described in the following description. It should be pointed out here that some of the adaptive filters 201a to 201 N can be trivial functions whenever required, for example with the output equal to the input. The output 204a, 204b, ..., 204N of each adaptive filter 201 a, 201 b, ... , 201 N is then band-pass filtered through a sub-band filter 202a, 202b, ... , 202N, respectively, and the post-processed decoded speech signal 113 is obtained by adding through a processor 203 the respective resulting outputs 205a, 205b, ... , 205N of Sub-band filters 202a, 202b,...,202N.

[0021] In one illustrative embodiment, a two-band decomposition is used and adaptive filtering is applied only to the lower band. This results in a total post-processing that is mostly targeted at frequencies near the first harmonics of the synthesized speech signal.

[0022] Figure 3 is a schematic block diagram of a two-band pitch enhancer, which constitutes a special case of the illustrative embodiment of Figure 2. More specifically, Figure 3 shows the basic functions of a two-band post-processor (see post-processor 108 of Figure 1). According to this illustrative embodiment, only pitch enhancement is considered as post-processing although other types of post-processing could be contemplated. In Figure 3, the decoded speech signal (assumed to be the output 112 of the source decoder 107 of Figure 1) is supplied through a pair of sub-branches 308 and 309.

[0023] In the higher branch 308, the decoded speech signal 112 is filtered by a high-pass filter 301 to produce the higher band signal 310 (S_H). In this specific example, no adaptive filter is used in the higher branch. In the lower branch 309, the decoded speech signal 112 is first processed through an adaptive filter 307 comprising an optional low-pass filter 302, a pitch tracking module 303, and a pitch enhancer 304, and then filtered through a low-pass filter 305 to obtain the lower band, post processed signal 311 (S_{LEF}). The post-processed decoded speech signal 113 is obtained by adding through an adder 306 the lower 311 and higher 312 band post-processed signals from the output of the low-pass filter

305 and high-pass filter 301, respectively. It should be pointed out that the low-pass 305 and high-pass 301 filters could be of many different types, for example Infinite Impulse Response (UR) or Finite Impulse Response (FIR). In this illustrative embodiment, linear phase FIR filters are used.

[0024] Therefore, the adaptive filter 307 of Figure 3 is composed of two, and possibly three processors, the optional low-pass filter 302 similar to low-pass filter 305, the pitch tracking module 303 and the pitch enhancer 304.

[0025] The low-pass filter 302 can be omitted, but it is included to allow viewing of the post-processing of Figure 3 as a two-band decomposition followed by specific filtering in each sub-band. After optional low-pass filtering (filter 302) of the decoded speech signal 112 in the lower- band, the resulting signal S_L is processed through the pitch enhancer 304. The abject of the pitch enhancer 304 is to reduce the inter-harmonic noise in the decoded speech signal. In the present illustrative embodiment, the pitch enhancer 304 is achieved by a time-varying linear filter described by the following , equation :

$$y[n]=\left(1-\frac{\alpha}{2}\right)x[n]+\frac{\alpha}{4}\{x[n-T]+x[n+T]\} \qquad (1)$$

where α is a coefficient that controls the inter-harmonic attenuation, T is the pitch period of the input signal X[n], and y [n] is the output signal of the pitch enhancer. A more general equation could also be used where the filter taps at *n-T* and *n+T* could be at different delays (for example *n-T1* and *n+T2).* Parameters T and α vary with time and are given by the pitch tracking module 303. With a value of α = 1, the gain of the filter described by Equation (1) is exactly 0 at frequencies 1/(27),3/(27), 5/(27), etc, i.e. at the mid-point between the harmonic frequencies *1/T, 3/T, 5/T*; etc. When α approaches 0, the attenuation between the harmonics produced by the filter of Equation (1) reduces. With a value of α = 0, the filter output is equal to its input. Figure 8 shows the frequency response (in dB) of the filter described by Equation (1) for the values α = 0.8 and 1, when the pitch delay is (arbitrarily) set at a value T = 10 samples. The value of a can be computed using several approaches. For example, the normalized pitch correlation, which is well-known by those of ordinary skill in the art, can be used to control the coefficient α : the higher the normalized pitch correlation (the closer to 1 it is), the higher the value of α A periodic signal x[n] with a period of T = 10 samples would have harmonics at the maxima of the frequency responses of Figure 8, i.e. at normalized frequencies 0.2, 0.4, etc. It is easy to understand from Figure 8 that the pitch enhancer of Equation (1) would attenuate the signal energy only between its harmonics, and that the harmonic components would not be altered by the filter. Figure 8 also shows that varying parameter α enables control of the amount of inter-harmonic attenuation provided by the filter of Equation (1). Note that the frequency response of the filter of Equation (1), shown in Figure 8, extends to all frequencies of the spectrum.

[0026] Since the pitch period of a speech signal varies in time, the pitch value T of the pitch enhancer 304 has to vary accordingly. The pitch tracking module 303 is responsible for providing the proper pitch value *T* to the pitch enhancer 304, for every frame of the decoded speech signal that has to be processed. For that purpose, the pitch tracking module 303 receives as input not only the decoded speech samples but also the decoded parameters 114 from the parameter decoder 106 of Figure 1.

[0027] Since a typical speech encoder extracts, for every speech subframe, a pitch delay which we call *To* and possibly a fractional value *To_frac* used to interpolate the adaptive codebook contribution to fractional sample resolution, the pitch tracking module 303 can then use this decoded pitch delay to focus the pitch tracking at the decoder. One possibility is to use *To* and *To_frac* directly in the pitch enhancer 304, exploiting the fact that the encoder has already performed pitch tracking. Another possibility, used in this illustrative embodiment, is to recalculate the pitch tracking at the decoder focussing on values around, and multiples or submultiples of, the decoded pitch value To. The pitch tracking module 303 then provides a pitch delay T to the pitch enhancer 304, which uses this value of Tin Equation (1) for the present frame of decoded speech signal. The output is signal S_{LE}.

[0028] Pitch enhanced signal S_{LE} is then low-pass filtered through filter 305 to isolate the low frequencies of the pitch enhanced signal S_{LE}, and to remove the high-frequency components that arise when the pitch enhancer filter of Equation (1) is varied in time, according to the pitch delay *T,* at the decoded speech frame boundaries. This produces the lower band post-processed signal S_{LEF}, which can now be added to the higher band signal S_H in the adder 306. The result is the post-processed decoded speech signal 113, with reduced inter-harmonic noise in the lower band. The frequency band where pitch enhancement Will be applied depends on the cut-off frequency of the low-pass filter 305 (and optionally in low-pass filter 302).

[0029] Figures 6a and 6b show an example signal spectrum illustrating the effect of the post-processing described in Figure 3. Figure 6a is the spectrum of the input signal 112 of the post-processor 108 of Figure 1 (decoded speech signal 112 in Figure 3). In this illustrative example, the input signal is composed of 20 harmonics, with fundamental frequency fo = 373 Hz chosen arbitrarily, with «noisy» components added at frequencies *fo*/2, 3*fo*/2 and 5 *fo*/2. These three noisy components can be seen between the low-frequency harmonics in Figure 6a. The sampling frequency is assumed to be 16 kHz in this example. The two-band pitch enhancer shown in Figure 3 and described above is then applied to the signal of Figure 6a. With a sampling frequency of 16 kHz and a periodic signal of fundamental frequency equal to 373 Hz as in Figure 6a, the pitch tracking module 303 should find a period of T= 16000/373 = 43 samples. This is the value that was used for the pitch enhancer filter of Equation (1), applied to the pitch enhancer 304 of Figure 3. A value of α = 0.5 was also used. The low-pass 305 and high-pass 301 filters are symmetric, linear phase FIR filters with 31 taps. The cut-off frequency for this example is chosen as 2000 Hz. These specific values are given only as an illustrative example.

[0030] The post-processed decoded speech signal 113 at the output of the adder 306 has a spectrum shown in Figure 6b. It can be seen that the three inter-harmonic sinusoids in Figure 6a have been completely removed, while the harmonics of the signal have been practically unaltered. Also it is noted that the effect of the pitch enhancer diminishes as the frequency approaches the low-pas filter cut-off frequency (2000 Hz in this example). Hence, only the lower band is affected by the post-processing. This is a key feature of this illustrative embodiment of the present invention. By varying the cut-off frequencies of the optional low-pass filter 302, low-pass filter 305 and high-pass filter 301, it is possible to control up to which frequency pitch enhancement is applied.

Application to the AMR-WB speech decoder

[0031] The present invention can be applied to any speech signal synthesized by a speech decoder, or even to any speech signal corrupted by inter-harmonic noise that needs to be reduced. This section will show a specific, exemplary implementation of the present invention to an AMR-WB decoded speech signal. The post-processing is applied to the low-band synthesized speech signal 712 of Figure 7, i.e. to the output of the speech decoder 702, which produces a synthesized speech at a sampling frequency of 12.8 kHz.

[0032] Figure 4 shows the block diagram of a pitch post-processor when the input signal is the AMR-WB low-band synthesized speech signal at the sampling frequency of 12.8 kHz. More precisely, the post-processor presented in Figure 4 replaces the up-sampling unit 703, which comprises processors 704, 705 and 706. The pitch post-processor of Figure 4 could also be applied to the 16 kHz up-sampled synthesized speech signal, but applying it prior to up-sampling reduces the number of filtering operations at the decoder, and thus reduces complexity.

[0033] The input signal *(AMR-WB low-band synthesized speech (12.8 kHz))* of Figure 4 is designated as signal s. In this specific example, signal s is the AMR-WB low-band synthesized speech signal at the sampling frequency of 12.8 kHz (output of processor 702). The pitch post-processor of Figure 4 comprises a pitch tracking module 401 to determine, for every 5 millisecond subframe, the pitch delay *T* using the received, decoded parameters 114 (Figure 1) and the synthesized speech signal s. The decoded parameters used by the pitch tracking module are *To*, the integer pitch value for the subframe, and *To_frac,* the fractional pitch value for subsample resolution. The pitch delay T calculated in the pitch tracking module 401 will be used in the next steps for pitch enhancement. It would be possible to use directly the received, decoded pitch parameters To and *To_frac* to form the delay T used by the pitch enhancer in the pitch filter 402. However, the pitch tracking module 401 is capable of correcting pitch multiples or submultiples, which could have a harmful effect on the pitch enhancement.

[0034] An illustrative embodiment of pitch tracking algorithm for the module 401 is the following (the specific thresholds and pitch tracked values are given only by way of example):

- First, the decoded pitch information (pitch delay To) is compared to a stored value of the decoded pitch delay *T_prev* of the previous frame. *T_prev* may have been modified by some of the following steps according to the pitch tracking algorithm. For example, if To < *1.16*T_prev* then go to case 1 below, else if *To> 1.16*T_prev,* then set *T_temp = To* and go to case 2 below.

  Case 1: First, calculate the cross-correlation *C2* (cross-product) between the last synthesized subframe and the synthesis signal starting at *To/2* samples before the beginning of the last subframe (look 'at correlation at half the decoded pitch value).
  Then, calculate the cross-correlation *C3* (cross-product) between the last synthesized subframe and the synthesis signal starting at *To/3* samples before the beginning of the last subframe (look at correlation at one-third the decoded pitch value).
  Then, Select the maximum value between *C2* and *C3* and calculate the normalized correlation *Cn* (normalized version of *C2* or *C3*) at the corresponding sub-multiple of *To* (at *To/2* if *C2* > *C3* and at *To/3* if *C3* > *C2*). Call *T_new* the pitch sub-multiple corresponding to the highest normalized correlation.
  If *Cn* > 0.95 (strong normalized correlation) the new pitch period is *T_new* (instead of *To*). Output the value *T = T_new* from the pitch tracking module 401. Save *T_prev = T* for next subframe pitch tracking and exit the pitch tracking module 401.
  If 0.7 < *Cn* < 0.95, then save *T_ temp = To*/2 or *To*/3 (according to *C2* or *C3* above) for comparisons in case 2

below. Otherwise, *if Cn < 0. 7* save *T_temp = To*.

Case 2: Calculate all possible values of the ratio *Tn = [T_ temp/n]* where *[x]* means the integer part of x and *n = 1,2,3, etc.* is an integer.
Calculate all cross correlations *Cn* at the pitch delay submultiples *Tn*. Retain *Cn_max* as the maximum cross correlation among all *Cn*. If *n* > 1 and *Cn* > 0.8, output *Tn* as the pitch period output T of the pitch tracking unit 401. Otherwise, output *T1 = T_temp*. Here, the value of *T_temp* will depend on the calculations in Case 1 above.

[0035] It should be noted that the above example of pitch tracking module 401 is given for the purpose of illustration only. Any other pitch tracking method or device could be implemented in module 401 (or 303 and 502) to ensure a better pitch tracking at the decoder.

[0036] Therefore, the output of the pitch tracking, module is the period *T* to be used in the pitch filter 402 which, in this preferred embodiment, is described by the filter of Equation (1). Again, a value of α = 0 implies no filtering (output of the pitch filter 402 is equal to its input), and a value of α = 1 corresponds to the highest amount of pitch enhancement.

[0037] Once the enhanced signal S_E (Figure 4) is determined, it is combined with the input signal s such that, as in Figure 3, only the lower band is subjected to pitch enhancement. In Figure 4, a modified approach is used compared to Figure 3. Since the pitch post-processor of Figure 4 replaces the up-sampling unit 703 in Figure 7, the sub-band filters 301 and 305 of Figure 3 30 are combined with the interpolation filter 705 of Figure 7 to minimize the number of filtering operations, and the filtering delay. More specifically, filters 404 and 407 of Figure 4 act both as band-pass filters (to separate the frequency bands) and as interpolation filters (for up-sampling from 12.8 to 16 kHz). These filters 404 and 407 could be further designed such that the band-pass filter 407 has relaxed constraints in 'its low-frequency stop band (i.e. it does not have to completely attenuate the signal at low frequencies). This could be achieved by using design constraints similar to those shown in Figure 9. Figure 9a is an example of frequency response for the low-pass filter 404. It should be noted that the DC (Direct Current) gain of this filter is 5 (instead of 1) since this filter also acts as interpolation filter, with a 5/4 interpolation ratio which implies that the filter gain must be 5 at 0 Hz. Then, Figure 9b shows the frequency response of the band-pass filter 407 making this filter 407 complementary, in the low band, to the low-pass filter 404. In this example, the filter 407 is a band-pass filter, not a high-pass filter such as filter 301, since it must act both as high-pass filter (such as filter 301) and low-pass filter (such as interpolation filter 705). Referring again to Figure 9, we see that the low-pass and band-pass filters 404 and 407 are complementary when considered in parallel, as in Figure 4. Their combined frequency response (when used in parallel) is shown in Figure 9c.

[0038] For completeness, the tables of filter coefficients used in this illustrative embodiment of the filters 404 and 407 are given below. Of course, these tables of filter coefficients are given by way of example only. It should be understood that these filters can be replaced without modifying the scope, spirit and nature of the present invention.

