



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

(21) International Application Number:
PCT/EP2009/061884

(22) International Filing Date:
14 September 2009 (14.09.2009)

(25) Filing Language: English

(26) Publication Language: English

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

(72) Inventors; and

(75) Inventors/Applicants (for US only): **ESCH, Thomas** [DE/DE]; Neptunstr. 35, 50823 Köln (DE). **VARY, Peter** [AT/DE]; Morinerweg 1, 52074 Aachen (DE).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ,

CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

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

Published:

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

(54) Title: SIGNAL ENHANCEMENT PROCESSING

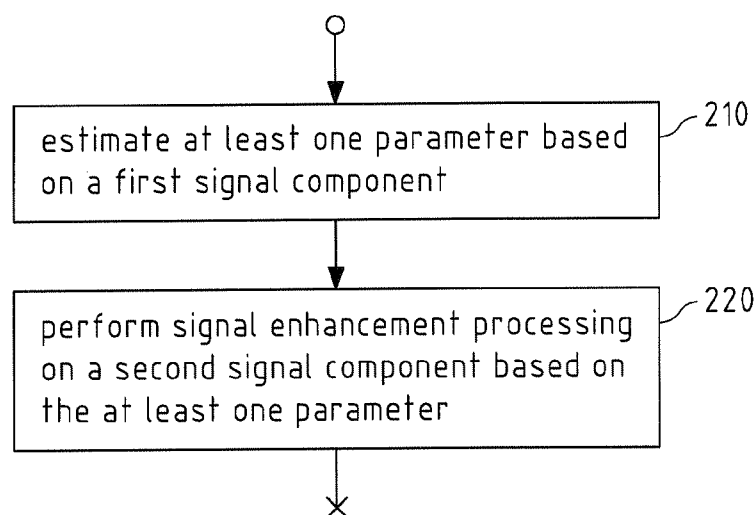


Fig.2a

(57) Abstract: It is disclosed to estimate at least one parameter based on a first signal component of an audio signal, the at least one parameter being associated with a second signal component of the audio signal; wherein the first signal component is associated with a first frequency band and the second signal component is associated with a second frequency band; and to perform a signal enhancement processing on the second signal component at least based on the at least one parameter.

Signal Enhancement Processing

FIELD OF THE DISCLOSURE

This invention relates to the field of signal enhancement processing of audio signals.

BACKGROUND

The quality of today's telephone speech was designed to achieve a sufficient intelligibility. The acoustic bandwidth in telephony systems is typically limited to the frequency range between 300 Hz and 3.4 kHz. However, this typical "telephone sound" cannot satisfy the increased demands as the perceived speech quality is considerably reduced compared to the full audio bandwidth. As a reasonable compromise, various wideband (50 Hz - 7 kHz) speech codecs have been developed in the past (e.g., the Adaptive Multi-Rate (AMR) Wideband Codec) and are about to be introduced in current mobile networks. Most of these codecs are mainly designed for nearly noise-free input speech signals and may not perform well when the input signal is degraded by acoustic background noise. In order to improve the listening comfort and to keep the high quality capability also in noisy environments, noise suppression techniques may be required for wideband communication systems.

When a speech communication device is used in environments with high levels of ambient noise, the noise picked up by the microphone may significantly impair the quality and the

intelligibility of the transmitted speech signal. In order to get a reliable separation from the noise signal (e.g., engine noise, street noise), noise reduction algorithms have become part of many digital speech coding and speech processing systems. They are used for example in mobile communications, in hearing aids and in hands-free devices.

One of the popular methods for enhancing degraded speech is based on modeling the noisy input coefficients in the short-time Fourier transform (STFT) domain and to apply individual adaptive gains for each frequency bin. In many occasions the processing applied to implement such techniques has been derived for narrowband signals under certain assumptions about the statistics of the speech and noise signals.

SUMMARY OF SOME EMBODIMENTS OF THE INVENTION

A method is described, which comprises estimating at least one parameter based on a first signal component of an audio signal, the at least one parameter being associated with a second signal component of the audio signal, wherein the first signal component is associated with a first frequency band and the second signal component is associated with a second frequency band, and performing a signal enhancement processing on the second signal component at least based on the at least one parameter.

Moreover, a first apparatus is described, which comprises means for estimating at least one parameter based on a first signal component of an audio signal, the at least one parameter being associated with a second signal component of

the audio signal, wherein the first signal component is associated with a first frequency band and the second signal component is associated with a second frequency band, and wherein the first apparatus comprises means for performing a signal enhancement processing on the second signal component at least based on the at least one parameter.

The means of this apparatus can be implemented in hardware and/or software. They may comprise for instance a processor for executing computer program code for realizing the required functions, a memory storing the program code, or both. Alternatively, they could comprise for instance a circuit that is designed to realize the required functions, for instance implemented in a chipset or a chip, like an integrated circuit.

Moreover, a second apparatus is described, which comprises at least one processor and at least one memory including computer program code, the at least one memory and the computer program code, with the at least one processor, configured to cause the apparatus at least to perform the actions of the presented method.

Moreover, a computer readable storage medium is described, in which computer program code is stored. The computer program code causes an apparatus to realize the actions of the presented method when executed by a processor.

The computer readable storage medium could be for example a disk or a memory or the like. As an example, the memory may represent a memory card such as SD and micro SD cards or any other well-suited memory cards or memory sticks. The computer

program code could be stored in the computer readable storage medium in the form of instructions encoding the computer-readable storage medium. The computer readable storage medium may be intended for taking part in the operation of a device, like an internal or external hard disk of a computer, or be intended for distribution of the program code, like an optical disc.

As an example, the signal enhancement processing may represent noise reduction processing, or a voice enhancement processing or any other well-suited signal enhancement processing.

The first frequency band differs from the second frequency band. For instance, the first frequency band may span a frequency range from f_1 to f_2 , and the second frequency band may span a frequency range from f_3 to f_4 , wherein $f_3 > f_1$ and $f_4 > f_2$ holds. As an example, the first and second frequency band may have an overlapping frequency range, i.e. $f_3 < f_2$ holds, but as another example, the first and second frequency band may not overlap, i.e. $f_3 > f_2$ may hold. For instance, the first frequency band may span a frequency range from 50 Hz to 4 kHz, and the second frequency band may span a frequency range from 4 to 7 kHz. But of course, any other well suited frequency assignments for the first and second frequency range may be applied.

The estimated at least one parameter may represent at least one parameter which may be beneficial to increase the quality of the signal enhancement processing. The type of the at least one parameter may depend on the kind of signal enhancement processing.

For instance, the first signal component may be extracted from the audio signal by means of filtering, e.g. by a lowpass filter or a bandpass filter, and the second signal component may be extracted from the audio signal by means of filtering, e.g. by a highpass filter or a bandpass filter.

As an example, at least one parameter of the at least one parameter may represent estimations of signal properties of the second signal component. For instance, these signal properties may comprise statistical information of the spectral characteristics of the second signal component. E.g., at least one parameter of the at least one parameter may represent a set of spectral envelope representatives of the second signal component. Thus, this set of spectral envelope representatives may comprise estimations of representations of the spectral envelope of the second signal component. For instance, this set of spectral envelope representatives may comprise at least one energy value, wherein each of the at least one energy value is associated with one frequency sub-band of the second signal component. Furthermore, at least one parameter of the at least one parameter may represent further information with respect to the second signal component.

For instance, the estimation of the at least one parameter is not based on the second signal component.

Based on this at least one parameter, which has been estimated based on the first signal component, a signal enhancement processing is performed to the second signal component. Thus, spectral dependencies may be exploited between the first

signal component and the second signal component in order to perform the signal enhancement on the second signal component. Accordingly, spectral dependencies between the first frequency band and the second frequency band may be used to improve the signal enhancement processing.

For instance, there may be a further signal processing with respect to the at least one parameter before signal enhancement processing is performed on the second signal component.

As an example, the signal enhancement processing on the second signal may be performed by means of at least one filter, e.g. in the time-domain or in the frequency domain. As an example, one filter of this at least one filter may be configured to be adapted to at least one value, i.e., this at least one value may be configured to be used to perform a signal enhancement processing to the second signal component. For instance, this at least one value may represent at least one filter coefficient. Accordingly, as an example, the signal enhancement processing to the second signal component may be performed based on the at least one parameter by means of the at least one filter, wherein the at least one value may be configured to be used to perform a signal enhancement processing at least based on the at least one parameter.

For instance, this at least one value may comprise at least one value configured to be used to perform a signal enhancement processing to the second signal component in the time-domain. As an example, this at least one value may represent at least one filter coefficient of a filter configured to filter the second signal component in the

time-domain. For instance, this filter may represent a Finite Impulse Response (FIR) filter or an Infinite Impulse Response (IIR) filter. As an example, this at least one value may be determined based on a Minimum Mean Squared Error (MMSE) criterion or a Zero Forcing (ZF) criterion or any other well-suited criterion. In this exemplary case, a signal enhancement processing to the second signal component may be performed by filtering the second signal component by means of the respective filter in the time-domain.

Furthermore, as another example, the at least one value may comprise at least one value configured to be used to perform a signal enhancement processing to the second signal component in the frequency-domain.

For instance, the second signal component in the frequency domain may be represented by M_F spectral coefficients at frequency bin k and frame λ given by:

$$Y_2(\lambda, k) = S_2(\lambda, k) + N_2(\lambda, k), \quad (1)$$

where $S_2(\lambda, k)$ and $N_2(\lambda, k)$ may represent the spectral coefficients of the audio and the noise signal of the second signal component, respectively. Correspondingly, the first signal component in the frequency domain may be represented by M_F spectral coefficients at frequency bin k and frame λ given by

$$Y_1(\lambda, k) = S_1(\lambda, k) + N_1(\lambda, k), \quad (2)$$

where $S_1(\lambda, k)$ and $N_1(\lambda, k)$ may represent the spectral coefficients of the audio and the noise signal of the first signal component, respectively. For the sake of brevity, the frame index λ is omitted in the following.

As an example, a converter may be configured to output M_F subcomponents representing the first signal component and a converter may be configured to output M_F subcomponents representing the second signal component in the frequency-domain, wherein each subcomponent of the M_F subcomponents is associated with one of M_F sub-bands in the first and second frequency band.

For instance, this converter may represent a Fourier Transformation, e.g. a FFT, a DFT, a STFT or any other well-suited transformation. The at least one filter may be adapted to the at least one value and may be configured to perform a signal enhancement processing to the second signal component in the frequency domain. Accordingly, the at least one filter may be configured to output an enhanced second signal component in the frequency domain, wherein the at least one filter is configured to perform this signal enhancement based on the at least one parameter.

For instance, the at least one value may comprise at least two gain factors, wherein each of the at least two gain factors is associated with a separate sub-band of the second frequency band. As an example, these at least two gain factors may be determined based on a MMSE criterion or a ZF criterion or any other well-suited criterion. In this exemplary case, a signal enhancement processing to the second signal component may be

performed based on weighting a frequency component of the second signal associated with a sub-band of the second frequency band with the respective gain factor of the at least two gain factors. Thus, for instance, each of the at least two gain factors is associated with one sub-band of the second frequency band.

As an example, a signal enhancement processing is performed on the first signal component in order to obtain an enhanced first signal component.

For instance, the signal enhancement processing performed on the first signal component may represent a conventional signal enhancement processing. This signal enhancement processing may be of the same type as the signal enhancement processing performed on the second signal component.

As an example, an enhanced second signal component obtained by signal enhancement processing performed on the second signal component may be combined with the enhanced first signal component to an enhanced audio signal. As an example, the enhanced audio signal may be fed to a further signal processing, e.g., to a speech encoder or any other well suited processing.

Furthermore, as another example, said performing a signal enhancement processing on the second signal may be based on a combined signal enhancement processing.

For instance, this combined signal enhancement processing may be based on a first signal enhancement processing and a second signal enhancement processing, both of the first and second

signal enhancement processing may be configured to be applied to the second signal component, wherein the first signal enhancement processing is based on the at least one parameter and the second signal enhancement processing is not based on the at least one parameter. For instance, the second signal enhancement processing may represent a conventional signal enhancement processing.

As an example, the first signal enhancement processing may be applied to the second signal component in order to obtain a first enhanced second signal component and the second signal enhancement processing may be applied to the second signal component in order to obtain a second enhanced second signal component, and the first enhanced second signal component and the second enhanced second signal component may be combined to the enhanced first signal component.

Or, as another example, the first signal enhancement processing may be combined with the second signal enhancement processing in order to be applied as combined signal enhancement processing on the second signal component. This example will be explained in more detail in the sequel of this description.

According to one embodiment, a signal enhancement processing is performed on the first signal component in order to obtain an enhanced first signal component, wherein said estimating at least one parameter is performed on basis of the enhanced first signal component.

