

(19)



Europäisches Patentamt
European Patent Office
Office européen des brevets



(11)

EP 0 714 088 B1

(12)

FASCICULE DE BREVET EUROPEEN

(45) Date de publication et mention
de la délivrance du brevet:
18.08.1999 Bulletin 1999/33

(51) Int Cl.⁶: **G10L 3/00**

(21) Numéro de dépôt: **95402589.6**

(22) Date de dépôt: **17.11.1995**

(54) **Détection d'activité vocale**

Sprachaktivitätsdetektion

Voice activity detection

(84) Etats contractants désignés:
AT BE CH DE DK ES FR GB IT LI NL SE

(30) Priorité: **22.11.1994 FR 9413962**

(43) Date de publication de la demande:
29.05.1996 Bulletin 1996/22

(73) Titulaire: **ALCATEL**
75008 Paris (FR)

(72) Inventeurs:
• **Chaqui, Jamil**
F-92400 Courbevoie (FR)
• **Bourmeyster, Ivan**
F-75014 Paris (FR)
• **Robbe, François**
F-95240 Corneilles-en-Parisis (FR)

(74) Mandataire: **Lamoureux, Bernard et al**
COMPAGNIE FINANCIERE ALCATEL
Département Propriété Industrielle
30, avenue Kléber
75116 Paris (FR)

(56) Documents cités:
EP-A- 0 123 349 **EP-A- 0 335 521**

- **IEEE INTERNATIONAL CONFERENCE ON SYSTEMS ENGINEERING, 24 Août 1989 FAIRBORN, OHIO, pages 537-540, K.S. RAFILA ET AL. 'Voiced/Unvoiced/Mixed excitation classification of speech using the autocorrelation of the output of an adpcm system'**

EP 0 714 088 B1

Il est rappelé que: Dans un délai de neuf mois à compter de la date de publication de la mention de la délivrance du brevet européen, toute personne peut faire opposition au brevet européen délivré, auprès de l'Office européen des brevets. L'opposition doit être formée par écrit et motivée. Elle n'est réputée formée qu'après paiement de la taxe d'opposition. (Art. 99(1) Convention sur le brevet européen).

Description

[0001] Le domaine de l'invention est celui de la détection d'activité vocale dans un signal audio.

[0002] En présence d'un signal audio qui est souvent issu d'un microphone, il est parfois nécessaire de savoir si ce signal contient de la parole ou bien s'il ne comporte que du bruit.

[0003] En effet, la détection d'activité vocale va souvent conditionner certains traitements que le signal audio est susceptible de subir. Au nombre des applications typiques qu'il convient d'activer en présence d'un signal de parole, on peut identifier la reconnaissance de la parole, l'annulation d'écho ou encore la fonction d'enregistrement.

[0004] Au contraire, si l'on considère un signal de téléphonie où seule la parole représente l'information utile, il est maintenant courant dans le domaine des radio-communications de ne pas transmettre ce signal si celui-ci ne comprend que du bruit, c'est que l'on appelle couramment la transmission discontinue.

[0005] Ainsi, des solutions ont déjà été proposées pour tenter de détecter l'activité vocale dans un signal audio.

[0006] Une première solution consiste à suivre l'évolution de l'énergie du signal. Si celle-ci augmente rapidement, cela peut correspondre à l'apparition d'une activité vocale mais cela peut aussi correspondre à une variation du bruit ambiant. Il s'ensuit que cette méthode, bien que très simple à mettre en oeuvre ne se présente pas comme très fiable dans les milieux relativement bruités comme c'est le cas par exemple dans un véhicule automobile.

[0007] On connaît également de nombreuses autres solutions qui ont été développées pour pallier le défaut de fiabilité de la précédente. C'est le cas notamment de celles qui mettent en oeuvre une transformée de Fourier du signal audio pour mesurer la distance spectrale le séparant d'un signal de bruit moyenné qui est mis à jour en l'absence de toute activité vocale. C'est également le cas des méthodes utilisant une analyse du signal en sous-bandes, méthodes qui sont proches de celles faisant appel à une transformée de Fourier. C'est encore le cas des méthodes faisant appel à l'analyse cepstrale.

[0008] Il s'agit là de techniques beaucoup plus complexes qui, si elles apportent bien un gain au niveau de la fiabilité, ne donnent cependant pas complète satisfaction sur ce point.