Table 1. Low-pass coefficients of filter 404

hlp[0]	0.04375000000000	hlp[30]	0.01998000000000
hlp[1]	0.04371500000000	hlp[31]	0.01882400000000
hlp[2]	0.04361200000000	hlp[32]	0.01768200000000
hlp[3]	0.04344000000000 -	hlp[33]	0.01655700000000
hlp[4]	0.04320000000000	hlp[34]	0.01545100000000
hip[5]	0.04289300000000	hlp[35]	0.01436900000000
hlp[6]	0.04252100000000	hlp[36]	0.01331200000000
hlp[7]	0.04208300000000	hlp[37]	0.01228400000000
hlp[8]	0.04158200000000	hlp[38]	0.01128600000000
hlp[9]	0.04102000000000	hlp[39]	0.01032300000000
hlp[10]	0.04039900000000	hlp[40]	0.00939500000000
hlp[11]	0.03972100000000	hlp[41]	0.00850500000000
hip[12]	0.03898800000000	hlp[42]	0.00765500000000
hlp[13]	0.03820200000000	hlp[43]	0.00684600000000
hlp[14]	0.03736700000000	hlp[44]	0.00608100000000

(continued)

hlp[15]	0.03648600000000	hlp[45]	0.00535900000000
hlp[16]	0.03556100000000	hip[46]	0.00468200000000
hlp[17]	0.03459600000000	hip[47]	0.00405100000000
hlp[18]	0.03359400000000	hlp[48]	0.00346700000000
hlp[19]	0.03255800000000	hlp[49]	0.00292900000000
hlp[20]	0.03149200000000	hlp[50]	0.00243900000000
hip[21]	0.03039900000000	hlp[51]	0.00199500000000
hlp[22]	0.02928400000000	hlp[52]	0.00159900000000
hlp[23]	0.02814900000000	hip[53]	0.00124800000000
hlp[24]	0.02699900000000	hlp[54]	0.00094400000000
hlp[25]	0.02583700000000	hlp[55]	0.00068400000000
hip[26]	0.02466700000000	hlp[56]	0.00046800000000
hlp[27]	0.02349300000000	hlp[57]	0.00029500000000
hlp[28]	0.02231800000000	hlp[58]	0.00016300000000
hlp[29]	0.02114600000000	hip[59]	0.00007100000000
		hlp[60]	0.00001800000000

Table 2. Band-pass coefficients of filter 407

hbp[0]	0.95625000000000	hbp[30]	-0.01998000000000
hbp[1]	0.89115400000000	hbp[31]	-0.00412400000000
hbp[2]	0.71120900000000	hbp[32]	0.00414300000000
hbp[3]	0.45810600000000	hbP[33]	0.00343300000000
hbp[4]	0.18819900000000	hbp[34]	-0.00416100000000
hbp[5]	-0.04289300000000	hbp[35]	-0.01436900000000
hbp[6]	-0.19474300000000	hbp[36]	-0.02267300000000
hbp[7]	-0.25136900000000	hbp[37]	-0.02601800000000
hbp[8]	-0.22287200000000	hbp[38]	-0.02370000000000
hbp[9]	-0.13948000000000	hbp[39]	-0.01723200000000
hbp[10]	-0.04039900000000	hbp[40]	-0.00939500000000
hbp[11]	0.03868100000000	hbp[41]	-0.00297000000000
hbp[12]	0.07548400000000	hbp[42]	0.00030500000000
hbp[13]	0.06566500000000	hbp[43]	0.00019000000000
hbp[14]	0.02113800000000	hbp[44]	-0.00226000000000
hbp[15]	-0.03648600000000	hbp[45]	-0.00535900000000
hbp[18]	-0.08465300000000	hbp[46]	-0.00756800000000
hbp[17]	-0.10763400000000	hbp[47]	-0.00805800000000
hbp[18]	-0.10087600000000	hbp[48]	-0.00687000000000
hbp[19]	-0.07091900000000	hbp[49]	-0.00469500000000
hbp[20]	-0.03149200000000	hbp[50]	-0.00243900000000

(continued)

hbp[21]	0.00234200000000	hbp[51]	-0.00080600000000
hbp[22]	0.01970000000000	hbp[52]	-0.00006300000000
hbp[23]	0.01715300000000	hbp[53]	-0.00005300000000
hbp[24]	-0.00110700000000	hbp[54]	-0.00038700000000
hbp[25]	-0.02583700000000	hbp[55]	-0.00068400000000
hbp[26]	-0.04678900000000	hbp[56]	-0.00074400000000
hbp[27]	-0.05654900000000	hbp[57]	-0.00057600000000
hbp[28]	-0.05281800000000	hbp[58]	-0.00031900000000
hbp[29]	-0.03851900000000	hbp[59]	-0.00011300000000
		hbp[60]	-0.00001800000000

[0039] The output of the pitch filter 402 of Figure 4 is called S_E. To be recombined with the signal of the Upper branch, it is first up-sampled by processor 403, low-pass filter 404 and processor 405, and added through an adder 409 to the up-sampled Upper branch signal 410. The up-sampling operation in the Upper branch is performed by processor 406, band-pass filter 407 and processor 408.

Alternate implementation of the proposed pitch enhancer

[0040] Figure 5 shows an alternative implementation of a two-band pitch enhancer according to an illustrative embodiment of the present invention. It should be noted that the Upper branch of Figure 5 does not process the input signal at all. This means that, in this particular case, the filters in the Upper branch of Figure 2 (adaptive filters 201a and 201 b) have trivial input-output characteristics (output is equal to input). In the lower branch, the input signal (signal to be enhanced) is processed first through an optional low-pass filter 501, then through a linear filter called inter-harmonic filter 503, defined by the following equation:

$$y[n]=\frac{1}{2}x[n]-\frac{1}{4}\{x[n-T]+x[n+T]\} \qquad (2)$$

It should be noted that the negative sign in front of the second term on the right hand side, compared to Equation (1). It should also be noted that the enhancement factor a! is not included in Equation (2), but rather it is introduced by means of an adaptive gain by the processor 504 of Figure 5. The inter-harmonic filter 503, described by Equation (2), has a frequency response such that it completely removes the harmonics of a periodic signal having a period of T samples, and such that a sinusoid at a frequency exactly between the harmonics passes through the filter unchanged in amplitude but with a phase reversal of exactly 180 degrees (same as sign inversion). For example, Figure shows the frequency response of the filter described by Equation (2) when the period is (arbitrarily) chosen at T = 10 samples. A periodic signal with period T = 10 samples would present harmonics at normalized frequencies 0.2, 0.4, 0.6, etc., and Figure 16 shows that the filter of Equation (2), with T = 10 samples, would completely remove these harmonics. On the other hand, the frequencies at the exact mid-point between the harmonics would appear at the output of the filter with the same amplitude but with a 180° phase shift. This is the reason why the filter described by Equation (2) and used as filter 503 is called inter-harmonic filter.

[0041] The pitch value T for use in the inter-harmonic filter. 503 is obtained adaptively by the pitch tracking module 502. Pitch tracking module 502 operates on the decoded speech signal and the decoded parameters, similarly to the previously disclosed methods as shown in Figures 3 and 4.

[0042] Then, the output 507 of the inter-harmonic filter 503 is a signal formed essentially of the inter-harmonic portion of the input decoded signal II 2, with 180° phase shift at mid-point between the signal harmonics. Then, the output 507 of the inter-harmonic filter 503 is multiplied by a gain α (processor 504) and subsequently low-pass filtered (filter 505) to obtain the low frequency band modification that is applied to the input decoded speech signal 112 of Figure 5, to obtain the post-processed decoded signal (enhanced signal) 509. The coefficient α in processor 504 controls the amount of pitch or inter-harmonic enhancement. The closer to 1 is a, the higher the enhancement is. When α is equal to 0, no enhancement is obtained, i.e. the output of adder 506 is exactly equal to the input signal (decoded speech in Figure 5).

The value of α can be computed using several approaches. For example, the normalized pitch correlation, which is well known to those of ordinary skill in the art, can be used to control coefficient a: the higher the normalized pitch correlation (the closer to 1 it is), the higher the value of α.

[0043] The final post-processed decoded speech signal 509 is obtained by adding through an adder 506 the output of low-pass filter 505 to the input signal (decoded speech signal 112 of Figure 5). Depending on the cut-off frequency of the low-pass filter 505, the impact of this post-processing will be limited to the low frequencies of the input signal 112, up to a given frequency. The higher frequencies Will be effectively unaffected by the post-processing.

One-band alternative using an adaptive high-pass filter

[0044] One last alternative for implementing sub-band post-processing for enhancing the synthesis signal at low frequencies is to use an adaptive high-pass filter, whose cut-off frequency is varied according to the input signal pitch value. Specifically, and without referring to any drawing, the low frequency enhancement using this illustrative embodiment would be performed, at each input signal frame, according to the following steps:

1. Determine the input signal pitch value (signal period) using the input signal and possibly the decoded parameters (output of speech decoder 105) if post-processing a decoded speech signal; this is a similar operation as the pitch tracking operation of modules 303, 401 and 502.

2. Calculate the coefficients of a high-pass filter such that the cut-off frequency is below, but close to, the fundamental frequency of the input signal; alternatively, interpolate between pre-calculated, stored high-pass filters of known cut-off frequencies (the interpolation can be done in the filtertaps domain, or in the pole-zero domain, or in some other transformed domain such as the LSF (Line Spectral Frequencies) of ISF (Immitance Spectral Frequencies) domain).

3. Filter the input signal frame with the calculated high-pass filter, to obtain the post-processed signal for that frame.

[0045] It should be pointed out that the present illustrative embodiment of the present invention is equivalent to using only one processing branch in Figure 2, and to define the adaptive filter of that branch as a pitch-controlled high-pass filter. The post-processing achieved with this approach will only affect the frequency range below the first harmonic and not the inter-harmonic energy above the first harmonic.

[0046] Although the present invention has been described in the foregoing description with reference to illustrative embodiments thereof, these embodiments can be modified at will, within the scope of the appended claims without departing from the nature of the present invention. For example, although the illustrative embodiments have been described in relation to a decoded speech signal, those of ordinary skill in the art will appreciate that the concepts of the present invention can be applied to other types of decoded signals, in particular but not exclusively to other types of decoded sound signals.

Claims

1. A method for post-processing a decoded sound signal (112) in view of enhancing a perceived quality of said decoded sound signal (112), comprising:

dividing the decoded sound signal (112) into a plurality of frequency sub-band signals; and
applying post-processing to at least one of the frequency sub-band signals;

characterized in that, for pitch enhancement, post-processing is applied to only a lower sub-band of the frequency sub-band signals.

2. A post-processing method as defined in claim 1, further comprising summing the frequency sub-band signals, after post-processing of said at least one frequency sub-band signal, to produce an output post-processed decoded sound signal (113).

3. A post-processing method as defined in claim 1, wherein applying post-processing to at least one of the frequency sub-band signals comprises adaptively filtering said at least one frequency sub-band signal.

4. A post-processing method as defined in claim 1, wherein dividing the decoded sound signal (112) into a plurality of frequency sub-band signals comprises sub-band filtering the decoded sound signal (112) to produce the plurality of frequency sub-band signals.

5. A post-processing method as defined in claim 1, wherein, for said at least one of the frequency sub-band signals:

applying post-processing comprises adaptively filtering the decoded sound signal (112); and
dividing the decoded sound signal (112) comprises sub-band filtering the adaptively filtered decoded sound signal.

6. A post-processing method as defined in claim 1, wherein:

dividing the decoded sound signal into a plurality of frequency sub-band signals comprises:

- a high-pass filtering of the decoded sound signal (112) to produce a frequency high-band signal (310); and
- a first low-pass filtering the decoded sound signal (112) to produce a frequency low-band signal (311); and

applying post-processing to at least one of the frequency sub-band signals comprises:

- applying post-processing to the decoded sound signal (112) prior to the first low-pass filtering of the decoded sound signal (112) to produce the frequency low-band signal (311).

7. A post-processing method as defined in claim 6, wherein applying post-processing to the decoded sound signal (112) comprises pitch enhancing said decoded sound signal (112) to reduce an inter-harmonic noise in the decoded sound signal (112).

8. A post-processing method as defined in claim 7, wherein applying post-processing to the decoded sound signal (112) further comprises a second low-pass filtering of the decoded sound signal (112) prior to pitch enhancing said decoded sound signal (112).

9. A post-processing method as defined in claim 6, further comprising summing the frequency high-band (310) and low-band signals (311) to produce an output post-processed decoded sound signal (113).

10. A post-processing method as defined in claim 1, wherein:

dividing the decoded sound signal (112) into a plurality of frequency sub-band signals comprises:

- band-pass filtering the decoded sound signal (112) to produce a frequency upper-band signal (410); and
- low-pass filtering the decoded sound signal (112) to produce a frequency lower-band signal; and

applying post-processing to at least one of the frequency sub-band signals comprises:

applying post-processing to the decoded sound signal (112) prior to low-pass filtering the decoded sound signal (112) to produce the frequency lower-band signal.

11. A post-processing method as defined in claim 10, wherein applying post-processing to the frequency lower-band signal comprises pitch enhancing the decoded sound signal (112) prior to low-pass filtering the decoded sound signal (112).

12. A post-processing method as defined in claim 10, further comprising summing the frequency upper-band and lower-band signals to produce an output post-processed decoded sound signal.

13. A post-processing method as defined in claim 1, wherein:

dividing the decoded sound signal (112) into a plurality of frequency sub-band signals comprises:

- low-pass filtering the decoded sound signal (112) to produce a frequency low-band signal; and

applying post-processing to at least one of the frequency sub-band signals comprises:

- applying post-processing to the frequency low-band signal.

14. A post-processing method as defined in claim 13, wherein applying post-processing to the frequency low-band signal comprises processing the decoded sound signal (112) through an inter-harmonic filter (503) for inter-harmonic attenuation of the decoded sound signal (112).

15. A post-processing method as defined in claim 14, wherein applying post-processing to the frequency low-band signal comprises multiplying the inter-harmonic filtered decoded sound signal (507) by an adaptive pitch enhancement gain (α).

16. A post-processing method as defined in claim 14, further comprising low-pass filtering the decoded sound signal (112) prior to processing the decoded sound signal (112) through the inter-harmonic filter (503).

17. A post-processing method as defined in claim 13, further comprising summing the decoded sound signal (112) and the frequency low-band signal to produce an output post-processed decoded sound signal (509).

18. A post-processing method as defined in claim 13, wherein applying post-processing to the frequency low-band signal comprises processing the decoded sound signal (112) through an inter-harmonic filter (503) having the following transfer function:

$$y[n] = \frac{1}{2}x[n] - \frac{1}{4}\{x[n-T] + x[n+T]\}$$

for inter-harmonic attenuation of the decoded sound signal, where x*[n]* is the decoded sound signal, y*[n]* is the inter-harmonic filtered decoded sound signal in a given sub-band, and T is a pitch delay of the decoded sound signal.

19. A post-processing method as defined in claim 18, further comprising summing the unprocessed decoded sound signal (112) and the inter-harmonic filtered frequency low-band signal (508) to produce an output post-processed decoded sound signal (509).

20. A post-processing method as defined in claim 1, wherein applying post-processing to at least one of the frequency sub-band signals comprises pitch enhancing the decoded sound signal (112) using the following equation:

$$y[n] = \left(1 - \frac{\alpha}{2}\right)x[n] + \frac{\alpha}{4}\{x[n-T] + x[n+T]\}$$

where *x[n]* is the decoded sound signal, *y[n]* is the pitch enhanced decoded sound signal in a given sub-band, T is a pitch delay of the decoded sound signal, and α is a coefficient varying between 0 and 1 to control an amount of inter-harmonic attenuation of the decoded sound signal.

21. A post-processing method as defined in claim 20, comprising receiving the pitch delay T through a bitstream.

22. A post-processing method as defined in claim 20, comprising decoding the pitch delay T from a received, encoded bitstream.

23. A post-processing method as defined in claim 20, comprising calculating the pitch delay T in response to the decoded sound signal (112) for an improved pitch tracking.

24. A post-processing method as defined in claim 1, wherein, during encoding, the sound signal is down-sampled from a higher sampling frequency to a lower sampling frequency, and wherein dividing the decoded sound signal (112) into a plurality of frequency sub-band signals comprises up-sampling the decoded sound signal from the lower sampling frequency to the higher sampling frequency.

25. A post-processing method as defined in claim 24, wherein dividing the decoded sound signal (112) into a plurality of frequency sub-band signals comprises sub-band filtering the decoded sound signal (112), and wherein the up-sampling of the decoded sound signal (112) from the lower sampling frequency to the higher sampling frequency is combined to the sub-band filtering.

26. A post-processing method as defined in claim 24, comprising:

band-pass filtering the decoded sound signal (112) to produce a frequency upper-band signal, said band-pass filtering of the decoded sound signal (112) being combined with up-sampling of the decoded sound signal (112) from the lower sampling frequency to the higher sampling frequency; and
post-processing the decoded sound signal (112) and low-pass filtering the post-processed decoded sound signal (112) to produce a frequency lower-band signal, said low-pass filtering of the post-processed decoded sound signal being combined with up-sampling of the post-processed decoded sound signal from the lower sampling frequency to the higher sampling frequency.

27. A post-processing method as defined in claim 26, further comprising adding the frequency upper-band signal with the frequency lower-band signal to form an output post-processed and up-sampled decoded sound signal.

28. A post-processing method as defined in claim 26, wherein post-processing of the decoded sound signal (112) comprises pitch enhancing the decoded sound signal (112) to reduce an inter-harmonic noise in the decoded sound signal (112).

29. A post-processing method as defined in claim 28, wherein pitch enhancing the decoded sound signal (112) comprises processing the decoded sound signal (112) by means of the following equation:

$$y[n]=\left(1-\frac{\alpha}{2}\right)x[n]+\frac{\alpha}{4}\{x[n-T]+x[n+T]\}$$

where *x[n]* is the decoded sound signal, *y[n]* is the pitch enhanced decoded sound signal in a given sub-band, *T* is a pitch delay of the decoded sound signal, and α is a coefficient varying between 0 and 1 to control an amount of inter-harmonic attenuation of the decoded sound signal.

30. A post-processing method as defined in claim 1, wherein:

dividing the decoded sound signal (112) into a plurality of frequency sub-band signals comprises dividing the decoded sound signal (112) into a frequency upper-band signal and a frequency lower-band signal; and
applying post-processing to at least one of the frequency sub-band signals comprises post-processing the frequency lower-band signal.

31. A post-processing method as defined in claim 1, wherein applying post-processing to said at least one of the frequency sub-band signals comprises:

determining a pitch value of the decoded sound signal;
calculating, in relation to the determined pitch value, a high-pass filter with a cut-off frequency below a fundamental frequency of the decoded sound signal; and
processing the decoded sound signal through the calculated high-pass filter.

32. A device for post-processing (108) a decoded sound signal (112) in view of enhancing a perceived quality of said decoded sound signal (112), comprising:

means for dividing (202a to 202N; 301, 305; 407, 404; 505) the decoded sound signal (112) into a plurality of frequency sub-band signals; and
means for post-processing (201 a to 201 N; 307; 401, 402; 503, 504 means for pitch enhancing a sub-band signal; and 502) at least one of the frequency sub-band signals;

characterized in that, the post-processing means is adapted to supply only a lower sub-band of the frequency sub-band signals to the pitch enhancing means.

33. A post-processing device (108) as defined in claim 32, further comprising adder means (203; 306; 409; 506) for summing the frequency sub-band signals, after post-processing of said at least one frequency sub-band signal, to produce an output post-processed decoded sound signal (113).

