

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
2 February 2006 (02.02.2006)

PCT

(10) International Publication Number
WO 2006/011104 A1

(51) International Patent Classification⁷: **G10L 21/02**,
H04B 3/23

(74) Agents: **SLENDERS, Petrus, J., W.** et al.; Prof. Holst-
laan 6, NL-5656 AA Eindhoven (NL).

(21) International Application Number:
PCT/IB2005/052377

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

(22) International Filing Date: 18 July 2005 (18.07.2005)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
04103509.8 22 July 2004 (22.07.2004) EP

(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, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT,
RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA,
GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(71) Applicant (for all designated States except US): **KONIN-
KLIJKE PHILIPS ELECTRONICS N.V.** [NL/NL];
Groenewoudseweg 1, NL-5621 BA Eindhoven (NL).

(72) Inventors; and

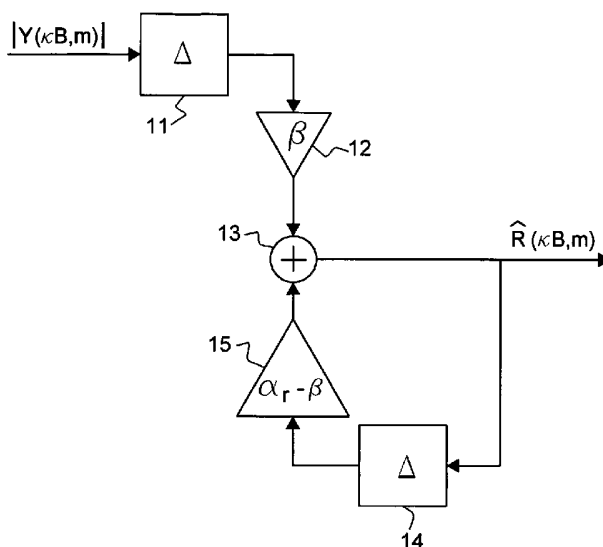
(75) Inventors/Applicants (for US only): **DERKX, Rene, M.**,
M. [NL/NL]; c/o Prof. Holstlaan 6, NL-5656 AA Eindh-
hoven (NL). **JANSE, Cornelis, P.** [NL/NL]; c/o Prof. Hol-
stlaan 6, NL-5656 AA Eindhoven (NL). **BOSCARINO,
Corrado** [IT/NL]; c/o Prof. Holstlaan 6, NL-5656 AA
Eindhoven (NL).

Published:

— with international search report

For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.

(54) Title: AUDIO SIGNAL DEREVERBERATION



(57) Abstract: A method of estimating the reverberations in an acoustic signal (y) comprises the steps of determining the frequency spectrum (Y) of the signal (y), providing a first parameter (α) indicative of the decay of the reverberations part (r) of the signal over time, and providing a second parameter (β) indicative of the amplitude of the direct part (d) of the signal relative to the reverberations part (r). An estimated frequency spectrum (\hat{R}) of the reverberations signal (r) is produced using the frequency spectrum (Y) of a previous frame, the first parameter (α), and the second parameter (β). The second parameter (β) is preferably inversely proportional to the early-to-late ratio of the signal (y).

WO 2006/011104 A1

Audio signal dereverberation

The present invention relates to the dereverberation of audio signals. More in particular, the present invention relates to a method and a device for estimating the reverberations in audio signals, in particular non-stationary audio signals such as speech.

5 It is well known that a signal, such as an acoustic signal, may contain reverberations or echoes from various surfaces. In a room, for example, an acoustic signal (such as speech or music) is reflected by the walls, the ceiling and the floor. A microphone present in the room will therefore receive the acoustic signal as a combination of a direct signal (received directly from the source) and an indirect signal (received via reflecting surfaces). This indirect signal is referred to as the reverberations part of the received signal.

10 Many attempts have been made to separate the desired (that is, direct) signal from its reverberations. The paper "A New Method Based on Spectral Subtraction for Speech Dereverberation" by K. Lebart, J.M. Boucher and P.N. Denbigh, *Acta Acustica*, Vol. 87, pages 359 – 366 (2001), for example, discloses a method for the suppression of late room reverberation from speech signals based on spectral subtraction. In this known method, the
15 frequency spectrum of the reverberations part of the received signal is estimated using the delayed frequency spectrum of the received signal and a (first) parameter that is indicative of the decay of the reverberations part over time. The frequency spectrum of the reverberations part may then be used to estimate, by spectral subtraction, the frequency spectrum of the direct part.

20 This known method works well for signals of which the amplitudes of the direct part and reverberations part are similar or, in other words, of which the energy content of the direct part is (much) smaller than the energy content of the reverberations part. However, when the amplitude (and hence the energy content) of the direct signal is significantly larger than the amplitude (and energy content) of the reverberations signal, the
25 known method introduces errors which result in signal distortion.

It is therefore an object of the present invention to overcome these and other problems of the Prior Art and to provide a method of and a device for estimating the reverberations part of an acoustic signal that takes any difference in energy contents of the direct part and the reverberations part into account.

Accordingly, the present invention provides a method of estimating the reverberations in a signal comprising a direct part and a reverberations part, the method comprising the step of providing an estimate of the frequency spectrum of the reverberations part using a first parameter, a second parameter and the frequency spectrum of the signal, wherein the first parameter is indicative of the decay of the reverberations part over time, and wherein the second parameter is indicative of the amplitude of the direct part relative to the reverberations part of the signal.

By providing a second parameter that is indicative of the amplitude of the direct part relative to the reverberations part, the relative amplitudes of these two signal parts is taken into account. As a result, it is possible to compensate for any difference in amplitudes and thereby to obtain a more accurate estimate.