Thus, the enhanced first signal component may be used to estimate the at least one parameter, which may lead to

increased quality of the estimated at least one parameter. In the sequel, it has to be understood, that any determining or estimation or extraction based on the first signal component may be based on the enhanced first signal component.

According to one embodiment, at least one parameter of said at least one parameter represents a set of spectral envelope representatives of the second signal component.

For instance, this set of spectral envelope representatives may be estimations of representations of the spectral envelope of the second signal component. For instance, this set of spectral envelope representatives may comprise at least one energy value, wherein each of the at least one energy value is associated with one frequency sub-band of the second signal component. Furthermore, at least one parameter of the at least one parameter may represent further information with respect to the second signal component.

According to one embodiment, at least one feature is determined based on the first signal component, the at least one feature representing at least signal parameter being associated with the first signal component, wherein said estimating the at least one parameter is performed on basis of the at least one feature.

It has to be understood, that determining the at least one feature may also be performed based on the enhanced first signal component, as mentioned above.

For instance, at least one feature of the at least one feature may represent a set of spectral envelope representatives of

the first signal component. For instance, the at least one feature of the at least one feature may be at least one energy representative, wherein each of the at least one energy representative is associated with the energy of the first signal component. As an example, at least one feature of the at least one feature may be N_c mel-frequency cepstral coefficients (MFCCs). Thus the frequency sub-bands associated with the first frequency band may be equally spaced on the mel scale.

Furthermore, at least one feature of the at least one feature may be associated with further information of the first signal component. For instance, at least one feature of the at least one feature may be associated with the zero-crossing rate (ZCR) of the first signal component. The zero-crossing rate may be the rate of sign-changes along a signal, i.e., the rate at which the signal changes from positive to negative or back.

As an example, the at least one feature may be denoted as vector $x(i)$, where 'i' denotes the time index. For instance, an extractor may configured to output the at least one feature may be configured to operate on a frame-by-frame basis, i.e., for instance, $X=\{x(1), \dots, x(\lambda)\}$ may represent a sequence of feature vectors from the first signal component of frames 1 to λ . For instance, λ designates the current frame.

Thus, for instance this at least one feature may be any well-suited features of the first signal component which may be used to exploit a correlation and/or spectral dependencies to a signal parameter of the second signal component in the second frequency band.

For instance, an estimator may be configured to estimate the at least one parameter at least based on the at least one feature.

As an example, the estimator may apply a MMSE criterion, a Maximum a Posteriori (MAP) criterion, a Maximum Likelihood (ML) criterion, a ZF criterion or any other well-suited criterion for estimating the at least one parameter at least based on the at least one feature. As an example, the estimator may not apply any information extracted from or based on the second signal component in order to estimate the at least one parameter. Thus, for instance, the at least one parameter is estimated based on features that are determined only from the first signal component.

As an example, the estimator may be configured to perform the estimation based on a Hidden Markov Model (HMM).

According to one embodiment, at least one feature of said at least one feature is a set of spectral envelope representatives of one out of: the first signal component, and the enhanced first signal component.

For instance, the at least one feature of said at least one feature may be represented by vector $x = \{x'(0), \dots, x'(M'_F - 1)\}$, wherein $x'(j)$ with $j \in \{0, \dots, M'_F - 1\}$ may be a representative of the energy of the first signal/enhanced first signal in the j th sub-band of a set of $M'_F - 1$ frequency sub-bands of the first frequency band. As an example, the sub-bands of this set of $M'_F - 1$ frequency sub-bands of the first frequency band may overlap in the frequency range.

According to one embodiment, said estimating at least one parameter is further based on at least one previous feature which has been previously determined based on the first signal component.

As an example, the estimating may be based on sequence $X=\{x(\lambda-m), \dots, x(\lambda-1), x(\lambda)\}$, wherein $x(\lambda)$ may represent the current at least one feature and $x(\lambda-m), \dots, x(\lambda-1)$ may represent the m previous at least one features. Thus, in this case $m \geq 1$ holds, m representing an integer.

As an example, under assumption of applying an MMSE estimation, the estimated at least one parameter, exemplarily denoted as vector $\hat{y}$, may be estimated by means of MMSE criterion for estimation of parameter vector y :

$$E\{\|y - \hat{y}\|^2 | X\} = \min \quad (3)$$

Equation (3) may be solved by the conditional expectation

$$\hat{y}_{MMSE} = E\{\hat{y}_i | X\}.$$

Of course, any other well-suited criterion may be used for estimating the at least one parameter based on sequence $X=\{x(\lambda-m), \dots, x(\lambda-1), x(\lambda)\}$.

It is now assumed, as an example, that at least one parameter of the at least one parameter is a set of spectral envelope representatives of the second signal component. For instance, the estimated at least one parameter may be represented by vector $\hat{y} = \left\{ |\hat{S}_2(0)|^2, \dots, |\hat{S}_2(M'_F - 1)|^2 \right\}$, wherein $|\hat{S}_2(j)|^2$ with

$j \in \{0, \dots, M'_F - 1\}$ may represent the estimated energy of the second signal in the j th sub-band of a set of $M'_F - 1$ frequency subbands of the second frequency band. As an example, the sub-bands of this set of $M'_F - 1$ frequency subbands of the second frequency band may overlap in the frequency range.

According to one embodiment, at least one parameter of said at least one parameter is a set of spectral envelope representatives of the second signal component, and said estimating at least one parameter is further based on a codebook, the codebook comprising a plurality of sets of spectral envelope representatives associated with the second signal component.

For instance, this codebook may represent a precomputed codebook $C = \{\hat{y}_1, \dots, \hat{y}_{M_c}\}$ for the parameter vectors y comprising M_c entries. For instance, this codebook may be obtained with the LBG algorithm presented by Y. Linde, A. Buzo, R.M. Gray in "An algorithm for vector quantizer design," IEEE Transactions on Communications, vol. 28, no. 1, pp. 84-95, Jan. 1980. Of course, any other well-suited codebook may also be used, e.g. depending on the type of signal enhancement processing.

As an example, using the codebook C , the MMSE estimate may be expressed as:

$$\hat{y}_{MMSE} = \sum_{\hat{y}_i \in C} \hat{y}_i \cdot P(\hat{y}_i | X), \quad (4)$$

which may represent a weighted sum over the M_c centroids of the codebook C . The weights $P(\hat{y}_i|X)$ represent a posteriori probabilities based on the determined sequence X . Thus, according to equation (4), an a posteriori probability based on the determined sequence X is determined for each of the plurality of sets of spectral envelope representatives of the codebook, and a weighted sum over the plurality of sets of spectral envelope representatives is determined, wherein each set of spectral envelope representatives $\hat{y}_i$ is weighted by the respective a posteriori probability $P(\hat{y}_i|X)$.

According to one embodiment, said estimating comprises: Determining for each of the plurality of sets of spectral envelope representatives of the codebook an a posteriori probability at least based on the at least one feature, and determining a weighted sum over the plurality of sets of spectral envelope representatives, wherein each set of spectral envelope representatives of the plurality of sets of spectral envelope representatives is weighted by the respective a posteriori probability.

For instance, the respective a posteriori probability may be determined based on an offline training phase. Equation (4) shows an exemplary determining of the weighted sum over the plurality of sets of spectral envelope representatives.

According to one embodiment, determining the at least one a posteriori probability is based on a Hidden Markov Model (HMM).

This HMM may be trained during an offline training phase. As an example, the HMM techniques described by B. Geiser, H. Taddei and P. Vary in "Artificial Bandwidth Extension without Side Information for ITU-T G.729.1," in Proceedings of European Conference on Speech Communication and Technology (INTERSPEECH), Antwerp, Belgium, August 2007 may be applied.

According to one embodiment, said performing a signal enhancement processing on the second signal is based on a combined signal enhancement processing comprising:
Determining a first set of at least one value configured to be used to perform a signal enhancement processing on the second signal component at least based on the at least one parameter, determining a second set of at least one value configured to be used to perform a signal enhancement processing on the second signal component not based on the at least one parameter, and combining the first set of at least one value and the second set of at least one value.

As an example, the combined set of at least one value may be used as the at least one value explained with respect to the at least one filter. Thus, the combined set of at least one value may be used to perform a signal enhancement processing on the second signal component, e.g. in time-domain or in the frequency domain. For instance, the first signal enhancement method and the second signal enhancement method may be combined in order to provide an increased signal enhancement of the second signal component.

For instance, the first set of at least one value may comprise M'_f values and the second set of at least one value may

comprise M'_F values, and a combiner may be configured to combine one value of the first set of at least one value with one value of the second set of at least one value to one combined value of a set of at least one combined value, wherein the set of at least one combined value comprises M'_F combined values.

For instance, the at least one value of the first set of at least one value may represent at least one filter coefficient, wherein a filter adapted to this at least one filter coefficient may be used to perform the signal enhancement processing on the second signal component, as explained above. Determining this first set of at least one value may depend on the applied signal enhancement processing.

For instance, the at least one value of the second set of at least one value may represent at least one filter coefficient, wherein a filter adapted to this at least one filter coefficient may be used to perform the signal enhancement processing on the second signal component, as explained above. Determining this second set of at least one value may depend on the applied signal enhancement processing.

For instance, the combining may be performed adaptively. As an example, this adaptively combining may be performed on the basis of determined signal and/or noise parameters.

According to one embodiment, the first set of at least one value represents a first set of at least two gain factors and said second set of at least one value represent a second set of at least two gain factors, wherein each of the at least

two gain factors of the first and second set is associated with a sub-band frequency component of the second signal component.

As an example, the first set of at least two gain factors may be configured to be used to perform a spectral weighting of the second signal based on weighting a sub-band frequency of the second signal associated with a sub-band of the second frequency band with the respective gain factor of the at least two gain factors of the first set. Further, as an example, the second set of at least two gain factors may be configured to be used to perform a spectral weighting of the second signal based on weighting a sub-band frequency of the second signal associated with a sub-band of the second frequency band with the respective gain factor of the at least two gain factors of the second set.

As an example, an SNR estimator may be configured to estimate at least one SNR representative based on the at least one parameter and the second signal component. This at least one SNR representative is associated with the second signal component. For instance, this at least one SNR representative may comprise a plurality of SNR representatives, wherein each of the plurality of SNR representatives is associated with one sub-band of a plurality of frequency sub-bands of the second frequency band. As an example, the plurality of frequency sub-bands may comprise $M'_F - 1$ frequency sub-bands of the second frequency band, and, for instance, the sub-bands of this set of $M'_F - 1$ frequency sub-bands of the second frequency band may overlap in the frequency range. As an example, the plurality of SNR representatives may comprise

a set of $M'_F - 1$ a priori SNR values $\hat{\xi}_2(j) = \frac{|\hat{S}_2(j)|^2}{|\hat{N}_2(j)|^2}$, wherein $|\hat{N}_2(j)|^2$ may represent the estimated noise power of the j th sub-band of the second signal component and wherein $|\hat{S}_2(j)|^2$ represents one of the at least one parameter.

For instance, an entity may be configured to estimate the j noise powers $|\hat{N}_2(j)|^2$ with $j \in \{0, \dots, M'_F - 1\}$. As an example, the estimation of the j th noise power may be written as

$$|\hat{N}_2(j)|^2 = \max\left(|Y_2(j)|^2 - |\hat{S}_2(j)|^2, 0\right), \quad (5)$$

wherein $|Y_2(j)|^2$ represents the energy of the audio signal in the j th sub-band of the second frequency band.

Furthermore, as an example, the plurality of SNR representatives may further comprise a set of $M'_F - 1$ a posteriori SNR values $\hat{\gamma}_2(j) = \frac{|Y_2(j)|^2}{|\hat{N}_2(j)|^2}$.

In this exemplary case, the first set of at least two gain factors is determined at least based on the plurality of SNR representatives. For instance, the at least two gain factors of the first set may be represented by $G_{bwe}(j)$ with $j \in \{0, \dots, M'_F - 1\}$, wherein $G_{bwe}(j)$ represents the gain factor associated with the j th sub-band of the $M'_F - 1$ frequency

sub-bands of the second frequency band. As an example, under the exemplary assumption that the signal enhancement processing is a noise reduction processing, $G_{bwe}(j)$ may be calculated as

$$G_{bwe}(j) = \frac{\hat{\xi}_2(j)}{1 + \hat{\xi}_2(j)}, \text{ or} \quad (6a)$$

$$G_{bwe}(j) = \frac{\hat{\gamma}_2(j)}{1 + \hat{\gamma}_2(j)}, \quad (6b)$$

or $G_{bwe}(j)$ may be calculated based on $\hat{\gamma}_2(j)$ and $\hat{\xi}_2(j)$, e.g. as described by Y. Ephraim and D. Malah in "Speech Enhancement Using a Minimum Mean-Square Error Short-Time Spectral Amplitude Estimator", IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. ASSP-32, no. 6, pp. 1109-1121, December 1984. For instance, the entity may be configured to determine the first set of at least two gain factors at least based on the plurality of SNR representatives based on any other well-suited method or as described by J. S. Lim and A. V. Oppenheim in "Enhancement and Bandwidth Compression of Noisy Speech", Proceedings of the IEEE, vol. 67, no. 12, pp. 1585-1604, December 1979.