34. A post-processing device (108) as defined in claim 32, wherein the post-processing means comprises adaptive filter means (201 a to 201 N; 307) supplied with the decoded sound signal (112).

35. A post-processing device (108) as defined in claim 32, wherein the dividing means comprises sub-band filter means (202a to 202N; 301, 305; 407, 404; 505) supplied with the decoded sound signal (112).

36. A post-processing device (108) as defined in claim 32, wherein, for said at least one of the frequency sub-band signals:

the post-processing means comprises an adaptive filter (201 a; 307) supplied with the decoded sound signal (112) to produce an adaptively filtered decoded sound signal (204 a; S_{LE}); and
the dividing means comprises a sub-band filter (202a) supplied with the adaptively filtered decoded sound signal (204 a; S_{LE}).

37. A post-processing device (108) as defined in claim 32, wherein:

the dividing means comprises:

- a high-pass filter (301) supplied with the decoded sound signal (112) to produce a frequency high-band signal (310); and
- a first low-pass filter (305) supplied with the decoded sound signal (112) to produce a frequency low-band signal (311); and

the post-processing means comprises:

- a post-processor (307) for post-processing the decoded sound signal (112) prior to low-pass filtering the decoded sound signal (112) through the first low-pass filter (305).

38. A post-processing device (108) as defined in claim 37, wherein the post processor (307) comprises a pitch enhancer (304) supplied with the decoded sound signal (112) to produce a pitch enhanced decoded sound signal (S_{LE}).

39. A post-processing device (108) as defined in claim 38, wherein the post-processor (307) further comprises a second low-pass filter (302) supplied with the decoded sound signal (112) to produce a low-pass filtered decoded sound signal (S_L) supplied to the pitch enhancer (304).

40. A post-processing device (108) as defined in claim 37, further comprising an adder (306) for summing the frequency high-band (310) and low-band signals (311) to produce an output post-processed decoded sound signal (113).

41. A post-processing device (108) as defined in claim 32, wherein:

the dividing means comprises:

- a band-pass filter (407) supplied with the decoded sound signal to produce a frequency upper-band signal (410); and
- a low-pass filter (404) supplied with the decoded sound signal to produce a frequency lower-band signal; and

the post-processing means comprises:

- a post-processor (402; 401) for post-processing the decoded sound signal prior to low-pass filtering the decoded sound signal through the low-pass filter (404) to produce the frequency lower-band signal.

42. A post-processing device (108) as defined in claim 41, wherein the post-processor comprises a pitch filter (402) supplied with the decoded sound signal (s) to produce a pitch enhanced decoded sound signal (S_E) supplied to the low-pass filter (404).

43. A post-processing device (108) as defined in claim 41, further comprising an adder (409) for summing the frequency upper-band and lower-band signals to produce an output post-processed decoded sound signal.

44. A post-processing device (108) as defined in claim 32, wherein:

the dividing means comprises:

- a low-pass filter (505) supplied with the decoded sound signal (112) to produce a frequency low-band signal (508); and

the post-processing means comprises:

- a post-processor (503; 504; 502) for post-processing the decoded sound signal (112) to produce a post-processed decoded sound signal supplied to the low-pass filter (505).

45. A post-processing device (108) as defined in claim 44, wherein the post-processor (503; 504; 502) comprises an inter-harmonic filter (503) supplied with the decoded sound signal (112) to produce an inter-harmonic, attenuated decoded sound signal (507).

46. A post-processing device (108) as defined in claim 45, wherein the post-processor (503; 504; 502) comprises a multiplier (504) for multiplying the inter-harmonic, attenuated decoded sound signal (507) by an adaptive pitch enhancement gain (α).

47. A post-processing device (108) as defined in claim 45, further comprising a low-pass filter (501) supplied with the decoded sound signal (112) to produce a low-pass filtered decoded sound signal (S_{LP}) supplied to the inter-harmonic filter (503).

48. A post-processing device (108) as defined in claim 44, further comprising an adder (506) for summing the decoded sound signal (112) and the frequency low-band signal (508) to produce an output post-processed decoded sound signal (509).

49. A post-processing device (108) as defined in claim 44, wherein the post-processor (503; 504; 502) comprises an inter-harmonic filter (503) having the following transfer function:

$$y[n] = \frac{1}{2}x[n] - \frac{1}{4}\{x[n-T] + x[n+T]\}$$

for inter-harmonic attenuating the decoded sound signal, where *x[n]* is the decoded sound signal, *y[n]* is the inter-harmonic filtered decoded sound signal in a given sub-band, and T is a pitch delay of the decoded sound signal.

50. A post-processing device (108) as defined in claim 49, further comprising an adder (506) for summing the unprocessed decoded sound signal (112) and the inter-harmonic filtered frequency low-band signal (508) to produce an output post-processed decoded sound signal (509).

51. A post-processing device (108) as defined in claim 32, wherein the post-processing means (307) comprises a pitch enhancer (304) of the decoded sound signal (112) using the following equation:

$$y[n] = \left(1 - \frac{\alpha}{2}\right)x[n] + \frac{\alpha}{4}\{x[n-T] + x[n+T]\}$$

where *x[n]* is the decoded sound signal, *y[n]* is the pitch enhanced decoded sound signal in a given sub-band, T is a pitch delay of the decoded sound signal, and α is a coefficient varying between 0 and 1 to control an amount of inter-harmonic attenuation of the decoded sound signal (112).

52. A post-processing device (108) as defined in claim 51, comprising means for receiving the pitch delay T through a bitstream.

53. A post-processing device (108) as defined in claim 51, comprising means for decoding the pitch delay *T* from a received, encoded bitstream.

54. A post-processing device (108) as defined in claim 51, comprising means for calculating the pitch delay T in response to the decoded sound signal for an improved pitch tracking.

55. A post-processing device (108) as defined in claim 32, wherein, during encoding, the sound signal is down-sampled from a higher sampling frequency to a lower sampling frequency, and wherein the dividing means comprises means for up-sampling (403, 404, 405; 406, 407, 408) the decoded sound signal from the lower sampling frequency to the higher sampling frequency.

56. A post-processing device (108) as defined in claim 55, wherein the dividing means comprises sub-band filter means (407) supplied with the decoded sound signal, and wherein the up-sampling means (406) is combined with the sub-band filter means (407).

57. A post-processing device (108) as defined in claim 55, wherein:

- the post-processing means comprises:

means for post-processing (402; 401) the decoded sound signal; and
- the dividing means comprises:

a band-pass filter (407) supplied with the decoded sound signal to produce a frequency upper-band signal, said band-pass filter (407) being combined with the up-sampling means (406, 407, 408); and
a low-pass filter (404) supplied with the post-processed decoded sound signal to produce a frequency lower-band signal, said low-pass filter (404) being combined with the up-sampling means (403, 404, 405).

58. A post-processing device (108) as defined in claim 57, further comprising an adder (409) for summing the frequency upper-band signal (410) with the frequency lower-band signal to form an output post-processed and up-sampled decoded sound signal.

59. A post-processing device (108) as defined in claim 57, wherein the means for post-processing the decoded sound signal comprises means for pitch enhancing (402) the decoded sound signal to reduce an inter-harmonic noise in the decoded sound signal.

60. A post-processing device (108) as defined in claim 59, wherein the pitch enhancing means (402) comprises means for processing the decoded sound signal by means of the following equation:

$$y[n]=\left(1-\frac{\alpha}{2}\right)x[n]+\frac{\alpha}{4}\{x[n-T]+x[n+T]\}$$

where *x[n]* is the decoded sound signal, *y[n]* is the pitch enhanced decoded sound signal in a given sub-band, T is a pitch delay of the decoded sound signal, and α is a coefficient varying between 0 and 1 to control an amount of inter-harmonic attenuation of the decoded sound signal.

61. A post-processing device (108) as defined in claim 32, wherein:

the dividing means comprises means for dividing the decoded sound signal into a frequency upper-band signal

(711) and a frequency lower-band signal (713); and
the post-processing means (703) comprises means for post-processing the frequency lower-band signal.

62. A post-processing device (108) as defined in claim 32, wherein the post-processing means comprises:

means (303; 401; 502) for determining a pitch value of the decoded sound signal;
means for calculating, in relation to the determined pitch value, a high-pass filter with a cut-off frequency below a fundamental frequency of the decoded sound signal; and
means for processing the decoded sound signal (112) through the calculated high-pass filter.

63. A sound signal decoder (105) comprising:

an input for receiving an encoded sound signal (110);
a parameter decoder (106) supplied with the encoded sound signal (110) for decoding sound signal encoding parameters;
a sound signal decoder (107) supplied with the decoded sound signal encoding parameters for producing a decoded sound signal (112); and
a post processing device (108) as recited in any of claims 32 to 62 for post-processing the decoded sound signal (112) in view of enhancing a perceived quality of said decoded sound signal (112).

Patentansprüche

1. Verfahren zur Nachverarbeitung eines dekodierten Tonsignals (112) in Hinblick auf eine Verbesserung einer wahrgenommenen Qualität des dekodierten Tonsignals (112), umfassend:

Teilen des dekodierten Tonsignals (112) in mehrere Frequenz-Sub-Band-Signale; und
Anwenden einer Nachverarbeitung an mindestens einem der Frequenz-Sub-Band-Signale;

dadurch gekennzeichnet, dass zur Tonhöhenverbesserung die Nachverarbeitung nur bei einem Lower-Sub-Band der Frequenz-Sub-Band-Signale angewendet wird.

2. Verfahren zur Nachverarbeitung nach Anspruch 1, des Weiteren umfassend das Summieren der Frequenz-Sub-Band-Signale nach der Nachverarbeitung des mindestens einen Frequenz-Sub-Band-Signals, um ein ausgegebenes, nachverarbeitetes, dekodiertes Tonsignal (113) zu erzeugen.

3. Verfahren zur Nachverarbeitung nach Anspruch 1, wobei die Nachverarbeitung an mindestens einem der Frequenz-Sub-Band-Signale das adaptive Filtern des mindestens einen Frequenz-Sub-Band-Signals umfasst.

4. Verfahren zur Nachverarbeitung nach Anspruch 1, wobei das Teilen des dekodierten Tonsignals (112) in mehrere Frequenz-Sub-Band-Signale das Sub-Band-Filtern des dekodierten Tonsignals (112) umfasst, um die mehreren Frequenz-Sub-Band-Signale zu erzeugen.

5. Verfahren zur Nachverarbeitung nach Anspruch 1, wobei für das mindestens eine der Frequenz-Sub-Band-Signale:
das Anwenden der Nachverarbeitung das adaptive Filtern des dekodierten Tonsignals (112) umfasst; und
das Teilen des dekodierten Tonsignals (112) das Sub-Band-Filtern des adaptiv gefilterten, dekodierten Tonsignals (112) umfasst.

6. Verfahren zur Nachverarbeitung nach Anspruch 1, wobei:

das Teilen des dekodierten Tonsignals in mehrere Frequenz-Sub-Band-Signale umfasst:

- ein Hochpassfiltern des dekodierten Tonsignals (112), um ein Frequenz-High-Band-Signal (310) zu erzeugen; und
- ein erstes Tiefpassfiltern des dekodierten Tonsignals (112), um ein Frequenz-Low-Band-Signal (311) zu erzeugen; und

das Anwenden der Nachverarbeitung an dem mindestens einen der Frequenz-Sub-Band-Signale umfasst:

- das Anwenden einer Nachverarbeitung an dem dekodierten Tonsignal (112) vor dem ersten Tiefpassfiltern des dekodierten Tonsignals (112), um das Frequenz-Low-Band-Signal (311) zu erzeugen.

7. Verfahren zur Nachverarbeitung nach Anspruch 6, wobei das Anwenden der Nachverarbeitung an dem dekodierten Tonsignal (112) die Tonhöhenverbesserung des dekodierten Tonsignals (112) umfasst, um ein zwischenharmonisches Rauschen in dem dekodierten Tonsignals (112) zu verringern.

8. Verfahren zur Nachverarbeitung nach Anspruch 7, wobei das Anwenden der Nachverarbeitung an dem dekodierten Tonsignal (112) des Weiteren ein zweites Tiefpassfiltern des dekodierten Tonsignals (112) vor der Tonhöhenverbesserung des dekodierten Tonsignals (112) umfasst.

9. Verfahren zur Nachverarbeitung nach Anspruch 6, des Weiteren umfassend das Summieren der Frequenz-High-Band- (310) und der Frequenz-Low-Band-Signale (311), um ein ausgegebenes, nachverarbeitetes, dekodiertes Tonsignal (113) zu erzeugen.

10. Verfahren zur Nachverarbeitung nach Anspruch 1, wobei:

das Teilen des dekodierten Tonsignals (112) in mehrere Frequenz-Sub-Band-Signale umfasst:

- ein Bandpassfiltern des dekodierten Tonsignals (112), um ein Frequenz-Upper-Band-Signal (410) zu erzeugen; und
- ein Tiefpassfiltern des dekodierten Tonsignals (112), um ein Frequenz-Lower-Band-Signal zu erzeugen; und

das Anwenden der Nachverarbeitung an dem mindestens einen der Frequenz-Sub-Band-Signale umfasst:

- das Anwenden einer Nachverarbeitung an dem dekodierten Tonsignal (112) vor dem Tiefpassfiltern des dekodierten Tonsignals (112), um das Frequenz-Lower-Band-Signal zu erzeugen.

11. Verfahren zur Nachverarbeitung nach Anspruch 10, wobei das Anwenden der Nachverarbeitung an dem Frequenz-Lower-Band-Signal die Tonhöhenverbesserung des dekodierten Tonsignals (112) vor dem Tiefpassfiltern des dekodierten Tonsignals (112) umfasst.

12. Verfahren zur Nachverarbeitung nach Anspruch 10, des Weiteren umfassend das Summieren der Frequenz-Upper-Band- und der Frequenz-Lower-Band-Signale, um ein ausgegebenes, nachverarbeitetes, dekodiertes Tonsignal zu erzeugen.

13. Verfahren zur Nachverarbeitung nach Anspruch 1, wobei:

das Teilen des dekodierten Tonsignals (112) in mehrere Frequenz-Sub-Band-Signale umfasst:

- ein Tiefpassfiltern des dekodierten Tonsignals (112), um ein Frequenz-Low-Band-Signal zu erzeugen; und

das Anwenden der Nachverarbeitung an dem mindestens einen der Frequenz-Sub-Band-Signale umfasst:

- das Anwenden einer Nachverarbeitung an dem Frequenz-Low-Band-Signal.

14. Verfahren zur Nachverarbeitung nach Anspruch 13, wobei das Anwenden der Nachverarbeitung an dem Frequenz-Low-Band-Signal das Verarbeiten des dekodierten Tonsignals (112) durch ein zwischenharmonisches Filter (503) zur Dämpfung der Zwischenharmonischen des dekodierten Tonsignals (112) umfasst.

15. Verfahren zur Nachverarbeitung nach Anspruch 14, wobei das Anwenden der Nachverarbeitung an dem Frequenz-Low-Band-Signal das Multiplizieren des zwischenharmonisch gefilterten, dekodierten Tonsignals (507) mit einer adaptiven Tonhöhenverstärkung (Pitch Enhancement Gain) (α) umfasst.

16. Verfahren zur Nachverarbeitung nach Anspruch 14, des Weiteren umfassend ein Tiefpassfiltern des dekodierten Tonsignals (112) vor der Verarbeitung des dekodierten Tonsignals (112) durch das zwischenharmonische Filter (503) .

17. Verfahren zur Nachverarbeitung nach Anspruch 13, des Weiteren umfassend das Summieren des dekodierten Tonsignals (112) und des Frequenz-Low-Band-Signals, um ein ausgegebenes, nachverarbeitetes, dekodiertes Tonsignal (509) zu erzeugen.

18. Verfahren zur Nachverarbeitung nach Anspruch 13, wobei das Anwenden der Nachverarbeitung an dem Frequenz-Low-Band-Signal das Verarbeiten des dekodierten Tonsignals (112) durch ein zwischenharmonisches Filter (503) mit der folgenden Übertragungsfunktion umfasst:

$$y[n] = \frac{1}{2}x[n] - \frac{1}{4}\{x[n-T] + x[n+T]\}$$

zur Dämpfung der Zwischenharmonischen des dekodierten Tonsignals, wobei *x[n]* das dekodierte Tonsignal ist, *y [n]* das zwischenharmonisch gefilterte, dekodierte Tonsignals in einem bestimmten Sub-Band ist, und T eine Tonhöhenverzögerung (Pitch Delay) des dekodierten Tonsignals ist.

19. Verfahren zur Nachverarbeitung nach Anspruch 18, des Weiteren umfassend das Summieren des unverarbeiteten dekodierten Tonsignals (112) und des zwischenharmonisch gefilterten Frequenz-Low-Band-Signals (508), um ein ausgegebenes, nachverarbeitetes, dekodiertes Tonsignal (509) zu erzeugen.

20. Verfahren zur Nachverarbeitung nach Anspruch 1, wobei das Anwenden der Nachverarbeitung an mindestens einem der Frequenz-Sub-Band-Signale die Tonhöhenverbesserung des dekodierten Tonsignals (112) unter Verwendung der folgenden Gleichung umfasst:

$$y[n] = \left(1 - \frac{\alpha}{2}\right)x[n] + \frac{\alpha}{4}\{x[n-T] + x[n+T]\}$$

wobei *x[n]* das dekodierte Tonsignal ist, *y[n]* das tonhöhenverbesserte dekodierte Tonsignal in einem bestimmten Sub-Band ist, und T eine Tonhöhenverzögerung des dekodierten Tonsignals ist, und α ein Koeffizient ist, der zwischen 0 und 1 variiert, um ein Ausmaß der Dämpfung der Zwischenharmonischen des dekodierten Tonsignals (112) zu steuern.