It will be understood that the frequency spectrum of the direct part, and hence the direct signal part itself, may be obtained from the estimated frequency spectrum of the reverberations part using well-known spectral subtraction techniques.

In a preferred embodiment, the second parameter is inversely proportional to the early-to-late ratio of the signal. The early-to-late ratio is a ratio that indicates the amplitude of the direct (early) part relative to the reverberations (late) part of the signal.

In an advantageous embodiment, the step of providing an estimate of the frequency spectrum of the reverberations part involves using a previous estimate of the frequency spectrum of the reverberations part. In this way, the estimate is determined using a previous value and an update term, the update term preferably comprising the frequency spectrum of the signal.

In a preferred embodiment, therefore, the estimate of the frequency spectrum of the reverberations part is equal to the first parameter times a previous absolute value of the frequency spectrum, minus a third parameter times the difference of the previous absolute value of the frequency spectrum and a previous estimate of the frequency spectrum of the reverberations signal, wherein the third parameter is equal to the first parameter minus the second parameter.

In a practical embodiment, the method of the present invention comprises the further steps of:

- defining frames containing time-limited segments of the signal, and
- determining, for each frame, the frequency spectrum of the signal.

The frames may overlap partially, so that some signal values are used more than once.

Although a single value of the second parameter can be determined for a situation in which the distance between the source and the microphone (and hence the relative amplitude of the direct signal) does not change, it is preferred that the second parameter is determined for each frame separately. In this way, a more accurate
5 determination of the second parameter is made possible, in particular when movement is involved. In accordance with the present invention, the second parameter is preferably determined using only the signal and derivatives thereof.

It is further preferred that for each time segment of the signal, the immediately preceding time segment is used to determine the frequency spectrum of the signal itself and
10 the estimated frequency spectrum of the direct part that are used for the estimation of the frequency spectrum of the reverberations part in the current time segment.

In order to obtain an even better estimate it is preferred that the frequency spectrum of the signal is determined, for each frame, per frequency band, and that the second parameter, the estimated frequency spectrum of the direct signal and the estimated frequency
15 spectrum of the reverberations part are also determined per frequency band. In this way, separate estimates can be made for individual frequency bands. The choice of frequency bands may be dictated by the particular signal.

As stated above, the frequency spectrum of the reverberations part is estimated using the frequency spectrum of a previous frame, the first parameter, the second parameter,
20 and the estimated frequency spectrum of the reverberations part of a previous frame (for the first frame, the estimate of the previous frame may be assumed to have a predetermined value, for example zero). The method may comprise the further step of dereverberating the signal using the estimate of the frequency spectrum of the reverberations part, preferably using spectral subtraction and reconstruction of the dereverberated signal on the basis of the
25 spectrum resulting from the subtraction.

The present invention also provides a computer program product for carrying out the method defined above. The computer program product may comprise a carrier, such as a CD, DVD, or a floppy disc, on which a computer program is stored in electronic or optical form. The computer program specifies the method steps to be carried out by a general
30 purpose computer or a special purpose computer.

The present invention additionally provides a device for carrying out the method as defined above, as well as an audio system, such as a speech recognition system comprising such a device.

The present invention will further be explained below with reference to exemplary embodiments illustrated in the accompanying drawings, in which:

Fig. 1 schematically shows a typical example of the impulse response of a room.

Fig. 2 schematically shows the energy decay curve of the room impulse response of Fig. 1.

Fig. 3 schematically shows a filter arrangement for estimating reverberations according to the present invention.

Fig. 4 schematically shows a device for estimating reverberations according to the present invention.

Fig. 5 schematically shows a system which incorporates the inventive device of Fig. 4.

15

The exemplary signal shown in Fig. 1 represents the impulse response of a room, that is, the signal received by a microphone in a room when an acoustic impulse (a pulse of an extremely short duration) is produced in the same room. This impulse response contains a first (signal delay) section I which contains no significant signal components, a second section II containing a direct signal part d and a third section III containing an indirect of reverberations signal part r . The direct signal part d and the reverberations signal part r together constitute the signal y , the amplitude y_i of which is plotted as a function of the sample index i . It will be understood that the sample index i of a sampled (digitized) signal corresponds with the time in an analog signal.

The direct signal part d is received via the direct signal path, that is, without being reflected, while the indirect or reverberations part r is received via walls and other reflecting surfaces. As can be seen, the amplitude y_i of the reverberations part r decays exponentially. This is also illustrated in Fig. 2 which represents the Energy Decay Curve EDC of the signal y as a function of the sample index i . The EDC is defined as:

30

$$EDC(i) = 10 \log \sum_{m=i}^{\infty} y_m^2 \quad (1)$$

Expressed in words, the energy decay curve EDC yields, for a given sample index i , ten times the logarithm of the energy of the remainder of the signal. It can be seen that in the third section III containing the EDC of the reverberations part r , the EDC is approximately linear (it is noted that the values of the EDC shown are negative as all values of y_i in the present example are much smaller than one). As the EDC is logarithmic, this linear decay of the EDC represents an exponential decay of the signal y . The slope of the EDC in the section III may be represented by a parameter α_r , where the subscript r refers to the reverberations.

It can further be seen in Fig. 2 that a step occurs in the EDC between the regions I and III. This step represents the energy of the direct signal part d relative to the reverberations part r (Fig. 1). In some situations, this step is relatively small, the reverberations part r having an amplitude similar to (or larger than) the direct part d . In the example shown, however, the step in the EDC is relatively large as the direct part d of Fig. 1 is significantly larger than the reverberations part r . Ignoring this step may lead to estimation errors when estimating the (frequency spectrum of the) reverberations part r . The present invention, however, takes this step into account using a further parameter β and thereby produces much improved results.

