



Europäisches Patentamt
European Patent Office
Office européen des brevets



(11) **EP 1 557 820 B1**

(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention
of the grant of the patent:
18.10.2006 Bulletin 2006/42

(51) Int Cl.:
G10L 11/02^(2006.01) G10L 19/00^(2006.01)

(21) Application number: **04425031.4**

(22) Date of filing: **22.01.2004**

(54) **Voice activity detection operating with compressed speech signal parameters**

Sprachaktivitätsdetektion unter Vewendung von komprimierten Sprachsignal-Parametern

Détection de l'activité vocale utilisant des paramètres d'une représentation compressée d'un signal de parole

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

(43) Date of publication of application:
27.07.2005 Bulletin 2005/30

(73) Proprietor: **Siemens S.p.A.
20126 Milano (IT)**

(72) Inventor: **Aldrovandi, Matteo
40122 Bologna (IT)**

(74) Representative: **Giustini, Delio
Siemens Mobile Communications S.p.A.
Postfach 22 16 34
80506 München (DE)**

(56) References cited:
**EP-A- 1 017 042 US-A- 5 579 435
US-A- 5 732 390**

EP 1 557 820 B1

Note: Within nine months from the publication of the mention of the grant of the European patent, any person may give notice to the European Patent Office of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

Description

Field of the invention

[0001] The present invention refers to digital radio communication systems, in particular mobile communication systems, and more specifically it concerns a method of and device for voice activity detection in received speech signals in one such system.

[0002] Preferably, but not exclusively, the method and the device are intended for use in connection with voice quality enhancement.

Background of the Invention

[0003] Voice activity detectors (VADs) are devices that are supplied with a signal to detect therein periods of speech and periods of silence, where only noise is present. Possibly, the VADs are also arranged to distinguish among voiced/unvoiced sounds in speech periods.

[0004] A class of VADs performs detection through an energetic analysis and a spectral analysis of the input signal, the analysis results being combined to provide the classification of an analysed speech segment. An algorithm for classifying a speech segment as voiced speech, unvoiced speech or silence based on energetic and spectral analyses is disclosed in "Application of an LPC Distance Measure to the VoicedUnvoiced-Silence Detection Problem", by L.R. Rabiner and M.R. Sambur, IEEE Transactions on Acoustics, Speech and Signal Processing, Vol. ASSP-25, No. 4, August 1977, pages 338 - 343.

[0005] In mobile communication systems, VADs are typically used at the mobile terminals, in association with the speech coder, to drive a discontinuous transmission in which coded speech signals are transmitted during active speech periods whereas in silence periods the speech transmitter is inhibited and the so-called comfort noise is transmitted. This helps in saving power.

[0006] It has now been found that it is advantageous to use a VAD also in the control part of a mobile communication system, in particular for improving the Noise Reduction (NR) feature of the so-called Voice Quality Enhancement (VQE) function. An example of VAD-assisted noise reduction for the uplink direction of communication of a digital radio communication system is disclosed in EP-A 1 017 042.

[0007] The Applicant, as well as other manufacturers, integrate the VQE function in the units (such as the Transcoding and Rate Adapting Units, or TRAU, of the GSM system) adapting the speech signals from the requirements of the radio part of the system to the requirements of the control part and, if necessary, of the fixed telephone network, and vice versa. In the Applicant's VQE, the VAD is intended to drive the noise suppression in the uplink direction of communication. For several reasons, including cost and size of the apparatus, the Applicant wishes that the addition of a VAD-driven VQE in a TRAU does

not entail changes in the hardware of the TRAU itself.

[0008] If the detection exploits both the spectrum and the energy characteristics of the received signal, the conventional approach, which is substantially as disclosed in the above-mentioned document by L.R. Rabiner and M.R. Sambur, entails that the speech signal is decoded and the spectral information (here the linear prediction coefficients LPCs) is recovered from the linear frames resulting from speech decoding. Yet, the spectral analysis of the linear signal to recover the LPCs is a heavy processing task. Taking into account that the TRAU processor generally operates in parallel on a plurality of channels, real time execution of the complete VAD algorithm on the same channels could compel to use a dedicated processor or a more powerful and hence more expensive processor than the one that would be used for the TRAU. Both solutions are in contrast with the goal of keeping the TRAU hardware unchanged.

[0009] To avoid the need for a VAD-dedicated or a more powerful processor, the spectral analysis could be dispensed with and the VAD could perform only the energetic analysis. Such a solution is disclosed in EP-A 1 017 042. The document teaches also that the energy estimation can be performed directly on the compressed signal, in order to dispense the speech decoder with the relevant processing tasks and to speed up the actual speech decoding.

[0010] Yet, by performing only the energetic analysis, only one feature of the received signal is exploited, and the detection, and hence the operation of the VAD-driven VQE, is less effective.

[0011] US-A 5 732 390 discloses a voice transmitting-receiving apparatus, in particular a portable telephone set, with a digital speech encoder and a digital speech decoder. The received sound volume is adjusted based on the results of the analysis of the performed by the speech encoder on a speech signal to be transmitted.

