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

(11) Numéro de publication : **0 518 742 A1**

(12) **DEMANDE DE BREVET EUROPEEN**

(21) Numéro de dépôt : **92401553.0**

(51) Int. Cl.^{5} : **G10L 3/00**

(22) Date de dépôt : **05.06.92**

(30) Priorité : **14.06.91 FR 9107323**

(43) Date de publication de la demande : **16.12.92 Bulletin 92/51**

(84) Etats contractants désignés : **BE CH DE FR GB IT LI LU NL SE**

(71) Demandeur : **SEXTANT AVIONIQUE**
5,7, rue Jeanne Braconnier Parc Tertiaire
F-92360 Meudon-la-Forêt (FR)

(72) Inventeur : **Pastor, Dominique**
THOMSON-CSF, SCPI, Cédex 67
F-92045 Paris la Défense (FR)

(74) Mandataire : **Chaverneff, Vladimir**
THOMSON-CSF SCPI B.P. 329
F-92402 COURBEVOIE CEDEX (FR)

(54) **Procédé de détection d'un signal utile bruité.**

(57) Pour détecter un signal utile bruité, on fait une mesure du rapport S/B attendu de ce signal sur une tranche temporelle, une mesure du bruit blanc seul estimé sur une autre tranche temporelle sans signal utile, on calcule l'énergie moyenne du bruit et du signal bruité, chacun dans sa tranche temporelle, on calcule le seuil de détection théorique, le rapport de ces deux énergies, et on compare le rapport au seuil calculé, ce seuil étant supérieur à 1 (seuil idéal).

Jouve, 18, rue Saint-Denis, 75001 PARIS

La présente invention se rapporte à un procédé de détection d'un signal utile bruité.

Un des problèmes importants en traitement du signal, simple quant à son énoncé, mais d'autant plus complexe quant à sa résolution, consiste à déterminer la présence ou l'absence d'un signal utile noyé dans un bruit additif.

Diverses solutions sont envisageables. On peut utiliser comme variable l'amplitude instantanée du signal reçu ou traité par référence à un seuil déterminé expérimentalement.

On peut aussi utiliser comme variable l'énergie du signal total sur une tranche temporelle de durée T, en seuillant, toujours expérimentalement, cette énergie.

Ces seuillages permettent une première présomption sur la présence ou l'absence du signal. Ils sont de plus applicables à tout signal. Aussi, sont ils complétés par des systèmes de "confirmation", définissant des critères "quasi-certains", propres au type de signal utile, lorsque la nature de celui-ci est connue a priori.

Un tel système complémentaire est largement utilisé en traitement de la parole et peut consister, par exemple, en une extraction de "pitch" ou en l'évaluation de l'énergie minimale d'une voyelle.

La présente invention a pour objet un procédé de détection d'un signal utile bruité, déterminant de façon la plus rigoureuse possible le seuil de détection, et pouvant fonctionner de façon autoadaptative.

Selon l'invention, on dispose du rapport signal/bruit attendu du signal à traiter, et on dispose d'une mesure du bruit seul estimé, mesure numérisée sur M points, ce bruit étant blanc ou rendu blanc, on calcule l'énergie moyenne du bruit sur ces M points, on prend une tranche de N points de signal bruité, on calcule l'énergie moyenne de ces N points, on calcule le seuil de détection théorique, on calcule le rapport des deux dites énergies moyennes, et on compare ce rapport audit seuil.

La présente invention sera mieux comprise à la lecture de la description détaillée d'un mode de réalisation pris à titre d'exemple non limitatif.

On va d'abord expliquer comment doit se faire théoriquement, dans le cas idéal, la détection d'un signal bruité.

On dispose d'une première information u(n) pour une première tranche temporelle telle que :

$$u(n) = s(n) + x(n)$$

n étant un nombre entier : $0 \leqq n \leqq N-1$, s(n) étant un signal utile et x(n) un bruit. En outre, on dispose d'une autre information y(n), avec $0 \leqq n \leqq M-1$, et M pouvant être égal à N ou différent de celui-ci. y(n) est une mesure du bruit x(n) sur une autre tranche temporelle exempte de signal utile.