The signal y of Fig. 1 can be dereverberated using a known technique, such as a single channel noise suppression technique, replacing the noise frequency spectrum with an estimate of the reverberations frequency spectrum. In the case of spectral subtraction this involves determining the frequency spectrum Y of the signal, estimating the frequency spectrum R of the reverberations, subtracting this estimated reverberations frequency spectrum from the signal frequency spectrum so as to obtain an estimated frequency spectrum of the direct signal part, and reconstructing the direct signal on the basis of its estimated frequency spectrum.

According to the Prior Art mentioned above this may be accomplished as follows. The signal y is divided into frames that each contain a number of samples of the digitized signal. Each frame may, for example, contain 128 or 256 samples. Zeroes may be added (so-called "zero padding") to arrive at a suitable number of samples per frame. The frames typically but not necessarily overlap partially, the term "block" is used to refer to the "new" samples of each frame.

Advantageously a window, for example a Hamming window known *per se*, is applied to alleviate the introduction of artifacts. For each frame the frequency spectrum Y is

determined using the well-known Fast Fourier Transform (FFT). For each frame κ an estimate \hat{R} of the reverberations frequency spectrum R is determined by:

$$\hat{R}(\kappa B, m) = \alpha_r \cdot |Y((\kappa - 1)B, m)| \quad (2)$$

5

where B is the block size (that is, the number of new samples in each frame), m is the frequency, the vertical bars indicate the absolute value, and α_r is the parameter indicating the decay speed of the reverberations part r . Mathematically, α_r may be defined as

$$\alpha_r = \left(e^{-\frac{2\Delta B}{F_s}} \right)^{\frac{1}{2}} \quad (3)$$

10

where F_s is the sampling frequency and

$$\Delta = \frac{3 \ln 10}{T_{60}} \quad (4)$$

15

In being the natural logarithm ($3 \ln 10$ is approximately equal to 6.9) and T_{60} being the reverberation time, that is, the length of time after which the signal level has dropped 60 dB (deciBel) relative to the initial signal level.

The estimate \hat{R} of the reverberations frequency spectrum may be used to determine a gain function:

20

$$G(\kappa B, m) = \max \left(\frac{|Y(\kappa B, m)| - \hat{R}(\kappa B, m)}{|Y(\kappa B, m)|}, \lambda \right) \quad (5)$$

where λ is the so-called spectral floor, a value which ensures that any severe distortion of the dereverberated signal is avoided. A typical value of λ is 0.1, although other values may also be used. The frequency spectrum Y of the original signal is multiplied by this gain factor $G(\kappa B, m)$ to yield the frequency spectrum D of the direct (dereverberated) signal d .

25

Although this known method is very effective, it leads to signal distortions when the direct part *d* of the signal *y* has a large amplitude (energy) relative to the reverberations part *r* or, in other words, when the signal shows a large step in region II of Fig. 2. This amplitude of the direct part relative to the reverberations part may be expressed using the Early-to-Late Ratio ELR, which may be written as:

$$ELR(k) = 10 \log\left(\sum_{i=0}^k y_i^2 / \sum_{i=k}^{\infty} y_i^2\right) \quad (6)$$

where *k* is the sample number separating regions II and III in Fig. 2, that is, the point where the drop in the EDC curve ends and the relatively straight part begins. The ELR therefore determines the (logarithm of the) energy ratio of the direct part (up to sample *k*) and the reverberations part (from sample *k*). The ELR is sometimes also referred to as Clarity Index.

When the Early-to-Late Ratio ELR is small (for example smaller than 0 dB, using the definition above), the energy content of the direct part *d* is small compared to the reverberations part *r* of the signal *y*, and the Prior Art method discussed above can effectively be used to dereverberate the signal *y*. However, when the ELR is large (for example larger than 0 dB, or larger than 5 dB, using the definition above), the known method introduces distortions as the step in region II of the EDC (Fig. 2) is ignored.

Accordingly, the present invention uses an improved estimation method in which the relative energy contents of the direct signal part *d* and the reverberations signal part *r* are taken into account.

Starting from equation (2) above, the present invention proposes to correct the estimation of the frequency spectrum *R* of the reverberations part *r* by subtracting a correction term $\gamma \cdot C(\kappa)$, where γ is a factor dependent on the ELR, and where *C* is a function of the frame number κ (that is, time), and possibly also of the block size *B* and the frequency *m*:

$$\hat{R}(\kappa B, m) = \alpha_r \cdot |Y((\kappa - 1)B, m)| - \gamma \cdot C((\kappa - 1)B, m) \quad (7)$$

The present invention further proposes to use the estimate \hat{D} of the direct part frequency spectrum *D* as the function *C*:

$$\hat{R}(\kappa B, m) = \alpha_r \cdot |Y((\kappa - 1)B, m)| - \gamma \cdot \hat{D}((\kappa - 1)B, m) \quad (8)$$

It will be understood that there may be other functions C that have the required properties.