[0009] On connaît aussi des solutions qui mettent à profit une certaine périodicité de la parole au nombre desquelles figure celle décrite dans la demande de brevet EP 0 123 349. En effet, les sons voisés présentent tous une périodicité déterminée alors que le bruit est normalement apériodique ou bien présente une périodicité distincte de celle de la parole.

[0010] On peut donc rechercher la valeur de cette périodicité déterminée (ou "pitch" en anglais) pour reconnaître la présence de sons voisés.

[0011] Pour ce faire, on calcule généralement les coefficients d'autocorrélation du signal audio pour rechercher le second maximum de ces coefficients, le premier maximum représentant l'énergie. Il s'agit là encore d'une technique relativement complexe qui ne donne pas complète satisfaction sur le plan de la fiabilité.

[0012] La demande de brevet EP 0 335 521 décrit un dispositif de détection d'activité vocale qui utilise de coefficients d'autocorrélation d'un signal, pondérés et combinés, pour obtenir une quantité qui dépend de l'énergie dans la partie du spectre qui ne contient pas de bruit, cette quantité étant comparée avec un seuil variable pour donner une indication d'activité vocale.

[0013] La présente invention propose donc une solution pour détecter l'activité vocale qui procure une fiabilité acceptable pour une complexité réduite.

[0014] Selon l'invention, un dispositif de détection d'activité vocale dans un signal audio comprend :

- des moyens pour calculer les coefficients d'autocorrélation de ce signal,
- des moyens pour identifier un premier vecteur d'autocorrélation ayant pour composantes une première série de coefficients d'autocorrélation,
- des moyens pour identifier un second vecteur d'autocorrélation ayant pour composantes une deuxième série de coefficients d'autocorrélation décalée par rapport à la première série d'une valeur de décalage prédéterminée,
- des moyens pour soustraire le premier vecteur d'autocorrélation du second vecteur d'autocorrélation afin d'obtenir un vecteur de différentiation,
- des moyens pour calculer une norme de ce vecteur de différentiation, cette norme représentant un premier indicateur d'activité vocale.

[0015] De plus, le dispositif comprend des moyens de réduction pour établir une norme réduite en divisant la norme du vecteur de différentiation par une valeur de réduction, cette norme réduite représentant un deuxième indicateur d'activité vocale.

[0016] A titre d'exemple, la valeur de réduction est égale à l'énergie du signal ou bien elle est égale à la somme de l'énergie du signal et d'une constante de compression.

[0017] Selon une caractéristique additionnelle du dispositif, celui-ci comprend des moyens de lissage de l'un de ces indicateurs d'activité vocale pour produire une combinaison linéaire de la valeur présente de cet indicateur et de sa valeur antérieure, cette combinaison linéaire représentant un troisième indicateur d'activité vocale.

[0018] Par ailleurs, le dispositif comprend des moyens de décision pour produire un signal d'activité vocale si l'un de ces indicateurs excède un seuil de détection.

[0019] On peut trouver un intérêt à établir ce seuil de détection à partir de l'énergie du signal audio en l'ab-

sence de signal d'activité vocale.

[0020] En outre, une solution avantageuse consiste à choisir la somme des valeurs absolues des composantes du vecteur de différentiation comme norme de ce vecteur.

[0021] L'invention concerne également une méthode de détection d'activité vocale dans un signal audio comprenant les opérations suivantes :

- calcul des coefficients d'autocorrélation de ce signal,
- identification d'un premier vecteur d'autocorrélation ayant pour composantes une première série de coefficients d'autocorrélation,
- identification d'un second vecteur d'autocorrélation ayant pour composantes une deuxième série de coefficients d'autocorrélation décalée par rapport à la première série d'une valeur de décalage prédéterminée,
- soustraction du premier vecteur d'autocorrélation du second vecteur d'autocorrélation afin d'obtenir un vecteur de différentiation,
- calcul d'une norme du vecteur de différentiation, cette norme représentant un premier indicateur d'activité vocale.

[0022] La présente invention apparaîtra maintenant de manière plus claire dans le cadre d'un exemple de réalisation donné à titre illustratif en se référant à la figure annexée qui représente le déroulement des opérations effectuées par le dispositif de détection d'activité vocale.

[0023] On se place dans le cas où un signal audio est de nature numérique, c'est-à-dire qu'il se présente sous la forme d'une suite d'échantillons qui correspondent à la valeur du signal à des instants successifs qui se répètent au rythme d'une fréquence d'échantillonnage.