Object of the Invention

[0012] it is an object of the invention to provide a method and a device for voice activity detection, which are suitable for mobile communication systems and in particular are intended to drive a voice quality enhancement of the uplink signals integrated in a unit performing speech rate and coding adaptation in parallel for a plurality of channels, which method and device perform a robust detection and allow using the same processor as would be required for performing said adaptation in parallel for said plurality of channels.

Summary of the Invention

[0013] According to the invention, there is provided a method according to claim 1 of detecting voice activity voice activity in a digitally coded speech signal propagating in uplink direction in a mobile communication system and received at a receiver belonging to a control part of

said system, in which said detection is performed by operating directly on the received speech signal and a signal representative of the presence or absence of voice activity is generated by submitting the speech signal to an energetic analysis: The detection also includes a spectral analysis performed directly on the received coded speech signal, by exploiting only the information contained in the received signal itself; and spectrum and energy information resulting from said spectral analysis and said energetic analysis, respectively, are jointly processed for generating said signal representative of the presence or absence of voice.

[0014] The invention also concerns a voice activity detector according to claim 5 the detector comprising means for performing an energetic analysis on the received speech signal and for supplying energy information to means for generating a signal representative of the presence or absence of voice activity. The detector also includes means for performing a spectral analysis on the received speech signal, which spectral analysis means are directly connected with a detector input where coded speech signals are present and are arranged to perform spectral analysis by exploiting only information contained in the received signal itself and to supply and to supply said generating means with the spectral information resulting from the spectral analysis, whereby said generating means are arranged to generate said signal representative of the presence or absence of voice activity by jointly processing spectral and energy information.

[0015] In the preferred application, the voice activity detector drives a noise reduction operation, within a voice quality enhancement function performed on speech signals propagating in the uplink communication direction in the mobile communication system and embodied in units, like the so-called TRAU (Transcoding and Rate Adapting Unit), which adapt the uplink directed speech signals to the requirements of the control part of the mobile system and possibly of the fixed telephone network and adapt downlink directed speech signals to the requirements of the radio part of the mobile system.

[0016] Therefore, the invention provides also a method of voice quality enhancement according to claim 10 in a mobile communication system, in which a voice quality enhancement including a noise reduction operation is performed at least for speech signals propagating in uplink direction, in which said noise reduction operation is driven by a signal representative of the presence or absence of voice activity generated by a method of and device for voice activity detection as defined above.

Brief description of the drawings

[0017] A preferred embodiment of the invention, given by way of non-limiting example, will now be described with reference to the accompanying drawings, in which:

- Fig. 1 is a schematic block diagram of a TRAU em-

bodiment of a VQE unit and of its connections inside the mobile communications system; and

- Fig. 2 is a schematic block diagram of the invention.

Description of the preferred embodiment

[0018] The preferred embodiment disclosed here concerns a VAD intended to drive the noise reduction feature in a voice quality enhancement performed in the uplink direction of communication in a mobile communication system, in case the VQE function is incorporated into the units performing transcoding and/or rate adaptation in the control part of such a system.

[0019] Referring to Fig. 1, there is schematically shown a Transcoding and Rate Adapting Unit (TRAU) 1 of a mobile communication system, for instance a GSM system. The TRAU is connected to the Mobile Switching Centre (MSC) 2 and the Base station Controller (BSC) 3 through interfaces A and Asub and embodies a VAD-driven Voice Quality Enhancement function performed in block 4 labelled VAD & VQE.

[0020] In the most general case, a VQE includes the well-known features of Acoustic Echo Cancellation, Noise Reduction and Acoustic Level Control. In the preferred application of the invention, all of said features are provided for the uplink direction of communication only and the VAD drives the Noise Reduction (NR) feature. In downlink direction, only the Acoustic Level Control is performed, which is not concerned by the present invention.

[0021] The drawing only shows the units that, in TRAU 1, are directly concerned with the transcoding function, namely a speech coder 5 and a speech decoder 6 on the Asubinterface side, and an A/ μ law expander 7 and an A/ μ law compander 8 on the A-interface side.

[0022] In downlink direction, the TRAU receives A-law PCM signals from MSC through a line 10, sends the expanded signals to the In_Down input of VAD&VQE block 4 through line 11. The signals outgoing from the Out_Down output of block 4 are fed to coder 5 through line 12, are coded according to the desired coding technique (full-rate, enhanced full-rate, half-rate or adaptive multi-rate) and the coded signals are then forwarded to base station controller 3 through line 13.

[0023] In uplink direction, the coded signals arriving from base station controller 3 through line 14 are fed to both decoder 6 and VAD&VQE block 4. The decoded signals are fed to the In_Down input of VAD&VQE block 4 through line 15. The decoded signals having undergone voice quality enhancement are fed from Out_Up output of VAD&VQE to A/ μ law compander 8 through line 16 and hence to MSC through line 17.

