

A1

**DEMANDE  
DE BREVET D'INVENTION**

(21)

**N° 79 24227**

---

(54) Dispositif de détection de signaux vocaux et système d'alternat comportant un tel dispositif.

(51) Classification internationale (Int. Cl. <sup>3</sup>). G 10 L 1/04.

(22) Date de dépôt..... 28 septembre 1979.

(33) (32) (31) Priorité revendiquée :

(41) Date de la mise à la disposition du  
public de la demande..... B.O.P.I. — « Listes » n° 15 du 10-4-1981.

---

(71) Déposant : Société dite : THOMSON-CSF, société anonyme, résidant en France.

(72) Invention de : François Dumont.

(73) Titulaire : *Idem* (71)

(74) Mandataire : Société dite Thomson-CSF, S.C.P.I.,  
173, bd Haussmann, 75360 Paris Cedex 06.

La présente invention concerne les dispositifs de détection de signaux vocaux dans un signal d'entrée, en particulier ceux qui équipent les systèmes d'alternat.

De tels dispositifs fournissent une information  
5 logique, présence ou absence de parole, pouvant être utilisée pour commander, par exemple, un interrupteur afin de discriminer les signaux vocaux du signal d'entrée.

Toute troncature de début ou de fin d'activité vocale est en général supprimée par l'adjonction d'une  
10 ligne à retard précédant l'interrupteur et retardant le signal d'entrée audiofréquence.

Il est actuellement connu de réaliser des dispositifs simples de commande d'alternat à la voix basés sur la comparaison de la seule énergie du signal d'entrée par rapport à un seuil fixe. Mais les performances de tels dis-  
15 positifs en milieu bruyant sont insuffisantes.

Des dispositifs plus performants, basés essentiellement sur l'analyse spectrale du signal sont également connus, mais ils donnent lieu à des réalisations complexes.

20 La présente invention a pour objet de remédier à ces inconvénients à l'aide de moyens simples.

Selon l'invention, un dispositif de détection de signaux vocaux dans un signal d'entrée en milieu bruyant, est caractérisé en ce qu'il comporte : des moyens d'asservissement non linéaire d'une tension de seuil sur la valeur  
25 de l'amplitude crête du signal d'entrée ; des moyens d'atténuation du signal d'entrée dans un rapport  $k$  ( $k$  entier positif non nul) ; des moyens de comparaison de l'amplitude du signal d'entrée atténué par rapport à la tension  
30 de seuil variable, ayant une sortie délivrant un premier signal de commande intermédiaire lorsque l'amplitude du signal d'entrée atténué est supérieure à la tension de seuil ; et des moyens de décision couplés à la sortie des moyens de comparaison et effectuant une analyse tem-

porelle des dépassements de la tension de seuil par l'amplitude du signal d'entrée atténué, afin de délivrer un signal de commande correspondant à une décision de présence de signaux vocaux.

5 L'invention sera mieux comprise et d'autres caractéristiques apparaîtront à l'aide de la description ci-après et des schémas s'y rapportant sur lesquels :

- la figure 1 est un exemple d'un mode de réalisation du dispositif selon l'invention ;

10 - la figure 2 représente le chronogramme des signaux permettant d'expliquer le fonctionnement du dispositif de la figure 1.

Sur la figure 1 une borne d'entrée 1, destinée à recevoir un signal d'entrée constitué par un signal de parole mêlé de bruit, est couplée à l'entrée d'un amplificateur 3 à travers un redresseur mono-alternance 2. La sortie de l'amplificateur 3 est couplée à l'entrée "+" d'un comparateur  $C_2$  à travers une résistance  $R_3$  et à l'entrée "+" d'un comparateur  $C_1$ . L'entrée "+" du comparateur  $C_2$  est en outre couplée à la masse à travers une résistance  $R_4$ . Les entrées "-" des comparateurs  $C_1$  et  $C_2$  sont couplées à la masse à travers un condensateur C et à une extrémité d'un commutateur 8. Ce commutateur 8 permet de coupler cette extrémité soit à une source de tension "+V" à travers une résistance  $R_1$ , soit à la masse à travers une résistance  $R_2$ . La sortie du comparateur  $C_1$  est couplée à l'entrée de commande de ce commutateur 8 à travers un premier monostable 4. La sortie du comparateur  $C_2$  est couplée à une borne de sortie 10 à travers un deuxième monostable 5, un troisième monostable 6 et un quatrième monostable 7 disposés en série.