5 Since the estimate \hat{D} of the direct part frequency spectrum D can be expressed as:

$$\hat{D}(\kappa B, m) = |Y(\kappa B, m)| - \hat{R}(\kappa B, m) \quad (9)$$

10 equation (8) may be rewritten as:

$$\hat{R}(\kappa B, m) = (\alpha_r - \gamma) \cdot |Y((\kappa - 1)B, m)| + \gamma \cdot \hat{R}((\kappa - 1)B, m) \quad (10)$$

Introducing a parameter $\beta(\kappa B) = \alpha_r - \gamma$, with $0 \leq \beta(\kappa B) \leq \alpha_r$, equation (10)
15 can be written as:

$$\hat{R}(\kappa B, m) = \beta(\kappa B) \cdot |Y((\kappa - 1)B, m)| + (\alpha_r - \beta(\kappa B)) \cdot \hat{R}((\kappa - 1)B, m) \quad (11)$$

Using equation (9), equation (11) may also be expressed as:

20

$$\hat{R}(\kappa B, m) = \beta(\kappa B) \cdot \hat{D}((\kappa - 1)B, m) + \alpha_r \cdot \hat{R}((\kappa - 1)B, m) \quad (12)$$

It can be shown that:

$$25 \quad \beta(\kappa B) = \frac{(1 - \alpha_r)}{ELR(\kappa B)} \quad (13)$$

in other words, the (second) parameter $\beta(\kappa B)$ is inversely proportional to the Early-to-Late Ratio ELR. It is further noted that both β and ELR are functions of time (that is, the frame index κ times the block index B), and that β (and ELR) may also depend on the frequency
30 (or sub-band) m: $\beta(\kappa B, m)$.

It can be seen from equation (11) that the estimate $\hat{R}(\kappa B, m)$ according to the present invention combines both the (absolute value of the) frequency spectrum Y of the previous frame and the previous estimate, while taking the ELR into account. It can further be seen from equations (11) and (12) that for a large ELR, $\beta(\kappa B)$ is small and the estimate

5 $\hat{R}(\kappa B, m)$ is effectively based on the previous estimate, suppressing the influence of the spectrum Y . For a small ELR, $\beta(\kappa B)$ is "large" and approximately equal to α_r , and $\hat{R}(\kappa B, m)$ is then approximately equal to $\alpha_r \cdot |Y((\kappa - 1)B, m)|$, as in the Prior Art.

It can therefore be seen that the method of the present invention is consistent with the Prior Art in case the Early-to-Late Ratio ELR is small, while providing an important

10 improvement when ELR is large.

It is noted that the method of the present invention may be carried out per sub-band, that is per frequency m , or independent of the frequency, using a single term for all frequencies.

The method of the present invention may be implemented in software or in

15 hardware. An exemplary hardware implementation is shown in Fig. 3, where a filter section 10 is shown. The filter section 10 comprises a first delay element 11, a first amplifier 12, a combination element 13, a second delay element 14 and a second amplifier 15.

In the example shown, the first delay element 11 receives the absolute value (that is, the magnitude) $|Y(\kappa B, m)|$ of the frequency spectrum Y and outputs the delayed

20 absolute value $|Y((\kappa - 1)B, m)|$. In the preferred embodiment, the delay Δ is equal to one frame. In the amplifier 12, this delayed absolute value is multiplied by the (second) parameter β and fed to the combination element 13 which is preferably constituted by an adder.

The combination element 13 also receives the output signal of the second multiplier 15 and outputs the estimate $\hat{R}(\kappa B, m)$. This estimate is received by the second

25 delay element 14 which outputs the delayed estimate $\hat{R}((\kappa - 1)B, m)$ to the second amplifier 15. This delayed estimate $\hat{R}((\kappa - 1)B, m)$ is multiplied with the factor $(\alpha_r - \beta)$ in the amplifier 15 and fed to the combination element 13. As can be seen, the filter section 10 produces the same result as equation (11) above.

The parameter β (or $\beta(\kappa B)$) may be predetermined. For example, a fixed

30 value of 0.1 or 0.2 could be used for a situation, provided an estimate of the ELR for that situation is known. It is preferred, however, to estimate $\beta(\kappa B)$ for each signal. Of course

$\beta(\kappa B)$ can be estimated on the basis of the Early-to-Late Ratio, using formula (13) above. However, starting from an initial value of $\beta(\kappa B)$ (which may be a predetermined value, for example zero), an update may be provided using the formula:

$$5 \quad \beta(\kappa B) = \beta((\kappa - 1)B) + f(|Y(\kappa B, m)|, \hat{R}(\kappa B, m), \lambda) \quad (14)$$

where the function $f()$ is an update function and the parameter λ is the spectral floor mentioned above. An example of an update function which uses the absolute value $|Y(\kappa B, m)|$ and the estimated spectrum $\hat{R}(\kappa B, m)$ is:

$$10 \quad \beta(\kappa B) = \beta((\kappa - 1)B) + \mu \frac{\sum_{m=0}^{M-1} (1 - \lambda) |Y(\kappa B, m)| - \hat{R}(\kappa B, m)}{\sum_{m=0}^{M-1} (1 - \lambda) |Y(\kappa B, m)| + \varepsilon} \quad (15)$$

where ε is an auxiliary parameter having a (very small) value to prevent division by zero, and where μ is a non-negative parameter controlling the speed and accuracy of the update of $\beta(\kappa B)$.

A preferred embodiment of a device for dereverberating a signal is schematically shown in Fig. 4. The exemplary device 1 shown in Fig. 4 comprises a filter unit 10, an FFT unit 20, a spectrum decomposition unit 30, a parameter estimation unit 40, a gain determination unit 50, a delay unit 60, a multiplication unit 70, a spectrum reconstruction unit 80 and an inverse FFT unit 90.

The FFT unit 20 receives the (digital) signal $y(k)$ and performs a well-known Fast Fourier Transform on a frame of signal samples. It will be understood that a A/D (analog / digital) converter may be present if the original signal is analog. The (complex) spectrum $Y(m)$ produced by the FFT unit 20 is fed to the decomposition unit 30 which decomposes the complex spectrum into a phase part ϕ and a magnitude part ρ . This magnitude ρ is equal to the absolute value $|Y(\kappa B, m)|$, where κ is the frame index (frame number), B is the block size and m is the frequency, as before. The phase part ϕ is fed directly to the spectrum reconstruction unit 80, while the magnitude part ρ is fed to the filter unit 10, the parameter unit 40, the gain unit 50 and the multiplication unit 70.