21. Verfahren zur Nachverarbeitung nach Anspruch 21, umfassend das Empfangen der Tonhöhenverzögerung *T* durch einen Bitstrom.

22. Verfahren zur Nachverarbeitung nach Anspruch 20, umfassend das Dekodieren der Tonhöhenverzögerung *T* aus einem empfangenen, kodierten Bitstrom.

23. Verfahren zur Nachverarbeitung nach Anspruch 20, umfassend das Berechnen der Tonhöhenverzögerung *T* als Reaktion auf das dekodierte Tonsignal (112) für eine verbesserte Tonhöhenerkennung (Pitch Tracking).

24. Verfahren zur Nachverarbeiturig nach Anspruch 1, wobei während des Kodierens das Tonsignal von einer höheren Abtastfrequenz auf eine tiefere Abtastfrequenz frequenzgesenkt wird, und wobei das Teilen des dekodierten Tonsignals (112) in mehrere Frequenz-Sub-Band-Signale das Frequenzanheben des dekodierten Tonsignals von der tieferen Abtastfrequenz auf die höhere Abtastfrequenz umfasst.

25. Verfahren zur Nachverarbeitung nach Anspruch 24, wobei das Teilen des dekodierten Tonsignals (112) in mehrere Frequenz-Sub-Band-Signale das Sub-Band-Filtern des dekodierten Tonsignals (112) umfasst, und wobei das Frequenzanheben des dekodierten Tonsignals (112) von der tieferen Abtastfrequenz auf die höhere Abtastfrequenz mit dem Sub-Band-Filtern kombiniert ist.

26. Verfahren zur Nachverarbeitung nach Anspruch 24, umfassend:

Bandpassfiltern des dekodierten Tonsignals (112), um ein Frequenz-Upper-Band-Signal zu erzeugen, wobei das Bandpassfiltern des dekodierten Tonsignals (112) mit dem Frequenzanheben des dekodierten Tonsignals (112) von der tieferen Abtastfrequenz auf die höhere Abtastfrequenz kombiniert ist; und
Nachverarbeiten des dekodierten Tonsignals (112) und Tiefpassfiltern des nachverarbeiteten, dekodierten Tonsignals (112), um ein Frequenz-Lower-Band-Signal zu erzeugen, wobei das Tiefpassfiltern des nachverarbeiteten dekodierten Tonsignals mit dem Frequenzanheben des nachverarbeiteten dekodierten Tonsignals von der tieferen Abtastfrequenz auf die höhere Abtastfrequenz kombiniert ist.

27. Verfahren zur Nachverarbeitung nach Anspruch 26, des Weiteren umfassend das Addieren des Frequenz-Upper-Band-Signals zu dem Frequenz-Lower-Band-Signal zur Bildung eines ausgegebenen, nachverarbeiteten und frequenzangehobenen dekodierten Tonsignals.

28. Verfahren zur Nachverarbeitung nach Anspruch 26, wobei die Nachverarbeitung des dekodierten Tonsignals (112) die Tonhöhenverbesserung des dekodierten Tonsignals (112) umfasst, um ein zwischenharmonisches Rauschen in dem dekodierten Tonsignal (112) zu verringern.

29. Verfahren zur Nachverarbeitung nach Anspruch 28, wobei die Tonhöhenverbesserung des dekodierten Tonsignals (112) eine Verarbeitung des dekodierten Tonsignals (112) unter Verwendung der folgenden Gleichung umfasst:

$$y[n] = \left(1 - \frac{\alpha}{2}\right) x[n] + \frac{\alpha}{4}\{x[n-T] + x[n+T]\}$$

wobei *x[n]* das dekodierte Tonsignal ist, *y[n]* das tonhöhenverbesserte dekodierte Tonsignal in einem bestimmten Sub-Band ist, *T* eine Tonhöhenverzögerung des dekodierten Tonsignals ist, und α ein Koeffizient ist, der zwischen 0 und 1 variiert, um ein Ausmaß der Dämpfung der Zwischenharmonischen des dekodierten Tonsignals zu steuern.

30. Verfahren zur Nachverarbeitung nach Anspruch 1, wobei:

das Teilen des dekodierten Tonsignals (112) in mehrere Frequenz-Sub-Band-Signale das Teilen des dekodierten Tonsignals (112) in ein Frequenz-Upper-Band-Signal und ein Frequenz-Lower-Band-Signal umfasst; und
das Anwenden der Nachverarbeitung an mindestens einem der Frequenz-Sub-Band-Signale das Nachverarbeiten des Frequenz-Lower-Band-Signals umfasst.

31. Verfahren zur Nachverarbeitung nach Anspruch 1, wobei das Anwenden der Nachverarbeitung an dem mindestens einen der Frequenz-Sub-Band-Signale umfasst:

Bestimmen eines Tonhöhenwertes des dekodierten Tonsignals;
Berechnen, im Verhältnis zu dem bestimmten Tonhöhenwert, eines Hochpassfilters mit einer Sperrfrequenz unter einer Grundfrequenz des dekodierten Tonsignals; und
Verarbeiten des dekodierten Tonsignals durch das berechnete Hochpassfilter.

32. Vorrichtung zur Nachverarbeitung (108) eines dekodierten Tonsignals (112) in Hinblick auf eine Verbesserung einer wahrgenommenen Qualität des dekodierten Tonsignals (112), umfassend:

ein Mittel zum Teilen (202a, bis 202N; 301, 305; 407, 404; 505) des dekodierten Tonsignals (112) in mehrere Frequenz-Sub-Band-Signale; und
ein Mittel zum Nachverarbeiten (201a bis 201N, 307; 401, 402; 503, 504, 502) mindestens eines der Frequenz-Sub-Band-Signale;
ein Mittel zur Tonhöhenverbesserung eines Sub-Band-Signals; und

dadurch gekennzeichnet, dass das Nachverarbeitungsmittel dazu ausgebildet ist, nur ein Lower Sub-Band der Frequenz-Sub-Band-Signale zu dem Tonhöhenverbesserungsmittel zu leiten.

33. Vorrichtung zur Nachverarbeitung (108) nach Anspruch 32, des Weiteren umfassend ein Additionsmittel (203; 306; 409; 506) zum Summieren der Frequenz-Sub-Band-Signale nach der Nachverarbeitung des mindestens einen

Frequenz-Sub-Band-Signals, um ein ausgegebenes, nachverarbeitetes, dekodiertes Tonsignal (113) zu erzeugen.

34. Vorrichtung zur Nachverarbeitung (108) nach Anspruch 32, wobei das Nachverarbeitungsmittel ein adaptives Filtermittel (201a bis 201N; 307) umfasst, dem das dekodierte Tonsignal (112) zugeleitet wird.

35. Vorrichtung zur Nachverarbeitung (108) nach Anspruch 32, wobei das Teilungsmittel ein Sub-Band-Filtermittel (202a bis 202N; 301, 305; 407, 404; 505) umfasst, dem das dekodierte Tonsignal (112) zugeleitet wird.

36. Vorrichtung zur Nachverarbeitung (108) nach Anspruch 32, wobei, für das mindestens eine der Frequenz-Sub-Band-Signale:

das Nachverarbeitungsmittel ein adaptives Filtermittel (201a; 307) umfasst, dem das dekodierte Tonsignal (112) zugeleitet wird, um ein adaptiv gefiltertes, dekodiertes Tonsignal (204a; S_{LE}) zu erzeugen; und
das Teilungsmittel ein Sub-Band-Filtermittel (202a) umfasst, dem das adaptiv gefilterte, dekodierte Tonsignal (204a; S_{LE}) zugeleitet wird.

37. Vorrichtung zur Nachverarbeitung (108) nach Anspruch 32, wobei:

das Teilungsmittel umfasst:

- ein Hochpassfilter (301, dem das dekodierte Tonsignal (112) zugeleitet wird, um ein Frequenz-High-Band-Signal (310) zu erzeugen; und
- ein erstes Tiefpassfilter (305), dem das dekodierte Tonsignal (112) zugeleitet wird, um ein Frequenz-Low-Band-Signal (311) zu erzeugen; und

das Nachverarbeitungsmittel umfasst:

- ein Nachverarbeitungsmittel (307) zum Nachverarbeiten des dekodierten Tonsignals (112) vor dem Tiefpassfiltern des dekodierten Tonsignals (112) durch das erste Tiefpassfilter (305).

38. Vorrichtung zur Nachverarbeitung (108) nach Anspruch 37, wobei das Nachverarbeitungsmittel (307) einen Tonhöhenverbesserer (304) umfasst, dem das dekodierte Tonsignal (112) zugeleitet wird, um ein tonhöhenverbessertes, dekodiertes Tonsignal (S_{LE}) zu erzeugen.

39. Vorrichtung zur Nachverarbeitung (108) nach Anspruch 38, wobei das Nachverarbeitungsmittel (307) des Weiteren ein zweites Tiefpassfilter (302) umfasst, dem das dekodierte Tonsignal (112) zugeleitet wird, um ein tiefpassgefiltertes, dekodiertes Tonsignal (S_L) zu erzeugen, das dem Tonhöhenverbesserer (304) zugeleitet wird.

40. Vorrichtung zur Nachverarbeitung (108) nach Anspruch 37, des Weiteren umfassend einen Addierer (306) zum Summieren der Frequenz-High-Band- (310) und der Frequenz-Low-Band-Signale (311), um ein ausgegebenes, nachverarbeitetes, dekodiertes Tonsignal (113) zu erzeugen.

41. Vorrichtung zur Nachverarbeitung (108) nach Anspruch 32, wobei:

das Teilungsmittel umfasst:

- ein Bandpassfilter (407), dem das dekodierte Tonsignal zugeleitet wird, um ein Frequenz-Upper-Band-Signal (410) zu erzeugen; und
- ein Tiefpassfilter (404), dem das dekodierte Tonsignal zugeleitet wird, um ein Frequenz-Lower-Band-Signal (410) zu erzeugen; und

das Nachverarbeitungsmittel umfasst:

- ein Nachverarbeitungsmittel (402 ; 401) zum Nachverarbeiten des dekodierten Tonsignals vor dem Tiefpassfiltern des dekodierten Tonsignals durch das Tiefpassfilter (404), um das Frequenz-Lower-Band-Signal zu erzeugen.

42. Vorrichtung zur Nachverarbeitung (108) nach Anspruch 41, wobei das Nachverarbeitungsmittel ein Tonhöhenfilter

(402) umfasst, dem das dekodierte Tonsignal (s) zugeleitet wird, um ein tonhöhenverbessertes, dekodiertes Tonsignal (S_E) zu erzeugen, das dem Tiefpassfilter (404) zugeleitet wird.

43. Vorrichtung zur Nachverarbeitung (108) nach Anspruch 41, des Weiteren umfassend einen Addierer (409) zum Summieren der Frequenz-Upper-Band- und der Frequenz-Lower-Band-Signale, um ein ausgegebenes, nachverarbeitetes, dekodiertes Tonsignal zu erzeugen.

44. Vorrichtung zur Nachverarbeitung (108) nach Anspruch 32, wobei:

das Teilungsmittel umfasst:

- ein Tiefpassfilter (505), dem das dekodierte Tonsignal (112) zugeleitet wird, um ein Frequenz-Low-Band-Signal (508) zu erzeugen; und

das Nachverarbeitungsmittel umfasst:

- ein Nachverarbeitungsmittel (503; 504; 503) zum Nachverarbeiten des dekodierten Tonsignals (112), um ein nachverarbeitetes, dekodiertes Tonsignal zu erzeugen, das dem Tiefpassfilter (505) zugeleitet wird.

45. Vorrichtung zur Nachverarbeitung (108) nach Anspruch 44, wobei das Nachverarbeitungsmittel (503; 504; 502) ein zwischenharmonisches Filter (503) umfasst, dem das dekodierte Tonsignal zugeleitet wird, um ein zwischenharmonisch gedämpftes, dekodiertes Tonsignal (507) zu erzeugen.

46. Vorrichtung zur Nachverarbeitung (108) nach Anspruch 45, wobei das Nachverarbeitungsmittel (503; 504; 502) einen Multiplizierer (504) umfasst, um das zwischenharmonisch gedämpfte, dekodierte Tonsignal (507) mit einer adaptiven Tonhöhenverstärkung (α) zu multiplizieren.

47. Vorrichtung zur Nachverarbeitung (108) nach Anspruch 45, des Weiteren umfassend ein Tiefpassfilter (501), dem das dekodierte Tonsignal (112) zugeleitet wird, um ein tiefpassgefiltertes, dekodiertes Tonsignal (S_{LP}) zu erzeugen, das dem zwischenharmonischen Filter (503) zugeleitet wird.

48. Vorrichtung zur Nachverarbeitung (108) nach Anspruch 44, des Weiteren umfassend einen Addierer (506) zum Summieren des dekodierten Tonsignals (112) und des Frequenz-Low-Band-Signals (508), um ein ausgegebenes, nachverarbeitetes, dekodiertes Tonsignal (509) zu erzeugen.

49. Vorrichtung zur Nachverarbeitung (108) nach Anspruch 44, wobei das Nachverarbeitungsmittel (503; 504; 503) ein zwischenharmonisches Filter (503) mit der folgenden Übertragungsfunktion umfasst:

$$y[n] = \frac{1}{2}x[n] - \frac{1}{4}\{x[n-T] + x[n+T]\}$$

zur Dämpfung der Zwischenharmonischen des dekodierten Tonsignals, wobei *x[n]* das dekodierte Tonsignal ist, *y [n]* das zwischenharmonisch gefilterte, dekodierte Tonsignal in einem bestimmten Sub-Band ist, und T eine Tonhöhenverzögerung des dekodierten Tonsignals ist.

50. Vorrichtung zur Nachverarbeitung (108) nach Anspruch 49, des Weiteren umfassend einen Addierer (506) zum Summieren des unverarbeiteten dekodierten Tonsignals (112) und des zwischenharmonisch gefilterten Frequenz-Low-Band-Signals (508), um ein ausgegebenes, nachverarbeitetes, dekodiertes Tonsignal (509) zu erzeugen.

51. Vorrichtung zur Nachverarbeitung (108) nach Anspruch 32, wobei das Nachverarbeitungsmittel (307) einen Tonhöhenverbesserer (304) des dekodierten Tonsignals (112) unter Verwendung der folgenden Gleichung umfasst:

$$y[n]=\left(1-\frac{\alpha}{2}\right)x[n]+\frac{\alpha}{4}\{x[n-T]+x[n+T]\}$$

wobei *x[n]* das dekodierte Tonsignal ist, *y[n]* das tonhöhenverbesserte dekodierte Tonsignal in einem bestimmten Sub-Band ist, T eine Tonhöhenverzögerung des dekodierten Tonsignals ist, und α ein Koeffizient ist, der zwischen 0 und 1 variiert, um ein Ausmaß der Dämpfung der Zwischenharmonischen des dekodierten Tonsignals (112) zu steuern.

52. Vorrichtung zur Nachverarbeitung (108) nach Anspruch 51, umfassend ein Mittel zum Empfangen der Tonhöhenverzögerung T durch einen Bitstrom.

53. Vorrichtung zur Nachverarbeitung (108) nach Anspruch 51, umfassend ein Mittel zum Dekodieren der Tonhöhenverzögerung T aus einem empfangenen, kodierten Bitstrom.

54. Vorrichtung zur Nachverarbeitung (108) nach Anspruch 51, umfassend ein Mittel zum Berechnen der Tonhöhenverzögerung T als Reaktion auf das dekodierte Tonsignal für eine verbesserte Tonhöhenerkennung.

55. Vorrichtung zur Nachverarbeitung (108) nach Anspruch 32, wobei während des Kodierens das Tonsignal von einer höheren Abtastfrequenz auf eine tiefere Abtastfrequenz frequenzgesenkt wird, und wobei das Teilungsmittel ein Mittel zum Frequenzanheben (403, 404, 405; 406, 407, 408) des dekodierten Tonsignals von der tieferen Abtastfrequenz auf die höhere Abtastfrequenz umfasst.

56. Vorrichtung zur Nachverarbeitung (108) nach Anspruch 55, wobei das Teilungsmittel ein Sub-Band-Filtermittel (407), dem das dekodierte Tonsignal zugeleitet wird, umfasst und wobei das Mittel zum Frequenzanheben (406) mit dem Sub-Band-Filtermittel (407) kombiniert ist.

57. Vorrichtung zur Nachverarbeitung (108) nach Anspruch 55, wobei:

- das Nachverarbeitungsmittel umfasst:

  ein Mittel zum Nachverarbeiten (402; 401) des dekodierten Tonsignals; und

- das Teilungsmittel umfasst:

  ein Bandpassfilter (407), dem das dekodierte Tonsignal zugeleitet wird, um ein Frequenz-Upper-Band-Signal zu erzeugen, wobei das Bandpassfilter (407) mit dem Mittel zum Frequenzanheben (406, 407, 408) kombiniert ist; und
  ein Tiefpassfilter (404), dem das nachverarbeitete, dekodierte Tonsignal zugeleitet wird, um ein Frequenz-Lower-Band-Signal zu erzeugen, wobei das Tiefpassfilter (404) mit dem Mittel zum Frequenzanheben (403, 404, 405) kombiniert ist.