L'ensemble formé par le comparateur  $C_1$ , le premier monostable 4, le commutateur 8, le condensateur C, les résistances  $R_1$  et  $R_2$  constituent des moyens 11 d'asser-

vissement non linéaire d'une tension de seuil.

De même, les deuxième, troisième et quatrième monos-tables 5, 6 et 7 forment des moyens de décision 12.

Avant de décrire le fonctionnement de ce dispositif,  
5 il est utile de rappeler tout d'abord que les différents sons constituant un signal vocal peuvent être classés suivant leur mode de production, en sons voisés et non voisés.

Les sons voisés sont constitués soit par des voyelles,  
10 soit par des consonnes liquides ou sonores, et possèdent des propriétés spectrales bien définies qui ne se retrouvent pas dans les sons non voisés constitués par les consonnes sourdes. Les sons voisés, dont l'amplitude est généralement beaucoup plus importante que celle des sons  
15 non voisés, présentent une périodicité très marquée au niveau du signal temporel. Cette périodicité qui n'existe pas pour les sons non voisés est due à la vibration des cordes vocales. La valeur de la fréquence correspondant à cette périodicité, appelée fréquence de mélodie, peut  
20 être comprise, suivant les personnes, entre environ 60 et 300 Hz.

Il est donc clair qu'il est beaucoup plus facile de discriminer du bruit les sons voisés plutôt que les sons non voisés qui, de par leurs caractéristiques spectrales  
25 et d'amplitudes, seront généralement "noyés" dans le bruit de fond.

Pour garantir des performances maximales, le dispositif représenté sur la figure 1 permet d'optimiser la discrimination pour les sons voisés. Pour cela, la déci-  
30 sion présence ou absence de parole est élaborée à partir de l'analyse de l'amplitude instantanée du signal vocal plutôt que son énergie et se fait en trois étapes :

a) Asservissement non linéaire d'une tension de seuil sur la valeur de l'amplitude crête du signal d'entrée

afin de s'affranchir de la valeur absolue du bruit de fond et de ne garder comme paramètre de discrimination que la valeur relative du rapport d'amplitude parole sur bruit. Cet asservissement est effectué par les moyens  
5 d'asservissement 11 qui délivrent une tension de seuil, soit exponentiellement croissante en fonction du temps suivant une constante de temps  $\tau_m$  lorsque l'amplitude instantanée du signal d'entrée audiofréquence est supérieure à la tension de seuil, soit exponentiellement  
10 décroissante suivant une constante de temps  $\tau_d$  lorsque l'amplitude instantanée du signal d'entrée est inférieure à la tension de seuil.

b) Comparaison de l'amplitude instantanée du signal d'entrée audiofréquence atténué d'un rapport  $k$  ( $k$  :  
15 entier non nul) par rapport à la tension de seuil générée par les moyens d'asservissement 11. Cette comparaison est effectuée par le comparateur  $C_2$  qui délivre un signal logique de niveau 1 lorsque cette amplitude est supérieure à la tension de seuil.

20 c) Analyse temporelle du signal de sortie du comparateur  $C_2$  de façon à exploiter les caractéristiques particulières des sons voisés, et extension par proximité de la décision d'activité vocale (procédure décrite dans le brevet THOMSON-CSF 78 03606) pour remédier aux tronca-  
25 tures des sons non voisés. En effet les sons non voisés que l'on peut à quelques exceptions près assimiler aux consonnes sont presque toujours directement précédés ou suivis par une voyelle, c'est-à-dire un son voisé. Dans ces conditions, il suffit pour éviter les troncatu-  
30 res de prolonger, de quelques dizaines de millisecondes de chaque côté, la décision prise sur un son voisé. En pratique, le signal d'entrée est retardé dans une ligne à retard et la décision d'activité prise sur un son voisé est pro-  
35 longée de deux fois le retard apporté par cette ligne à retard.

En outre, l'analyse temporelle du signal de sortie du comparateur  $C_2$  se fait suivant un double critère temporel qui permet d'exploiter d'une part la fréquence de mélodie par test du temps maximal séparant deux transitions de l'état "0" à l'état "1" du signal analysé, et d'autre part la fréquence syllabique par test du temps pendant lequel ces transitions rythmées existent.

L'analyse temporelle du signal de sortie du comparateur  $C_2$  et l'extension par proximité de la détection d'activité vocale sont effectuées par les moyens de décision 12.