On pose :

$$U = (u(0)^2 + u(1)^2 + ... + u(N)^2)/N$$

et

$$V = (y(0)^2 + y(1)^2 + ... + y(M)^2)/M$$

et

$$Z = U/V$$

Ainsi, dans un cas idéal et irréaliste, on aurait, en notant RSB = rapport signal à bruit :

$$Z = 1 + RSB$$

et le simple critère de détection serait :

$Z > 1$: présence de signal utile

$Z < 1$: absence de signal utile

Selon la présente invention, on remplace le seuil théorique de 1 par un seuil μ, calculé de la façon expliquée ci-dessous, qui tient compte du fait que les signaux dont on dispose ne sont pas parfaitement ergodiques et que U et V ne sont que des estimées des valeurs vraies des variances $\sigma_u 2$ et $\sigma_x 2$.

Pour effectuer ce calcul de μ, on procède de la façon suivante.

On part du fait que les variables U et V sont de nature aléatoire, et que par conséquent Z l'est aussi. On calcule alors la densité de probabilité de Z (qui dépend du rapport signal sur bruit).

Il s'agit ensuite, en faisant appel au principe du maximum de vraisemblance, de déterminer la meilleure estimation du rapport signal sur bruit après avoir calculé la variable Z.

A cet effet, on mesure sur une tranche temporelle la variable U(n) précitée, et on mesure la variable y(n) sur une autre tranche temporelle où l'on est sûr qu'il n'y a pas de signal utile, mais uniquement du bruit (indépendant et décorrélé de s(n)).

Pour déterminer la densité de la variable aléatoire Z (que l'on peut qualifier de variable observée), on procède de la façon suivante. Soient X_1 appartenant à N (m_1 ; σ_1^2) et X_2 appartenant à N (m_2 ; σ_2^2) deux variables aléatoires gaussiennes indépendantes pour lesquelles les probabilités $P_r\{X_1 < o\}$ et $P_r\{X_2 < o\}$ sont pratiquement nulles.

On pose :

$$m = m_1/m_2, \sigma^2 = \sigma_1^2/\sigma_2^2, \alpha = m_2/\sigma_2.$$

La densité de probabilité f_x (x) de X est alors :

$$f_X(x) = \left(\frac{1}{2\pi}\right)^{-1/2} \frac{mx + \sigma^2}{(x^2 + \sigma^2)^{3/2}} \cdot e^{-\frac{\alpha^2}{2} \cdot \frac{(x-m)^2}{x^2+\sigma^2}} \cdot U(x)$$

où $U(x) = 1$ si $x \geqq o$ et $U(x) = o$ si $x < o$.

Si

$$h(x) = \alpha \cdot \frac{x - m}{(x^2 + \sigma^2)^{1/2}}$$

on a : $P(x) = P_r\{X < x\} = F[h(x)]$, expression dans laquelle $F(x)$ désigne la fonction caractéristique de la variable gaussienne normalisée.

On suppose maintenant que les signaux s(n), x(n) et y(n) sont blancs, gaussiens et centrés.

On pose

$$\sigma_s^2 = E\left[s(n)^2\right]$$

$$\sigma_x^2 = E\left[x(n)^2\right] = E\left[y(n)^2\right] \cdot u(n) = s(n) + x(n).$$

Ce dernier terme est donc, lui aussi, blanc, gaussien et centré ;

et on pose

$$\sigma_u^2 = E\left[u(n)^2\right] = \sigma_s^2 + \sigma_x^2.$$

Puisque l'on définit σ_s^2 et σ_x^2, on suppose implicitement que le calcul de la densité de probabilité se fait à σ_s^2 et σ_x^2 connus. On évalue donc la densité de Z en connaissant σ_s^2 et σ_x^2. Dans ce cas, U et V suivent des lois du chi-2, et, pour N et M suffisamment grands, U et V sont approximées par des lois gaussiennes pratiquement toujours positives :

$$U \text{ appartient à } N\left(\sigma_u^2\ ;\ 2\sigma_u^4/N\right) \text{ et } V \text{ à } N\left(\sigma_x^2\ ;\ 2\sigma_x^4/M\right)$$

Z est donc le rapport de deux variables gaussiennes indépendantes. On peut facilement démontrer que U et V sont indépendantes.