[0024] Lorsque le signal à analyser est de nature analogique, s'il est issu d'un microphone par exemple, il est d'abord soumis à un convertisseur analogique-numérique qui fonctionne à la cadence de cette fréquence d'échantillonnage pour produire le signal audio.

[0025] Le signal audio étant numérique, il apparaît naturel de réaliser le dispositif de détection d'activité vocale au moyen d'un processeur de signal numérique. Ce processeur pourra bien entendu être utilisé à d'autres fins.

[0026] On comprend donc que ce dispositif de détection ne sera pas décrit dans sa structure car il met en oeuvre des opérations élémentaires bien connues de l'homme du métier telles que additions, multiplications, comparaisons. C'est donc une description fonctionnelle qui a été retenue, car elle semble de loin préférable pour expliciter la mise en oeuvre de l'invention avec la plus grande clarté.

[0027] En référence à la figure unique, le dispositif reçoit donc le signal audio et on considère une série d'échantillons $S(i)$ où i varie de 0 à N .

[0028] La première opération qu'effectue le dispositif

est le calcul des coefficients d'autocorrélation $R(k)$ du signal pour toutes les valeurs de k comprises entre 0 et N :

$$R(k) = \sum_{i=0}^{N-k} S(i) S(i+k)$$

[0029] A partir de ces coefficients d'autocorrélation $R(k)$ on peut définir un premier R_0 et un second R_q vecteurs d'autocorrélation en considérant de plus une valeur de décalage q qui est un entier strictement positif. Le premier vecteur d'autocorrélation R_0 a pour composants les $(N-q+1)$ premiers coefficients d'autocorrélation $R(k)$:

$$R_0 = (R(0), R(1), \dots, R(N-q))$$

[0030] Le second vecteur d'autocorrélation R_q a pour composants les $(N-q+1)$ derniers coefficients d'autocorrélation $R(k)$:

$$R_q = (R(q), R(q+1), \dots, R(N))$$

[0031] Le dispositif de détection calcule alors un vecteur de différentiation ΔR en soustrayant le premier vecteur d'autocorrélation R_0 du second vecteur d'autocorrélation R_q :

$$\Delta R = R_q - R_0$$

[0032] Si l'on note $\Delta R(k)$ la $(k+1)$ ème composante de ce vecteur de différenciation, celle-ci vaut alors pour tout k compris entre 0 et $N-q$:

$$\Delta R(k) = R(k+q) - R(k)$$

[0033] On s'aperçoit que les premiers R_0 et deuxième R_q vecteurs d'autocorrélation n'ont pas d'utilité en eux-mêmes. Ils ont été introduits dans le simple but de clarifier la présentation. Le point important est le calcul du vecteur de différenciation. Ainsi, ce vecteur se définit par la valeur de ces composantes telle que définie ci-dessus.

[0034] Dès lors, le dispositif de détection calcule une norme $\|\Delta R\|$ du vecteur de différentiation ΔR . De manière avantageuse, cette norme est égale à la somme des valeurs absolues des composantes du vecteur :

$$\|\Delta R\| = \sum_{k=0}^{N-q} |\Delta R(k)|$$

[0035] Il va sans dire que l'invention s'applique également si l'on choisit de retenir une autre norme telle que, notamment, la norme euclidienne ou la valeur maximale des valeurs absolues de chacune des composantes.

[0036] Cette norme, quelle qu'elle soit, constitue un premier indicateur d'activité vocale.

[0037] Une première option consiste à comparer cet indicateur à un seuil pour établir qu'il y a présence d'activité vocale dans le signal audio si l'indicateur est supérieur au seuil.

[0038] Selon une seconde option, le dispositif de détection calcule une norme réduite P en divisant la norme $\|\Delta R\|$ du vecteur de différentiation par une valeur de réduction. A titre d'exemple, cette valeur de réduction peut être choisie égale à l'énergie R(0) du signal audio, ce qui va tendre à comprimer la dynamique de la norme $\|\Delta R\|$. Une autre solution qui procure ses avantages propres consiste à affecter à cette valeur de réduction la somme de l'énergie R(0) du signal audio et d'une constante que l'on nommera valeur plancher C.

[0039] Cette norme réduite P, en tout état de cause constitue un deuxième indicateur d'activité vocale que l'on peut également comparer à un seuil pour établir l'absence ou la présence d'activité vocale dans ce signal.

