

A1

**DEMANDE
DE BREVET D'INVENTION**

(21)

N° 81 19123

(54) Procédé et dispositif de synthèse vocale avec compression de l'information binaire stockée en mémoire et application à la transmission d'alarmes par le réseau téléphonique autocommuté.

(51) Classification internationale (Int. Cl.³). G 10 L 1/08.

(22) Date de dépôt..... 6 octobre 1981.

(33) (32) (31) Priorité revendiquée :

(41) Date de la mise à la disposition du
public de la demande..... B.O.P.I. — « Listes » n° 14 du 8-4-1983.

(71) Déposant : DURANTON René. —

(72) Invention de : René Duranton.

(73) Titulaire : *Idem* (71)

(74) Mandataire : Cabinet Moutard,
35, av. Victor-Hugo, 78180 Voisins-le-Bretonneux.

- 1 -

Procédé et dispositif de synthèse vocale avec compression de l'information binaire stockée en mémoire et application à la transmission d'alarmes par le réseau téléphonique autocommuté.

L'invention se rapporte à la synthèse vocale.

Dans la demande de brevet français N° 80 12420 déposée le 4 Juin 1980 par le Demandeur, on a décrit l'application, à la
5 transmission d'alarmes par le réseau téléphonique autocommuté, d'un dispositif de synthèse vocale comportant une mémoire morte qui contient les codes binaires nécessaires à la génération de messages vocaux et un circuit encodeur de voix. Une mémoire vive contient les informations spécifiques
10 relatives aux messages particuliers à transmettre et un microprocesseur effectue la gestion des échanges entre ces mémoires et les moyens de transmission des messages vocaux sur la ligne téléphonique. Les codes binaires sont enregistrés dans la mémoire morte en démodulant un signal vocal
15 naturel. Le prix de la mémoire morte devient important dès que l'on veut engendrer des messages d'une certaine longueur.

Pour permettre l'utilisation d'une mémoire morte de capacité
20 réduite dans la synthèse de messages vocaux de longueur donnée, l'invention propose de découper le signal binaire, obtenu par échantillonnage, dont on part pour l'enregistrement des codes binaires, en une succession de tranches d'égale longueur, de prélever une fraction entière

- 2 -

prédéterminée de chaque tranche de signal binaire pour la stocker en mémoire morte et, au moment où s'effectue la synthèse d'un message vocal, d'effectuer des lectures successives de celles desdites fractions, qui correspondent 5 au message, lesdites lectures étant en nombre tel que la longueur initiale desdites tranches soit restituée et de procéder finalement à l'encodage de la voix par modulation à l'aide des signaux binaires ainsi lus.

10 De préférence, le prélèvement porte sur la moitié de chaque tranche et deux lectures successives des moitiés ainsi prélevées sont effectuées.

Selon un mode d'exécution préféré de l'invention, la 15 fréquence d'échantillonnage de la voix est comprise entre 8 et 16 KHz et, avantageusement, voisine de 12 KHz et la durée de chaque demi-tranche est comprise entre 5 et 20 ms et, avantageusement, voisine de 10 ms.

20 L'échantillonnage du signal vocal et sa reconstitution à partir des signaux binaires s'effectuent avantageusement au moyen d'un modulateur-démodulateur Δ .

Les diverses particularités, ainsi que les avantages de 25 l'invention, apparaîtront clairement à la lumière de la description ci-après.

Au dessin annexé :

30 La figure 1 est un dispositif de synthèse vocale avec compression de l'information binaire et enregistrement en mémoire morte, conforme à un mode d'exécution préféré de l'invention.

35 La figure 2 représente des formes d'ondes qui servent à illustrer le fonctionnement de ce dispositif.

Le signal à fréquences vocales (forme d'onde A de la figure 2) fourni par un microphone est appliqué sur un modulateur-

démodulateur MD, avantageusement du type Δ . Comme on le sait, dans son fonctionnement en démodulateur, un tel dispositif effectue une sorte d'échantillonnage, en ce sens qu'il engendre des impulsions de niveau 0 ou de niveau 1 à une
5 fréquence déterminée par un signal de commande F_b qui est ici fourni par un microprocesseur P et, dans l'exemple préféré décrit, a une fréquence voisine de 12 KHz.

Ces impulsions (forme d'onde B_E , figure 2) ont un niveau 1
10 ou un niveau 0 suivant que la pente du signal d'entrée intégré, par le démodulateur est positive ou négative.

Elles sont appliquées à une mémoire RAM dans laquelle elles sont stockées, à des adresses déterminées par des signaux
15 fournis par le microprocesseur et transmis sur un bus B, lorsqu'un signal de commande C approprié est fourni par le microprocesseur.