Of course, depending on the type of signal enhancement processing, another calculation of gain factors $G_{bwe}(j)$ may be applied.

The first set of at least one value may be associated with a first signal enhancement processing which is based on the at least one parameter determined on basis of the first signal component.

Furthermore, for instance, the second set of at least two gain factors may be configured to be used to perform a spectral weighting of the second signal based on weighting a sub-band frequency of the second signal associated with a sub-band of the second frequency band with the respective gain factor of the at least two gain factors of the second set. For instance, the at least two gain factors of the second set may be represented by $G_{conv}(j)$ with $j \in \{0, \dots, M'_F - 1\}$, wherein $G_{conv}(j)$ represents the gain factor associated with the j th sub-band of the $M'_F - 1$ frequency sub-bands of the second frequency band.

As an example, the second set of at least one value may be associated with a second signal enhancement processing which is not based on the at least one parameter determined on basis of the first signal component. For instance, this second signal enhancement processing may represent a conventional signal enhancement processing.

As an example, under the assumption that the signal enhancement processing represents a noise reduction processing, an entity may be configured to perform a noise power estimation, e.g. as explained by R. Martin in "Noise Power Spectral Density Estimation Based on Optimal Smoothing and Minimum Statistics", IEEE Transactions on Speech and Audio Processing, vol. 9, no. 5, pp. 501-512, 2001, and to perform SNR estimation, e.g. as mentioned above, and to calculate the second set of at least one value, e.g., as mentioned above with respect to calculating the first set of at least one value.

According to one embodiment, the sub-bands associated with the gain factors are associated with a first frequency resolution, wherein the combined first set of at least two gain factors and second set of at least two gain factor are expanded to a second frequency resolution, the second frequency resolution being higher than the first frequency resolution.

As an example, the sub-bands associated with the gain factors are associated with a first frequency resolution may overlap in the frequency band.

For instance, as explained with respect to equation (1), the second signal component in the frequency domain may be represented by M_F spectral coefficients. Furthermore, with respect to the signal processing explained with respect to the M'_F sub-bands throughout this specification, $M_F > M'_F$ may hold, i.e., the signal processing explained with respect to the M'_F sub-bands throughout this specification may represent a signal processing performed with the first frequency resolution representing a lower frequency resolution compared to the frequency solution of the spectral components of first and second signal components (assuming a frequency domain signal processing). The reduction of frequency resolution may allow for an increased suppression of musical tones and may correspond to the properties of the human auditory system where the frequency selectivity decreases with higher frequencies.

For instance, M_F gain factors of the second set may be determined being associated with M_F sub-bands of the second frequency band, and an entity may be configured to decrease the frequency resolution from M_F to M'_F , i.e. this entity may be configured to output M'_F gain factors. This decrease of frequency resolution may be performed based on combining frequency-bins using overlapping Hann windows. Thus, the variance of the gain factors over time may be reduced.

For instance, an entity may be configured to expand the frequency resolution of the combined set of gain factors from M'_F to M_F . For instance, this may be performed based on using overlap-add of scaled Hann windows, e.g., the same Hann windows as mentioned above. Based on the frequency expanded combined set of gain factors the combined signal enhancement processing may be performed, e.g. by means of spectral weighting.

According to one embodiment, said combining is performed based on at least one cross-fading factor, wherein one of the at least one cross-fading factor combines one value of the at least one value of the first set of at least one value with one value of the at second set of at least one value.

As an example, the first set of at least one value may be represented by $G_1(j)$ with $j \in \{0, \dots, M-1\}$ and the second set of at least one value may be represented by $G_2(j)$ with $j \in \{0, \dots, M-1\}$, wherein $M \geq 1$ holds. For instance, $G_1(j)$ may correspond to $G_{bwe}(j)$, $G_2(j)$ may correspond to $G_{com}(j)$ and M may correspond to M'_F .

A combiner may be configured to combine the first set of at least one value and the second set of at least one value by means of a set of at least one cross-fading factor, wherein this set of at least one cross-fading factor may be represented by $\alpha(j)$ with $j \in \{0, \dots, M-1\}$. For instance, the j th combined value $c(j)$ may be a function of the j th value $G_1(j)$ of the first set of at least one value and the j th value $G_2(j)$ of the second set of at least one value and at least of the j th cross-fading factor $\alpha(j)$.

As an example, the j th combined value $c(j)$ of the set of combined values may be determined as follows:

$$c(j) = \alpha(j) \cdot G_1(j) + (1 - \alpha(j)) \cdot G_2(j) \quad (7)$$

Any other well-suited combining based on the set of at least one cross-fading factor may also be applied.

Thus, for instance, a weighting between the influence of the first signal enhancement processing and the second signal processing may be performed.

According to one embodiment, each of the at least one cross-fading factor is calculated on the basis of one value of the first set of at least one value, of one value of the second set and of at one reference value of a set of at least one reference value, the set of at least one reference value representing at least one reference value configured to be used to perform a reference enhanced signal processing on the second signal component.

As an example, the first set of at least one value may be represented by $G_1(j)$ with $j \in \{0, \dots, M-1\}$, the second set of at least one value may be represented by $G_2(j)$, and the set of at least one reference value may be represented by $G_r(j)$. The at least one reference value is configured to be used to perform a reference enhanced signal processing on the second signal component. For instance, the at least one reference value may represent reference filter values or weighting factor configured to perform a reference filtering in order to perform the enhanced signal processing on the second signal component. E.g., each of the at least one reference value may be determined during a training process for each frame. The method of determining each of the at least one reference value may depend on the signal enhancement processing. For instance, during this training process reference signal parameters may be determined in order to obtain the at least one reference value.

For instance, a cross-fading factor $\alpha(j)$ of the set of at least one cross fading factors may be determined based on the respective value $G_1(j)$ of the first set of at least one value, based on the respective value $G_2(j)$ of the second set of at least one value and based on the respective cross-fading factor $\alpha(j)$ of the set of at least one cross fading factor. For instance, this may be performed as follows:

$$\alpha(j) = \frac{(G_r(j) - G_1(j))^2}{(G_r(j) - G_1(j))^2 + (G_r(j) - G_2(j))^2} \quad (8)$$

As an example, assuming that the signal enhancement processing represents a noise reduction on the audio signal as mentioned above, the at least one reference value $G_r(j)$ may represent at least one reference weighting factor which may be determined from a clean audio and noise signal. For instance, this determining may be performed on a reference

a posteriori SNR $\gamma_r(j) = \frac{|Y_2(j)|^2}{|N_r(j)|^2}$ and/or on a reference a priori

SNR $\xi_r(j) = \frac{|S_2(j)|^2}{|N_2(j)|^2}$, wherein $|N_2(j)|^2$ may represent the noise

power of the j th sub-band of the second signal component and wherein $|S_2(j)|^2$ represents power of j th sub-component of second signal, associated with the j th sub-band.

According to one embodiment, said at least one cross-fading vector is determined based on a plurality of reference cross-fading factors and estimations of signal parameters associated with the first and second frequency band, wherein the plurality of reference cross-fading factors has been determined during a training process, wherein each reference cross-fading factor of the plurality of reference cross-fading factors is associated with at least one signal parameter of the first frequency band and at least one signal parameter of the second frequency band.

For instance, during a training process, which may be performed for every frame λ and/or for every sub-band of the j sub-bands of the second frequency band, a plurality of reference cross-fading vectors may be determined, wherein

each reference cross-fading factor of the plurality of reference cross-fading factors is associated with respective at least one signal parameter of the first frequency band and a respective at least one signal parameter of the second frequency band.

The at least one signal parameter of the first frequency band and the at least one signal parameter of the second frequency band may represent any well-suited signal parameter associated with the first signal component and the second signal component.

As an example, based on training data, a look-up table for the estimation of cross-fading values $\alpha(j)$ may be generated for each of the j sub-bands. For instance, p_1 may represent the signal parameter of the first frequency band and p_2 may represent the signal parameter of the second frequency band, wherein p_1 and p_2 are determined for each cross-fading value $\alpha(j)$ obtained during a training process based on training data. p_1 and p_2 may be quantized, and the final look-up table may provide one reference cross-fading value $\alpha(j)$ for each quantized combination of p_1 and p_2 .

For instance, p_1 may represent a first SNR associated with the first frequency band and p_2 may represent a second SNR associated with the second frequency band.

As an example, the at least one signal parameter of the first frequency band may represent the averaged SNR

$$\bar{\xi}_{1,r} = \frac{1}{M_F} \sum_{j=0}^{M_F-1} \frac{|Y_1(j)|^2}{|N_{1,r}(j)|^2} \text{ of the first signal component in the first}$$

frequency band, and the at least one signal parameter of the second frequency band may represent the reference a priori

SNR $\xi_r(j) = \frac{|S_2(j)|^2}{|N_2(j)|^2}$ of the respective jth sub-band.

For instance, as a non-limiting example, a first reference cross-fading value $\alpha_r(j)$ is obtained in a training process, e.g. for every frame λ and/or for every sub-band j , e.g. as mentioned above. In addition, the reference a priori SNR

$\xi_r(j) = \frac{|S_2(j)|^2}{|N_2(j)|^2}$ of the respective jth sub-band and the averaged

SNR $\bar{\xi}_{1,r} = \frac{1}{M_F} \sum_{j=0}^{M_F-1} \frac{|Y_1(j)|^2}{|N_{1,r}(j)|^2}$ of the first signal component in the

first frequency band may be determined. On this basis, the plurality of reference cross-fading value $\alpha_r(j)$ may be determined being associated with the respective reference a

priori SNRs $\xi_r(j) = \frac{|S_2(j)|^2}{|N_2(j)|^2}$ of the respective jth sub-band and

the respective averaged SNRs $\bar{\xi}_{1,r} = \frac{1}{M_F} \sum_{j=0}^{M_F-1} \frac{|Y_1(j)|^2}{|N_{1,r}(j)|^2}$ of the first signal component in the first frequency band.

As an example, based on training data, a look-up table for the estimation of cross-fading values $\alpha(j)$ may be generated for each of the j sub-bands. Therefore, $\xi_r(j)$ and $\bar{\xi}_{1,r}$ may be quantized (e.g., 1 dB step size or another well-suited step size) and the associated reference cross-fading values $\alpha_r(j)$ of the j sub-bands define the at least one reference cross-fading value. Thus, the final look-up table may provide

one reference cross-fading value $\alpha(j)$ for each quantized combination of $\xi_r(j)$ and $\bar{\xi}_{1,r}$.

Accordingly, the first and second signal enhancement processing can be combined in an adaptive way based on the at least one reference cross-fading value.

For instance, a cross-fading value $\alpha(j)$ may be determined by means of interpolation between at least two reference cross-fading values.

Furthermore, as an example, the respective SNR estimates of the SNR associated with the first frequency band and the SNR associated with the second frequency band may be determined based on the above mentioned conventional techniques.

According to one embodiment, said signal enhancement processing represents noise reduction processing.

The features of the present invention and of its exemplary embodiments as presented above shall also be understood to be disclosed in all possible combinations with each other.

It is to be noted that the above description of embodiments of the present invention is to be understood to be merely exemplary and non-limiting.

Further aspects of the invention will be apparent from and elucidated with reference to the detailed description presented hereinafter.

BRIEF DESCRIPTION OF THE FIGURES

In the figures show:

Fig. 1a is a schematic block diagram which illustrates a first exemplary apparatus;

Fig. 1b is a schematic block diagram which illustrates a second exemplary apparatus;

Fig. 2a is a flow chart illustrating a first exemplary method;

Fig. 2b is a flow chart illustrating a second exemplary method;

Fig. 3a is a schematic block diagram which illustrates a first exemplary filtering;

Fig. 3b is a schematic block diagram which illustrates a second exemplary filtering;

Fig. 4a is a schematic block diagram which illustrates an exemplary entity configured to estimate the at least one parameter;

Fig. 4b is a schematic block diagram which illustrates a first exemplary determining of a first set of at least one value configured to be used to perform a signal enhancement processing on the second signal component;

Fig. 4c is a schematic block diagram which illustrates a first exemplary entity configured to determine a second set of at least one value configured to be used to perform a signal enhancement processing on the second signal component;

Fig. 4d is a schematic block diagram which illustrates a second exemplary entity configured to determine a second set of at least one value configured to be used

to perform a signal enhancement processing on the second signal component;

Fig. 5a is a schematic block diagram which illustrates an exemplary combiner; and

Fig. 5b is a schematic block diagram which illustrates a second exemplary combiner.

DETAILED DESCRIPTION OF EXEMPLARY EMBODIMENTS OF THE INVENTION

In the following detailed description, exemplary embodiments of the present invention will be described in the context of exemplary methods and apparatuses.