[0040] Selon une troisième option, le dispositif de détection procède à un lissage de cette norme réduite. Ainsi, si l'on considère plusieurs séries successives de N échantillons du signal audio, une norme réduite P_i correspond à la ième série. La valeur lissée $\bar{P}_i$ de cette norme réduite sera une combinaison linéaire de la valeur lissée $\bar{P}_{i-1}$ de la norme réduite P_{i-1} associée à la série précédente et de cette norme réduite P_i :

$$\bar{P}_i = \alpha \bar{P}_{i-1} + \beta P_i$$

[0041] On peut choisir α et β de sorte que leur somme soit égale à l'unité.

[0042] De plus, il convient d'initialiser $\bar{P}_0$ à l'aide d'une constante quelconque, 0 par exemple.

[0043] Cette valeur lissée $\bar{P}_i$ constitue un troisième indicateur d'activité vocale que l'on peut aussi comparer à un seuil pour établir si le signal audio présente ou non une activité vocale.

[0044] Quel que soit l'indicateur d'activité vocale retenu, le dispositif de détection le compare donc à un seuil de détection T. La solution la plus simple consiste à affecter une valeur constante à ce seuil de détection.

[0045] Cependant, une solution avantageuse consiste à adapter ce seuil au niveau de la norme réduite P lorsque le signal audio est dépourvu d'activité vocale.

[0046] On peut donc calculer la valeur moyenne de la norme réduite sur plusieurs séries successives d'échantillons du signal audio pour lesquelles aucune activité vocale n'a été détectée et multiplier cette valeur moyenne par un coefficient constant pour obtenir le seuil de détection T. Il s'agit là d'une technique analogue à celle du lissage bien connue de l'homme du métier et elle ne sera donc pas plus détaillée.

[0047] Outre le dispositif de détection proprement dit, l'invention concerne naturellement la méthode de détection d'activité vocale qui est mise en oeuvre par ce dispositif.

[0048] A titre d'application numérique et pour présenter un cas concret d'utilisation de l'invention, on prendra pour illustration le système paneuropéen de radiocommunication cellulaire numérique dit système GSM. Dans ce système le signal analogique à traiter est échantillonné à la fréquence de 8 kHz. Les échantillons ainsi obtenus sont regroupés en séries de 160 qui correspondent donc chacune à 20 ms.

[0049] Ainsi, N, le nombre d'échantillons, vaut 160 et l'on choisira de manière avantageuse de fixer la valeur de décalage q égale à l'unité.

[0050] Les composantes du vecteur de différentiation s'écrivent alors pour tout k compris entre 1 et 160 :

$$\Delta R(k) = R(k+1) - R(k)$$

[0051] La norme de ce vecteur peut donc s'écrire :

$$\|\Delta R\| = \sum_{k=0}^{159} |\Delta R(k)|$$

Revendications

1. Dispositif de détection d'activité vocale dans un signal audio comprenant :

- des moyens pour calculer les coefficients d'autocorrélation (R(k)) de ce signal,
- des moyens pour identifier un premier vecteur d'autocorrélation (R₀) ayant pour composantes une première série (k=0,...,N-q) de coefficients d'autocorrélation (R(k)),
- des moyens pour identifier un second vecteur d'autocorrélation (R_q) ayant pour composantes une deuxième série (k=q,...,N) de coefficients d'autocorrélation (R(k)) décalée par rapport à ladite première série d'une valeur de décalage (q) prédéterminée,
- des moyens pour soustraire ledit premier vecteur d'autocorrélation (R₀) dudit second vecteur d'autocorrélation (R_q) afin d'obtenir un vecteur de différentiation (ΔR),