Ces moyens de décision 12 décident qu'il y a présence de sons voisés quand les trois conditions suivantes sont remplies simultanément :

- 1) L'amplitude instantanée du signal audiofréquence dépasse la tension de seuil variable
- 2) Le temps séparant deux transitions de l'état "0" à l'état "1" du signal de sortie du comparateur ne dépasse pas un majorant  $T_m$  de la période de mélodie
- 3) Ces transitions durent un temps au moins égal à un minorant  $T_s$  de la période syllabique.

Ces conditions ont été établies de façon à être satisfaites en présence de sons voisés d'énergie suffisante, à condition toutefois que la tension de seuil ne varie que très peu pendant la détection.

Ceci entraîne que la constante de temps  $\tau_m$  d'accroissement de la tension de seuil variable doit être supérieure à la durée syllabique. Pour cela la valeur de cette constante de temps  $\tau_m$  à la montée est choisie de l'ordre de 5s. D'autre part, pour garder les mêmes performances de détection sur un débit vocal continu, il est nécessaire que la tension de seuil redescende rapidement au niveau du bruit entre deux syllabes et donc que la constante de temps de descente  $\tau_d$  soit beaucoup plus faible que celle

de montée. Pour cela la valeur de cette constante  $\tau_d$  est choisie de l'ordre de 200ms.

De plus, compte tenu des critères de détection précédents, l'immunité aux bruits repose sur les moyens d'asservissement de la tension de seuil. La tension de seuil doit avoir une valeur telle qu'en présence du bruit au moins une des trois conditions de détection précédentes ne soit pas vérifiée.

Il peut être mis en évidence trois type de bruits différents facilement rejectables en imposant des conditions d'asservissement de la tension de seuil.

Ainsi pour les bruits du type "stationnaire modulé" dont les variations d'énergie moyenne sont très lentes, mais qui présentent une modulation importante à fréquence supérieure ou égale à la fréquence de mélodie, les conditions de décision ② et ③ sont en général vérifiées. Afin d'éviter que la condition ① soit aussi vérifiée en présence d'un tel bruit, la tension de seuil doit être régulée sur les crêtes d'amplitude dont la fréquence de récurrence peut être confondue avec la fréquence de mélodie. Cette condition peut être réalisée en prolongeant la commande d'accroissement de la tension de seuil d'un temps au moins égal à  $T_m$  (période maximale de mélodie), après que l'amplitude instantanée du signal audiofréquence soit resdescendue en dessous de la tension de seuil. Ce rôle est rempli par le monostable 4 de la figure 1.

Les bruits du type "stationnaire blanc", dont les variations d'énergie moyenne, comme ceux de la catégorie précédente, sont assez lentes, ne présentent pas de périodicité mélodique mais une répartition d'amplitudes dont la densité de probabilité décroît rapidement quand l'amplitude augmente (type gaussien par exemple). Pour de tels bruits, l'amplitude instantanée pouvant prendre de très grandes valeurs, la condition de décision ① est en général véri-

fiée. Il est toutefois possible de rejeter un tel bruit si la probabilité d'avoir pendant le temps  $T_s$  plusieurs dépassements de la tension de seuil par l'amplitude du signal séparés par un temps inférieur à  $T_m$ , est quasi nulle (c'est-à-dire que les conditions ② et ③ ne soient pas vérifiées). Ceci peut être réalisé en atténuant le signal qui est comparé à la tension de seuil pour la décision d'activité vocale, par rapport au signal qui sert à générer la tension de seuil. En effet, en présence d'un bruit de ce type, le système d'asservissement régule la valeur moyenne de la tension de seuil en fonction de la probabilité de dépassement de la tension de seuil par l'amplitude du signal d'entrée non atténué. Cette probabilité est fixée par les paramètres  $T_m$  et  $\tau_m/\tau_d$ . Le choix du facteur d'atténuation  $k$  permet donc de faire tendre vers zéro la probabilité de dépassement de la tension de seuil par le signal atténué quand il s'agit d'un bruit, sans pratiquement modifier cette probabilité d'intersection en présence de sons voisés. Ce rôle d'atténuation est rempli par le diviseur potentiométrique composé des deux résistances  $R_3$  et  $R_4$ .

Enfin, les bruits de type "impulsionnel" dont la durée est inférieure à  $T_s$  quelles que soient leurs énergies ou les fréquences qu'ils contiennent, produisent généralement des dépassements de la tension de seuil par l'amplitude instantanée du signal, et satisfont ainsi aux conditions de décision ① et ②. Seule la condition ③ permet de rejeter ces bruits.