58. Vorrichtung zur Nachverarbeitung (108) nach Anspruch 57, des Weiteren umfassend einen Addierer (409) zum Summieren des Frequenz-Upper-Band-Signals (410) mit dem Frequenz-Lower-Band-Signal zur Bildung eines ausgegebenen, nachverarbeiteten und frequenzangehobenen, dekodierten Tonsignals.

59. Vorrichtung zur Nachverarbeitung (108) nach Anspruch 57, wobei das Mittel zur die Nachverarbeitung des dekodierten Tonsignals ein Mittel zur Tonhöhenverbesserung (402) des dekodierten Tonsignals umfasst, um ein zwischenharmonisches Rauschen in dem dekodierten Tonsignal zu verringern.

60. Vorrichtung zur Nachverarbeitung (108) nach Anspruch 59, wobei das Tonhöhenverbesserungsmittel (402) ein Mittel zum Verarbeiten des dekodierten Tonsignals unter Verwendung der folgenden Gleichung umfasst:

$$y[n]=\left(1-\frac{\alpha}{2}\right)x[n]+\frac{\alpha}{4}\{x[n-T]+x[n+T]\}$$

wobei *x[n]* das dekodierte Tonsignal ist, *y[n]* das tonhöhenverbesserte dekodierte Tonsignal in einem bestimmten Sub-Band ist, *T* eine Tonhöhenverzögerung des dekodierten Tonsignals ist, und α ein Koeffizient ist, der zwischen 0 und 1 variiert, um ein Ausmaß der Dämpfung der Zwischenharmonischen des dekodierten Tonsignals zu steuern.

61. Vorrichtung zur Nachverarbeitung (108) nach Anspruch 32, wobei:

das Teilungsmittel ein Mittel zum Teilen des dekodierten Tonsignals in ein Frequenz-Upper-Band-Signal (711) und ein Frequenz-Lower-Band-Signal (713) umfasst; und
das Nachverarbeitungsmittel (703) ein Mittel zum Nachverarbeiten des Frequenz-Lower-Band-Signals umfasst.

62. Vorrichtung zur Nachverarbeitung (108) nach Anspruch 32, wobei das Mittel zur Nachverarbeitung umfasst:

ein Mittel (301; 401; 502) zum Bestimmen eines Tonhöhenwertes des dekodierten Tonsignals;
ein Mittel zum Berechnen, im Verhältnis zu dem bestimmten Tonhöhenwert, eines Hochpassfilters mit einer Sperrfrequenz unter einer Grundfrequenz des dekodierten Tonsignals; und
ein Mittel zum Verarbeiten des dekodierten Tonsignals (112) durch das berechnete Hochpassfilter.

63. Tonsignaldekodierer (105), umfassend:

einen Eingang zum Empfangen eines kodierten Tonsignals (110);
einen Parameterdekodierer (106), dem das kodierte Tonsignal (110) zugeleitet wird, um die Tonsignal-Kodierungsparameter zu dekodieren;
einen Tonsignaldekodierer (107), dem die Kodierungsparameter des dekodierten Tonsignals zugeleitet werden, um ein dekodiertes Tonsignal (112) zu erzeugen; und
eine Nachverarbeitungsvorrichtung (108) nach einem der Ansprüche 32 bis 62, um das dekodierte Tonsignal (112) in Hinblick auf eine Verbesserung einer wahrgenommenen Qualität des dekodierten Tonsignals (112) nachzubearbeiten.

Revendications

1. Procédé de post-traitement d'un signal sonore décodé (112) en vue d'améliorer une qualité perçue dudit signal sonore décodé (112), comprenant :

diviser le signal sonore décodé (112) en une multitude de signaux de sous-bande de fréquences ; et
appliquer un post-traitement à au moins un des signaux de sous-bande de fréquences ;

caractérisé en ce que, pour accroître le pas, un post-traitement est appliqué à seulement une sous-bande inférieure des signaux de sous-bande de fréquences.

2. Procédé de post-traitement selon la revendication 1, comprenant en outre l'addition des signaux de sous-bande de fréquences après le post-traitement dudit au moins un signal de sous-bande de fréquences afin de produire un signal sonore décodé à valeur de sortie post-traitée (113).

3. Procédé de post-traitement selon la revendication 1, dans lequel l'application d'un post-traitement à au moins un des signaux de sous-bande de fréquences comprend le filtrage adaptatif dudit au moins un signal de sous-bande de fréquences.

4. Procédé de post-traitement selon la revendication 1, dans lequel la division du signal sonore décodé (112) en une multitude de signaux de sous-bande de fréquences comprend le filtrage de sous-bande du signal sonore décodé (112) afin de produire la multitude de signaux de sous-bande de fréquences.

5. Procédé de post-traitement selon la revendication 1, dans lequel, pour l'au moins un des signaux de sous-bande de fréquences :

   l'application d'un post-traitement comprend le filtrage adaptatif du signal sonore décodé (112) ; et
   la division du signal sonore décodé (112) comprend le filtrage de sous-bande du signal sonore décodé filtré de manière adaptative.

6. Procédé de post-traitement selon la revendication 1, dans lequel :

   la division du signal sonore décodé en une multitude de signaux de sous-bande de fréquences comprend :

   - un filtrage passe-haut du signal sonore décodé (112) afin de produire un signal de haute bande de fréquence (310) ; et
   - un filtrage passe-bas du signal sonore décodé (112) afin de produire un signal de basse bande de fréquence (311) ; et

   l'application d'un post-traitement à au moins un des signaux de sous-bande de fréquences comprend :

   - l'application d'un post-traitement au signal sonore décodé (112) avant le premier filtrage passe-bas du signal sonore décodé (112) afin de produire le signal de basse bande de fréquence (311).

7. Procédé de post-traitement selon la revendication 6, dans lequel l'application d'un post-traitement au signal sonore décodé (112) comprend l'accroissement de pas dudit signal sonore décodé (112) afin de réduire un bruit d'inter-harmonique dans le signal sonore décodé (112).

8. Procédé de post-traitement selon la revendication 7, dans lequel l'application d'un post-traitement au signal sonore décodé (112) comprend en outre un deuxième filtrage passe-bas du signal sonore décodé (112) avant d'accroître le pas dudit signal sonore décodé (112).

9. Procédé de post-traitement selon la revendication 6, comprenant en outre l'addition des signaux de haute bande (310) et de basse bande (311) de fréquence afin de produire un signal sonore décodé à valeur de sortie post-traitée (113).

10. Procédé de post-traitement selon la revendication 1, dans lequel :

    la division du signal sonore décodé (112) en une multitude de signaux de sous-bande de fréquences comprend :

    - le filtrage passe-bande du signal sonore décodé (112) afin de produire un signal de bande supérieure de fréquence (410) ; et
    - le filtrage passe-bas du signal sonore décodé (112) afin de produire un signal de bande inférieure de fréquence ; et

    l'application d'un post-traitement à au moins un des signaux de sous-bande de fréquence comprend:

    - l'application d'un post-traitement au signal sonore décodé (112) avant le filtrage passe-bas du signal sonore décodé (112) afin de produire le signal de bande inférieure de fréquence.

11. Procédé de post-traitement selon la revendication 10, dans lequel l'application d'un post-traitement au signal de bande inférieure de fréquence comprend l'accroissement de pas du signal sonore décodé (112) avant le filtrage passe-bas du signal sonore décodé (112).

12. Procédé de post-traitement selon la revendication 10, comprenant en outre l'addition des signaux de bande supérieure et de bande inférieure de fréquence afin de produire un signal sonore décodé à valeur de sortie post-traitée.

13. Procédé de post-traitement selon la revendication 1, dans lequel la division du signal sonore décodé (112) en une multitude de signaux de sous-bande de fréquences comprend:

    - le filtrage passe-bas du signal sonore décodé (112) afin de produire un signal de basse bande de fréquence ; et

l'application d'un post-traitement à au moins un des signaux de sous-bande de fréquence comprend :
- l'application d'un post-traitement au signal de basse bande de fréquence.

14. Procédé de post-traitement selon la revendication 13, dans lequel l'application d'un post-traitement au signal de basse bande de fréquence comprend le traitement du signal sonore décodé (112) au moyen d'un filtre d'inter-harmonique (503) en vue d'une atténuation d'inter-harmonique du signal sonore décodé (112) .

15. Procédé de post-traitement selon la revendication 14, dans lequel l'application d'un post-traitement au signal de basse bande de fréquence comprend la multiplication du signal sonore décodé à interharmonique filtré (507) par un gain d'accroissement de pas adaptatif (α).

16. Procédé de post-traitement selon la revendication 14, comprenant en outre le filtrage passe-bas du signal sonore décodé (112) avant le traitement du signal sonore décodé (112) au moyen du filtre d'inter-harmonique (503).

17. Procédé de post-traitement selon la revendication 13, comprenant en outre l'addition du signal sonore décodé (112) et du signal de basse bande de fréquence afin de produire un signal sonore décodé à valeur de sortie post-traitée (509).

18. Procédé de post-traitement selon la revendication 13, dans lequel l'application d'un post-traitement au signal de basse bande de fréquence comprend le traitement du signal sonore décodé (112) au moyen d'un filtre d'inter-harmonique (503) ayant la fonction de transfert suivante :

$$y[n] = \frac{1}{2}x[n] - \frac{1}{4}\{x[n-T] + x[n+T]\}$$

pour l'atténuation d'inter-harmonique du signal sonore décodé, *x[n]* étant le signal sonore décodé, *y[n]* étant le signal sonore décodé à inter-harmonique filtré dans une sous-bande donnée et *T* étant un retard de pas du signal sonore décodé.

19. Procédé de post-traitement selon la revendication 18, comprenant en outre l'addition du signal sonore décodé non traité (112) et du signal de basse bande de fréquence à inter-harmonique filtré (508) afin de produire un signal sonore décodé à valeur de sortie post-traitée (509).

20. Procédé de post-traitement selon la revendication 1, dans lequel l'application d'un post-traitement à au moins un des signaux de sous-bande de fréquence comprend l'accroissement de pas du signal sonore décodé (112) en utilisant l'équation suivante :

$$y[n] = \left(1 - \frac{\alpha}{2}\right)x[n] + \frac{\alpha}{4}\{x[n-T] + x[n+T]\}$$

x[n] étant le signal sonore décodé, *y[n]* étant le signal sonore décodé à pas accru dans une sous-bande donnée, *T* étant un retard de pas du signal sonore décodé et α étant un coefficient variant entre 0 et 1 destiné à contrôler une valeur d'atténuation d'inter-harmonique du signal sonore décodé.

21. Procédé de post-traitement selon la revendication 20, comprenant la réception du retard de pas *T* par le biais d'un train de bits.

22. Procédé de post-traitement selon la revendication 20, comprenant le décodage du retard de pas *T* provenant d'un train de bits codé reçu.

23. Procédé de post-traitement selon la revendication 20, comprenant le calcul du retard de pas *T* en réponse au signal sonore décodé (112) pour améliorer le suivi du pas.

24. Procédé de post-traitement selon la revendication 1, dans lequel, pendant le codage, le signal sonore est sous-échantillonné depuis une fréquence d'échantillonnage supérieure vers une fréquence d'échantillonnage inférieure et dans lequel la division du signal sonore décodé (112) en une multitude de signaux de sous-bande de fréquences comprend le sur-échantillonnage du signal sonore décodé depuis la fréquence d'échantillonnage inférieure vers la fréquence d'échantillonnage supérieure.

25. Procédé de post-traitement selon la revendication 24, dans lequel la division du signal sonore décodé (112) en une multitude de signaux de sous-bande de fréquences comprend le filtrage de sous-bande du signal sonore décodé (112) et dans lequel le sur-échantillonnage du signal sonore décodé (112) depuis la fréquence d'échantillonnage inférieure vers la fréquence d'échantillonnage supérieure est combiné au filtrage de sous-bande.

26. Procédé de post-traitement selon la revendication 24, comprenant :

le filtrage passe-bande du signal sonore décodé (112) afin de produire un signal de bande supérieure de fréquence, ledit filtrage passe-bande du signal sonore décodé (112) étant combiné au sur-échantillonnage du signal sonore décodé (112) depuis la fréquence d'échantillonnage inférieure vers la fréquence d'échantillonnage supérieure ; et

le post-traitement du signal sonore décodé (112) et le filtrage passe-bas du signal sonore décodé post-traité (112) afin de produire un signal de bande inférieure de fréquence, ledit filtrage passe-bas du signal sonore décodé post-traité étant combiné au sur-échantillonnage du signal sonore décodé post-traité depuis la fréquence d'échantillonnage inférieure vers la fréquence d'échantillonnage supérieure.

27. Procédé de post-traitement selon la revendication 26, comprenant en outre l'addition du signal de bande supérieure de fréquence au signal de bande inférieure de fréquence afin de former un signal sonore décodé à valeur de sortie post-traitée et sur-échantillonné.

28. Procédé de post-traitement selon la revendication 26, dans lequel le post-traitement du signal sonore décodé (112) comprend l'accroissement de pas du signal sonore décodé (112) afin de réduire un bruit d'inter-harmonique dans le signal sonore décodé (112).

29. Procédé de post-traitement selon la revendication 28, dans lequel l'accroissement de pas du signal sonore décodé (112) comprend le traitement du signal sonore décodé (112) au moyen de l'équation suivante:

$$y[n] = \left(1 - \frac{\alpha}{2}\right)x[n] + \frac{\alpha}{4}\{x[n-T] + x[n+T]\}$$

x[n] étant le signal sonore décodé, *y[n]* étant le signal sonore décodé à pas accru dans une sous-bande donnée, *T* étant un retard de pas du signal sonore décodé et α étant un coefficient variant entre 0 et 1 destiné à contrôler une valeur d'atténuation d'inter-harmonique du signal sonore décodé.

30. Procédé de post-traitement selon la revendication 1, dans lequel :

la division du signal sonore décodé (112) en une multitude de signaux de sous-bande de fréquences comprend la division du signal sonore décodé (112) en un signal de bande supérieure de fréquence et un signal de bande inférieure de fréquence ; et

l'application d'un post-traitement à au moins un des signaux de sous-bande de fréquences comprend le post-traitement du signal de bande inférieure de fréquence.

31. Procédé de post-traitement selon la revendication 1, dans lequel l'application d'un post-traitement à au moins un des signaux de sous-bande de fréquences comprend:

la détermination d'une valeur de pas du signal sonore décodé ;

le calcul, en relation avec la valeur de pas déterminée, d'un filtre passe-haut avec une fréquence de coupure inférieure à une fréquence fondamentale du signal sonore décodé ; et

le traitement du signal sonore décodé au moyen du filtre passe-haut calculé.

32. Dispositif de post-traitement (108) d'un signal sonore décodé (112) en vue d'améliorer une qualité perçue dudit signal sonore décodé (112), comprenant:

des moyens pour diviser (202a à 202N ; 301, 305 ; 407, 404 ; 505) le signal sonore décodé (112) en une multitude de signaux de sous-bande de fréquences ; et
des moyens pour post-traiter (201a à 201N ; 307 ; 401, 402 ; 503, 504, 502) au moins un des signaux de sous-bande de fréquences ;
un moyen d'accroissement de pas d'un signal de sous-bande ; et

caractérisé en ce que, pour accroître le pas, un post-traitement est appliqué à seulement une sous-bande inférieure des signaux de sous-bande de fréquences.

33. Dispositif de post-traitement (108) selon la revendication 32, comprenant en outre des moyens sommateurs (203 ; 306 ; 409 ; 506) pour additionner les signaux de sous-bande de fréquences après le post-traitement dudit au moins un signal de sous-bande de fréquences afin de produire un signal sonore décodé à valeur de sortie post-traitée (113).

34. Dispositif de post-traitement (108) selon la revendication 32, dans lequel le moyen de post-traitement comprend un moyen de filtrage adaptatif (201a à 201N ; 307) auquel est fourni le signal sonore décodé (112).

35. Dispositif de post-traitement (108) selon la revendication 32, dans lequel le moyen de division comprend un moyen de filtrage de sous-bande (202a à 202N ; 301, 305 ; 407, 404 ; 505) auquel est fourni le signal sonore décodé (112).

36. Dispositif de post-traitement (108) selon la revendication 32, dans lequel, pour ledit au moins un des signaux de sous-bande de fréquences :

le moyen de post-traitement comprend un filtre adaptatif (201a ; 307) auquel est fourni le signal sonore décodé (112) afin de produire un signal sonore décodé (204a ; S_{LE}) filtré de manière adaptative ; et
le moyen diviseur comprend un filtre de sous-bande (202a) auquel est fourni le signal sonore décodé (204a ; S_{LE}) filtré de manière adaptative.

37. Dispositif de post-traitement (108) selon la revendication 32, dans lequel :

le moyen diviseur comprend :

- un filtre passe-haut (301) auquel est fourni le signal sonore décodé (112) afin de produire un signal de haute bande de fréquence (310) ; et
- un filtre passe-bas (305) auquel est fourni le signal sonore décodé (112) afin de produire un signal de basse bande de fréquence (311) ; et

le moyen de post-traitement comprend :

- un post-processeur (307) servant à post-traiter le signal sonore décodé (112) avant le filtrage passe-bas du signal sonore décodé (112) au moyen du premier filtre passe-bas (305).