[0024] It is not necessary to provide here details on the organisation of the coded speech signals in a mobile communication system, which depends on the kind of system and on the chosen coding rate. On the other hand, for any given system and rate, such organisation is well known to the skilled in the art and can be found in the

relevant standards. It is sufficient here to recall that the coded speech signals include spectral information, such as the LPCs or a representation thereof.

[0025] In Fig. 2 block VAD&VQE is decomposed into its constituent blocks, namely VAD 40 and VQE 50. VAD 40 has been schematised by a spectral analyser 41 determining the LPC coefficients, an energy analyser 42 and Joint Processing Means including including an Hard Decision Unit 43 and a Soft Decision Unit 45. Said Joint Processing Means being adapted to combine the results of the two analyses and emitting on line 44 a signal indicating the nature of the received speech frame (the so-called VAD flag), which is an input to the Noise Reduction feature of VQE 50.

[0026] According to the invention, LPC analyser 41 is directly fed with the coded speech signal frames present on line 14, whereas energy estimator 42 is fed with the decoded signal outgoing from decoder 6 through line 15. The LPC analysis of course depends on the manner in which the LPCs are represented in the coded signal. The energy evaluation and the decision may be performed according to any technique known in the art, for instance as disclosed in the above-mentioned paper of L.R. Rabiner et al.

[0027] Performing the LPC analysis directly on the coded signal affords a number of advantages in terms of processing power requirements. In particular, there is no need of dedicating processing power to the reconstruction of the LPC coefficients from the decoded signal: it is sufficient to extract them from the relevant information included in the coded speech signal, which is available on the same board. Besides the greater processing simplicity, also a reduction to at least of one fifth or even less of the information amount to be processed is achieved: indeed, at most 244 bits are to be processed to obtain the LPC coefficients from the coded signal, whereas 1280 bits are to be processed when the linear signal is used.

[0028] Under such conditions, the same processor used on the TRAU board for performing all TRAU functions and for managing the so-called tandem free operation (i.e. for dispensing with the transcoding in case of communication between two mobile terminals), for a plurality of speech channels in parallel (for instance 12 channels), can perform in real time, for the same channels, also the voice activity detection by exploiting both the spectral and the energy information, and the subsequent VQE. The resulting detection is more accurate than when only the energy information is exploited and hence also the noise suppression operation is more accurate.

[0029] It is to be appreciated that, according to the existing GSM standards, the LPC information in the coded signal is updated every 5 ms. The energy information is computed on the same interval of 5 ms and then the two contributions are jointly processed to take a decision on the nature of audio segment. This is denoted by the presence of said Joint Processing Means 43 and 45. For voice quality enhancement the high rate of decisions available

at the output of said Hard decision Unit 43 is not necessary and has often a negative impact on audition. Therefore these "hard" decisions are softened through a smoothing process which aims to redefine as "voice" eventual isolated segments of noise among a group of segments of voice and to redefine as "noise" eventual isolated segments of voice among a group of segments of noise. This is the aim of Soft Decision Unit 45 sited immediately after said Hard Decision Unit 43.

[0030] It is clear that the above description is given only by way of non-limiting example and that variations and modifications are possible without departing from the scope of the invention. In particular, even if reference has been made to a TRAU unit in a GSM system, what has been said can be applied also to mobile communication systems operating according to other standards. In such case, the energy analysis and spectral analysis should be adapted to the specific requirements of that system. The invention could be used also in other radio communication signals in which a digital coding of speech is adopted and the digitally coded speech signals include spectral information. Moreover, also the energy analysis could be performed on the coded signal, as disclosed in the above-mentioned EP-A 1017042.

Claims

1. A method of detecting voice activity in a digitally coded speech signal propagating in uplink direction for a mobile communication system, said coded speech signal comprising a representation of a spectral information and received at a receiver (4, 6) belonging to a control part (1, 3) of said system, in which said detection is performed by operating directly on the received speech signal and a signal representative of the presence or absence of voice activity is generated by submitting the speech signal to an energetic analysis, **characterised in that** said detection also includes a spectral analysis performed directly on the received coded speech signal, by exploiting only the spectral information contained in the received signal itself; and spectrum and energy information resulting from said spectral analysis and said energetic analysis, respectively, are jointly processed for generating said signal representative of the presence or absence of voice.
2. A method as claimed in claim 1, **characterised in that** said energetic analysis is performed on decoded received speech signals.
3. A method as claimed in claim 1 or 2, wherein spectral information in said coded signal is periodically updated, **characterised in that** said joint processing includes:

- taking a preliminary hard decision, at each up-

dating period of said spectral information, about the presence or absence of voice activity; and
 - generating said signal representative of the presence or absence of voice activity through a soft decision step adapted to smooth the decision rate.