- des moyens pour calculer une norme ($\|\Delta R\|$) dudit vecteur de différentiation, cette norme représentant un premier indicateur d'activité vocale.
2. Dispositif selon la revendication 1, caractérisé en ce qu'il comprend de plus des moyens de réduction pour établir une norme réduite en divisant ladite norme ($\|\Delta R\|$) du vecteur de différentiation par une valeur de réduction, cette norme réduite représentant un deuxième indicateur d'activité vocale. 5 10
3. Dispositif selon la revendication 2 caractérisé en ce que ladite valeur de réduction est égale à l'énergie du signal audio.
4. Dispositif selon la revendication 2 caractérisé en ce que ladite valeur de réduction est égale à la somme de l'énergie du signal audio et d'une valeur plancher (C). 20
5. Dispositif selon l'une quelconque des revendications 1 à 4, caractérisé en ce qu'il comprend des moyens de lissage de l'un desdits indicateurs d'activité vocale pour produire une combinaison linéaire de la valeur présente de cet indicateur et de sa valeur antérieure, ladite combinaison linéaire représentant un troisième indicateur d'activité vocale. 25
6. Dispositif selon l'une quelconque des revendications 1 à 5, caractérisé en ce qu'il comprend des moyens de décision pour produire un signal d'activité vocale si l'un desdits indicateurs excède un seuil de détection. 30
7. Dispositif selon la revendication 6, caractérisé en ce que ledit seuil de détection est établi à partir de la valeur de la norme réduite dudit signal audio en l'absence dudit signal d'activité vocale. 35
8. Dispositif selon l'une quelconque des revendications 1 à 7, caractérisé en ce que ladite norme ($\|\Delta R\|$) du vecteur de différentiation est égale à la somme des valeurs absolues des composantes de ce vecteur. 40
9. Méthode de détection d'activité vocale dans un signal audio comprenant les opérations suivantes : 45
- calcul des coefficients d'autocorrélation ($R(k)$) de ce signal,
 - identification d'un premier vecteur d'autocorrélation (R_0) ayant pour composantes une première série ($k=0, \dots, N-q$) de coefficients d'autocorrélation ($R(k)$),
 - identification d'un second vecteur d'autocorrélation (R_q) ayant pour composantes une deuxième série ($k=q, \dots, N$) de coefficients d'autocorrélation ($R(k)$) décalée par rapport à

- ladite première série d'une valeur de décalage (q) prédéterminée,
- soustraction dudit premier vecteur d'autocorrélation (R_0) dudit second vecteur d'autocorrélation (R_q) afin d'obtenir un vecteur de différentiation (ΔR),
- calcul d'une norme ($\|\Delta R\|$) dudit vecteur de différentiation, cette norme représentant un premier indicateur d'activité vocale.

Patentansprüche

1. Vorrichtung zur Erfassung der Stimmaktivität in einem Tonsignal, welche umfasst: 15
- Mittel zum Berechnen der Autokorrelationskoeffizienten ($R(k)$) dieses Signals,
 - Mittel zum Identifizieren eines ersten Autokorrelationsvektors (R_0) mit einer ersten Folge ($k=0, \dots, N-q$) von Autokorrelationskoeffizienten ($R(k)$) als Komponenten,
 - Mittel zum Identifizieren eines zweiten Autokorrelationsvektors (R_q) mit einer zweiten Folge ($k=q, \dots, N$) von Autokorrelationskoeffizienten ($R(k)$) als Komponenten, die in bezug auf die erste Folge um einen vorbestimmten Verschiebungswert (q) verschoben ist,
 - Mittel zum Subtrahieren des ersten Autokorrelationsvektors (R_0) vom zweiten Autokorrelationsvektor (R_q), um einen Differenzierungsvektor (ΔR) zu erhalten,
 - Mittel zum Berechnen einer Norm ($\|\Delta R\|$) des Differenzierungsvektors, wobei diese Norm einen ersten Indikator für die Stimmaktivität darstellt.
2. Vorrichtung nach Anspruch 1, dadurch gekennzeichnet, dass sie außerdem Reduktionsmittel umfasst, um eine reduzierte Norm aufzustellen, indem die Norm ($\|\Delta R\|$) des Differenzierungsvektors durch einen Reduktionswert dividiert wird, wobei diese reduzierte Norm einen zweiten Indikator für die Stimmaktivität darstellt.
3. Vorrichtung nach Anspruch 2, dadurch gekennzeichnet, dass der Reduktionswert gleich der Energie des Tonsignals ist.
4. Vorrichtung nach Anspruch 2, dadurch gekennzeichnet, dass der Reduktionswert gleich der Summe der Energie des Tonsignals und eines Deckwertes (C) ist.
5. Vorrichtung nach einem der Ansprüche 1 bis 4, dadurch gekennzeichnet, dass sie Glättungsmittel für den einen der Indikatoren für die Stimmaktivität aufweist, um eine Linearkombination des gegenwärtigen

gen Wertes des Indikators und seines vorhergehenden Wertes zu erzeugen, wobei die Linearkombination einen dritten Indikator für die Stimmaktivität darstellt.

6. Vorrichtung nach einem der Ansprüche 1 bis 5, dadurch gekennzeichnet, dass sie Entscheidungsmittel umfasst, um ein Signal der Stimmaktivität zu erzeugen, wenn einer der Indikatoren eine Erfassungsschwelle übersteigt.