Figure 1a is a schematic block diagram which illustrates a first exemplary apparatus. This first exemplary apparatus will be described in conjunction with the flow chart of a first exemplary method depicted in figure 2a.

The method comprises estimating 210 at least one parameter 150 based on a first signal component 130 of an audio signal 101, the at least one parameter 150 being associated with a second signal component 140 of the audio signal 210, and the method comprises performing 220 a signal enhancement processing on the second signal component 140 at least based on the at least one parameter.

As an example, the signal enhancement processing may represent noise reduction processing, or a voice enhancement processing or any other well-suited signal enhancement processing.

The first signal component is associated with a first frequency band and the second signal component is associated with a second frequency band. The first frequency band differs from the second frequency band.

For instance, the first frequency band may span a frequency range from f_1 to f_2 , and the second frequency band may span a frequency range from f_3 to f_4 , wherein $f_3 > f_1$ and $f_4 > f_2$ holds. As an example, the first and second frequency band may have an overlapping frequency range, i.e. $f_3 < f_2$ holds, but as another example, the first and second frequency band may not overlap, i.e. $f_3 > f_2$ holds. For instance, the first frequency band may span a frequency range from 50 Hz to 4 kHz, and the second frequency band may span a frequency range from 4 to 7 kHz. But of course, any other well suited frequency assignments for the first and second frequency range may be applied.

The estimated at least one parameter may represent at least one parameter which may be beneficial to increase the quality of the signal enhancement processing. The type of the at least one parameter may depend on the kind of signal enhancement processing.

For instance, the first signal component 130 may be extracted from the audio signal 101 by means of filtering, e.g. by a lowpass filter or a bandpass filter, and the second signal component 140 may be extracted from the audio signal 101 by means of filtering, e.g. by a highpass filter or a bandpass filter. Such a filtering is depicted in the exemplary systems depicted in figures 1c and 1d, respectively, wherein the audio signal 101 is filtered by first filter 103 in order to output

the first signal component 130 and filtered by second filter 104 in order to output the second signal component 140. For instance, the first and second filters 103, 104 may further be configured to perform a downsampling. As an example, the first filter 103 and the second filter 104 may represent a 2-channel Finite Impulse Response (FIR) Quadrature Mirror Filter (QMF) bank.

Entity 110 is configured to estimate the at least one parameter 150 at least based on the first signal component 130. For instance, this estimation is not based on the second signal component 140.

As an example, at least one parameter of the at least one parameter 150 may be estimations of signal properties of the second signal component 140. For instance, these signal properties may comprise statistical information of the spectral characteristics of the second signal component. E.g., at least one parameter of the at least one parameter may represent a set of spectral envelope representatives of the second signal component. Thus, this set of spectral envelope representatives may be estimations of representations of the spectral envelope of the second signal component 140. For instance, this set of spectral envelope representatives may comprise at least one energy value, wherein each of the at least one energy value is associated with one frequency sub-band of the second signal component. Furthermore, at least one parameter of the at least one parameter may represent further information with respect to the second signal component.

Based on this at least one parameter 150, which has been estimated based on the first signal component 130, a signal enhancement processing is performed to the second signal component 140, as indicated by entity 120 in figure 1a, wherein entity 120 is configured to output an enhanced second signal component 142. For instance, in case the signal enhancement processing represents a noise reduction processing, the enhanced second signal component 142 may represent a noise reduced second signal component 142 or the weighting gains to perform the noise reduction in the frequency domain.

Thus, spectral dependencies may be exploited between the first signal component 130 and the second signal component 140 in order to perform the signal enhancement processing on the second signal component 140. Accordingly, spectral dependencies between the first frequency band and the second frequency band may be used to improve the signal enhancement processing.

The dashed arrow depicted in figure 1a illustrates that there may be a further signal processing with respect to the at least one parameter 150 before signal enhancement processing to the second signal component 140 is performed.

For instance, entity 120 may comprise at least one filter 145 configured to perform a signal enhancement processing to the second signal component 140, as depicted in figure 1b. As an example, one filter 145 of this at least one filter may be configured to be adapted to at least one value 150', i.e. this at least one value 150' may be configured to be used to perform a signal enhancement processing to the second signal

component 140. For instance, this at least one value 150' may represent at least one filter coefficient. Accordingly, as an example, the signal enhancement processing to the second signal component 140 may be performed based on the at least one parameter 150 by means of filter 145, wherein the at least one value 150' may be determined at least based on the at least one parameter 150.

For instance, this at least one value 150' may comprise at least one value configured to be used to perform a signal enhancement processing to the second signal component 140 in the time-domain. As an example, this at least one value may represent at least one filter coefficient of a filter configured to filter the second signal component 140 in the time-domain. For instance, this filter may represent a Finite Impulse Response (FIR) filter or an Infinite Impulse Response (IIR) filter. As an example, this at least one value may be determined based on a Minimum Mean Squared Error (MMSE) criterion or a Zero Forcing (ZF) criterion or any other well-suited criterion. In this exemplary case, a signal enhancement processing to the second signal component is performed by filtering the second signal component by means of the respective filter in the time-domain.

Figure 1c depicts such a filtering in the time-domain, wherein filter 141 is adapted to the at least one value 150' and configured to perform a signal enhancement processing to the second signal component 140. Accordingly, filter 141 is configured to output an enhanced second signal component 142, wherein filter 141 is configured to perform this signal enhancement based on the at least one parameter 150.

Similarly, filter 131 is configured to perform a signal enhancement processing to the first signal component 130 and to output an enhanced first signal component 132. This filter 131 may be adapted to at least one value 116 configured to be used to perform a signal enhancement processing to the first signal component 130. For instance, this at least one value 116 may be determined based on the first signal component 130. The system depicted in figure 1c comprises a third and fourth filter 193, 194 and a combiner 195 configured to combine the enhanced first signal component 132 and the enhanced second signal component 142 to a combined enhanced output audio signal 199 having the same bandwidth like the input audio signal 101. For instance, the third filter and fourth filter 193, 194 may be configured to perform an upsampling to the respective input signal. Furthermore, as an example, the third and fourth filter may represent a FIR QMF bank configured to combine the enhanced first and second signal components 132, 142 to the wideband output signal 199.

Furthermore, as another example, the at least one value 150' may comprise at least one value configured to be used to perform a signal enhancement processing to the second signal component 140 in the frequency-domain. Figure 1d depicts such a filtering in the frequency-domain, wherein converters 170 are configured to perform a conversion from time-domain to frequency-domain. Thus, converters 170 may convert the time domain first and second signal components 130, 140 into frequency domain first and second signal components 130', 140', respectively. As an example, the converter 170 may be configured to output M_F subcomponents representing the respective first and second signal component 130', 140' in

the frequency-domain, wherein each subcomponent of the M_F subcomponents is associated with one of M_F sub-bands in the first and second frequency band.

For instance, converter 170 may represent a Fourier Transformation, e.g. a FFT, a DFT, STFT or any other well-suited transformation. Filter 141' is adapted to the at least one value 150' and configured to perform a signal enhancement processing to the second signal component 140' in the frequency domain. Accordingly, filter 141' is configured to output an enhanced second signal component 142' in the frequency domain, wherein filter 141' is configured to perform this signal enhancement based on the at least one parameter 150.

Similarly, filter 131' is configured to perform a signal enhancement processing to the first signal component 130' in the frequency domain and to output an enhanced first signal component 132' in the frequency domain. This filter 131' may be adapted to at least one value 116 configured to be used to perform a signal enhancement processing to the first signal component 130. For instance, this at least one value 116 may be determined based on the first signal component 130. The system depicted in figured 1d further comprises converters 175 configured to perform a conversion from frequency-domain to time-domain. Thus, converters 175 may convert the enhanced frequency-domain first and second signal components 132', 142' to enhanced time-domain first and second signal components 132, 142, respectively. Of course, the converters 170 and 175 may be re-arranged. For instance, time-domain to frequency-domain conversion may be applied to the audio

signal 101 before being filtered by the first and second filter 103, 104, and/or frequency-domain to time-domain conversion may be applied to output signal 199 after being combined by combiner 195. In this exemplary case, the first, second, third and fourth filter 103, 104, 193, 194 and combiner 195 are configured to operate in the frequency domain.

For instance, the at least one value 150' may comprise at least two gain factors, wherein each of the at least two gain factors is associated with a separate sub-band of the second frequency band. As an example, these at least two gain factors may be determined based on a MMSE criterion or a ZF criterion or any other well-suited criterion. In this exemplary case, a signal enhancement processing to the second signal component 140' may be performed based on weighting a frequency component of the second signal 140' associated with a sub-band of the second frequency band with the respective gain factor of the at least two gain factors. Thus, for instance, each of the at least two gain factors is associated with one sub-band of the second frequency band.

Figure 3a is a schematic block diagram which illustrates a first exemplary filtering in the frequency domain based on the at least one value 150' configured to be used to perform a signal enhancement processing to the second signal component 140' in the frequency-domain. As an example, the second signal component 140' comprises M_F sub-components 301, 302 ... 303 representing the second signal component 140' in the frequency-domain, wherein each subcomponent 301, 302 ... 303 of the M_F subcomponents is associated with one of M_F

sub-bands in the second frequency band. The at least two gain factors 150 are represented by M_F gain factors 321, 322 ... 323 which are fed to filter 141'. For instance, a first multiplier 331 multiplies the first subcomponent 301 with the first gain factor 321, a second multiplier 332 multiplies the second subcomponent 302 with the second gain factor 322, and so on until the M_F th multiplier 333 multiplies the M_F th subcomponent 301 with the M_F th gain factor 333. Thus, the filter 141' is configured to output M_F enhanced sub-components 311, 312 ... 313 which have been weighted by the respective gain factors of the M_F gain factors, wherein these M_F enhanced sub-components 311, 312 ... 313 may represent the enhanced second signal component 142'. As an example, the filter 131' depicted in figure 1d may be realized in the same way as filter 141' depicted in figure 3. Of course, any other weighting implementation may be applied in order to perform the signal enhancement processing to the second (or first) signal component.

Thus, in the time-domain as well as in the frequency domain a signal enhancement processing on the second signal component 140 may be performed based on the at least one parameter determined on the basis of the first signal component 130.

Figure 1a depicts that the at least one parameter 150 is estimated based on the first signal component 130. This has to be understood, for instance, that the at least one parameter may be estimated based on the first signal component 130 in the time-domain, or based on the first signal component 130' in the frequency-domain, or based on the enhanced signal

component 132 in the time-domain, or based on the enhanced signal component 132' in the frequency domain.

Fig. 2b depicts a flow chart illustrating a second exemplary method based on the first exemplary method.

This second exemplary method comprises performing 205 a signal enhancement processing to the first signal component 130, 130' in order to obtain an enhanced first signal component 132, 132'. For instance, this signal enhancement processing to the first signal component 130, 130' may be performed as explained with respect to figures 1b and 1c, but any other well-suited signal enhancement may also be applied. Then, estimating 210 the at least one parameter is performed based on the enhanced first signal component 132, 132'. Accordingly, the input of entity 110 depicted in figure 1a may be replaced with an enhanced first signal component 132, 132'. Thus, regarding the estimation of the at least one parameter 150, it has to be understood that the term based on the first signal component may include based on the enhanced first signal component.

Fig. 4a depicts a schematic block diagram which illustrates an exemplary entity 110' configured to estimate the at least one parameter 150 based on the first signal component 130. For instance, this exemplary entity 110' may represent the entity 110 depicted in figure 1a.

For instance, the input 430 depicted in figure 4a may represent the first signal component 130, 130' or the enhanced first signal component 132, 132'. In the sequel, it is referred to the first signal component 130, 130', but the

enhanced first signal component 132, 132' may also be used as input.

Entity 110' comprises a unit 410 and an estimator 420. The unit 410 is configured to determine at least one feature 440 based on the first signal component 130, 130' representing at least one signal parameter being associated with the first signal component 130, 130'. At least one feature of the least one feature 440 may represent a set of spectral envelope representatives of the first signal component 130, 130'. For instance, the at least one feature of the at least one feature may be at least one energy representative, wherein each of the at least one energy representative is associated with the energy of the first signal component or the enhanced first signal component in a frequency sub-band of the first frequency band. As an example, at least one feature of the at least one feature 440 may be N_c mel-frequency cepstral coefficients (MFCCs). Thus the frequency sub-bands associated with the first frequency band are equally spaced on the mel scale.

Furthermore, at least one feature of the at least one feature may comprise further information of the first signal component 130, 130'. For instance, at least one feature of the at least one feature may be associated with the zero-crossing rate (ZCR) of the first signal component 130, 130'. The zero-crossing rate may be the rate of sign-changes along a signal, i.e., the rate at which the signal changes from positive to negative or back.