4. A method as claimed in any preceding claim, wherein said voice activity detection is for assisting a voice quality enhancement function for the uplink direction of said system, **characterised in that** said detection is performed in parallel on a plurality of speech channels, along with an adaptation of the uplink speech signals of said channels to the requirements of the control part of the mobile system and possibly of a fixed network and with an adaptation of the downlink speech signals of said channels to the requirements of the radio part of the mobile system and by using the same processor as would be provided for performing speech signal adaptation and voice quality enhancement in parallel for said plurality of speech channels.
5. A voice activity detector (40) for detecting voice activity in a digitally coded speech signal propagating in uplink direction for a mobile communication system, said coded speech signal comprising a representation of a spectral information and received at a receiver (4, 6) belonging to a control part (1, 3) of said system, the detector (40) comprising means (42) for performing an energetic analysis on the received speech signal and for supplying energy information to means (43, 45) for generating a signal representative of the presence or absence of voice activity, **characterised in that** said detector (40) also includes means (41) for performing a spectral analysis on the received speech signal, said spectral analysis means (41) being connected directly with a detector input (14) where coded speech signals are present and being arranged to perform spectral analysis by exploiting only the spectral information contained in the received signal itself and to supply said generating means (43, 45) with spectral information resulting from said spectral analysis, whereby said generating means (43, 45) are arranged to generate said signal representative of the presence or absence of voice activity by a joint processing of the spectral and energy information.
6. A voice activity detector as claimed in claim 5, **characterised in that** said means (42) for performing an energetic analysis on the received speech signal are connected downstream a speech decoder (6).
7. A voice activity detector as claimed in claim 5 or 6, wherein spectral information in said coded signals is periodically updated, **characterised in that** said

means (43, 45) for generating a signal representative of the presence or absence of voice activity include:

- a hard decision unit (43) connected to the output of said means (42) for performing an energetic analysis and said means (41) for performing a spectral analysis, said unit being adapted to joint processing, at each updating period, an input speech segment and to output a hard noise or voice decision;
- a soft decision unit (45) connected to the output of said hard decision unit (43) and adapted to perform a smoothing process in order to redefine as "voice" possible isolated noise segments among a group of voice segments and to redefine as "noise" possible isolated voice segments among a group of noise segments.

8. A voice activity detector (40) as claimed in any of claims 5 to 7, wherein said detector (40) is located upstream of means (50) performing a voice quality enhancement on uplink speech signals and supplies said voice quality enhancement means (50) with said signal representative of the presence or absence of voice activity, **characterised in that** it is arranged to perform detection in parallel for the uplink speech signals transmitted over a plurality of speech channels.
9. A voice activity detector as claimed in claim 8, **characterised in that** said detector (40) is part, together with said voice quality enhancement means (50), of a unit (1) performing, for said plurality of channels, an adaptation of the uplink directed speech signals to the requirements of the control part (1, 2) of the mobile system and possibly of a fixed network and an adaptation of the downlink directed speech signals to the requirements of the radio part of the mobile system, and is implemented by the same processor that would be provided for performing speech signal adaptation and voice quality enhancement in parallel for said plurality of speech channels.
10. A method of voice quality enhancement for a mobile communication system, in which a voice quality enhancement including a noise reduction operation is performed at least for speech signals propagating in uplink direction, **characterised in that** said noise reduction operation is driven by a signal representative of the presence or absence of voice activity generated by a method of voice activity detection as claimed in any of claims 1 to 4 and/or by a voice activity detector as claimed in any of claims 5 to 9.