7. Vorrichtung nach Anspruch 6, dadurch gekennzeichnet, dass die Erfassungsschwelle ausgehend vom Wert der reduzierten Norm des Tonsignals in Abwesenheit des Signals der Stimmaktivität aufgestellt wird.

8. Vorrichtung nach einem der Ansprüche 1 bis 7, dadurch gekennzeichnet, dass die Norm ($\|\Delta R\|$) des Differenzierungsvektors gleich der Summe der Absolutwerte der Komponenten dieses Vektors ist.

9. Verfahren zur Erfassung einer Stimmaktivität in einem Tonsignal, das die folgenden Vorgänge umfasst:

- Berechnung der Autokorrelationskoeffizienten ($R(k)$) dieses Signals,
- Identifizierung eines ersten Autokorrelationsvektors (R_0) mit einer ersten Folge ($k=0, \dots, N-q$) von Autokorrelationskoeffizienten ($R(k)$) als Komponenten,
- Identifizierung eines zweiten Autokorrelationsvektors (R_q) mit einer zweiten Folge ($k=q, \dots, N$) von Autokorrelationskoeffizienten ($R(k)$) als Komponenten, die in bezug auf die erste Folge um einen vorbestimmten Verschiebungswert (q) verschoben ist,
- Subtraktion des ersten Autokorrelationsvektors (R_0) vom zweiten Autokorrelationsvektor (R_q), um einen Differenzierungsvektor (ΔR) zu erhalten,
- Berechnung einer Norm ($\|\Delta R\|$) des Differenzierungsvektors, wobei diese Norm einen ersten Indikator für die Stimmaktivität darstellt.

Claims

1. Apparatus for detecting voice activity in an audio signal, the apparatus comprising:

means for computing the autocorrelation coefficients ($R(k)$) of the signal;

means for identifying a first autocorrelation vector (R_0) whose components comprise a first series ($k=0, \dots, N-q$) of autocorrelation coefficients ($R(k)$);

means for identifying a second autocorrelation vector (R_q) whose components comprise a second series ($k=q, \dots, N$) of autocorrelation coefficients ($R(k)$) offset from said first series by a predetermined offset value (q);

means for subtracting said first autocorrelation vector (R_0) from said second autocorrelation vector (R_q) to obtain a differentiation vector (ΔR); and

means for computing a norm ($\|\Delta R\|$) of said differentiation vector, which differentiation vector norm represents a first indicator of voice activity.

2. Apparatus according to claim 1, characterized in that it further comprises reduction means for establishing a reduced norm by dividing said differentiation vector norm ($\|\Delta R\|$) by a reduction value, said reduced norm representing a second indicator of voice activity.

3. Apparatus according to claim 2, characterized in that said reduction value is equal to the energy of the audio signal.

4. Apparatus according to claim 2, characterized in that said reduction value is equal to the sum of the energy of the audio signal plus a floor value (C).

5. Apparatus according to any one of claims 1 to 4, characterized in that it includes means for smoothing one of said voice activity indicators to produce a linear combination of the present value of said indicator and its preceding value, said linear combination representing a third indicator of voice activity.

6. Apparatus according to any one of claims 1 to 5, characterized in that it includes decision means for producing a voice activity signal if any one of said indicators exceeds a detection threshold.

7. Apparatus according to claim 6, characterized in that said detection threshold is established on the basis of the value of the reduced norm of said audio signal in the absence of said voice activity signal.

8. Apparatus according to any one of claims 1 to 7, characterized in that said differentiation vector norm ($\|\Delta R\|$) is equal to the sum of the absolute values of the components of said vector.

9. A method of detecting voice activity in an audio signal, the method comprising the following operations:

computing the autocorrelation coefficients ($R(k)$) of the signal;

identifying a first autocorrelation vector (R_0)

whose component comprise a first series
($k=0, \dots, N-q$) of autocorrelation coefficients (R
(k));
identifying a second autocorrelation vector (R_q)
whose components comprise a second series 5
($k=q, \dots, N$) of autocorrelation coefficients (R
(k)) offset from said first series by a predeter-
mined offset value (q);
subtracting said first autocorrelation vector (R_0)
from said second autocorrelation vector (R_q) to 10
obtain a differentiation vector (ΔR); and
computing a norm ($\|\Delta R\|$) of said differentiation
vector, which differentiation vector norm repre-
sents a first indicator of voice activity.

15

20

25

30

35

40

45

50

55