The filter unit 10, which may be identical to the filter unit 10 of Fig. 3, also receives the parameters α , and $\beta(\kappa B)$ to produce the estimated reverberations spectrum $\hat{R}(\kappa B, m)$ which is fed to the units 40 and 50. The unit 40 produces an updated value of $\beta(\kappa B)$, for example in accordance with equation 15, using the estimate $\hat{R}(\kappa B, m)$, the parameter λ , and the delayed value $\beta((\kappa - 1)B)$ output by delay (Δ) unit 60.

The unit 50 produces a gain factor $G(\kappa B, m)$, for example in accordance with equation (5). This gain factor is fed to multiplier 70 where it is multiplied with the absolute value spectrum $|Y(\kappa B, m)|$ to produce the dereverberated spectral magnitude ρ' . The reconstruction unit 80 reconstructs the dereverberated spectrum $|\hat{D}(\kappa B, m)|$ on the basis of ϕ and ρ' . This spectrum $|\hat{D}(\kappa B, m)|$ is then converted into the dereverberated time signal $\hat{d}(k)$ by the IFFT (Inverse Fast Fourier Transform) unit 90.

An audio system incorporating the device 1 of Fig. 4 is schematically shown in Fig. 5. The system 9 comprises a dereverberation device 1, a signal processing device 2, a microphone 3 and a loudspeaker 4. The dereverberation device 1 is preferably the device 1 of Fig. 4 or its equivalent. The signal processing device 2 may comprise an amplifier and/or any other suitable audio signal processing means. The microphone 3 may be substituted with a set of microphones. Similarly, the loudspeaker 4 may be substituted with a set of loudspeakers or other suitable transducers. The order of the dereverberation device 1 and the signal processing device 2 may be reversed, and other devices (not shown) may be added. The system 9 may, for example, constitute a conference system, a hands-free telephony system, or a speech recognition system.

The present invention is based upon the insight that the energy content ratio of the direct signal part and the reverberations signal part has to be taken into account when dereverberating a signal. By introducing a parameter related to this energy content ratio, a better dereverberation which introduces less signal distortion is achieved.

The present invention could be summarized as a method of estimating the reverberations in a signal comprising a direct part and a reverberations part, the method comprising the steps of estimating the spectrum of the reverberations part using a first parameter and the spectrum of the signal, and correcting the estimated spectrum using correction term involving a second parameter, wherein the second parameter is indicative of the amplitude of the direct part relative to the reverberations part of the signal.

It is noted that any terms used in this document should not be construed so as to limit the scope of the present invention. In particular, the words “comprise(s)” and “comprising” are not meant to exclude any elements not specifically stated. Single (circuit) elements may be substituted with multiple (circuit) elements or with their equivalents.

5 The term computer program product should be understood to include any physical realization, e.g. an article of manufacture, of a collection of commands enabling a processor -generic or special purpose-, after a series of loading steps to get the commands into the processor, to execute any of the characteristic functions of an invention. In particular the computer program product may be realized as program code, processor adapted code
10 derived from this program code, or any intermediate translation of this program code, on a carrier such as e.g. a disk or other plug-in component, present in a memory, temporarily present on a network connection -wired or wireless- , or program code on paper. Apart from program code, invention characteristic data required for the program may also be embodied as a computer program product.

15 It will be understood by those skilled in the art that the present invention is not limited to the embodiments illustrated above and that many modifications and additions may be made without departing from the scope of the invention as defined in the appending claims.

CLAIMS:

1. A method of estimating the reverberations in a signal (y) comprising a direct part (d) and a reverberations part (r), the method comprising the step of providing an estimate (\hat{R}) of the frequency spectrum (R) of the reverberations part (r) using a first parameter (α), a second parameter (β) and the frequency spectrum (Y) of the signal (y),
5 wherein the first parameter (α) is indicative of the decay of the reverberations part (r) over time, and wherein the second parameter (β) is indicative of the amplitude of the direct part (d) relative to the reverberations part (r) of the signal (y).
2. The method according to claim 1, wherein the second parameter (β) is
10 inversely proportional to the early-to-late ratio (ELR) of the signal (y).
3. The method according to claim 1, wherein the step of providing an estimate (\hat{R}) of the frequency spectrum (R) of the reverberations part (r) involves using a previous estimate (\hat{R}) of the frequency spectrum (R) of the reverberations part (r).
15
4. The method according to claim 1, wherein the estimate (\hat{R}) of the frequency spectrum (R) of the reverberations part (r) is equal to the first parameter (α) times a previous absolute value of the frequency spectrum (Y), minus a third parameter (γ) times the difference of the previous absolute value of the frequency spectrum (Y) and a previous
20 estimate (\hat{R}) of the frequency spectrum (R) of the reverberations signal (r), wherein the third parameter (γ) is equal to the first parameter (α) minus the second parameter (β).
5. The method according to claim 1, comprising the further steps of:
- defining frames (κ) containing time-limited segments of the signal (y), and
25 - determining, for each frame (κ), the frequency spectrum (Y) of the signal (y).
6. The method according to claim 5, wherein the frames overlap partially.