Patentansprüche

1. Verfahren zur Sprachaktivitätsdetektion in einem di-

- gital codierten Sprachsignal, welches in der Aufwärtsrichtung für eine Mobilfunkkommunikationssystem übertragen wird, wobei das besagte codierte Sprachsignal eine Darstellung einer Spektralinformation beinhaltet, und welches an einem Empfänger (4, 6) empfangen wird, der zu einem Steuerteil (1, 3) des besagten Systems gehört, wobei die besagte Detektion durch unmittelbare Bearbeitung des empfangenen Sprachsignals erfolgt und ein Signal, das das Vorhandensein oder Nichtvorhandensein von Sprachaktivität repräsentiert, erzeugt wird, indem das Sprachsignal einer energetischen Analyse unterzogen wird, **dadurch gekennzeichnet, dass** die besagte Detektion außerdem eine Spektralanalyse umfasst, welche unmittelbar an dem empfangenen codierten Sprachsignal durchgeführt wird, indem ausschließlich die Spektralinformation, die in dem empfangenen Signal enthalten ist, ausgelesen wird; und Spektral- und Energieinformationen, die aus der besagten Spektralanalyse bzw. der besagten energetischen Analyse hervorgehen, gemeinsam verarbeitet werden, um das besagte Signal zu erzeugen, welches das Vorhandensein bzw. Nichtvorhandensein von Sprachaktivität repräsentiert.
2. Verfahren gemäß Anspruch 1, **dadurch gekennzeichnet, dass** die besagte energetische Analyse an decodierten empfangenen Sprachsignalen durchgeführt wird.
3. Verfahren gemäß Anspruch 1 oder 2, wobei die Spektralinformationen in dem besagten codierten Signal in regelmäßigen Abständen aktualisiert werden, **dadurch gekennzeichnet, dass** die besagte gemeinsame Verarbeitung Folgendes umfasst :
- Treffen einer vorläufigen harten Entscheidung ("Hard-Decision") in jedem Aktualisierungszeitraum der besagten spektralen Informationen bezüglich des Vorhandenseins bzw. Nichtvorhandenseins von Sprachaktivität; und Erzeugen des besagten Signals, welches das Vorhandensein bzw. Nichtvorhandensein von Sprachaktivität repräsentiert, durch einen weichen Entscheidungsschritt ("Soft-Decision"), der dafür ausgelegt ist, die Entscheidungsrate zu glätten.
4. Verfahren gemäß einem der vorstehenden Ansprüche, wobei die besagte Sprachaktivitätsdetektion der Unterstützung einer Sprachqualitäts-Verbesserungsfunktion für die Aufwärtsrichtung des besagten Systems dient, **dadurch gekennzeichnet, dass** die besagte Detektion parallel für eine Mehrzahl von Sprachkanälen durchgeführt wird, zusammen mit einer Anpassung der Sprachsignale in der Aufwärtsrichtung der besagten Kanäle an die Anforderungen des Steuerteils des Mobilfunkkommunikationssystems sowie gegebenenfalls eines Festnetzes und mit einer Anpassung der Sprachsignale in der Abwärtsrichtung der besagten Kanäle an die Anforderungen des Funkteils des Mobilfunkkommunikationssystems und unter Verwendung desselben Prozessors, der bereitgestellt würde, um die Anpassung der Sprachsignale und die Verbesserung der Sprachqualität für die besagte Mehrzahl von Sprachkanälen parallel durchzuführen.
5. Sprachaktivitätsdetektor (40) zum Erkennen von Sprachaktivität in einem digital codierten Sprachsignal, welches in der Aufwärtsrichtung übertragen wird, für ein Mobilfunkkommunikationssystem, wobei das besagte codierte Sprachsignal eine Darstellung einer Spektralinformation beinhaltet, und welches an einem Empfänger (4, 6) empfangen wird, der zu einem Steuerteil (1, 3) des besagten Systems gehört, wobei der Detektor (40) Mittel (42) enthält, um eine energetische Analyse des empfangenen Sprachsignals durchzuführen und um Energieinformationen für die Mittel (43, 45) zum Erzeugen eines Signals, das das Vorhandensein bzw. Nichtvorhandensein von Sprachaktivitäten repräsentiert, bereitzustellen, **dadurch gekennzeichnet, dass** der besagte Detektor (40) außerdem ein Mittel (41) zur Durchführung einer Spektralanalyse des empfangenen Sprachsignals enthält, wobei das besagte Spektralanalysemittel (41) direkt mit einem Detektoreingang (14) verbunden ist, an dem die codierten Sprachsignale anliegen, und dafür ausgelegt ist, eine Spektralanalyse durchzuführen, indem ausschließlich die Spektralinformationen ausgelesen werden, die in dem empfangenen Signal selbst enthalten sind, und den besagten Erzeugungsmitteln (43, 45) Spektralinformationen bereitzustellen, die aus der besagten Spektralanalyse hervorgehen, wobei die besagten Erzeugungsmittel (43, 45) dafür ausgelegt sind, das besagte Signal zu erzeugen, welches das Vorhandensein bzw. Nichtvorhandensein von Sprachaktivitäten repräsentiert, indem die Spektral- und Energieinformationen gemeinsam verarbeitet werden.
6. Sprachaktivitätsdetektor gemäß Anspruch 5, **dadurch gekennzeichnet, dass** die besagten Mittel (42) zur Durchführung einer energetischen Analyse des empfangenen Sprachsignals einem Sprach-Decodierer (6) in Abwärtsrichtung nachgeschaltet sind.
7. Sprachaktivitätsdetektor gemäß Anspruch 5 oder 6, wobei die Spektralinformationen in den besagten codierten Signalen in regelmäßigen Abständen aktualisiert werden, **dadurch gekennzeichnet, dass** die besagten Mittel (43, 45) zum Erzeugen eines Signals, welches das Vorhandensein oder Nichtvorhandensein von Sprachaktivität repräsentiert, Folgendes umfassen:

- eine "Hard-Decision"-Einheit (43), die mit dem Ausgang des besagten Mittels (42) zum Durchführen einer energetischen Analyse und des besagten Mittels (41) zum Durchführen einer Spektralanalyse verbunden ist, wobei die besagte Einheit dafür ausgelegt ist, in jedem Aktualisierungszeitraum ein eingegebenes Sprachsegment zu verarbeiten und eine harte Entscheidung bezüglich "Geräusch" oder "Sprache" zu treffen;
- eine "Soft-Decision"-Einheit (45), die mit dem Ausgang der besagten "Hard-Decision"-Einheit (43) verbunden ist und die dafür ausgelegt ist, einen Glättungsprozess auszuführen, um eventuelle isolierte "Geräusch"-Segmente in einer Gruppe von Sprach-Segmenten als "Sprache" umzudefinieren und um eventuelle isolierte "Sprach"-Segmente in einer Gruppe von Geräusch-Segmenten als "Geräusch" umzudefinieren.
8. Sprachaktivitätsdetektor (40) gemäß einem der Ansprüche 5 bis 7, wobei der besagte Detektor (40) in Aufwärtsrichtung dem Mittel (50) zur Durchführung einer Sprachqualitätsverbesserung bei Sprachsignalen in der Aufwärtsstrecke vorgeschaltet ist und dem besagten Mittel (50) zur Durchführung einer Sprachqualitätsverbesserung das besagte Signal bereitstellt, welches das Vorhandensein bzw. Nichtvorhandensein von Sprachaktivität repräsentiert, **dadurch gekennzeichnet, dass** es dafür ausgelegt ist, die Detektion für die in der Aufwärtsstrecke über eine Mehrzahl von Sprachkanälen übertragenen Sprachsignale parallel durchzuführen.
9. Sprachaktivitätsdetektor gemäß Anspruch 8, **dadurch gekennzeichnet, dass** der besagte Detektor (40), zusammen mit dem besagten Mittel (50) zur Durchführung einer Sprachqualitätsverbesserung, Bestandteil einer Einheit (1) ist, welche für die besagte Mehrzahl von Kanälen eine Anpassung der Sprachsignale in der Aufwärtsrichtung an die Anforderungen des Steuerteils (1, 2) des Mobilfunkkommunikationssystems und gegebenenfalls eines Festnetzes sowie eine Anpassung der Sprachsignale in der Abwärtsrichtung an die Anforderungen des Funkteils eines Mobilfunkkommunikationssystems vornimmt, und durch denselben Prozessor implementiert ist, der bereitgestellt würde, um die Anpassung der Sprachsignale und die Verbesserung der Sprachqualität für die besagte Mehrzahl von Sprachkanälen parallel durchzuführen.
10. Verfahren zur Sprachqualitätsverbesserung für ein Mobilfunkkommunikationssystem, bei dem eine Verbesserung der Sprachqualität einschließlich einer Operation zur Geräuschminderung mindestens für Sprachsignale durchgeführt wird, die in der Auf-

wärtsrichtung übertragen werden, **dadurch gekennzeichnet, dass** die besagte Geräuschminderungsoperation von einem Signal gesteuert wird, das das Vorhandensein bzw. Nichtvorhandensein von Sprachaktivität repräsentiert und das von einem Verfahren zur Sprachaktivitätsdetektion, wie es in den Ansprüchen 1 bis 4 beansprucht wird, und/oder durch einen Sprachaktivitätsdetektor gemäß den Ansprüchen 5 bis 9 erzeugt wird.

Revendications

1. Procédé de détection de l'activité vocale dans un signal de parole codé numériquement se propageant dans une direction de liaison montante pour un système de communication mobile, ledit signal de parole codé comprenant une représentation d'une information spectrale et étant reçu au niveau d'un récepteur (4, 6) appartenant à une partie de commande (1, 3) dudit système, dans lequel ladite détection est effectuée en opérant directement sur le signal de parole reçu et un signal représentatif de la présence ou de l'absence d'activité vocale est généré en soumettant le signal de parole à une analyse énergétique, **caractérisé en ce que** ladite détection comprend également une analyse spectrale exécutée directement sur le signal de parole codé reçu, en exploitant seulement l'information spectrale contenue dans le signal reçu lui-même; et les informations de spectre et d'énergie résultant respectivement de ladite analyse spectrale et de ladite analyse énergétique sont traitées conjointement pour générer ledit signal représentatif de la présence ou de l'absence de voix.
2. Procédé selon la revendication 1, **caractérisé en ce que** ladite analyse énergétique est exécutée sur des signaux de parole reçus décodés.
3. Procédé selon la revendication 1 ou 2, dans lequel l'information spectrale dans ledit signal codé est mise à jour périodiquement, **caractérisé en ce que** ledit traitement conjoint comprend :
 - prendre une décision ferme préliminaire, à chaque période de mise à jour de ladite information spectrale, concernant la présence ou l'absence d'activité vocale; et
 - générer ledit signal représentatif de la présence ou de l'absence d'activité vocale par l'intermédiaire d'une étape de décision souple adaptée pour lisser le taux de décision.
4. Procédé selon l'une quelconque des revendications précédentes, dans lequel ladite détection d'activité vocale est destinée à aider une fonction d'amélioration de qualité vocale pour la direction de liaison

montante dudit système, **caractérisé en ce que** ladite détection est exécutée en parallèle sur une pluralité de voies de conversation, en même temps qu'une adaptation des signaux de parole de liaison montante desdites voies aux exigences de la partie de commande du système mobile et éventuellement d'un réseau fixe et qu'une adaptation des signaux de parole de liaison descendante desdites voies aux exigences de la partie radio du système mobile et en utilisant le même processeur que celui qui serait prévu pour exécuter l'adaptation de signal de parole et l'amélioration de qualité vocale en parallèle pour ladite pluralité de voies de conversation.