Avec :

$$m_1 = \sigma_u^2,\ \sigma_1^2 = 2\sigma_u^4/N,\ m_2 = \sigma_x^2,\ \sigma_2^2 = 2\sigma_x^4/M$$

il vient :

$$m = \sigma_u^2/\sigma_x^2,\ \sigma^2 = (M/N)\,(\sigma_u^2/\sigma_x^2)^2,\ \alpha = (M/2)^{1/2}.$$

Or : $\sigma_u^2/\sigma_x^2 = 1 + r$ où $r = \sigma_s^2/\sigma_x^2$ est le rapport signal à bruit. Soit $k = M/N$, il vient : $m = r+1$, $\sigma^2 = k(r+1)^2$.

La densité de probabilité de Z, connaissant σ_s^2 et σ_x^2, s'exprime donc par :

1°) $x \geqq 0$

$$f_Z(z:\sigma_s^2, \sigma_x^2) = (M/4\pi)^{1/2}\,(r+1)\,\frac{z+k(1+r)}{[z^2+k(r+1)^2]^{3/2}}\ e^{-\frac{M[z-(r+1)]^2}{4[z^2+k(r+1)^2]}}$$

2°) $x \leqq 0$ d'où: $f_z(z : \sigma_s^2, \sigma_x^2) = 0$

On posera :

$$f_{k,M}(z,r)=(M/4\pi)^{1/2}\ (r+1)\ \frac{z+k(1+r)}{[z^2+k(r+1)^2]^{3/2}}\ e^{-\frac{M[z-(r+1)]^2}{4[z^2+k(r+1)^2]}}\ U(z)$$

de sorte que : $f_z(z:\sigma_s^2, \sigma_x^2) = f_{k,M}(z,\sigma_s^2/\sigma_x^2)$

D'après les résultats ci-dessus relatifs à la densité de probabilité $f_x(x)$, on déduit la probabilité

$$Pr\{Z < z : \sigma_s^2, \sigma_x^2\}.$$

Soit :

$$h_{k.M}(x,r) = (M/2)^{1/2}\frac{x-(r+1)}{[x^2+k(r+1)^2]^{1/2}}$$

Il vient :

$$Pr\{Z < z : \sigma_s^2 ; \sigma_x^2\} = F\{h_{k,M}(x,r)\}.$$

On va maintenant examiner le cas d'un signal quelconque s(n) et d'un bruit blanc gaussien.

On suppose toujours que les bruits x(n) et y(n) sont blancs, gaussiens avec $\sigma_x^2 = E[x(n)^2] = E[y(n)^2]$. Le signal utile s(n) est supposé quelconque, indépendant du bruit.

L'hypothèse nouvelle faite ici est de supposer que s(n) et x(n) sont non corrélés au sens temporel du terme, c'est-à-dire que :

$$c = \frac{\Sigma\ 0 \leqq n \leqq N-1\ s(n)\,x(n)}{(\Sigma\ 0 \leqq n \leqq N-1\ s(n)^2)^{1/2}\ (\Sigma\ 0 \leqq n \leqq N-1\ x(n)^2)^{1/2}} = 0$$

On montre alors que U peut être approximée par :

$$U = \mu_s^2 + (1/N)\ \Sigma\ 0 \leqq n \leqq N-1\ x(n)^2 \text{ et Z par :}$$

$$Z = \frac{\mu_s^2 + (1/N)\ \Sigma\ 0 \leqq n \leqq N-1\ x(n)^2}{(1/M)\ \Sigma\ 0 \leqq n \leqq M-1\ y(n)^2}$$

De même que ci-dessous, le calcul de la densité de Z s'est fait en connaissant σ_s^2et σ_x^2, ici le calcul se fera en connaissant μ_s^2 et σ_x^2. La densité à calculer sera notée par $f_z(z:\mu_s^2, \sigma_x^2)$.

Connaissant μ_s^2, $U = \mu_s^2 + (1/N)\ \Sigma\ 0 \leqq n \leqq N-1\ x(n)^2$ appartient à $\mathbb{N}\ (\mu_s^2 + \sigma_x^2; (2/N)\ \sigma_x^4)$. V appartient à $\mathbb{N}\ (\sigma_x^2 ; (2/M)\ \sigma_x^4)$.

Z = U/V est donc approchée par le rapport de deux lois gaussiennes indépendantes. Comme U et V sont indépendantes, on applique donc le résultat concernant la densité de probabilité de X, avec :

$$m_1 = \mu_s^2 + \sigma_x^2,\ \sigma_1^2 = (2/N)\ \sigma_x^4,\ m_2 = \sigma_x^2,\ \sigma_2^2 = (2/M)\sigma_x^4$$

Donc : $m = r+1$, $\sigma^2 = k$, $\alpha = (M/2)^{1/2}$, avec $k = M/N$ et $r = \mu_s^2/\sigma_x^2$.