7. The method according to claim 5, comprising the further step of:
- providing, for each frame (κ), an estimated frequency spectrum (\hat{R}) of the reverberations part (r) using the frequency spectrum (Y) of a previous frame, the first parameter (α), the second parameter (β), and the estimated frequency spectrum (\hat{R}) of the reverberations part (r) of a previous frame.
8. The method according to claim 7, wherein the said previous frame is the immediately preceding frame ($\kappa-1$).
9. The method according to claim 5, wherein the second parameter (β) is determined for each frame (κ) separately, preferably using only the signal (y) and derivatives thereof.
10. The method according to claim 1, further comprising the step of dereverberating the signal (y) using the estimate (\hat{R}) of the frequency spectrum (R) of the reverberations part (r).
11. A computer program product for carrying out the method according to any of claims 1 - 10.
12. A device (1) for estimating the reverberations in a signal (y) comprising a direct part (d) and a reverberations part (r), the device comprising means for providing an estimate (\hat{R}) of the frequency spectrum (R) of the reverberations part (r) using a first parameter (α), a second parameter (β) and the frequency spectrum (Y) of the signal (y),
wherein the first parameter (α) is indicative of the decay of the reverberations part (r) over time, and wherein the second parameter (β) is indicative of the amplitude of the direct part (d) relative to the reverberations part (r) of the signal (y).
13. The device according to claim 12, wherein the second parameter (β) is inversely proportional to the early-to-late ratio (ELR) of the signal (y).

14. The device according to claim 12, wherein the means (10) for providing an estimate (\hat{R}) of the frequency spectrum (R) of the reverberations part (r) are arranged for using a previous estimate (\hat{R}) of the frequency spectrum (R) of the reverberations part (r).
- 5 15. The device according to claim 12, wherein the estimate (\hat{R}) of the frequency spectrum (R) of the reverberations part (r) is equal to the first parameter (α) times a previous absolute value of the frequency spectrum (Y), minus a third parameter (γ) times the difference of the previous absolute value of the frequency spectrum (Y) and a previous estimate (\hat{R}) of the frequency spectrum (R) of the reverberations signal (r), wherein the third
10 parameter (γ) is equal to the first parameter (α) minus the second parameter (β).
16. The device according to claim 12, further comprising means (20) for:
- defining frames (κ) containing time-limited segments of the signal (y), and
- determining, for each frame (κ), the frequency spectrum (Y) of the signal (y).
- 15
17. The device according to claim 16, wherein the frames overlap partially.
18. The device according to claim 16, further comprising means (10) for:
- providing, for each frame (κ), an estimated frequency spectrum (\hat{R}) of the
20 reverberations part (r) using the frequency spectrum (Y) of a previous frame, the first parameter (α), the second parameter (β), and the estimated frequency spectrum (\hat{R}) of the reverberations part (r) of a previous frame.
19. The device according to claim 18, wherein said previous frame is the
25 immediately preceding frame ($\kappa-1$).
20. The device according to claim 12, further comprising means (40, 60) for determining the second parameter (β).
- 30 21. The device according to claim 20, wherein the means (40, 60) for providing the second parameter (β) are arranged for determining the second parameter (β) for each frame (κ) separately, preferably using only the signal (y) and derivatives thereof.

22. The device according to claim 12, further comprising means (50) for determining a gain factor $G(\kappa B, m)$ using the frequency spectrum (Y) of a previous frame and the estimated frequency spectrum (\hat{R}) of the reverberations part (r).
- 5 23. The device according to claim 12, further comprising means (70) for multiplying the gain factor $G(\kappa B, m)$ and the frequency spectrum (Y) so as to produce an estimate (\hat{D}) of the spectrum of the direct signal (d).
24. The device according to claim 12, further comprising decomposition means
10 (30) for separating the phase (ϕ) and the magnitude (ρ) of a frequency spectrum, and/or reconstruction means (80) for reconstructing a frequency spectrum using a phase (ϕ) and a magnitude (ρ).
25. The device according to claim 12, further comprising means (90) for
15 determining the direct signal (d) using the estimate (\hat{D}) of its frequency spectrum.
26. An audio processing system (9), comprising a device (1) according to any of claims 12 - 25.
- 20 27. The audio processing system (9) of claim 26 which is a speech recognition system, a conference system, or a hands-free telephony system.