La forme d'onde C, représentée à la figure 2, correspond à
20 un découpage du signal B_E en demi-tranches 1-2-3 ayant chacune une durée voisine de 10 ms, séparées par des intervalles identiques.

Le signal stocké dans la RAM, représenté par la forme d'onde
25 RAM, ne comporte donc que la moitié de l'information binaire totale extraite du signal A.

Pour vérifier que cette information est suffisante pour reproduire normalement la voix, on procède à la lecture de
30 la RAM : la commutation de celle-ci du mode "enregistrement" au mode "lecture" (ou vice-versa) est provoquée par un signal K fourni par le microprocesseur.

La commande (non figurée) fournie par le microprocesseur est
35 alors telle que chacune des tranches 1-2-3 etc.. est lue deux fois successives, ce qui donne un signal sans discontinuité B_2 dont la forme d'onde enveloppe est représentée à la figure 2. Ce signal est converti en signal vocal par MD,

fonctionnant alors en modulateur.

L'expérience a montré que ce signal vocal est de bonne qualité et difficile à discerner, à l'audition, du signal initial, lorsque la fréquence d'échantillonnage et la durée des tranches ne s'écartent pas sensiblement des valeurs indiquées.

Si la fréquence d'échantillonnage tombe en dessous de 8KHz, la qualité de la reproduction se détériore. Pour des fréquences supérieures à 12 KHz et surtout, à 16 KHz, la qualité de la reproduction n'est pas sensiblement améliorée, tandis que la capacité de mémoire nécessaire augmente.

On notera qu'un découpage du signal B_E suivi d'une simple lecture ne comportant pas la répétition de chaque tranche ne donne pas une bonne reproduction de la voix, un effet d'écho gênant étant constaté.

Si on augmente la durée des tranches au-delà de 20 ms, cet effet d'écho se manifeste et devient rapidement gênant. En deçà de 5ms, la voix reproduite est caverneuse.

Il convient de souligner que la capacité de mémoire nécessaire à la mise en oeuvre du procédé ne dépasse pas 5000 bits par seconde de parole ; à titre de comparaison, les procédés de synthèse vocale les plus sophistiquées permettent seulement, au prix d'une complexité considérable, de la réduire à 2000 ou 3000 bits par seconde de parole.

En vue de son application à la transmission d'alarmes, telle qu'elle est décrite dans la demande de brevet susvisé, il est nécessaire de recopier le contenu de la RAM dans une mémoire morte. La figure 1 représente, entre RAM et EPROM, le dispositif de recopie R, bien connu en soi, nécessaire à cet effet.

Une fois que l'information binaire voulue a été stockée dans

- 5 -

l'EPROM, celle-ci est dissociée du reste du dispositif de la figure 1 et associée au dispositif décrit dans la demande de brevet français susvisée, dont le microprocesseur doit toutefois être programmé pour effectuer la lecture
5 répétitive des tranches d'information binaire stockées dans l'EPROM.

On pourra envisager d'autres applications du procédé de synthèse et de stockage de parole décrit, sans s'écarter de
10 l'esprit de l'invention. On pourra également envisager de modifier la loi de découpage.

Revendications de brevet.

1. Procédé de synthèse vocale comportant la conversion d'un signal électrique à fréquences vocales (A) en signal binaire 5 (B_E) et le stockage dudit signal binaire en mémoire, en vue de sa lecture ultérieure et de sa conversion numérique-analogique pour reproduire le signal vocal initial, caractérisé par un découpage dudit signal binaire, avant son stockage en mémoire, en tranches successives d'égale 10 longueur, par le prélèvement d'une fraction entière prédéterminée (1, 2, 3) de chaque tranche, par le stockage en mémoire de ladite fraction et par la lecture répétée de chacune desdites fractions, le nombre des lectures étant tel que la longueur initiale desdites tranches soit restituée.

15

2. Procédé selon la revendication 1, caractérisé en ce que le prélèvement porte sur la moitié de chaque tranche et que deux lectures successives des moitiés ainsi prélevées sont effectuées.

20 *

3. Procédé selon la revendication 2, caractérisé en ce que la fréquence de récurrence desdits signaux binaires est comprise entre 8 et 16 KHz, de préférence voisine de 12 KHz, et que la durée de chaque 25 demi-tranche est comprise entre 5 et 20 ms, avantageusement voisine de 10 ms.

4. Dispositif pour la mise en oeuvre du procédé selon l'une des revendications 1 à 3, 30 caractérisé en ce qu'il comporte un modulateur-démodulateur Δ (Ms) effectuant la connexion du signal électrique à fréquences vocales en signal binaire et vice-versa.