La densité de probabilité de Z, connaissant μ_s^2 et σ_x^2 vaut donc :

$$f_z(z:\mu_s^2, \sigma_x^2) = (M/4\pi)^{1/2}\ \frac{(1+r)z+k}{[z^2+k]^{3/2}}\ e^{-\frac{M}{4}\ \frac{[z-(r+1)]^2}{z^2+k}}\ .\ U(z)$$

On posera :

$$f_{k,M}\ (z) = (M/4\pi)^{1/2}\ \frac{(1+r)z+k}{[z^2+k]^{3/2}}\ e^{-\frac{M}{4}\ \frac{[z-(r+1)]^2}{z^2+k}}\ .\ U(z)$$

de sorte que : $f_z(z:\sigma_s^2, \sigma_x^2) = f_{k,M}(z, \sigma_s^2/\sigma_x^2)$

D'après les résultats ci-dessus concernant la densité de probabilité de X, on en déduit la probabilité

$$Pr\{Z < z : \mu_s^2, \sigma_x^2\}.$$

Soit :

$$h_{k,M}(x,r) = (M/2)^{1/2} \frac{x - (r+1)}{[x^2 + k]^{1/2}}$$

il vient :

$$Pr\{ Z < z : \mu_s, \sigma_x^2 \} = F(h_{k,M}(x,r))$$

Selon la présente invention, on met en oeuvre la détection d'activité en faisant appel au maximum de vraisemblance.

Dans les cas de signaux traités, la densité de probabilité de la variable Z, connaissant les énergies du signal utile et du bruit, s'exprime par une fonction de la forme :

$f_{k,M}(z,r)$ où r désigne le rapport signal à bruit. Cette probabilité dépend donc du rapport signal sur bruit. Aussi, la règle de décision ne peut se donner qu'à rapport signal sur bruit attendu. Soit donc r_o ce rapport signal à bruit attendu.

On suppose que la probabilité d'absence de s(n) est π_o et que la probabilité de présence de s(n) est π_1.

Puisqu'on connaît la densité de probabilité $f_{k,M}(z,r)$ la règle de décision optimale est fournie par la théorie générale de la détection et s'exprime par :

$$\frac{\pi_1 f_{k,M}(z,r_o)}{\pi_o f_{k,M}(z,0)} > 1 \Rightarrow D = 1$$

$$\frac{\pi_1 f_{k,M}(z,r_o)}{\pi_o f_{k,M}(z,0)} < 1 \Rightarrow D = 0$$

On peut aussi exprimer cette régle de décision sous la forme :

$$(Z < \mu \Rightarrow D = 0) \text{ et } (Z > \mu \Rightarrow D = 1).$$

Il faut alors déterminer μ et résoudre l'équation :

$$\ln[f_{k,M}(z,r_o)] - \ln[f_{k,M}(z, 0)] - \ln(\pi_o/\pi_1) = 0.$$

On démontre alors que la probabilité d'errreur vaut :

$$Pe = \pi_o [1 - F(h_{k,M}(\mu,0))] + \pi_1 F(h_{k,M}(\mu,r_o)).$$

On va maintenant examiner le cas de la détection d'un signal blanc gaussien dans un bruit lui-même blanc gaussien.

Les signaux s(n), x(n) et y(n) sont supposés blancs, gaussiens, centrés. Soit r_o le rapport signal à bruit attendu, et k = M/N. la probabilité d'absence de s(n) est π_o et la probabilité de présence de s(n) est π_1.

La règle de décision est alors :

Décision D = 1 lorsque :

$$\ln\left[(r+1)\frac{[z+k(r_0+1)]\ (z+k)}{(z+k)[z^2+k(r_o+1)^2]^{3/2}}\right] > (M/4)\left[\frac{[z-(r_o+1)]}{z^2+k(r_o+1)^2} - \frac{(z-1)}{z^2+k}\right] + \ln\frac{\pi_o}{\pi_1}$$

-

Décision D = 0 lorsque :

$$\ln\left[(r+1)\frac{[z+k(r_0+1)]\ (z^2+k)^{3/2}}{(z+k)[z^2+k(r_o+1)^2]^{3/2}}\right] < (M/4)\left[\frac{[z-(r_o+1)]^2}{z^2+k(r_o+1)^2} - \frac{(z-1)^2}{z^2+k}\right] + \ln\frac{\pi_o}{\pi_1}$$

Le seuil étant déterminé pour l'égalité (au lieu d'inégalité) entre les termes de ces deux expressions.