As an example, the at least one feature 440 may be denoted as vector $x(i)$, where 'i' denotes the time index. For

instance, unit 410 may operate on a frame-by-frame basis, i.e., for instance, $X=\{x(1), \dots, x(\lambda)\}$ may represent a sequence of feature vectors from the first signal component 130, 130' of frames 1 to λ . For instance, λ designates the current frame.

The estimator 420 is configured to estimate the at least one parameter 150 at least based on the at least one feature 440. For instance, the estimator 420 may be configured to perform the estimation based on a Hidden Markov Model (HMM).

For instance, the estimator may apply a MMSE criterion, a Maximum a Posteriori (MAP) criterion, a Maximum Likelihood (ML) criterion, a Zero Forcing (ZF) criterion or any other well-suited criterion for estimating the at least one parameter at least based on the at least one feature 440. As an example, the estimator 420 may not apply any information extracted from or based on the second signal component 140, 140' in order to estimate the at least one parameter 150. Thus, for instance, the at least one parameter is estimated based on features that are determined only from the first signal component 130, 130'.

The estimation may be based on at least one feature which has been previously determined by unit 410. As an example, the estimating may be based on sequence $X=\{x(\lambda-m), \dots, x(\lambda-1), x(\lambda)\}$, wherein $x(\lambda)$ may represent the current at least one feature and $x(\lambda-m) \dots x(\lambda-1)$ may represent the m previous at least one features. Thus, in this case $m \geq 1$ holds, m representing an integer.

As an example, under assumption of applying an MMSE estimation, the at least one estimated parameter 150,

exemplarily denoted as vector $\hat{y}$, may be estimated by means of MMSE criterion for estimation of parameter vector y :

$$E\{\|y - \hat{y}\|^2 | X\} = \min \quad (9)$$

For instance, equation (9) may be solved by the conditional expectation $\hat{y}_{MMSE} = E\{\hat{y}_i | X\}$.

It is now assumed, as an example, that at least one parameter of the at least one parameter represents a set of spectral envelope representatives of the second signal component 140, 140'. For instance, the estimated at least one parameter may be represented by vector $\hat{y} = \left\{ |\hat{S}_2(0)|^2, \dots, |\hat{S}_2(M'_F - 1)|^2 \right\}$, wherein

$|\hat{S}_2(j)|^2$ with $j \in \{0, \dots, M'_F - 1\}$ may represent the estimated energy of the second signal in the j th sub-band of a set of $M'_F - 1$ frequency sub-bands of the second frequency band. As an example, the sub-bands of this set of $M'_F - 1$ frequency sub-bands of the second frequency band may overlap in the frequency range.

Furthermore, the estimation may further be based on a codebook, the codebook comprising a plurality of sets of spectral envelope representatives associated with the second signal component 140. This codebook may represent a precomputed codebook $C = \{\hat{y}_1, \dots, \hat{y}_{Mc}\}$ for the parameter vectors y comprising Mc entries. For instance, this codebook may be obtained with the LBG algorithm presented by Y. Linde, A. Buzo, R.M. Gray in "An algorithm for vector quantizer design," IEEE Transactions on Communications, vol. 28, no.

1, pp. 84-95, Jan. 1980. Of course, any other well-suited codebook may also be used.

Using the codebook C , the MMSE estimate may be expressed as:

$$\hat{y}_{MMSE} = \sum_{\hat{y}_i \in C} \hat{y}_i \cdot P(\hat{y}_i | X), \quad (10)$$

which may represent a weighted sum over the M centroids of the codebook C . The weights $P(\hat{y}_i | X)$ represent a posteriori probabilities based on the determined sequence X . Thus, according to equation (10), an a posteriori probability based on the determined sequence X is determined for each of the plurality of sets of spectral envelope representatives of the codebook, and a weighted sum over the plurality of sets of spectral envelope representatives is determined, wherein each set of spectral envelope representatives $\hat{y}_i$ is weighted by the respective a posteriori probability $P(\hat{y}_i | X)$.

For instance, determining the at least one a posteriori probability may be based on a Hidden Markov Model (HMM). This HMM may be trained during an offline training phase. As an example, the HMM techniques described by B. Geiser, H. Taddei and P. Vary in "Artificial Bandwidth Extension without Side Information for ITU-T G.729.1," in Proceedings of European Conference on Speech Communication and Technology (INTERSPEECH), Antwerp, Belgium, August 2007 may be applied.

Thus, the entity 110' depicted in figures 4a is configured to estimate at least one parameter 150 based on the first signal component 130, 130'.

Figure 4b is a schematic block diagram which illustrates a first exemplary determining of a first set of at least one value 450 configured to be used to perform a signal enhancement processing on the second signal component 140, 140'. For instance, this at least one value 450 may represent at least one filter coefficient, wherein a filter adapted to this at least one filter coefficient may be used to perform the signal enhancement processing on the second signal component 140, 140', as explained above. As an example, the at least one value 450 may represent the first set of at least two gain factors configured to be used to perform a spectral weighting of the second signal 140' based on weighting a sub-band frequency of the second signal associated with a sub-band of the second frequency band with the respective gain factor of the at least two gain factors, as exemplarily depicted in figure 3a or 3b.

As an example, under the assumption that the at least one value 450 represents the first set of at least two gain factors, entity 460 may represent an SNR estimator 460 which is configured to estimate at least one SNR representative 465 based on the at least one parameter 150 and the second signal component 140, 140'. This at least one SNR representative is associated with the second signal component 140, 140'. For instance, this at least one SNR representative may comprise a plurality of SNR representatives, wherein each of the plurality of SNR representatives is associated with one sub-band of a plurality of frequency sub-bands of the second frequency band. As an example, the plurality of frequency sub-bands may comprise $M'_F - 1$ frequency sub-bands of the second frequency band, and, for instance, the sub-bands of

this set of $M'_F - 1$ frequency sub-bands of the second frequency band may overlap in the frequency range. As an example, the plurality of SNR representatives may comprise a set of $M'_F - 1$

a priori SNR values $\hat{\xi}_2(j) = \frac{|\hat{S}_2(j)|^2}{|\hat{N}_2(j)|^2}$, wherein $|\hat{N}_2(j)|^2$ may

represent the estimated noise power of the j th sub-band of the second signal component and wherein $|\hat{S}_2(j)|^2$ represents one of the at least one parameter 150. For instance, the entity 460 may be configured to estimate the j noise powers $|\hat{N}_2(j)|^2$ with $j \in \{0, \dots, M'_F - 1\}$. For instance, the estimation of the j th noise power may be written as

$$|\hat{N}_2(j)|^2 = \max(|Y_2(j)|^2 - |\hat{S}_2(j)|^2, 0), \quad (11)$$

wherein $|Y_2(j)|^2$ represents the energy of the audio signal in the j th sub-band of the second frequency band.

Furthermore, as an example, the plurality of SNR representatives may further comprise a set of $M'_F - 1$ a

posteriori SNR values $\hat{\gamma}_2(j) = \frac{|Y_2(j)|^2}{|\hat{N}_2(j)|^2}$.

In this exemplary case, i.e., under the assumption that the at least one value 450 represents the first set of at least two gain factors, entity 470 may be configured to determine the first set of at least two gain factors at least based on the plurality of SNR representatives. For instance, the at

least two gain factors of the first set may be represented by $G_{bwe}(j)$ with $j \in \{0, \dots, M'_F - 1\}$, wherein $G_{bwe}(j)$ represents the gain factor associated with the j th sub-band of the $M'_F - 1$ frequency sub-bands of the second frequency band. As an example, under the exemplary assumption that the signal enhancement processing is a noise reduction processing, $G_{bwe}(j)$ may be calculated as

$$G_{bwe}(j) = \frac{\hat{\xi}_2(j)}{1 + \hat{\xi}_2(j)}, \text{ or} \quad (12a)$$

$$G_{bwe}(j) = \frac{\hat{\gamma}_2(j)}{1 + \hat{\gamma}_2(j)} \quad (12b)$$

or $G_{bwe}(j)$ may be calculated based on $\hat{\gamma}_2(j)$ and $\hat{\xi}_2(j)$, e.g. as described by Y. Ephraim and D. Malah in "Speech Enhancement Using a Minimum Mean-Square Error Short-Time Spectral Amplitude Estimator", IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. ASSP-32, no. 6, pp. 1109-1121, December 1984. For instance, entity 470 may be configured to determine the first set of at least two gain factors at least based on the plurality of SNR representatives based on any other well-suited method or as described by J. S. Lim and A. V. Oppenheim in "Enhancement and Bandwidth Compression of Noisy Speech", Proceedings of the IEEE, vol. 67, no. 12, pp. 1585-1604, December 1979.

As another example, under the assumption that the at least one value represents at least one filter coefficient of a time-domain filter, the at least one value 465 may comprise a plurality of SNR representatives or other statistical representatives of the second signal component which may be

used by entity 470 to determine the at least one filter coefficient.

The first set of at least one value 450 may be associated with a first signal enhancement processing which is based on the at least one parameter 150 determined on basis of the first signal component 130, 130'.

Figure 4c is a schematic block diagram which illustrates a first exemplary entity 480 configured to determine a second set of at least one value 490 configured to be used to perform a signal enhancement processing on the second signal component 140, 140'. For instance, this at least one value 490 may represent at least one filter coefficient, wherein a filter adapted to this at least one filter coefficient may be used to perform the signal enhancement processing on the second signal component 140, 140', as explained above. As an example, the at least one value 490 may represent second set of at least two gain factors configured to be used to perform a spectral weighting of the second signal 140' based on weighting a sub-band frequency of the second signal associated with a sub-band of the second frequency band with the respective gain factor of the at least two gain factors, as exemplarily depicted in figure 3a or 3b. For instance, the at least two gain factors of the second set may be represented by $G_{conv}(j)$ with $j \in \{0, \dots, M'_F - 1\}$, wherein $G_{conv}(j)$ represents the gain factor associated with the j th sub-band of the $M'_F - 1$ frequency sub-bands of the second frequency band.

The second set of at least one value 490 may be associated with a second signal enhancement processing which is not based

on the at least one parameter 150 determined on basis of the first signal component 130, 130'. For instance, this second signal enhancement processing may represent a conventional signal enhancement processing.

As an example, under the assumption that the signal enhancement processing represents a noise reduction processing, entity 480 may be configured to perform a noise power estimation, e.g. as explained by R. Martin in "Noise Power Spectral Density Estimation Based on Optimal Smoothing and Minimum Statistics", IEEE Transactions on Speech and Audio Processing, vol. 9, no. 5, pp. 501-512, 2001, and to perform SNR estimation, e.g. as mentioned above with respect to figure 4b, and to calculate the set of at least one value 490, e.g. as mentioned above with respect to figure 4b.

Figure 4d is a schematic block diagram which illustrates a second exemplary entity 480' configured to determine a second set of at least one value 490' configured to be used to perform a signal enhancement processing on the second signal component 140' in the frequency domain. For instance, the second signal component 140' in the frequency domain, e.g. determined as explained with respect to figure 1d, may be represented by M_F spectral coefficients at frequency bin k and frame λ given by:

$$Y_2(\lambda, k) = S_2(\lambda, k) + N_2(\lambda, k), \quad (13)$$

where $S_2(\lambda, k)$ and $N_2(\lambda, k)$ represent the spectral coefficients of the audio and the noise signal. For the sake of brevity, the frame index λ is omitted in the following. Furthermore,

$M_F > M'_F$ may hold, i.e., the signal processing explained with respect to the M'_F sub-bands throughout this specification may represent a signal processing performed with a lower frequency resolution compared to the frequency resolution of the first and second signal components 130', 140' outputted by converters 170 depicted in figure 1d. In case the signal enhancement processing comprises noise reduction, the reduction of frequency resolution may allow for an increased suppression of musical tones and may correspond to the properties of the human auditory system where the frequency selectivity decreases with higher frequencies. For instance, entity 480' may be configured to output M_F gain factors being associated with M_F sub-bands of the second frequency band, and an entity 495 is configured to decrease the frequency resolution from M_F to M'_F , i.e. the entity 495 may be configured to output M'_F gain factors. This decrease of frequency resolution may be performed based on combining frequency-bins using overlapping Hann windows. Thus, the variance of the gain factors may be reduced.

For instance, under the assumption that $M_F > M'_F$ holds, the second exemplary filtering depicted in figure 3b may be used for performing the spectral weighting based on the at least one value 150', wherein the at least one value 150'' represent at least two gain factors associated with frequency resolution M'_F . Entity 350 may be configured to expand the frequency resolution from M'_F to M_F . For instance, this may be performed based on using overlap-add of scaled Hann windows, e.g. the same Hann windows as mentioned above. Thus, as an example, entity 350 is configured to output a M_F gain

factors 150'' which are representatives of inputted M'_F gain factors 150'. The remaining signal processing of the second exemplary filtering corresponds to the first exemplary filtering.