La réalisation pratique de moyens de décision mettant en oeuvre les critères de décision ①, ② et ③, et de moyens d'asservissement d'une tension de seuil présentant les caractéristiques définies précédemment, peut se faire de façon simple en utilisant les éléments décrits sur la figure 1. Le chronogramme de fonctionnement de ces éléments est donné sur la figure 2.



Sur cette figure 2 sont représentés des signaux BF ,  
 A , B ,  $BF_a$  , D , E , F et G respectivement présents à  
 la sortie de l'amplificateur 3, à la sortie du comparateur  
 $C_1$ , à la sortie du monostable 4, à l'entrée "+" du com-  
 5 parateur  $C_2$  , à la sortie du comparateur  $C_2$  , à la sortie  
 du monostable 5, à la sortie du monostable 6, et à la  
 sortie du monostable 7.

L'amplitude de la tension de seuil (notée S) à  
 été superposée aux signaux BF et  $BF_a$  .

10 Le signal BF, présent, à la sortie de l'amplificateur  
 3, correspond au signal d'entrée audiofréquence redressé  
 mono-alternance par le redresseur 2, puis amplifié par  
 l'amplificateur 3 pour couvrir la plage d'amplitude ma-  
 ximale. L'amplitude de ce signal BF est ensuite comparée  
 15 à la tension de seuil variable S par le comparateur  $C_1$  .  
 Ce comparateur fournit le signal logique A qui passe à  
 l'état haut quand l'amplitude instantanée du signal BF est  
 supérieure à la tension de seuil et à l'état bas dans le  
 cas contraire. Ce signal est appliqué à l'entrée du monos-  
 20 table 4 dont le signal logique de sortie B passe à l'état  
 haut en même temps que son signal d'entrée, mais ne repasse  
 à l'état bas que  $T_m$  millisecondes après que son signal  
 d'entrée soit stabilisé à l'état bas.  $T_m$  qui est égal à  
 16 ms représente un majorant de la période de mélodie pré-  
 25 sente dans le signal vocal.

Le signal de sortie B du monostable 4 commande un  
 intégrateur dissymétrique constitué des résistances  $R_1$  et  
 $R_2$  , du commutateur 8 et du condensateur C. La sortie 13  
 de cet intégrateur délivre la tension de seuil variable  
 30 de telle façon que lorsque le signal de commande B est  
 à l'état haut, la valeur de la tension de seuil s'accroisse  
 selon une loi exponentielle de constante de temps  $R_1 C$  de  
 l'ordre de 5s et, lorsque le signal de commande est à  
 l'état bas la valeur de la tension de seuil diminue sui-

vant la même loi, mais avec une constante de temps  $R_2C$  de l'ordre de 200 ms.

Le diviseur potentiométrique formé par les résistances  $R_3$  et  $R_4$  permet d'atténuer d'un rapport  $k$  le signal de sortie BF de l'amplificateur 3. Le signal atténué  $BF_a$  est ensuite comparé à la tension de seuil variable par le comparateur  $C_2$ . Le signal de sortie D de ce comparateur  $C_2$  est appliqué à l'entrée d'un monostable 5 dont le signal logique de sortie E passe à l'état haut quand la valeur de l'amplitude du signal  $BF_a$  devient supérieure à la tension de seuil, et ne repasse à l'état bas que  $T_m$  millisecondes après que l'amplitude du signal  $BF_a$  soit redescendue en dessous du seuil. Le signal logique E est appliqué à l'entrée du monostable 6. Le signal logique de sortie F de ce monostable 6 ne passe à l'état haut que  $T_s$  millisecondes après que son signal d'entrée soit stabilisé à l'état haut, et ne repasse à l'état bas qu'en synchronisme avec son signal d'entrée. Le dernier monostable 7 fournit un signal logique G qui passe à l'état haut en même temps que le signal de sortie F du monostable 6 et indique ainsi la présence de sons voisés. Ce signal G indique une fin de ces sons voisés quand il repasse à l'état bas,  $T_p$  millisecondes après le signal de sortie F du monostable 6,  $T_p$  est égal à  $T_s$  plus deux fois 20 ms constituant l'extension par proximité. Pour discriminer les signaux de parole du signal BF, il suffira, par exemple, de commander un interrupteur (non représenté) par le signal G.