5. Détecteur d'activité vocale (40) pour détecter l'activité vocale dans un signal de parole codé numériquement se propageant dans une direction de liaison montante pour un système de communication mobile, ledit signal de parole codé comprenant une représentation d'une information spectrale et étant reçu au niveau d'un récepteur (4, 6) appartenant à une partie de commande (1, 3) dudit système, le détecteur (40) comprenant un moyen (42) pour exécuter une analyse énergétique sur le signal de parole reçu et pour délivrer une information d'énergie à des moyens (43, 45) pour générer un signal représentatif de la présence ou de l'absence d'activité vocale, **caractérisé en ce que** ledit détecteur (40) comprend également un moyen (41) pour exécuter une analyse spectrale sur le signal de parole reçu, ledit moyen d'analyse spectrale (41) étant relié directement à une entrée de détecteur (14) où des signaux de parole codés sont présents et configuré pour exécuter une analyse spectrale en exploitant seulement l'information spectrale contenue dans le signal reçu lui-même et pour délivrer auxdits moyens de génération (43, 45) l'information spectrale résultant de ladite analyse spectrale, moyennant quoi lesdits moyens de génération (43, 45) sont configurés pour générer ledit signal représentatif de la présence ou de l'absence d'activité vocale par un traitement conjoint des informations spectrales et d'énergie.
6. Détecteur d'activité vocale selon la revendication 5, **caractérisé en ce que** ledit moyen (42) pour exécuter une analyse énergétique sur le signal de parole reçu est relié en aval d'un décodeur de parole (6).
7. Détecteur d'activité vocale selon la revendication 5 ou 6, dans lequel l'information spectrale dans lesdits signaux codés est mise à jour périodiquement, **caractérisé en ce que** lesdits moyens (43, 45) pour générer un signal représentatif de la présence ou de l'absence d'activité vocale comprennent :
 - une unité de décision ferme (43) reliée à la sortie dudit moyen (42) pour exécuter une analyse énergétique et dudit moyen (41) pour exé-

cuter une analyse spectrale, ladite unité étant adaptée pour traiter conjointement, à chaque période de mise à jour, un segment de parole d'entrée et pour délivrer en sortie une décision ferme de bruit ou vocale;

- une unité de décision souple (45) reliée à la sortie de ladite unité de décision ferme (43) et adaptée pour exécuter un traitement de lissage afin de redéfinir comme "voix" des segments de bruit isolés possibles parmi un groupe de segments vocaux et de redéfinir comme "bruit" des segments vocaux isolés possibles parmi un groupe de segments de bruit.

8. Détecteur d'activité vocale (40) selon l'une quelconque des revendications 5 à 7, dans lequel ledit détecteur (40) est situé en amont d'un moyen (50) exécutant une amélioration de qualité vocale sur les signaux de parole de liaison montante et délivre audit moyen d'amélioration de qualité vocale (50) ledit signal représentatif de la présence ou de l'absence d'activité vocale, **caractérisé en ce qu'il** est configuré pour exécuter une détection en parallèle pour les signaux de parole de liaison montante transmis sur une pluralité de voies de conversation.
9. Détecteur d'activité vocale selon la revendication 8, **caractérisé en ce que** ledit détecteur (40) fait partie, avec ledit moyen d'amélioration de qualité vocale (50), d'une unité (1) exécutant, pour ladite pluralité de voies, une adaptation des signaux de parole dirigés sur la liaison montante aux exigences de la partie de commande (1, 2) du système mobile et éventuellement d'un réseau fixe et une adaptation des signaux de parole dirigés sur la liaison descendante aux exigences de la partie radio du système mobile, et est mis en oeuvre par le même processeur que celui qui serait prévu pour exécuter une adaptation du signal de parole et une amélioration de la qualité vocale en parallèle pour ladite pluralité de voies de conversation.
10. Procédé d'amélioration de la qualité vocale pour un système de communication mobile, dans lequel une amélioration de qualité vocale comprenant une opération de réduction du bruit est exécutée au moins pour les signaux de parole se propageant dans une direction de liaison montante, **caractérisé en ce que** ladite opération de réduction de bruit est activée par un signal représentatif de la présence ou de l'absence d'activité vocale généré par un procédé de détection d'activité vocale selon l'une quelconque des revendications 1 à 4 et/ou par un détecteur d'activité vocale selon l'une quelconque des revendications 5 à 9.