5. Dispositif selon la revendication 4, 35 caractérisé en ce qu'il comporte une mémoire vive (RAM) associée à un microprocesseur pour effectuer ledit prélèvement et ledit stockage.

- 7 -

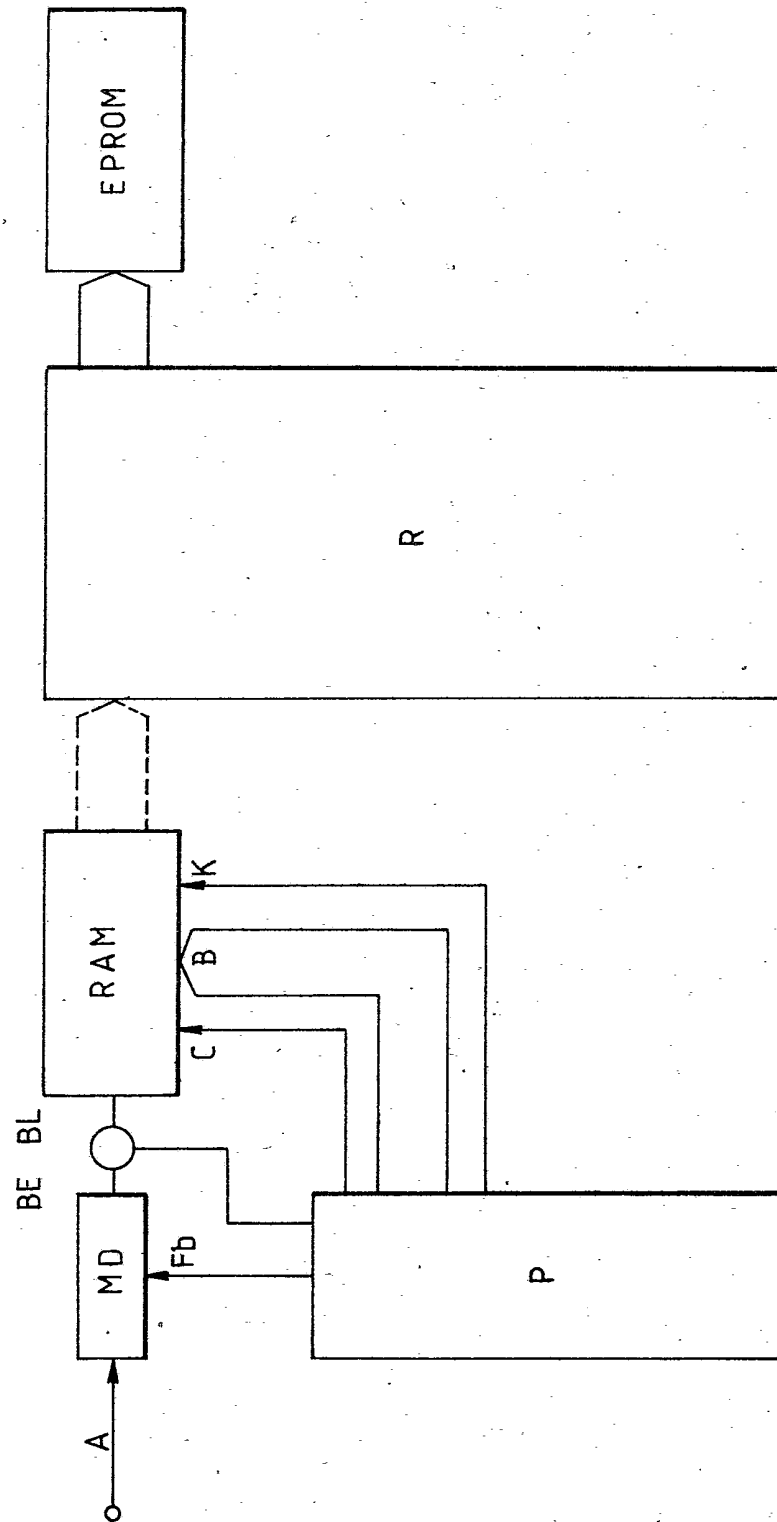
6. Dispositif selon la revendication 5, caractérisé en ce qu'il comporte une mémoire morte (EPROM) et des moyens (R) de transférer l'information binaire de la mémoire vive à la mémoire morte.

5

7. Application du dispositif selon la revendication 6 à la réalisation d'un dispositif de transmission d'alarmes par le réseau téléphonique autocommuté.

1/2

FIG. 1



2/2

FIG. 2