Figure 5a is a schematic block diagram which illustrates an exemplary combiner 510 configured to combine the first set of at least one value 450 and the second set of at least one value 490 to a combined set of at least one value 550. As an example, the combined set of at least one value 550 may be used as the at least one value 150' explained with respect to one of figures 1b, 1c, 1d, 3a, 3b and 4b. Thus, the combined set of at least one value 550 may be used to perform a signal enhancement processing on the second signal component 140, 140', e.g. in time-domain or in the frequency domain. For instance, the first signal enhancement method and the second signal enhancement method may be combined in order to provide an increased signal enhancement of the second signal component 140, 140'.

For instance, the first set of at least one value 450 may comprise M'_F values and the second set of at least one value 490 may comprise M'_F values, and the combiner 510 may be configured to combine one value of the first set of at least one value with one value of the second set of at least one value to one combined value of the set of at least one combined value 550, wherein the set of at least one combined value comprises M'_F combined values.

This combining may be performed adaptively. As an example, this adaptively combining may be performed on the basis of determined signal and/or noise parameters.

Figure 5b is a schematic block diagram which illustrates a second exemplary combiner 510' configured to combine the first set of at least one value 450 and the second set of at least one value 490 to a combined set of at least one value 550, wherein the second exemplary combiner 510' is based on the first exemplary combiner 510.

As an example, the first set of at least one value 450 may be represented by $G_1(j)$ with $j \in \{0, \dots, M-1\}$ and the second set of at least one value 490 may be represented by $G_2(j)$ with $j \in \{0, \dots, M-1\}$, wherein $M \geq 1$ holds. For instance, $G_1(j)$ may correspond to $G_{bwc}(j)$, $G_2(j)$ may correspond to $G_{conv}(j)$ and M may correspond to M'_F .

The combiner 550' may be configured to combine the first set of at least one value 450 and the second set of at least one value 490 by means of a set of at least one cross-fading factor, wherein this set of at least one cross-fading factor may be represented by $\alpha(j)$ with $j \in \{0, \dots, M-1\}$. For instance, the j th combined value $c(j)$ may be a function of the j th value $G_1(j)$ of the first set of at least one value 450 and the j th value $G_2(j)$ of the second set of at least one value 490 and at least of the j th cross-fading factor $\alpha(j)$.

Thus, for instance, said combining may be performed based on at least one cross-fading factor, wherein one of the at least

one cross-fading factor combines one value of the at least one value of the first set of at least one value 450 with one value of the at least second set of at least one value 490.

As an example, as exemplarily depicted in figure 5b, the j th combined value $c(j)$ 550' of the set of combined values 550 may be determined as follows:

$$c(j) = \alpha(j) \cdot G_1(j) + (1 - \alpha(j)) \cdot G_2(j) \quad (14)$$

For example, reference sign 545 may be associated with the set of cross-fading factors $\alpha(j)$, and reference sign 535 may be associated with a second set of cross-fading factors $\alpha_2(j)$, wherein in this exemplary case $\alpha_2(j) = 1 - \alpha(j)$ holds. The multipliers 535 and 540 and the adder 560 may perform the mathematical operations and may be implemented j -times.

Of course, the second set of cross-fading factors 535 may be independent from the set of cross-fading factors 545, or there may exist another correlation between the set of cross-fading factors 545 and the second set of cross-fading factors 535.

For instance, each of the at least one cross-fading factor is calculated on the basis of one value of the first set of at least one value 450, of one value of the second set of at least one value 490 and of at one reference value of a set of at least one reference value, the set of at least one reference value representing at least one reference value configured to be used to perform a reference enhanced signal processing on the second signal component.

As an example, the first set of at least one value 450 may be represented by $G_1(j)$ with $j \in \{0, \dots, M-1\}$, the second set of at least one value 490 may be represented by $G_2(j)$, and the set of at least one reference value may be represented by $G_r(j)$. The at least one reference value is configured to be used to perform a reference enhanced signal processing on the second signal component. E.g., each of the at least one reference value may be determined during a training process for each frame. The method of determining each of the at least one reference value may depend on the signal enhancement processing. For instance, during this training process reference signal parameters may be determined in order to obtain the at least one reference value.

For instance, a cross-fading factor $\alpha(j)$ of the set of at least on cross fading factors may be determined based on the respective value $G_1(j)$ of the first set of at least on value 450, the respective value $G_2(j)$ of the second set of at least on value 490 and the respective cross-fading factor $\alpha(j)$ of the set of at least on cross fading factor, e.g. as follows:

$$\alpha(j) = \frac{(G_r(j) - G_1(j))^2}{(G_r(j) - G_1(j))^2 + (G_r(j) - G_2(j))^2} \quad (15)$$

As an example, assuming that the signal enhancement processing represents a noise reduction on the audio signal as mentioned above, the at least one reference value $G_r(j)$ may represent at least one reference weighting factor which may be determined from a clean audio and noise signal. For

instance, this determining may be performed on a reference

a posteriori SNR $\gamma_r(j) = \frac{|Y_2(j)|^2}{|N_r(j)|^2}$ and/or on a reference a priori

SNR $\xi_r(j) = \frac{|S_2(j)|^2}{|N_2(j)|^2}$, wherein $|N_2(j)|^2$ may represent the noise

power of the jth sub-band of the second signal component and

wherein $|S_2(j)|^2$ represents power of jth sub-component of

second signal, associated with the jth sub-band.

For instance, during a training process, which may be performed for every frame λ and/or for every sub-band of the j sub-bands of the second frequency band, a plurality of reference cross-fading vectors may be determined, wherein each reference cross-fading factor of the plurality of reference cross-fading factors is associated with a first SNR associated with the first frequency band and with a second SNR associated with the second frequency sub-band.

As an example, a first reference cross-fading value $\alpha_r(j)$ is obtained in a training process, e.g. for every frame λ and/or for every sub-band j, e.g. as mentioned above. In addition,

the reference a priori SNR $\xi_r(j) = \frac{|S_2(j)|^2}{|N_2(j)|^2}$ of the respective jth

sub-band and the averaged SNR $\bar{\xi}_{1,r} = \frac{1}{M_F} \sum_{j=0}^{M_F-1} \frac{|Y_1(j)|^2}{|N_{1,r}(j)|^2}$ of the first

signal component in the first frequency band may be determined.

Based on training data, a look-up table for the estimation of cross-fading values $\alpha(j)$ may be generated for each of the j sub-bands. Therefore, $\xi_r(j)$ and $\bar{\xi}_{1,r}$ may be quantized (e.g., 1 dB step size or another well-suited step size) and the associated reference cross-fading values $\alpha_r(j)$ of the j sub-bands define the at least one reference cross-fading value. Thus, the final look-up table may provide one reference cross-fading value $\alpha(j)$ for each quantized combination of $\xi_r(j)$ and $\bar{\xi}_{1,r}$, $\xi_r(j)$ representing the SNR associated with the second frequency band and $\bar{\xi}_{1,r}$ representing the SNR associated with the first frequency band. Thus, the SNR associated with the second frequency band may represent a SNR associated with a sub-band of the second frequency band.

Accordingly, the first and second signal enhancement processing can be combined in an adaptive way based on the at least one reference cross-fading value.

A cross-fading value $\alpha(j)$ may be determined by means of interpolation between at least two reference cross-fading values.

Furthermore, the respective SNR estimates of the SNR associated with the first frequency band and the SNR associated with the second frequency band may be determined based on the above mentioned conventional techniques.

Furthermore, it is readily clear for a person skilled in the art that the logical blocks in the schematic block diagrams as well as the flowchart and algorithm steps presented in the

above description may at least partially be implemented in electronic hardware and/or computer software, wherein it may depend on the functionality of the logical block, flowchart step and algorithm step and on design constraints imposed on the respective devices to which degree a logical block, a flowchart step or algorithm step is implemented in hardware or software. The presented logical blocks, flowchart steps and algorithm steps may for instance be implemented in one or more digital signal processors (DSPs), application specific integrated circuits (ASICs), field programmable gate arrays (FPGAs) or other programmable devices. The computer software may be stored in a variety of computer-readable storage media of electric, magnetic, electro-magnetic or optic type and may be read and executed by a processor, such as for instance a microprocessor. To this end, the processor and the storage medium may be coupled to interchange information, or the storage medium may be included in the processor.

Any presented connection in the described embodiments is to be understood in a way that the involved components are operationally coupled. Thus, the connections can be direct or indirect with any number or combination of intervening elements, and there may be merely a functional relationship between the components.

Any of the processors mentioned in this text could be a processor of any suitable type. Any processor may comprise but is not limited to one or more microprocessors, one or more processor (s) with accompanying digital signal processor (s), one or more processor (s) without accompanying digital signal processor (s), one or more special-purpose computer chips,

one or more field-programmable gate arrays (FPGAs), one or more controllers, one or more application-specific integrated circuits (ASICs), or one or more computer(s). The relevant structure/hardware has been programmed in such a way to carry out the described function.

Any of the memories mentioned in this text could be implemented as a single memory or as a combination of a plurality of distinct memories, and may comprise for example a read-only memory, a random access memory, a flash memory or a hard disc drive memory etc.

Moreover, any of the actions described or illustrated herein may be implemented using executable instructions in a general-purpose or special-purpose processor and stored on a computer-readable storage medium (e.g., disk, memory, or the like) to be executed by such a processor. References to 'computer-readable storage medium' should be understood to encompass specialized circuits such as FPGAs, ASICs, signal processing devices, and other devices.

It will be understood that all presented embodiments are only exemplary, that features of these embodiments may be omitted or replaced and that other features may be added. Any mentioned element and any mentioned method step can be used in any combination with all other mentioned elements and all other mentioned method step, respectively. It is the intention, therefore, to be limited only as indicated by the scope of the claims appended hereto.

CLAIMS

1. A method comprising:
 - estimating at least one parameter based on a first signal component of an audio signal, the at least one parameter being associated with a second signal component of the audio signal; wherein
 - the first signal component is associated with a first frequency band and the second signal component is associated with a second frequency band; and
 - performing a signal enhancement processing on the second signal component at least based on the at least one parameter.
2. The method according to claim 1, comprising performing a signal enhancement processing on the first signal component in order to obtain an enhanced first signal component, wherein said estimating at least one parameter is performed on basis of the enhanced first signal component.
3. The method of one of the preceding claims, wherein at least one parameter of said at least one parameter represents a set of spectral envelope representatives of the second signal component.
4. The method according to one of the preceding claims, comprising determining at least one feature based on the first signal component representing at least signal

parameter being associated with the first signal component, wherein said estimating the at least one parameter is performed on basis of the at least one feature.

5. The method according to claim 4, wherein at least one feature of said at least one feature is a set of spectral envelope representatives of one out of:
 - the first signal component; and
 - the enhanced first signal component.
6. The method according to one of claims 4 and 5, wherein said estimating at least one parameter is further based on at least one previous feature which has been previously determined based on the first signal component.
7. The method of one of claims 4 to 6, wherein at least one parameter of said at least one parameter represents a set of spectral envelope representatives of the second signal component, and wherein said estimating at least one parameter is further based on a codebook, the codebook comprising a plurality of sets of spectral envelope representatives associated with the second signal component.
8. The method of claim 7, wherein said estimating comprises:
 - determining for each of the plurality of sets of spectral envelope representatives of the codebook an a posteriori probability at least based on the at least one feature; and
 - determining a weighted sum over the plurality of sets of spectral envelope representatives, wherein each set of

spectral envelope representatives of the plurality of sets of spectral envelope representatives is weighted by the respective a posteriori probability.

9. The method of claim 8, wherein determining the at least one a posteriori probability is based on a Hidden Markov Model.
10. The method according to one of claims 1 to 9, wherein said performing a signal enhancement processing on the second signal is based on a combined signal enhancement processing comprising:
 - determining a first set of at least one value configured to be used to perform a signal enhancement processing on the second signal component at least based on the at least one parameter,
 - determining a second set of at least one value configured to be used to perform a signal enhancement processing on the second signal component not based on the at least one parameter, and
 - combining the first set of at least one value and the second set of at least one value.
11. The method according to claim 10, wherein said first set of at least one value represents a first set of at least two gain factors and said second set of at least one value represent a second set of at least two gain factors, wherein each of the at least two gain factors of the first and second set is associated with a sub-band frequency component of the second signal component.