← Downlink Uplink →

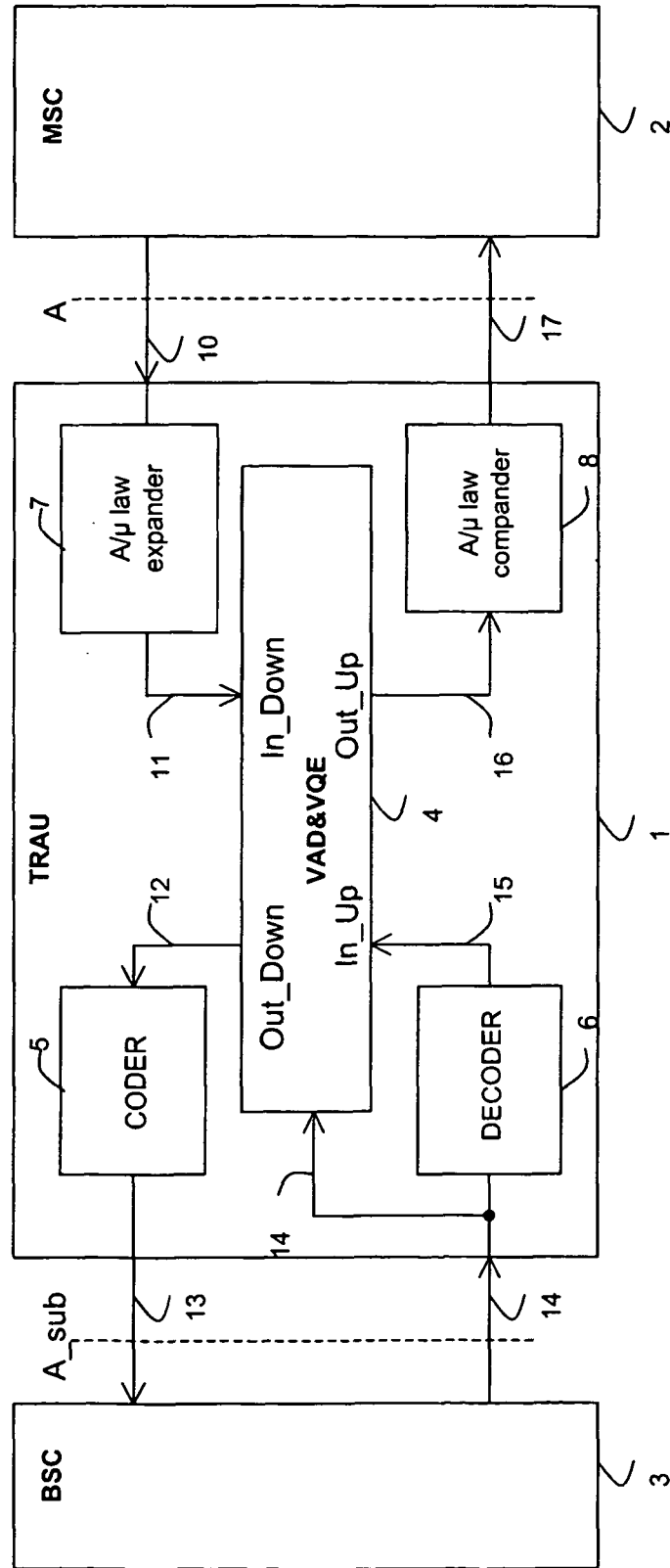


FIG. 1

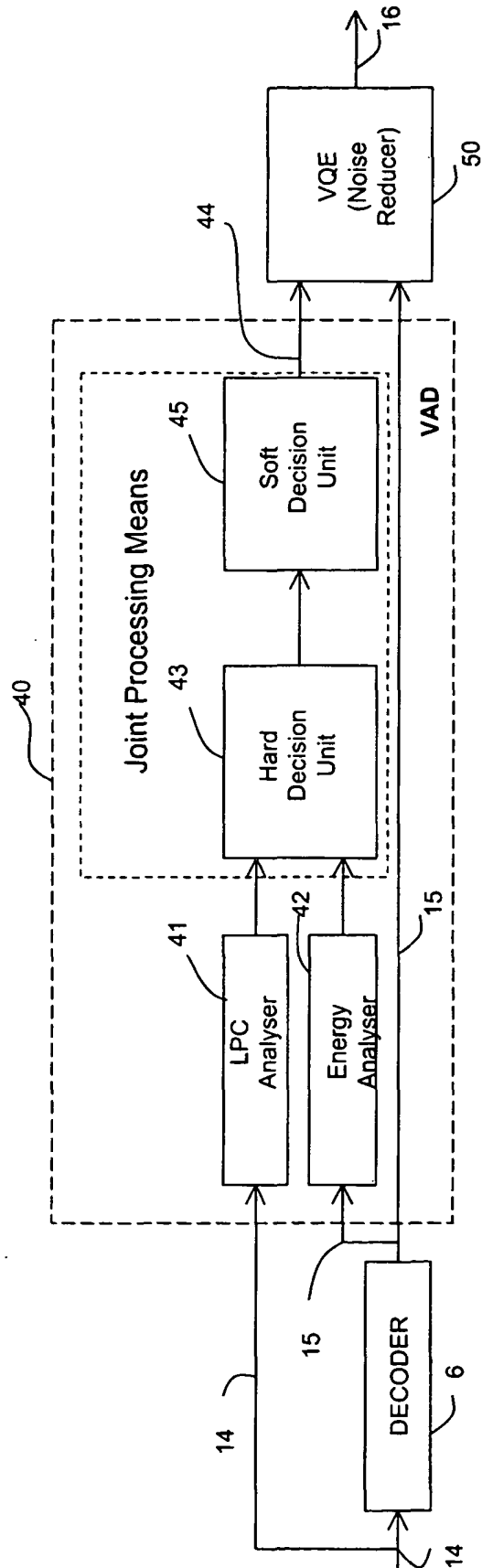


FIG. 2