38. Dispositif de post-traitement (108) selon la revendication 37, dans lequel le post-processeur (307) comprend un amplificateur de pas (304) auquel est fourni le signal sonore décodé (112) afin de produire un signal sonore décodé à pas accru (S_{LE}).

39. Dispositif de post-traitement (108) selon la revendication 38, dans lequel le post-processeur (307) comprend en outre un deuxième filtre passe-bas (302) auquel est fourni le signal sonore décodé (112) afin de produire un signal sonore décodé filtré par filtre passe-bas (S_L) fourni à l'amplificateur de pas (304).

40. Dispositif de post-traitement (108) selon la revendication 37, comprenant en outre un sommateur (306) pour additionner les signaux de haute bande de fréquence (310) et de basse bande de fréquence (311) afin de produire un signal sonore décodé à valeur de sortie post-traitée (113).

41. Dispositif de post-traitement (108) selon la revendication 32, dans lequel :

le moyen diviseur comprend :

- un filtre passe-bande (407) auquel est fourni le signal sonore décodé afin de produire un signal de bande supérieure de fréquence (410) ; et
- un filtre passe-bas (404) auquel est fourni le signal sonore décodé afin de produire un signal de bande inférieure de fréquence (410) ; et

le moyen de post-traitement comprend :

- un post-processeur (402 ; 401) servant à post-traiter le signal sonore décodé avant le filtrage passe-bas du signal sonore décodé au moyen du filtre passe-bas (404) afin de produire le signal de bande inférieure de fréquence.

42. Dispositif de post-traitement (108) selon la revendication 41, dans lequel le post-processeur comprend un filtre de pas (402) auquel sont fournis le ou les signaux sonores décodés afin de produire un signal sonore décodé à pas accru (S_E) fourni au filtre passe-bas (404).

43. Dispositif de post-traitement (108) selon la revendication 41, comprenant en outre un sommateur (409) servant à additionner les signaux de bande supérieure et de bande inférieure de fréquence afin de produire un signal sonore décodé à valeur de sortie post-traitée.

44. Dispositif de post-traitement (108) selon la revendication 32, dans lequel:

le moyens diviseurs comprennent:

- un filtre passe-bas (505) auquel est fourni le signal sonore décodé (112) afin de produire un signal de basse bande de fréquence (508) ; et

les moyens de post-traitement comprennent :

- un post-processeur (503 ; 504 ; 502) servant à post-traiter le signal sonore décodé (112) afin de produire un signal sonore décodé post-traité fourni au filtre passe-bas (505).

45. Dispositif de post-traitement (108) selon la revendication 44, dans lequel le post-processeur (503 ; 504 ; 502) comprend un filtre d'inter-harmonique (503) auquel est fourni le signal sonore décodé (112) afin de produire un signal sonore décodé à inter-harmonique atténué (507).

46. Dispositif de post-traitement (108) selon la revendication 45, dans lequel le post-processeur (503 ; 504 ; 502) comprend un multiplicateur (504) servant à multiplier le signal sonore décodé à inter-harmonique atténué (507) par un gain d'accroissement de pas adaptatif (α).

47. Dispositif de post-traitement (108) selon la revendication 45, comprenant en outre un filtre passe-bas (501) auquel est fourni le signal sonore décodé (112) afin de produire un signal sonore décodé filtré par filtre passe-bas (S_{LP}) fourni au filtre d'inter-harmonique (503).

48. Dispositif de post-traitement (108) selon la revendication 44, comprenant en outre un sommateur (506) servant à additionner le signal sonore décodé (112) et le signal de basse bande de fréquence (508) afin de produire un signal sonore décodé à valeur de sortie post-traitée (509).

49. Dispositif de post-traitement (108) selon la revendication 44, dans lequel le post-processeur (503 ; 504 ; 502) comprend un filtre d'inter-harmonique (503) ayant la fonction de transfert suivante :

$$y[n] = \frac{1}{2}x[n] - \frac{1}{4}\{x[n-T] + x[n+T]\}$$

pour l'atténuation d'inter-harmonique du signal sonore décodé, *x[n]* étant le signal sonore décodé, *y[n]* étant le signal sonore décodé à inter-harmonique filtré dans une sous-bande donnée et *T* étant un retard de pas du signal sonore décodé.

50. Dispositif de post-traitement (108) selon la revendication 49, comprenant en outre un sommateur (506) servant à additionner le signal sonore décodé non traité (112) et le signal de basse bande de fréquence à inter-harmonique filtré (508) afin de produire un signal sonore décodé à valeur de sortie post-traitée (509).

51. Dispositif de post-traitement (108) selon la revendication 32, dans lequel le moyen de post-traitement (307) comprend un amplificateur de pas (304) du signal sonore décodé (112) utilisant l'équation suivante :

$$y[n] = \left(1 - \frac{\alpha}{2}\right)x[n] + \frac{\alpha}{4}\{x[n-T] + x[n+T]\}$$

x[n] étant le signal sonore décodé, *y[n]* étant le signal sonore décodé à pas accru dans une sous-bande donnée, *T* étant un retard de pas du signal sonore décodé et α étant un coefficient variant entre 0 et 1 destiné à contrôler une valeur d'atténuation d'inter-harmonique du signal sonore décodé (112).

52. Dispositif de post-traitement (108) selon la revendication 51, comprenant un moyen de réception du retard de pas *T* par le biais d'un train de bits.

53. Dispositif de post-traitement (108) selon la revendication 51, comprenant un moyen de décodage du retard de pas *T* provenant d'un train de bits codé reçu.

54. Dispositif de post-traitement (108) selon la revendication 51, comprenant un moyen de calcul du retard de pas *T* en réponse au signal sonore décodé pour améliorer le suivi du pas.

55. Dispositif de post-traitement (108) selon la revendication 32, dans lequel, pendant le codage, le signal sonore est sous-échantillonné depuis une fréquence d'échantillonnage supérieure vers une fréquence d'échantillonnage inférieure et dans lequel les moyens diviseurs comprennent des moyens de sur-échantillonnage (403, 404, 405 ; 406, 407, 408) du signal sonore décodé depuis la fréquence d'échantillonnage inférieure vers la fréquence d'échantillonnage supérieure.

56. Dispositif de post-traitement (108) selon la revendication 55, dans lequel le moyen diviseur comprend un moyen de filtrage de sous-bande (407) auquel est fourni le signal sonore décodé et dans lequel le moyen de sur-échantillonnage (406) est combiné au moyen de filtrage de sous-bande (407).

57. Dispositif de post-traitement (108) selon la revendication 55, dans lequel :

- le moyen de post-traitement comprend :

  un moyen de post-traitement (402 ; 401) du signal sonore décodé ; et

- le moyen diviseur comprend:

  un filtre passe-bande (407) auquel est fourni le signal sonore décodé afin de produire un signal de bande supérieure de fréquence, ledit filtre passe-bande (407) étant combiné au moyen de sur-échantillonnage (406, 407, 408) ; et
  un filtre passe-bas (404) auquel est fourni le signal sonore décodé post-traité afin de produire un signal de bande inférieure de fréquence, ledit filtre passe-bas (404) étant combiné au moyen de sur-échantillonnage (403, 404, 405).

58. Dispositif de post-traitement (108) selon la revendication 57, comprenant en outre un sommateur (409) servant à additionner le signal de bande supérieure de fréquence (410) au signal de bande inférieure de fréquence afin de former un signal sonore décodé à valeur de sortie post-traitée et sur-échantillonné.

59. Dispositif de post-traitement (108) selon la revendication 57, dans lequel le moyen de post-traitement du signal sonore décodé comprend un moyen d'amplification de pas (402) du signal sonore décodé afin de réduire un bruit d'inter-harmonique dans le signal sonore décodé.

60. Dispositif de post-traitement (108) selon la revendication 59, dans lequel le moyen d'amplification de pas (402) comprend un moyen de traitement du signal sonore décodé au moyen de l'équation suivante :

$$y[n] = \left(1 - \frac{\alpha}{2}\right)x[n] + \frac{\alpha}{4}\{x[n-T] + x[n+T]\}$$

x[n] étant le signal sonore décodé, *y[n]* étant le signal sonore décodé à pas accru dans une sous-bande donnée, T étant un retard de pas du signal sonore décodé et α étant un coefficient variant entre 0 et 1 destiné à contrôler une valeur d'atténuation d'inter-harmonique du signal sonore décodé.

61. Dispositif de post-traitement (108) selon la revendication 32, dans lequel le moyen diviseur comprend un moyen de division du signal sonore décodé en un signal de bande supérieure de fréquence (711) et un signal de bande inférieure de fréquence (713) ; et le moyen de post-traitement (703) comprend un moyen de post-traitement du signal de bande inférieure de fréquence.

62. Dispositif de post-traitement (108) selon la revendication 32, dans lequel le moyen de post-traitement comprend :

des moyens (303 ; 401; 502) pour déterminer une valeur de pas du signal sonore décodé ;
des moyens pour calculer, en relation avec la valeur de pas déterminée, un filtre passe-haut avec une fréquence de coupure inférieure à une fréquence fondamentale du signal sonore décodé ; et
des moyens pour traiter le signal sonore décodé (112) au moyen du filtre passe-haut calculé.

63. Décodeur de signal sonore (105) comprenant:

une entrée destinée à recevoir un signal sonore codé (110);
un décodeur de paramètres (106) auquel est fourni le signal sonore codé (110) et servant à décoder les paramètres de codage du signal sonore ;
un décodeur de signal sonore (107) auquel sont fournis les paramètres de codage du signal sonore décodé afin de produire un signal sonore décodé (112) ; et
un dispositif de post-traitement (108) selon l'une quelconque des revendications 32 à 62 pour post-traiter le signal sonore décodé (112) en vue d'améliorer une qualité perçue dudit signal sonore décodé (112).

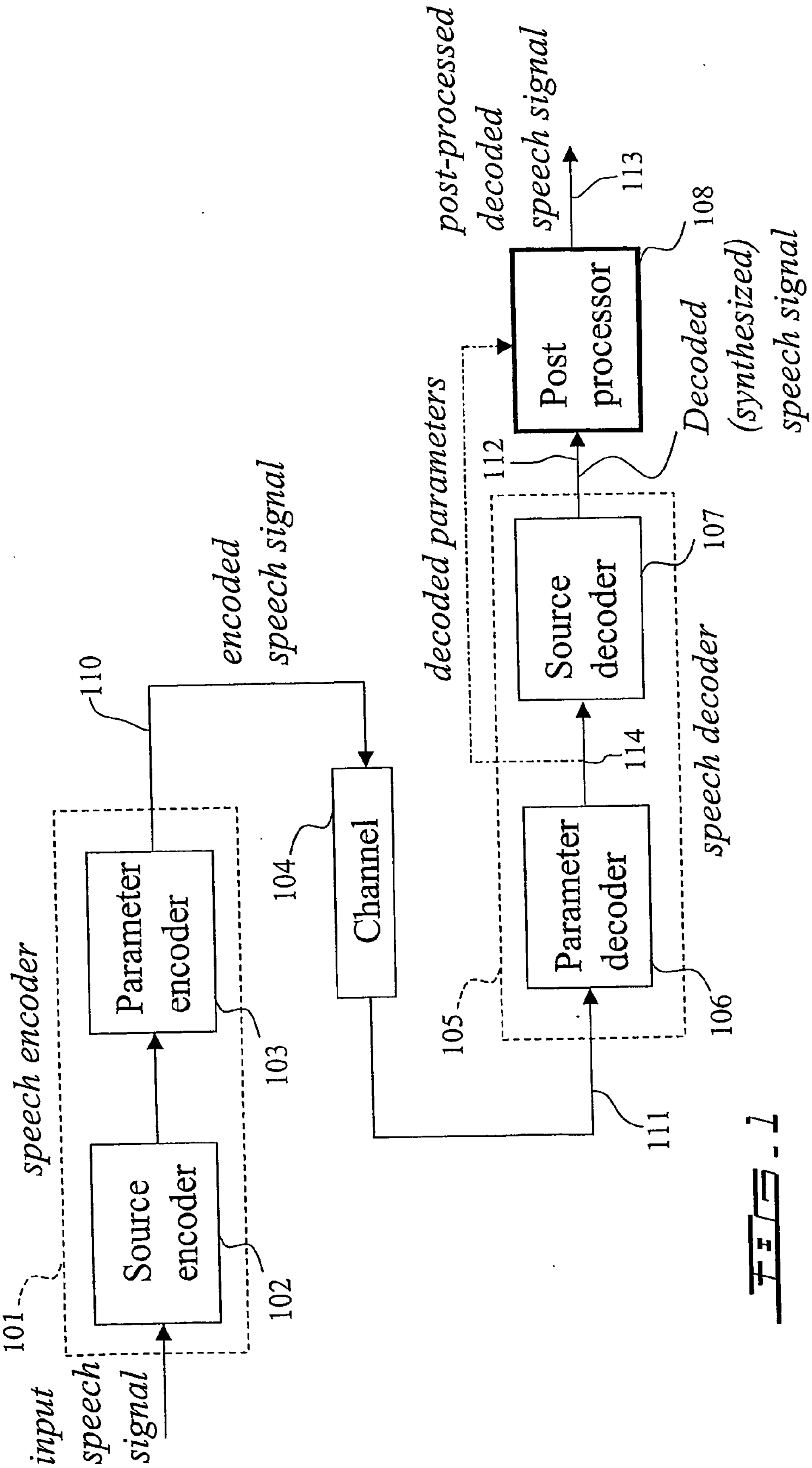
input speech signal
101
speech encoder
Source encoder
102
Parameter encoder
103
110
encoded speech signal
104
Channel
111
105
speech decoder
Parameter decoder
106
114
Source decoder
107
decoded parameters
112
Decoded (synthesized) speech signal
Post processor
108
113
post-processed decoded speech signal

FIG. 1

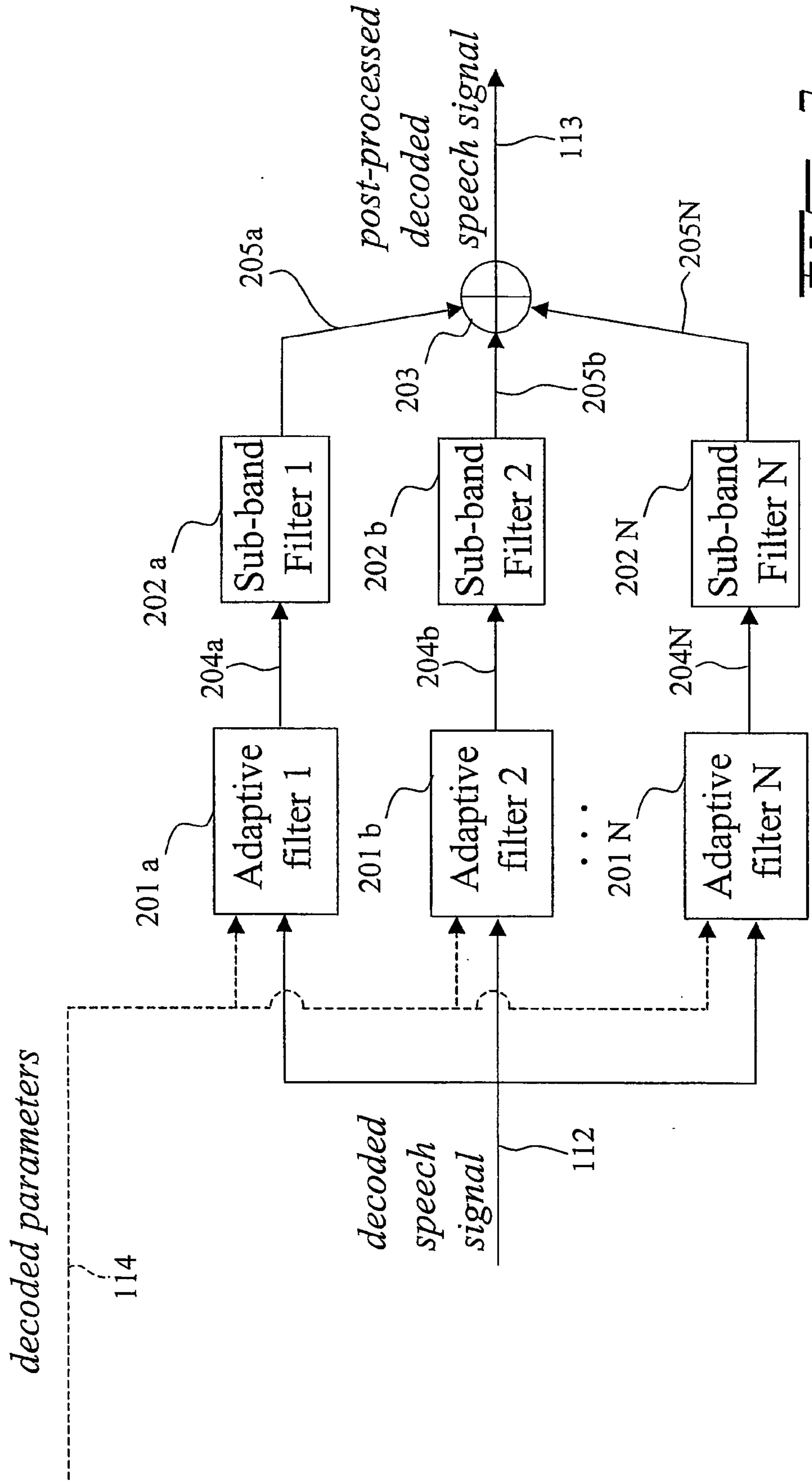
post-processed decoded speech signal
113
205a
205N
203
205b
Sub-band Filter 1
Sub-band Filter 2
Sub-band Filter N
202 a
202 b
202 N
204a
204b
204N
Adaptive filter 1
Adaptive filter 2
Adaptive filter N
201 a
201 b
201 N
decoded parameters
114
decoded speech signal
112