12. The method according to claim 11, wherein the sub-bands associated are associated with a first frequency resolution, the method comprising expanding the combined first set of at least two gain factors and second set of at least two gain factor to a second frequency resolution, the second frequency resolution being higher than the first frequency resolution.
13. The method according to one of claims 10 to 12, wherein said combining is performed based on at least one cross-fading factor, wherein one of the at least one cross-fading factor combines one value of the at least one value of the first set of at least one value with one value of the second set of at least one value.
14. The method according to claim 13, wherein each of the at least one cross-fading factor is calculated on the basis of one value of the first set of at least one value, of one value of the second set and of at one reference value of a set of at least one reference value, the set of at least one reference value representing at least one reference value configured to be used to perform a reference enhanced signal processing on the second signal component.
15. The method according to one of claim 13 and 14, comprising determining said at least one cross-fading vector based on a plurality of reference cross-fading factors and estimations of signal parameters associated with the first and second frequency band, wherein the plurality of reference cross-fading factors has been determined during a training process, wherein each reference

cross-fading factor of the plurality of reference cross-fading factors is associated with at least one signal parameter of the first frequency band and at least one signal parameter of the second frequency band.

16. The method according to one of the preceding claims, wherein said signal enhancement processing represents noise reduction processing.
17. An apparatus comprising:
 - means for estimating at least one parameter based on a first signal component of an audio signal, the at least one parameter being associated with a second signal component of the audio signal; wherein the first signal component is associated with a first frequency band and the second signal component is associated with a second frequency band; and
 - means for performing a signal enhancement processing on the second signal component at least based on the at least one parameter.
18. The apparatus according to claim 17, comprising means for performing a signal enhancement processing on the first signal component in order to obtain an enhanced first signal component, wherein said estimating at least one parameter is performed on basis of the enhanced first signal component.
19. The apparatus according to one of claims 17 to 18, wherein the apparatus is one of:
 - a chip;
 - an integrated circuit; and

- an audio device.

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

- estimating at least one parameter based on a first signal component of an audio signal, the at least one parameter being associated with a second signal component of the audio signal; wherein
- the first signal component is associated with a first frequency band and the second signal component is associated with a second frequency band; and
- performing a signal enhancement processing on the second signal component at least based on the at least one parameter.

21. The apparatus according to claim 20, wherein the at least one memory and the computer program code, with the at least one processor, are configured to cause the apparatus to perform a signal enhancement processing on the first signal component in order to obtain an enhanced first signal component, wherein said estimating at least one parameter is performed on basis of the enhanced first signal component.

22. The apparatus according to one of claims 20 to 21, wherein the apparatus is one of:

- a chip;
- an integrated circuit; and
- an audio device.

23. A computer program code causing an apparatus to perform the following when executed on a processor:
- estimating at least one parameter based on a first signal component of an audio signal, the at least one parameter being associated with a second signal component of the audio signal; wherein
 - the first signal component is associated with a first frequency band and the second signal component is associated with a second frequency band; and
 - performing a signal enhancement processing on the second signal component at least based on the at least one parameter.
24. A computer readable storage medium in which the computer program code according to claim 23 is stored.

1/6

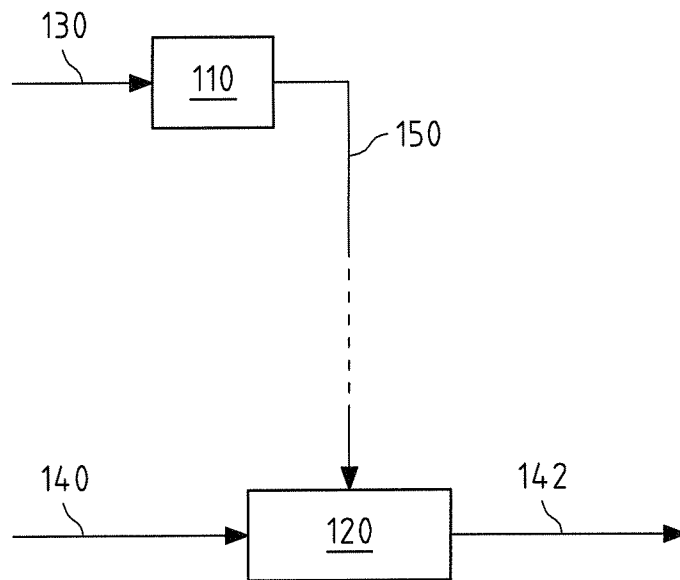


Fig.1a

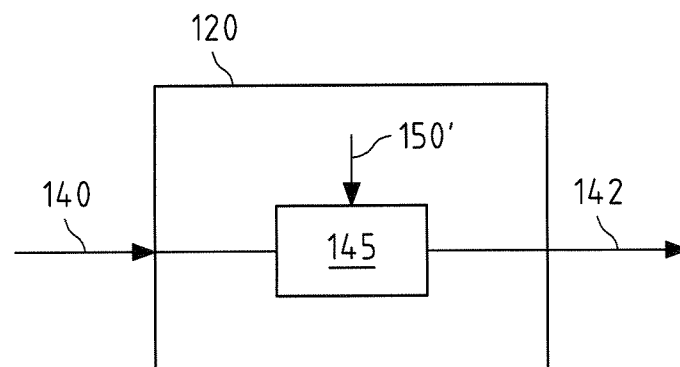
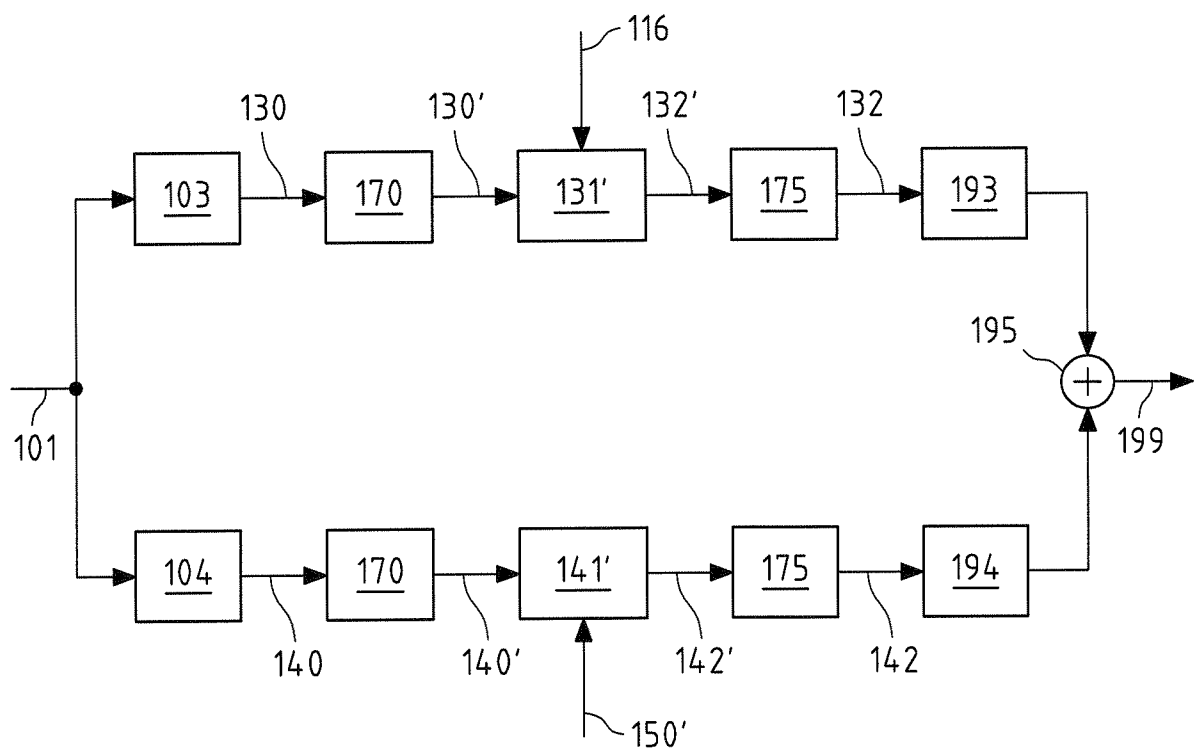
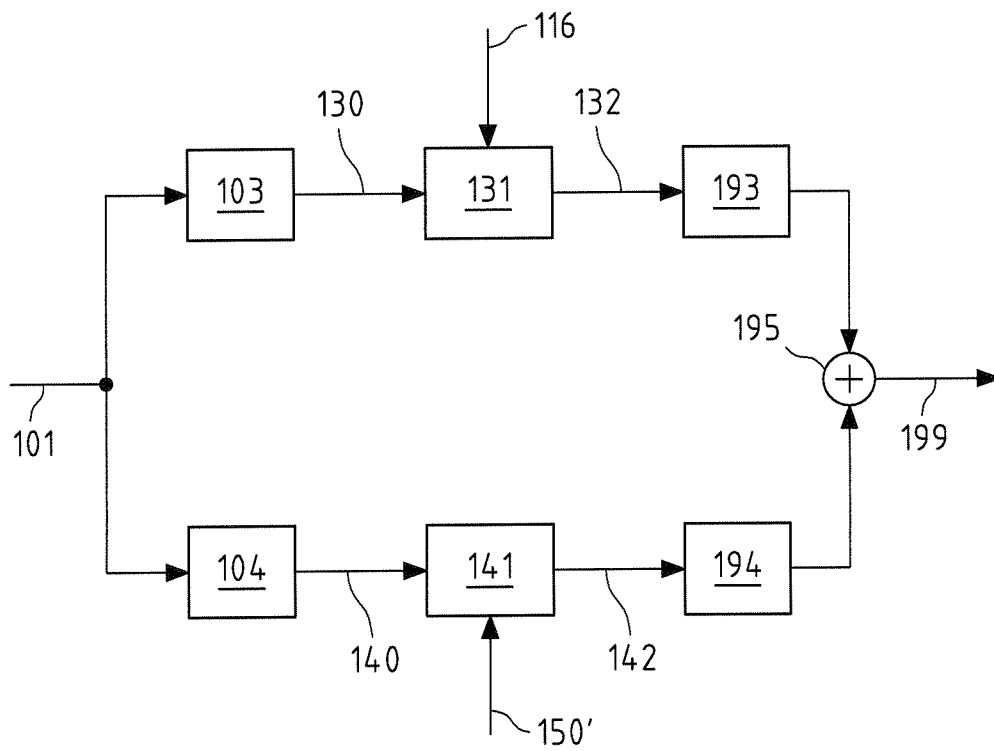