On peut aussi exprimer cette règle de décision sous la forme :

$$(Z < \mu \Rightarrow D = 0) \text{ et } (Z > \mu \Rightarrow D = 1).$$

On obtient par exemple pour μ, à M = N = 128, $\pi_o = \pi_1 = 1/2$:

r_o en dB	μ
-2	1,27
-1	1,34
0	1,41
1	1,50
2	1,68

La probabilité d'erreur est :

$$Pe = \pi_o [1 - F(h_{k,M}(\mu,0))] + \pi_1 F(h_{k,M}(\mu,r_o))$$

avec :

$$h_{k,M}(x,r) = (M/2)^{1/2} \frac{x - (r+1)}{[x^2 + k(r+1)^2]^{1/2}}$$

Nous donnons ci-après quelques valeurs de Pe fonction de r_o . π_o et π_1 sont prises égales à 0,5.

r_o en dB	Pe
-2	0,086
-1	0,052
0	0,028
1	0,013
2	0,005

Dans un exemple de simulation, on a généré un bruit blanc gaussien de variance unité. Pour chaque trame de 128 points (N = M = 128), on a décidé aléatoirement de générer un bruit s(n) additif, présentant un rapport signal sur bruit défini préalablement. Les probabilités d'apparition et d'absence (π_o et π_1) sont égales à 0,5. On a généré un second bruit blanc gaussien de variance unité, qui a servi à calculer la variable aléatoire V. Pour chaque trame, on a calculé Z. On a appliqué alors la règle de décision et l'on a compté le nombre d'erreurs.

r_o en dB	Nombre d'erreurs sur 1000 itérations
-2	73
-1	43
0	18
1	10
2	2

Ces résultats corroborent ceux prévus par le calcul théorique.

On va maintenant examiner le cas d'un signal quelconque s(n) et d'un bruit blanc gaussien.

On suppose toujours que les bruits x(n) et y(n) sont blancs, gaussiens avec $\sigma_x^2 = E[x(n)^2]=E[y(n)^2]$. Le signal utile s(n) est supposé quelconque, indépendant du bruit. Soit r_o le rapport signal à bruit attendu, k = M/N. La probabilité d'absence de s(n) est π_o et celle de présence de s(n) est π_1.

La règle de décision est alors :

Décision D = 1 lorsque :

$$\ln \frac{(r+1)\,z+k}{z+k} > (M/4)\,\frac{[z-(r_o+1)]^2 - (z-1)^2}{z^2+k} + \ln \frac{\pi_o}{\pi_1}$$

Décision D = 0 lorsque :

$$\ln \frac{(r+1)\,z+k}{z+k} < (M/4)\,\frac{[z-(r_o+1)]^2 - (z-1)^2}{z^2+k} + \ln \frac{\pi_o}{\pi_1}$$

On peut aussi exprimer cette règle de décision sous la forme :

$$(Z < \mu \Rightarrow D = 0) \text{ et } (Z > \mu \Rightarrow D = 1).$$

On obtient pour μ les valeurs suivantes en fonction de r_o, pour M = N = 128, $\pi_o = \pi_1 = 1/2$.

r_o en dB	μ
-2	1,30
-1	1,38
0	1,48
1	1,60
2	1,76

De plus, on obtient :

$$Pe = \pi_o [1 - F (h_{k,M}(\mu,0))] + \pi_1 F (h_{k,M}(\mu, r_o))$$

avec :

$$h_{k,M}(x,r) = (M/2)^{1/2} \frac{x - (r+1)}{[x^2 + k]^{1/2}}$$

Nous donnons ci-après quelques valeurs de Pe en fonction de r_o. Les probabilités π_o et $\pi_o$sont prises égales à 0,5.

r_o en dB	Pe
-2	0,062
-1	0,032
0	0,013
1	0,004
2	0,001

Dans un exemple de simulation, pour chaque trame de 128 points de bruit blanc généré (N = M = 128), on a décidé aléatoirement d'y ajouter s(n), qui est ici, une sinusoïde, présentant un rapport signal sur bruit défini préalablement. π_1 et π_o sont prises égales à 0,5.