FIG. 2

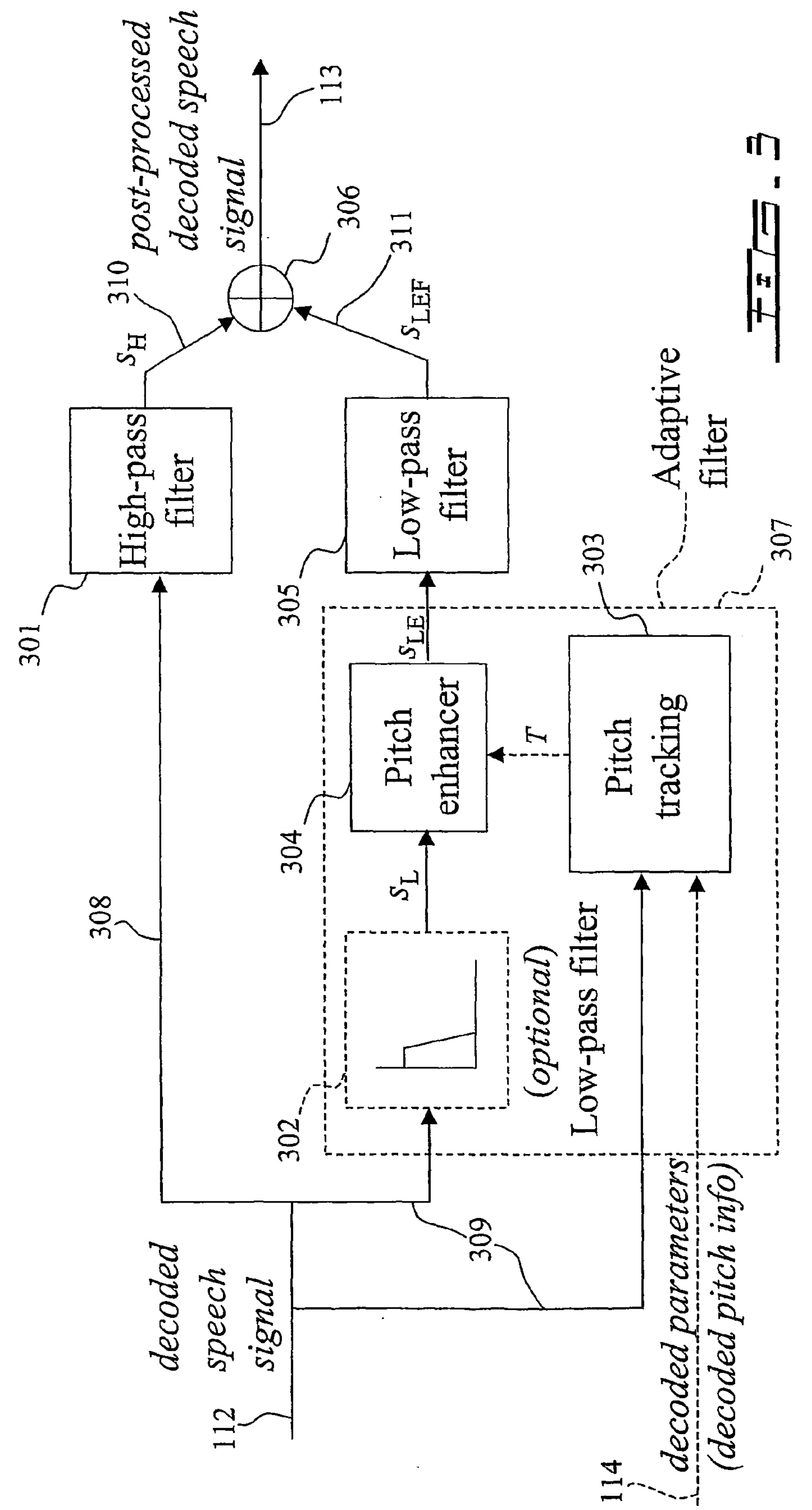
post-processed decoded speech signal
113
306
311
310
s_H
s_{LEF}
High-pass filter
Low-pass filter
Adaptive filter
307
301
305
303
s_{LE}
Pitch enhancer
T
Pitch tracking
304
s_L
308
(optional) Low-pass filter
302
309
decoded speech signal
112
decoded parameters (decoded pitch info)
114

FIG. 3

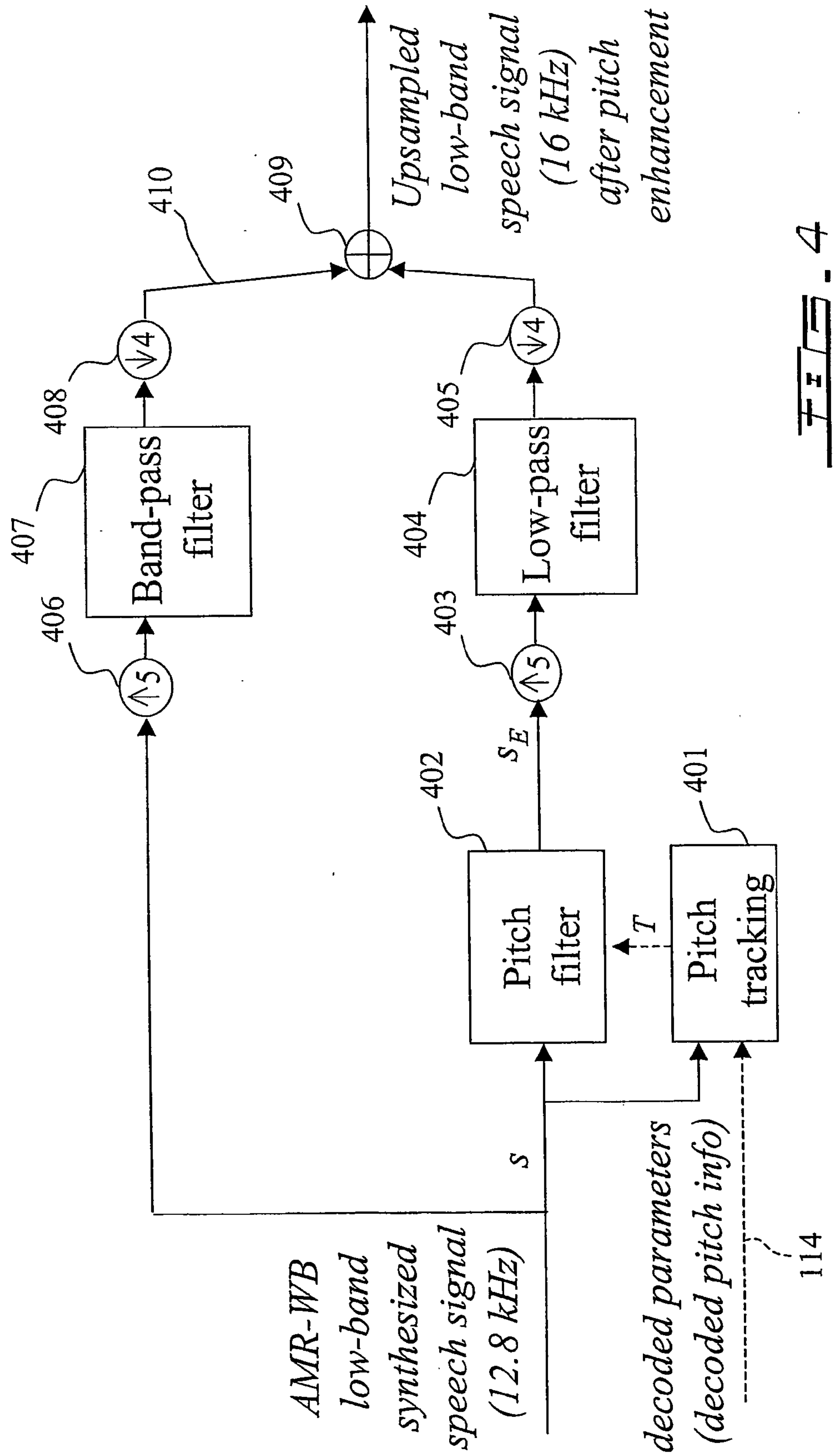
AMR-WB low-band synthesized speech signal (12.8 kHz)
decoded parameters (decoded pitch info)
114
s
Pitch filter
Pitch tracking
T
402
401
s_E
↑5
403
Low-pass filter
404
↓4
405
↑5
406
Band-pass filter
407
↓4
408
409
410
Upsampled low-band speech signal (16 kHz) after pitch enhancement

FIG. 4

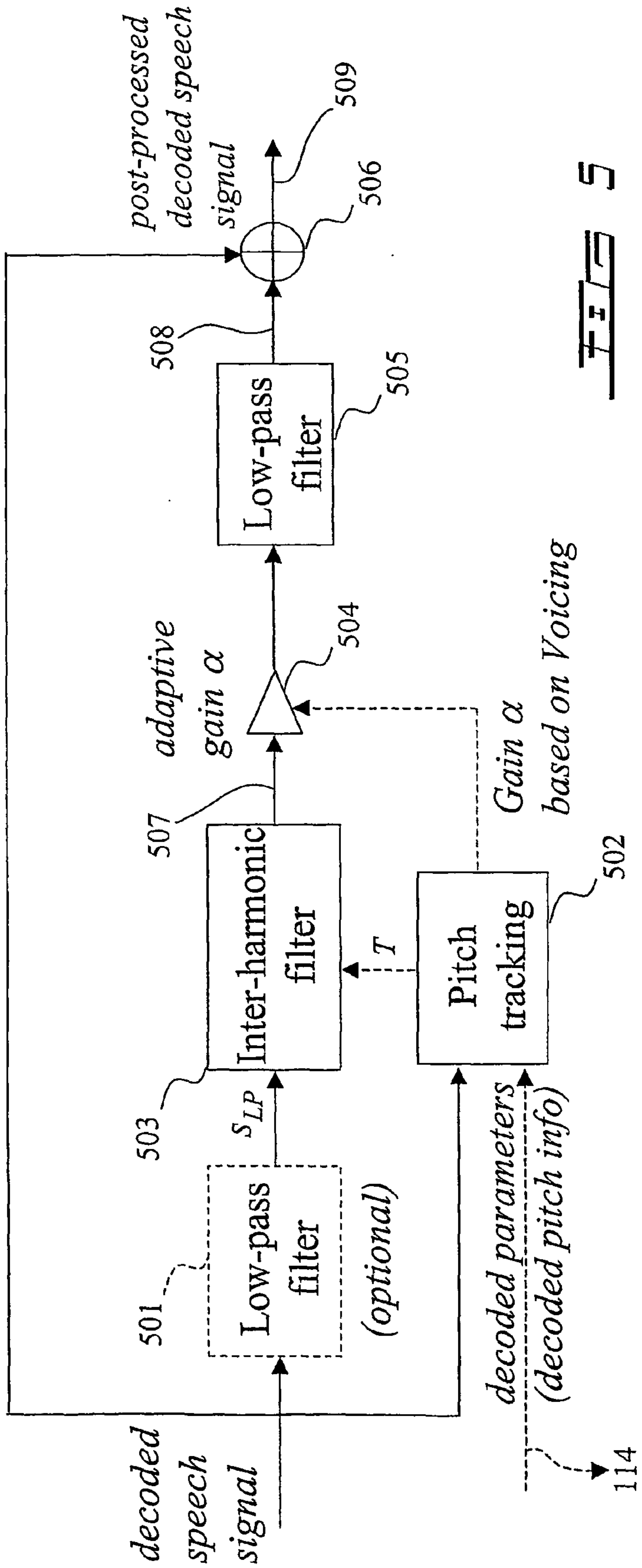
post-processed decoded speech signal
509
506
508
Low-pass filter
505
adaptive gain α
504
507
Inter-harmonic filter
T
Pitch tracking
502
Gain α based on Voicing
503
s_{LP}
501
Low-pass filter
(optional)
decoded speech signal
decoded parameters (decoded pitch info)
114

FIG 5

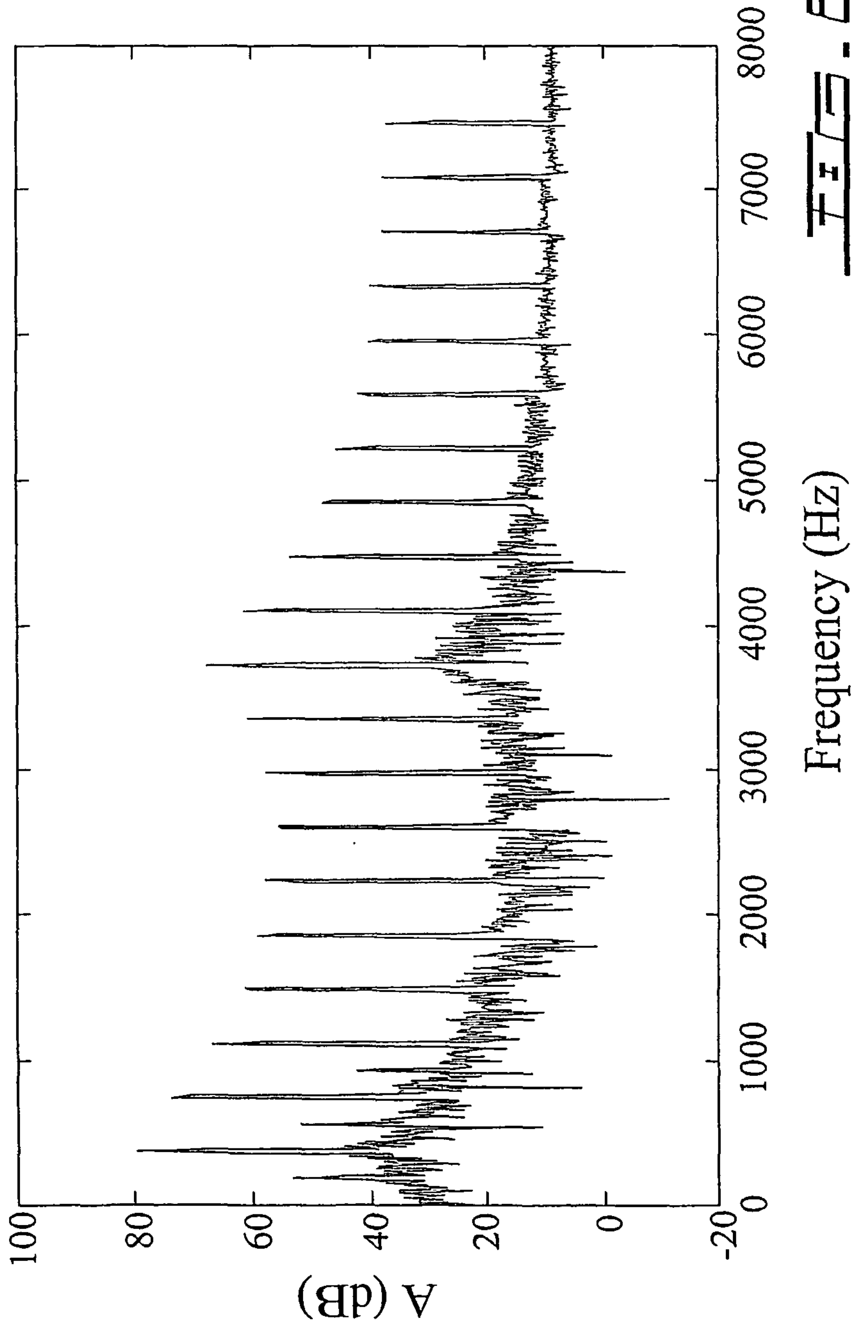
A (dB)
100
80
60
40
20
0
-20
Frequency (Hz)
0
1000
2000
3000
4000
5000
6000
7000
8000