De plus une ligne à retard (non représentée) associée au détecteur retardera le signal audiofréquence BF de  $T_s + 20$  ms, afin de supprimer le décalage entre le signal de décision G et le signal audiofréquence BF.

Le dispositif qui vient d'être décrit présente l'avantage d'une immunité quasi totale à tout bruit présentant des variations lentes quel qu'en soit le niveau et à tout

bruit impulsionnel de durée inférieure à environ 50 ms  
quelle qu'en soit l'amplitude.

De plus la détection de sons voisés se fait avec  
une bonne précision temporelle malgré un rapport signal  
5 sur bruit faible.

Enfin le dispositif décrit présente la possibilité  
de supprimer toute troncature de début ou de fin d'acti-  
vité vocale grâce à l'adjonction d'une ligne à retard  
permettant de retarder le signal audiofréquence.

10 L'invention n'est pas limitée au mode de réalisation  
décrit et représenté, en particulier, il est à la portée  
de l'homme de l'art de réaliser des moyens d'asservisse-  
ment non linéaires qui génèrent une tension de seuil pré-  
sésentant les caractéristiques exposées.

REVENDEICATIONS

1. Dispositif de détection de signaux vocaux dans un signal d'entrée en milieu bruyant, caractérisé en ce qu'il comporte : des moyens d'asservissement non linéaire d'une tension de seuil sur la valeur de l'amplitude crête du signal d'entrée ; des moyens d'atténuation du signal d'entrée dans un rapport  $k$  ( $k$  entier positif non nul) ; des moyens de comparaison de l'amplitude du signal d'entrée atténué par rapport à la tension de seuil variable ayant une sortie délivrant un premier signal de commande intermédiaire lorsque l'amplitude du signal d'entrée atténué est supérieure à la tension de seuil ; et des moyens de décision couplés à la sortie des moyens de comparaison et effectuant une analyse temporelle des dépassements de la tension de seuil par l'amplitude du signal d'entrée atténué, afin de délivrer un signal de commande correspondant à une décision de présence de signaux vocaux.

2. Dispositif de détection selon la revendication 1, caractérisé en ce que les moyens d'asservissement non linéaire comportent : un générateur de tension ayant une entrée de commande, et une sortie délivrant une tension soit exponentiellement croissante en fonction du temps suivant une première constante de temps supérieure à la durée syllabique maximum, soit exponentiellement décroissante suivant une deuxième constante de temps supérieure à la période maximale de mélodie, la tension délivrée par ce générateur servant de tension de seuil ; des moyens de commande du générateur ayant une première et une deuxième entrée recevant respectivement le signal d'entrée et la tension de seuil, et une sortie couplée à l'entrée de commande du générateur et délivrant un signal de commande auxiliaire pour commander le sens de variation de la tension de seuil générée par le générateur.

3. Dispositif de détection selon la revendication 1, caractérisé en ce que les moyens de décision compor-

tent : un premier temporisateur ayant une entrée couplée à la sortie des moyens de comparaison et une sortie délivrant un deuxième signal de commande intermédiaire correspondant au premier signal prolongé d'une durée égale à la période maximale de mélodie, un deuxième temporisateur ayant une entrée couplée à la sortie du premier temporisateur et une sortie délivrant le signal de commande lorsque la durée du premier signal est supérieure à la durée syllabique minimum, ce signal de commande correspondant à la présence de sons voisés dans le signal d'entrée.

4. Dispositif de détection selon la revendication 2, caractérisé en ce que les moyens de commande comportent : un comparateur ayant une première entrée recevant le signal d'entrée et une deuxième entrée recevant la tension de seuil délivrée par le générateur et une sortie délivrant un signal de commande auxiliaire intermédiaire lorsque l'amplitude du signal d'entrée est supérieure à la tension de seuil ; et un temporisateur ayant une entrée couplée à la sortie du comparateur, et une sortie délivrant le signal de commande auxiliaire correspondant au signal de commande auxiliaire intermédiaire prolongé d'une durée supérieure à la période maximale de mélodie.

5. Dispositif de détection selon la revendication 2, caractérisé en ce que le générateur de tension comporte : un commutateur ayant une entrée de commande correspondant à l'entrée de commande du générateur ; une capacité ayant une première extrémité couplée à la masse et une deuxième extrémité couplée soit à la masse à travers une première résistance, soit à une source de tension à travers une deuxième résistance, cette deuxième extrémité constituant la sortie du générateur.

6. Système d'alternat comportant un dispositif de détection selon l'une des revendications 1 à 5.