On a généré un second bruit blanc gaussien de variance unité, servant à calculer V. Pour chaque trame, on a calculé Z et on a appliqué la règle de décision précitée. On a compté le nombre d'erreurs.

On a obtenu les résultats suivants

r_o en dB	Nombre d'erreurs sur 1000 itérations
-2	70
-1	37
0	12
1	6
2	3

Ces résultats corroborent ceux prévus par le calcul théorique.

Les résultats précédents, parce que très généraux, permettent la détection de signaux noyés dans du bruit additif, même lorsque le rapport signal sur bruit est faible, voisin de 0 dB.

On va décrire ci-dessous une application dans laquelle ce type de détection peut se révéler très utile.

Les algorithmes présentés s'appliquent au cas de la parole, comme pré-système de détection d'activité vocale.

Le choix du seuil de détection dépend du contexte.

En ce qui concerne les bandes audio utilisées, une caractérisation préalable du bruit et de la parole, à l'aide de mesures basées sur l'estimation par maximum de vraisemblance montre que le signal vocal à détecter présente un rapport signal sur bruit d'au moins 6 dB.

D'autre part, le système de traitement utilise des trames de signal de 128 points, la fréquence d'échantillonnage étant de 10 kHz.

Les variables U et V sont toutes les deux évaluées sur 128 points de sorte que M = N = 128.

D'après ce qui précède, on déduit le seuil théorique de détection à 3.

Cependant, on ne peut pas se contenter de cet unique seuil. En effet, si le bruit est relativement stationnaire, il présente des instationnarités à prendre en compte pour renouveler la variable V, ce qui permet de rendre l'algorithme partiellement adaptatif.

On introduit donc un second seuil, qui permet de décider si la variable V va être renouvelée ou non.

Ce second seuil est choisi à 1,25, ce qui correspond à un bruit additif au bruit stationnaire présentant un rapport signal à bruit de -2 dB.

La règle de décision est alors :

<u>Si Z < 1,25</u> :

Alors la trame traitée est composée du même bruit que celle utilisée comme référence. La variable V est remplacée par la valeur de l'énergie de la trame traitée.

On notera que, puisque la décision est de considérer la trame traitée comme du bruit représentatif, on pourrait renouveler la variable V en faisant la moyenne de l'ancienne valeur de V et de l'énergie de la trame considérée. Ce qui amène à changer la valeur de M (nombre de points sur lequel est évalué V) mais cette opération peut induire un mauvais fonctionnement de l'algorithme.

<u>Si 1,25 < z > 3</u> :

La trame est considérée comme contenant une non-stationnarité du bruit, et exempte de parole.

<u>Si 3 < Z</u> :

La trame est considéré comme de la parole.

Des essais effectués sur des échantillons de signaux bruités ont validé cette détection.

Cependant, rappelons que cette détection vocale peut être améliorée par l'utilisation de critères propres au signal de parole, tel que le calcul de "pitch".

L'algorithme proposé ici concerne l'étude de quelques exemples de signaux. Il est évident que pour d'autres signaux de parole présentant des rapports signal à bruit différents, un nouveau choix de seuils est nécessaire.

L'utilisation de deux seuils est généralement préférable.

Une application de cet algorithme permet de créer des fichiers de référence corrects pour le système de reconnaissance vocale étudié. Une segmentation précise des élocutions est alors nécessaire.

Dans une application, on a utilisé un alternat micro (ouverture et fermeture micro) qui fournit une segmentation grossière des élocutions.

L'algorithme précédent a été utilisé pour affiner cet alternat. Une première passe de l'algorithme a permis de préciser le début de l'élocution. Une seconde passe a consisté à lire le fichier de parole "à l'envers", c'est-à-dire en partant de la fermeture micro vers l'ouverture micro. Ce qui a permis alors de préciser la fin de l'élocution.

Cette utilisation non causale de l'algorithme est nécessaire, car la détection d'activité est suffisamment précise pour détecter, à l'intérieur des mots, la présence de silences, ce qui est préjudiciable à une mise en place d'une segmentation pour les apprentissages.

Le même type d'application permet aussi de segmenter les fichiers de parole sur lesquels on effectue une reconnaissance.