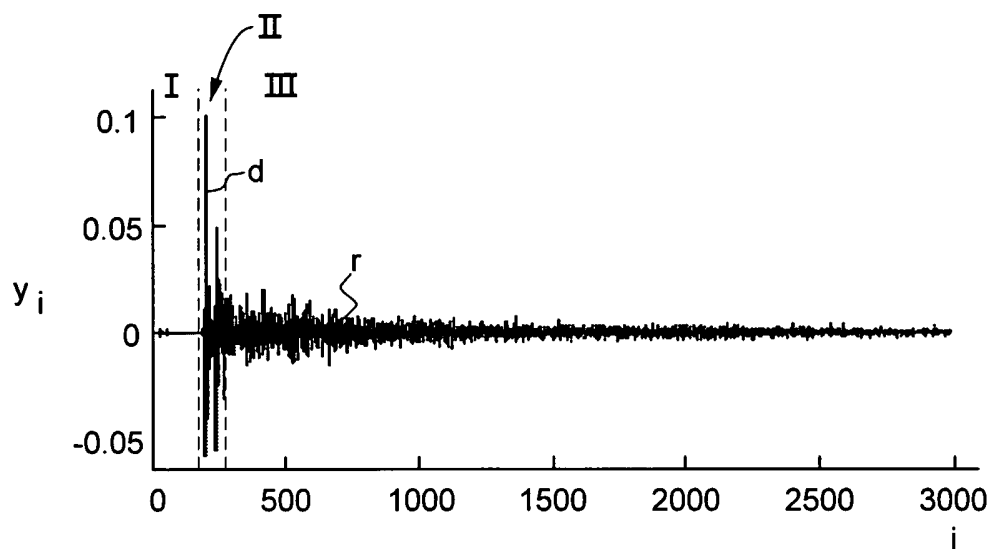


FIG. 1

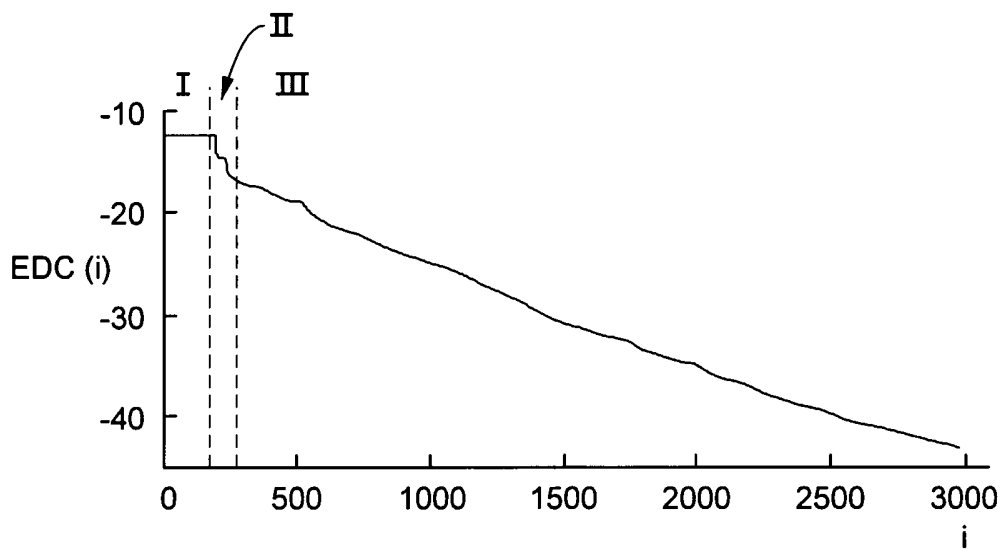


FIG. 2

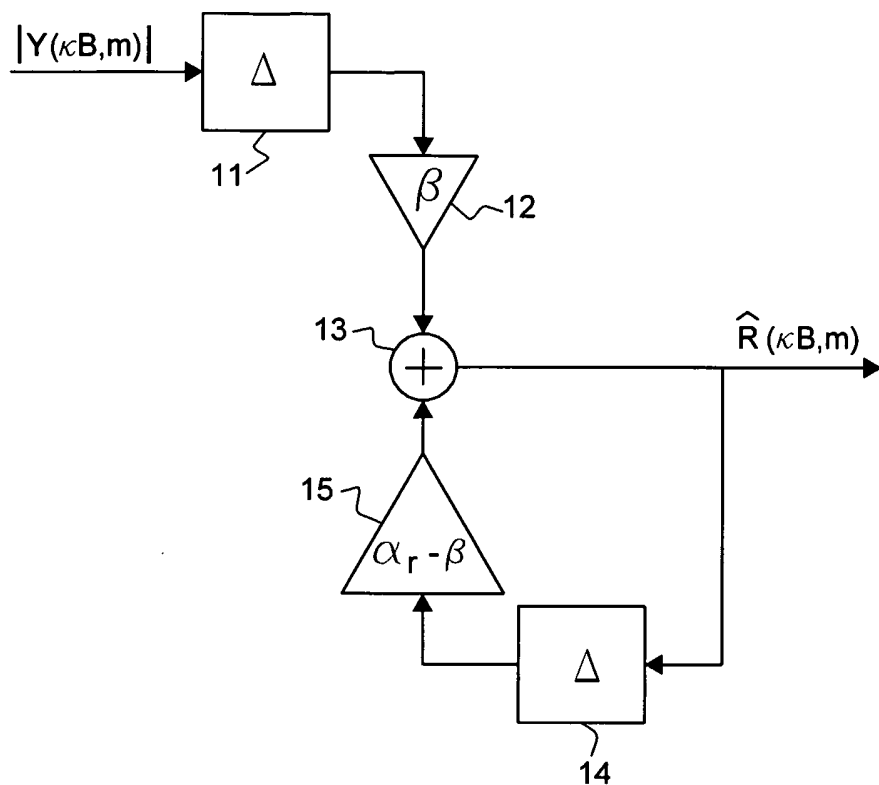


FIG. 3

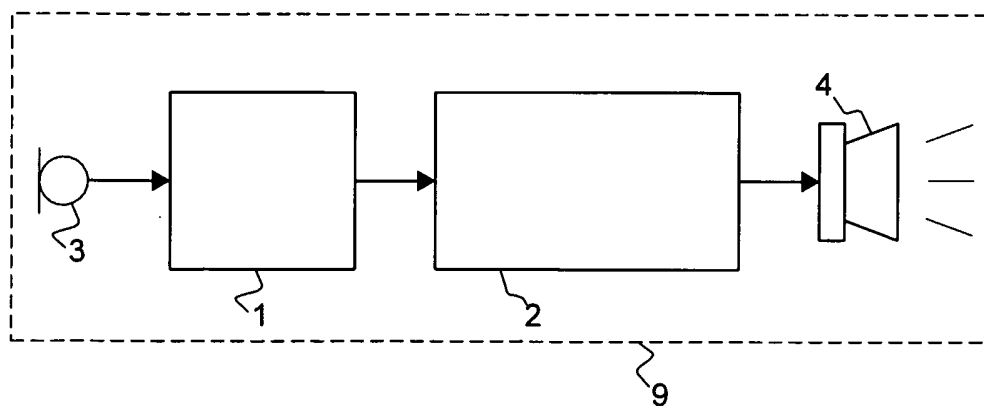


FIG. 5

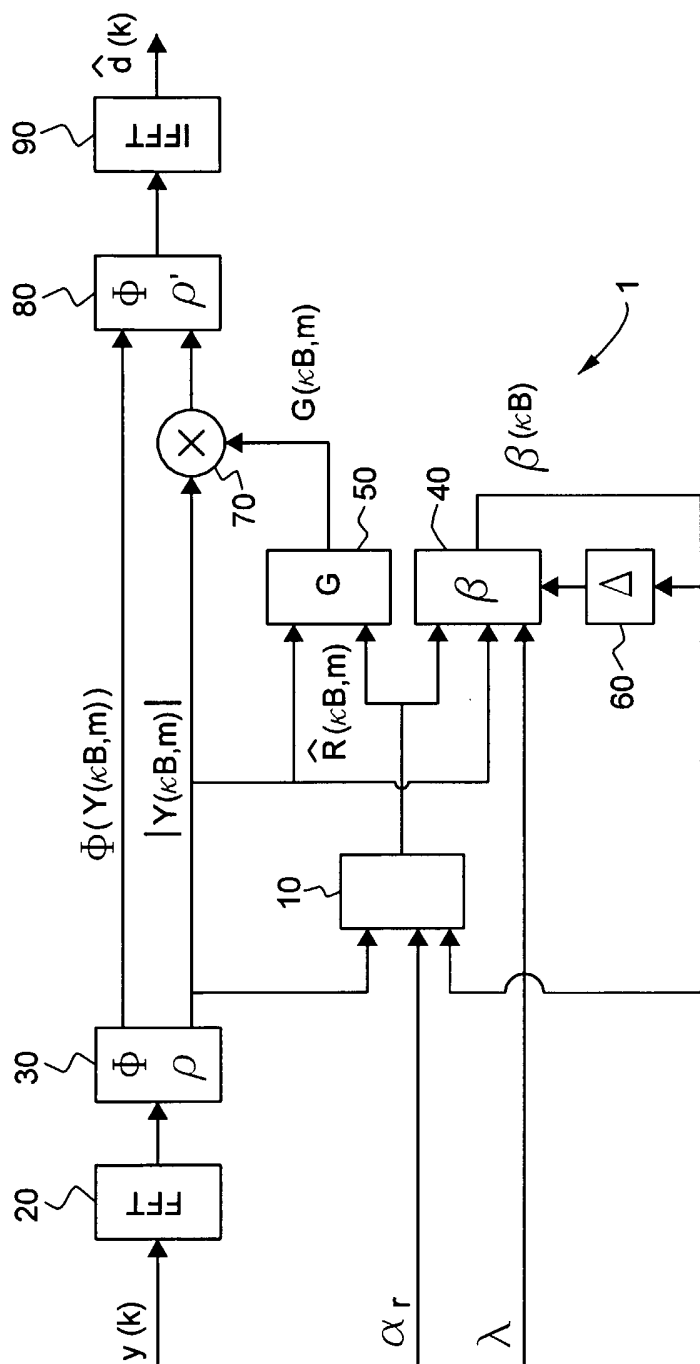


FIG. 4

INTERNATIONAL SEARCH REPORT

IB/05/52377

A. CLASSIFICATION OF SUBJECT MATTER
 IPC 7 G10L21/02 H04B3/23

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
 IPC 7 G10L H04B

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, INSPEC, WPI Data, PAJ, IBM-TDB

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	LEBART K ET AL: "A new method based on spectral subtraction for speech dereverberation" ACUSTICA - ACTA ACUSTICA S. HIRZEL VERLAG GERMANY, vol. 87, no. 3, May 2001 (2001-05), pages 359-366, XP009053193 ISSN: 0001-7884 cited in the application	1-3, 5-14, 16-27
Y	page 361, left-hand column, line 16 - page 362, right-hand column, line 12; figures 1,2	1,2,5, 10-13, 16,20, 22,23, 25-27
	----- -/--	

Further documents are listed in the continuation of box C.

Patent family members are listed in 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

5 September 2005

Date of mailing of the international search report

10/10/2005

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
 NL - 2280 HV Rijswijk
 Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
 Fax: (+31-70) 340-3016

Authorized officer

Dobler, E

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	<p>US 6 622 030 B1 (ROMESBURG ERIC DOUGLAS ET AL) 16 September 2003 (2003-09-16)</p> <p>column 5, line 35 - column 8, line 42; figure 3 column 10, line 19 - line 26</p> <p>-----</p>	<p>1,2,5, 10-13, 16,20, 22,23, 25-27</p>
P,X	<p>E.A.P. HABETS: "Single-Channel Speech Dereverberation based on Spectral Subtraction"</p> <p>PROCEEDINGS OF PRORISC 2004, 'Online! 25 November 2004 (2004-11-25), pages 250-254, XP002343438 Veldhoven, The Netherlands Retrieved from the Internet: URL:http://www.stw.nl/prorisc/proc-2004/signal_and_image_processing/habets.pdf 'retrieved on 2005-09-01! Sections II, III and IV on pages 250 to 252</p> <p>-----</p>	<p>1-3, 5-14, 16-27</p>
A	<p>MARTIN R: "SPECTRAL SUBTRACTION BASED ON MINIMUM STATISTICS"</p> <p>SIGNAL PROCESSING: THEORIES AND APPLICATIONS, vol. 2, 13 September 1994 (1994-09-13), pages 1182-1185, XP009026531 Sections 2.1, 2.2, 2.3, 3, 3.1, 3.2</p> <p>-----</p>	<p>1-27</p>

INTERNATIONAL SEARCH REPORT

IB/05/52377

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 6622030	B1 16-09-2003	AU 6539501 A	14-01-2002
		EP 1295406 A1	26-03-2003
		WO 0203563 A1	10-01-2002