FIG. 6A

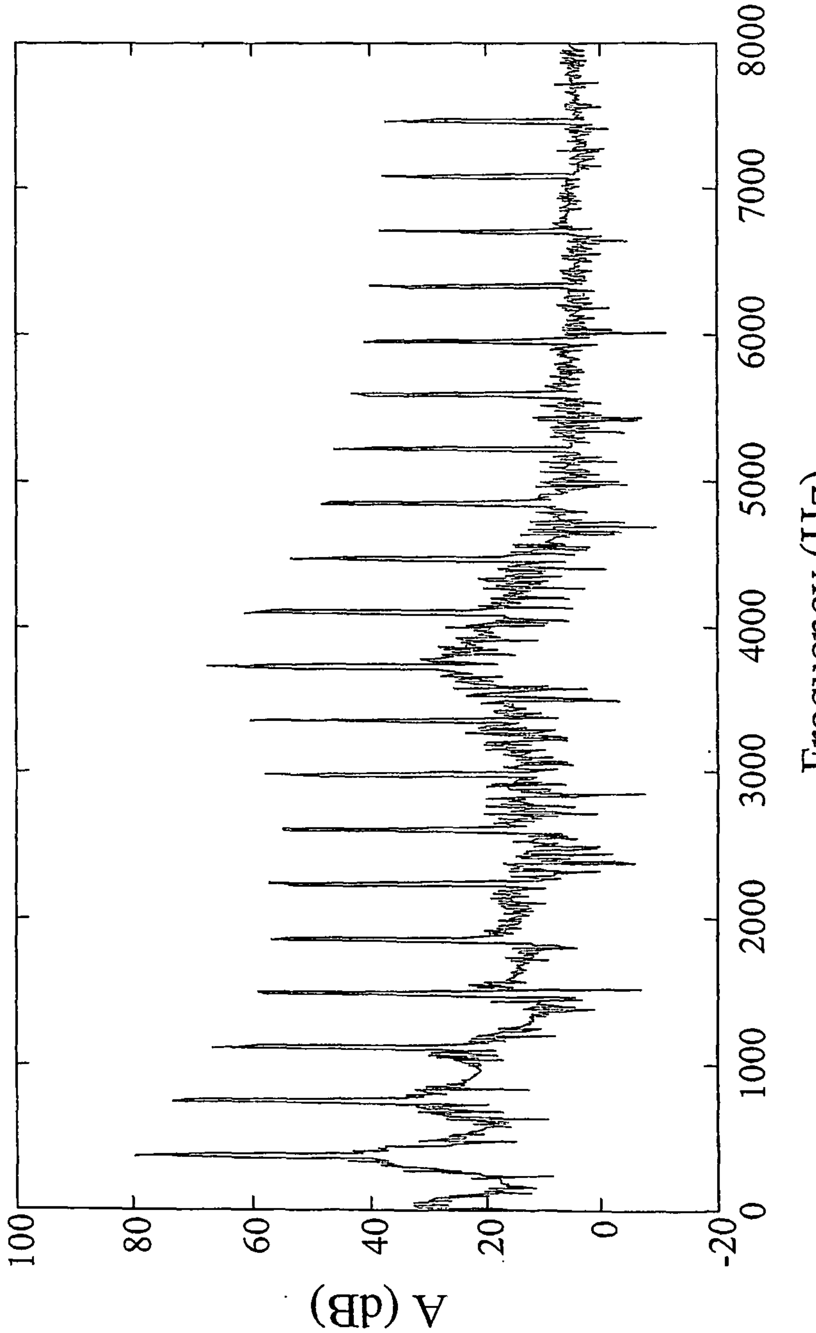
A (dB)
100
80
60
40
20
0
-20
0
1000
2000
3000
4000
5000
6000
7000
8000
Frequency (Hz)

FIG. 6B

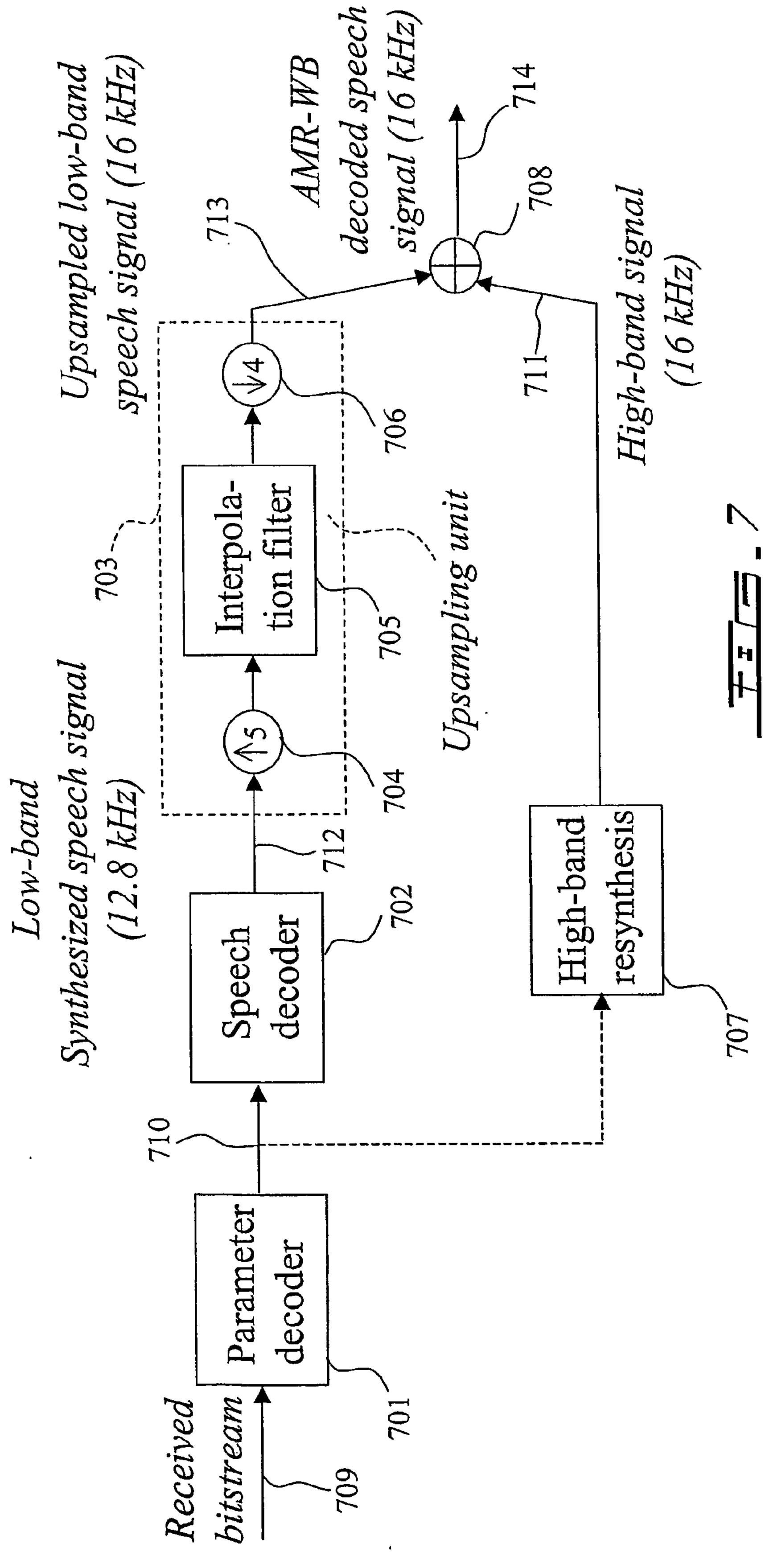
Received bitstream
709
Parameter decoder
701
710
Speech decoder
702
Low-band Synthesized speech signal (12.8 kHz)
712
↑5
704
Interpola-tion filter
705
703
Upsampling unit
↓4
706
Upsampled low-band speech signal (16 kHz)
713
AMR-WB decoded speech signal (16 kHz)
708
714
High-band resynthesis
707
711
High-band signal (16 kHz)

FIG. 7

$\alpha = 1$

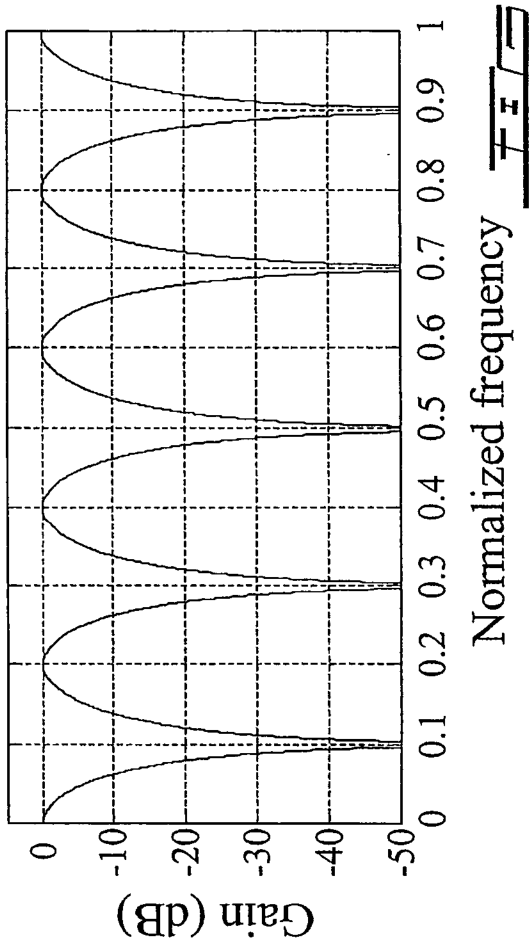

FIG. 8A

$\alpha = 0.8$

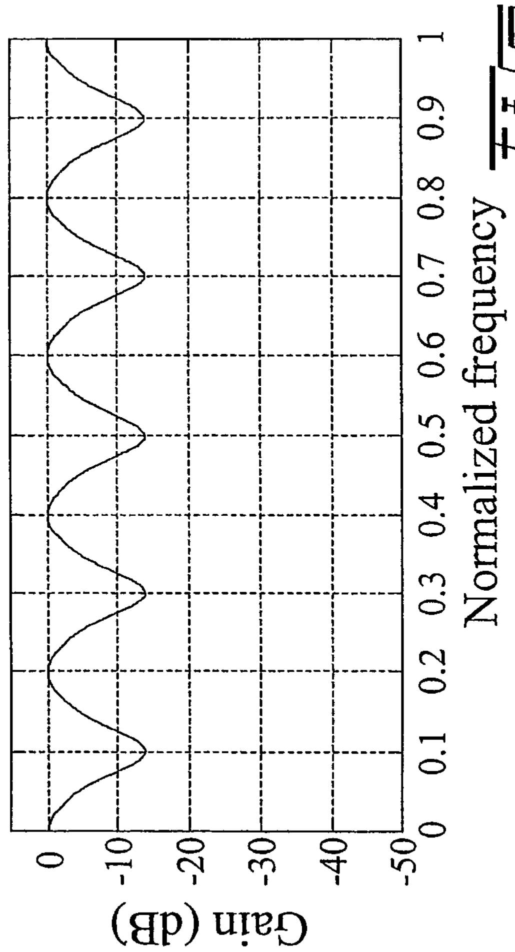

FIG. 8B

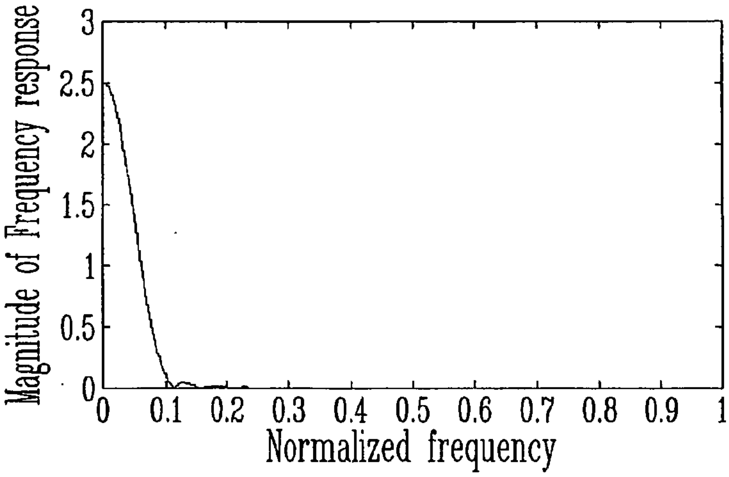

FIG. 9A

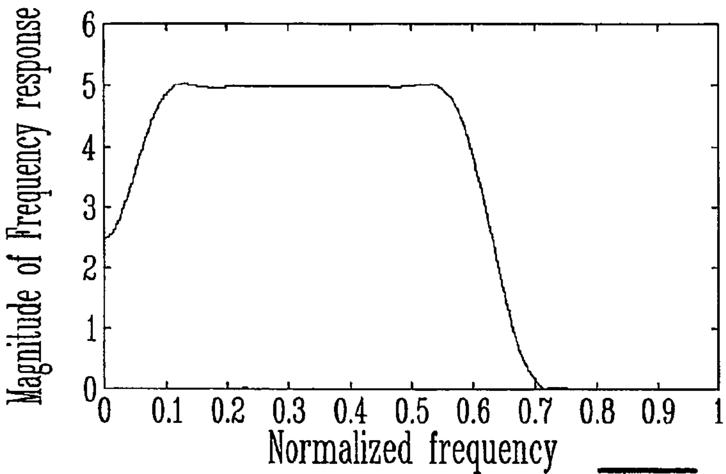

FIG. 9B

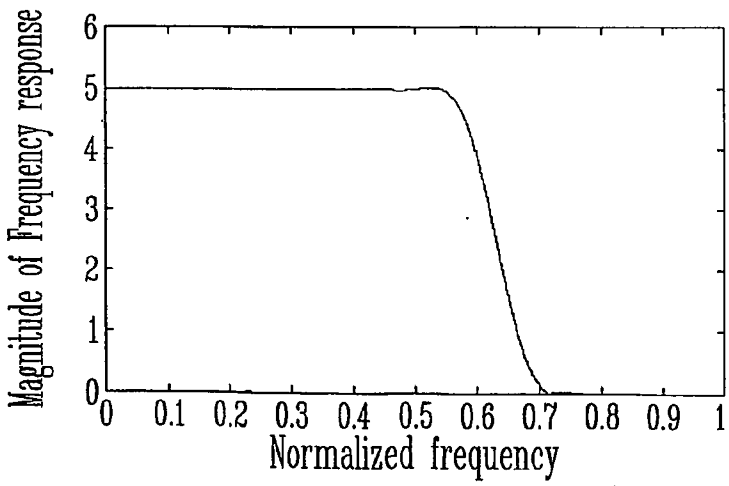

FIG. 9C

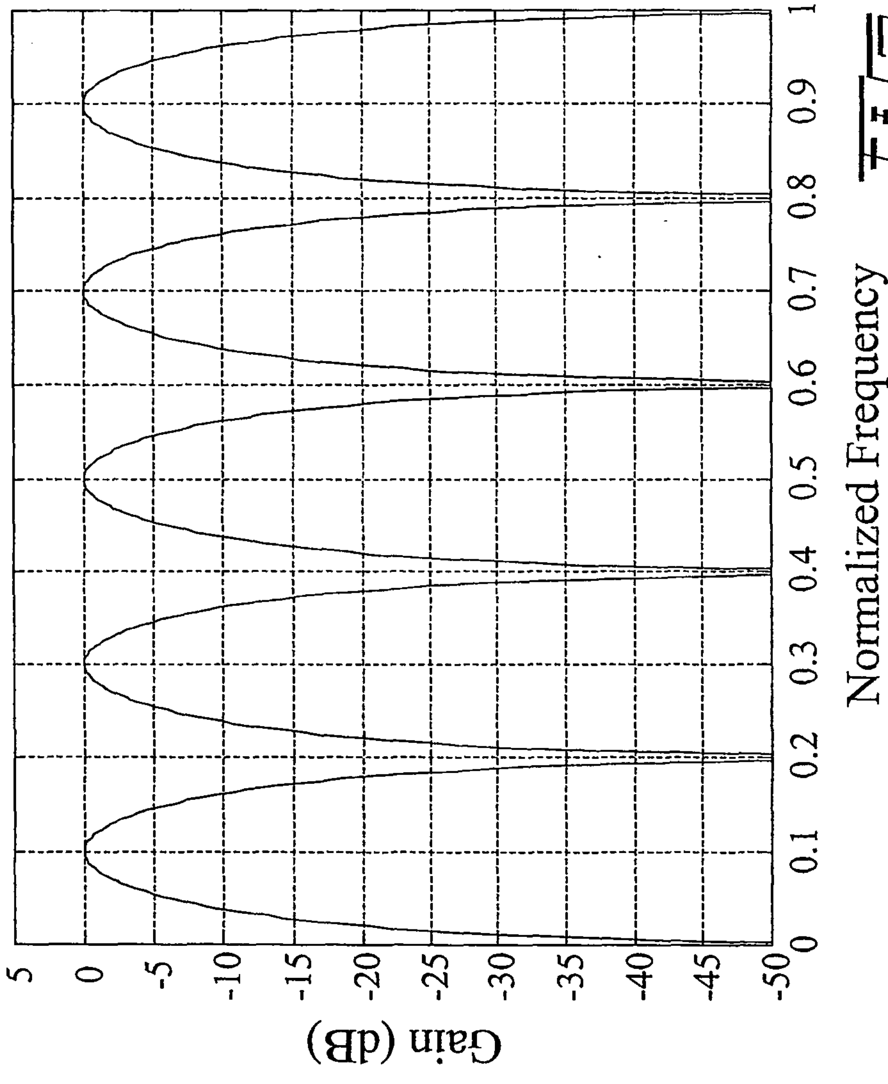
Gain (dB)
5
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Normalized Frequency

FIG. 10

REFERENCES CITED IN THE DESCRIPTION

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

Patent documents cited in the description

- US 5806025 A **[0009]**

Non-patent literature cited in the description

- **R. SALAMI et al.** Design and description of CS-ACELP: a toll quality 8 kb/s speech code. *IEEE Trans. on Speech and Audio Proc.,* March 1998, vol. 6 (2), 116-130 **[0006]**
- Speech Coding and Synthesis. Elsevier, 1995 **[0010]**