Cependant, cet algorithme n'est évidemment pas causal, ce qui est préjudiciable à une utilisation temps réel. D'où la nécessité de compléter cet algorithme par un calcul propre au traitement de la parole.

Nous avons démontré l'existence de seuils optimaux de détection, ce qui permet d'avoir une approche théorique du problème de l'estimation du rapport signal sur bruit et, surtout de la détection, dans le cas d'un bruit blanc et d'un signal connu seulement par son énergie sur N points lorsque celle-ci reste relativement stationnaire.

Revendications

1. Procédé de détection d'un signal utile bruité, pour lequel on dispose du rapport signal/bruit attendu du

signal à traiter, et d'une mesure du bruit seul estimé, mesure numérisée sur M points, caractérisé par le fait qu'on calcule l'énergie moyenne du bruit sur ces M points, on prend une tranche de N points de signal bruité, on calcule l'énergie moyenne de ces N points, on calcule le seuil de détection théorique, on calcule le rapport (Z) des deux dites énergies moyennes, et on compare ce rapport audit seuil.

2. Procédé selon la revendication 1, caractérisé par le fait que le bruit seul estimé est blanc ou rendu blanc.

3. Procédé selon l'une des revendications précédentes, caractérisé par le fait que le seuil de détection théorique est déterminé pour :

$$\ln \frac{(r+1)z+k}{z+k} = (M/4)\,\frac{[z-(r_o+1)]^2 - (z-1)^2}{z^2+k} + \ln \frac{\pi_o}{\pi_1}$$

r_o étant le rapport signal à bruit attendu, $k = M/N$, π_o étant la probabilité d'absence du signal utile et π_1 sa probabilité de présence.

4. Procédé selon l'une des revendications 1 ou 2 pour la détection d'un signal blanc gaussien, caractérisé par le fait que le seuil de détection théorique est déterminé pour :

$$\ln\left[(r+1)\,\frac{[z+k(r_o+1)]\,(z^2+k)^{3/2}}{(z+k)\,[z^2+k(r_o+1)^2]^{3/2}}\right] = (M/4)\left[\frac{[z-(r_o+1)]^2}{z^2+k(r_o+1)^2} - \frac{(z-1)^2}{z^2+k}\right] + \ln\frac{\pi_o}{\pi_1}$$

5. Procédé selon l'une des revendications 1 à 3, pour la détection de la parole, caractérisé par le fait qu'en plus du seuil de détection théorique on utilise un second seuil de décision de mise à jour de la tranche mesurée du bruit seul estimé, afin de tenir compte des instationnarités du bruit.

Office européen des brevets

RAPPORT DE RECHERCHE EUROPEENNE

Numero de la demande

EP 92 40 1553

DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes	Revendication concernée	CLASSEMENT DE LA DEMANDE (Int. Cl.5)
A	US-A-4 410 763 (L. STRAWCZYNSKI et al.) * Résumé * ---	1	G 10 L 3/00
A	IEEE TRANSACTIONS ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING, vol. ASSP-31, no. 3, juin 1983, pages 678-684, New York, US; P. DE SOUZA: "A statistical approach to the design of an adaptive self-normalizing silence detector" * Paragraphe III: "Training and adaptation" * ---	1	
A	IBM TECHNICAL DISCLOSURE BULLETIN, vol. 29, no. 12, May 1987, pages 5606-5609, Armonk, New York, US; "Digital signal processing algorithm for microphone input energy detection having adaptive sensitivity" * Page 5607, lignes 22-35 * ---	1	DOMAINES TECHNIQUES RECHERCHES (Int. Cl.5)
A	US-A-4 052 568 (J.A. JANKOWSKI) * Résumé * -----	1,5	G 10 L 3/00

Le présent rapport a été établi pour toutes les revendications

Lieu de la recherche	Date d'achèvement de la recherche	Examinateur
LA HAYE	04-05-1992	ARMSPACH J.F.A.M.

CATEGORIE DES DOCUMENTS CITES

X : particulièrement pertinent à lui seul
Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie
A : arrière-plan technologique
O : divulgation non-écrite
P : document intercalaire

T : théorie ou principe à la base de l'invention
E : document de brevet antérieur, mais publié à la date de dépôt ou après cette date
D : cité dans la demande
L : cité pour d'autres raisons

& : membre de la même famille, document correspondant

EPO FORM 1503 03.82 (P0402)