Fig.1b

2/6



3/6

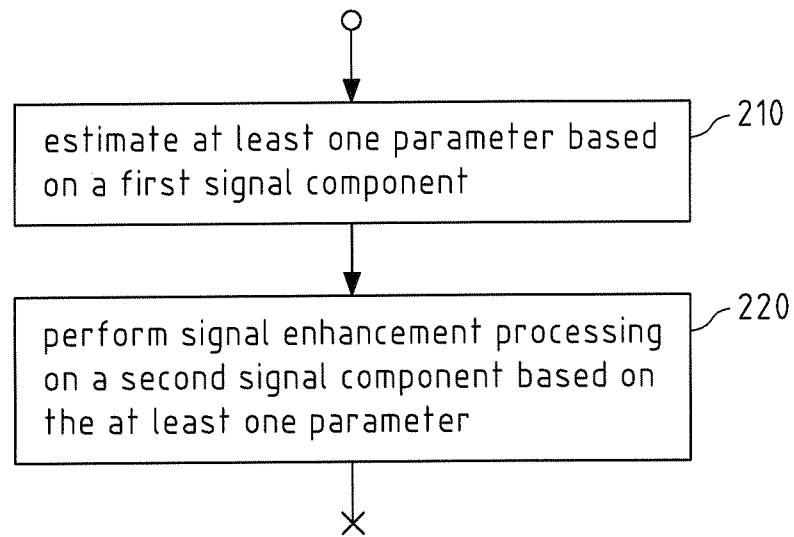


Fig.2a

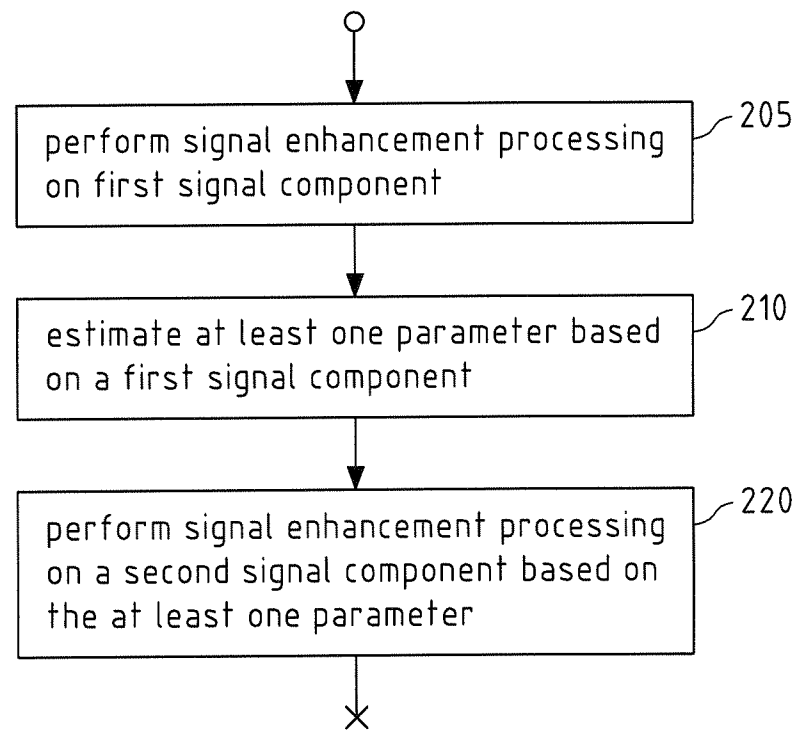


Fig.2b

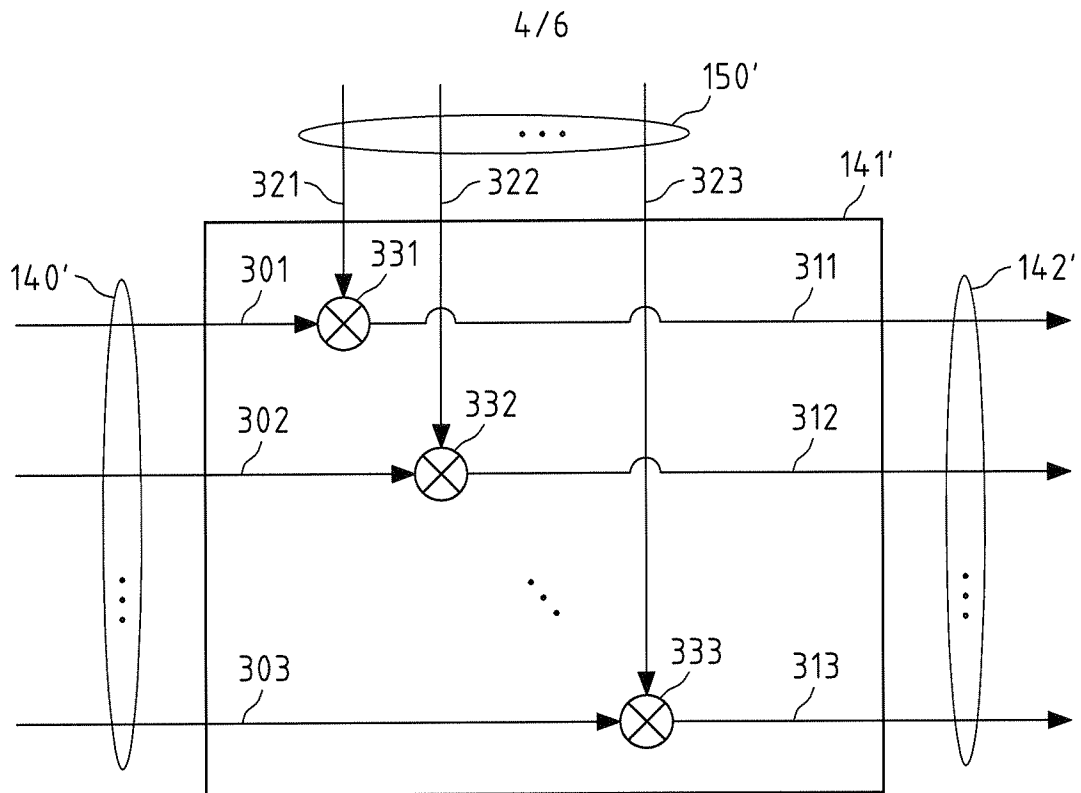


Fig.3a

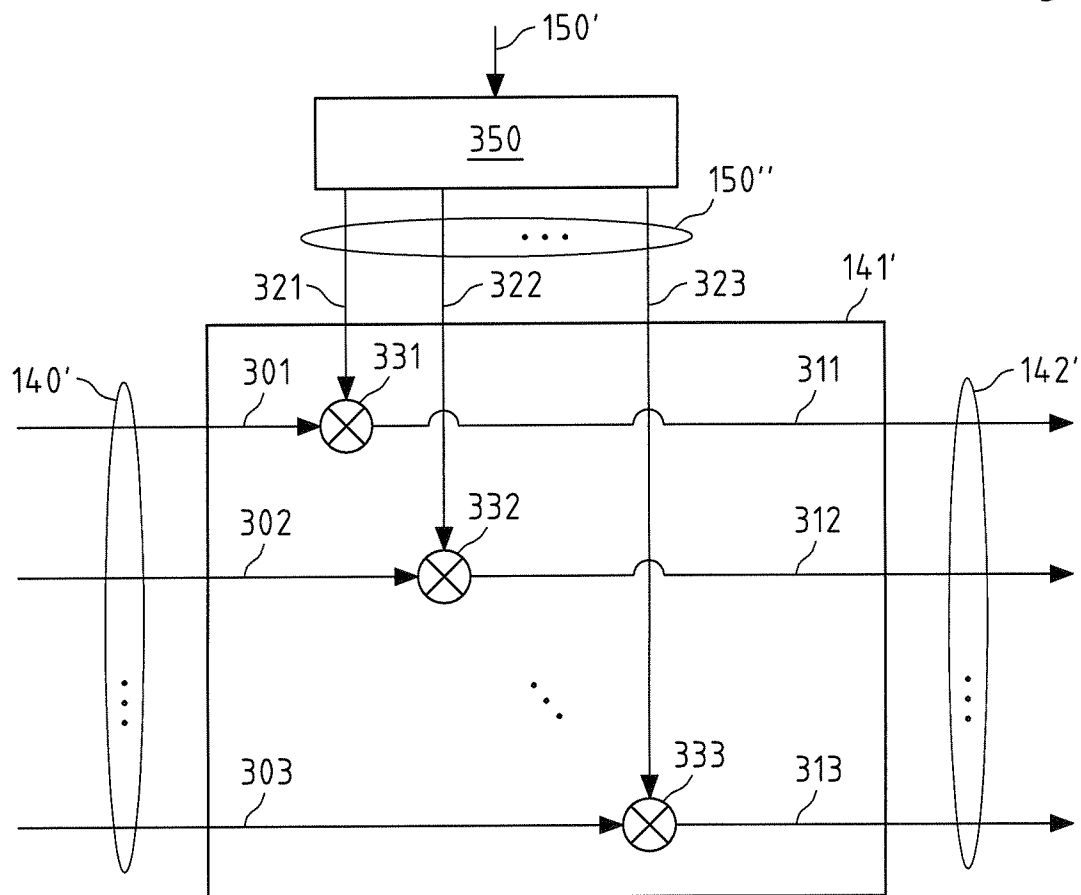


Fig.3b

5/6

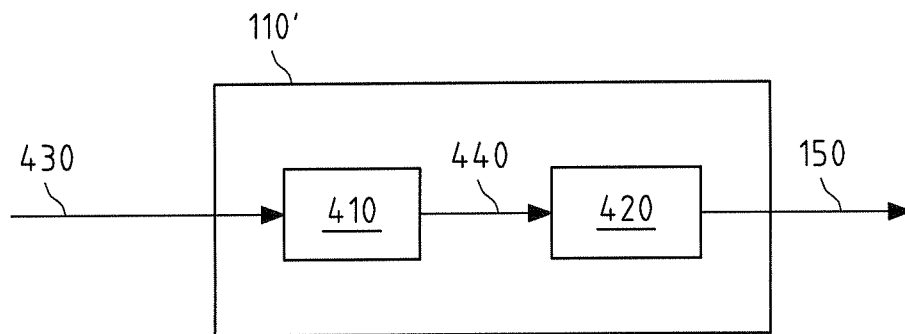


Fig.4a

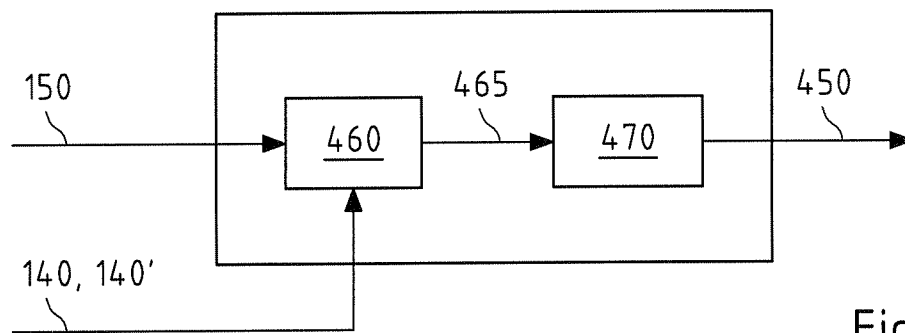


Fig.4b

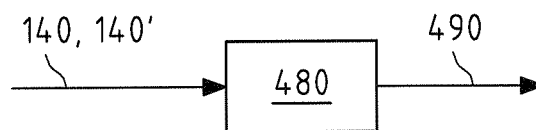


Fig.4c

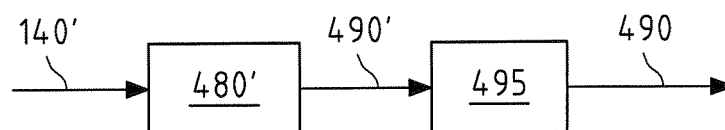


Fig.4d

6/6

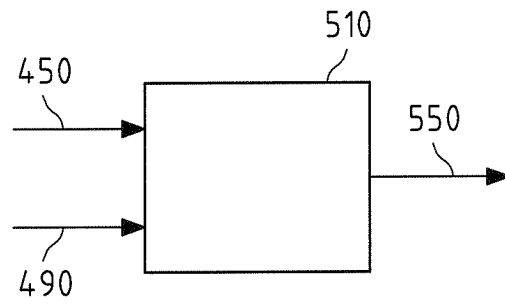


Fig.5a

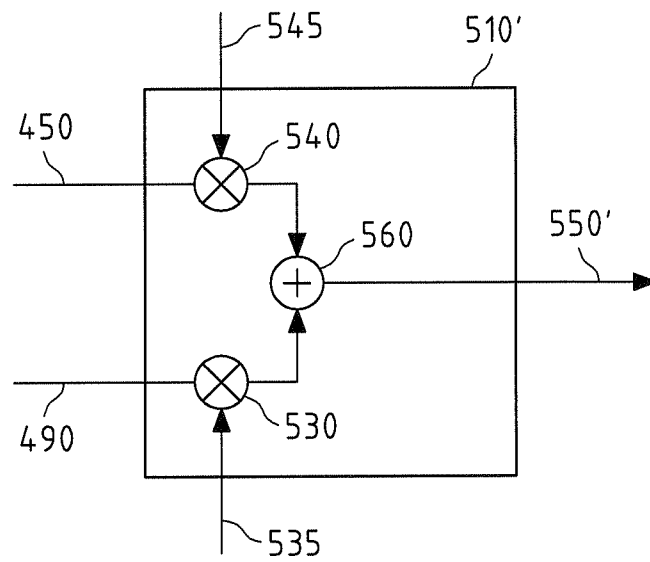


Fig.5b

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2009/061884

A. CLASSIFICATION OF SUBJECT MATTER

INV. G10L21/02
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G10L

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

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

EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2007/088541 A1 (VOS KOEN B [US] ET AL) 19 April 2007 (2007-04-19)	1,16,17, 19,20, 22-24
A	page 2, right-hand column, paragraph 41 - page 3, left-hand column, paragraph 43 page 7, left-hand column, paragraph 78 - paragraph 80 page 9, left-hand column, paragraph 99	2-15,18, 21
A	US 2008/243496 A1 (WANG YOUHUA [JP]) 2 October 2008 (2008-10-02) the whole document	1-24
A	US 2004/024596 A1 (CARNEY LAUREL H [US] ET AL) 5 February 2004 (2004-02-05) the whole document	1-24
	----- -/--	

☒ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

31 March 2010

Date of mailing of the international search report

14/04/2010

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Burchett, Stefanie

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2009/061884

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	Peter Vary: "Speech Enhancement by Conditional Estimation - Noise Reduction, Error Concealment & Bandwidth Extension, what makes the difference?" Proceedings of International Workshop on Acoustic Echo and Noise Control (IWAENC) 30 September 2008 (2008-09-30), XP002576239 Seattle, WA, USA Retrieved from the Internet: URL:http://www.ind.rwth-aachen.de/de/veroeffentlichungen/download/publabel/vary08a/ [retrieved on 2010-03-31] the whole document -----	1-24
T	Thomas Esch, Florian Heese, Bernd Geiser und Peter Vary: "Wideband Noise Suppression Supported by Artificial Bandwidth Extension Techniques" Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) 15 March 2010 (2010-03-15), pages 4790-4793, XP002576240 Dallas, TX, USA Retrieved from the Internet: URL:http://www.ind.rwth-aachen.de/de/veroeffentlichungen/download/publabel/esch10/ [retrieved on 2010-03-31] the whole document -----	

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2009/061884

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2007088541	A1	19-04-2007	NONE
US 2008243496	A1	02-10-2008	NONE
US 2004024596	A1	05-02-2004	NONE