

(19)



Europäisches Patentamt
European Patent Office
Office européen des brevets



(11)

EP 1 032 933 B1

(12)

FASCICULE DE BREVET EUROPEEN

(45) Date de publication et mention
de la délivrance du brevet:

02.01.2004 Bulletin 2004/01

(21) Numéro de dépôt: **98955709.5**

(22) Date de dépôt: **20.11.1998**

(51) Int Cl.7: **G10L 19/00**

(86) Numéro de dépôt international:
PCT/FR1998/002486

(87) Numéro de publication internationale:
WO 1999/027521 (03.06.1999 Gazette 1999/22)

(54) **PROCEDE DE CODAGE DE LA PAROLE ET TERMINAUX POUR LA MISE EN OEUVRE DU
PROCEDE**

SPRACH-KODIERUNGS-VERFAHREN UND ENDGERÄTE ZUM IMPLEMENTIEREN BESAGTEN
VERFAHRENS

SPEECH CODING METHOD AND TERMINALS FOR IMPLEMENTING SAID METHOD

(84) Etats contractants désignés:
DE FR GB

(30) Priorité: **21.11.1997 FR 9714667**

(43) Date de publication de la demande:
06.09.2000 Bulletin 2000/36

(73) Titulaire: **SAGEM S.A.**
75116 Paris (FR)

(72) Inventeur: **NICAISE, Pierre**
F-60580 Coye la Forêt (FR)

(74) Mandataire: **Bloch, Gérard et al**
2, square de l'Avenue du Bois
75116 Paris (FR)

(56) Documents cités:

EP-A- 0 223 014 WO-A-94/18668
DE-A- 3 416 238

- **ROUCOS ET AL.: "Speaker normalization algorithms for very-low-rate speech coding" PROCEEDINGS OF THE IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING (ICASSP 84), vol. 1, 19 - 21 mars 1984, pages 1.1.1-1.1.4, XP002073267 SAN DIEGO, CA, US**
- **PATENT ABSTRACTS OF JAPAN vol. 010, no. 240 (E-429), 19 août 1986 & JP 61 071730 A (NEC), 12 avril 1986**

EP 1 032 933 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] La transmission de la parole sur le réseau téléphonique commuté RTC nécessite de disposer d'une bande-passante suffisante pour que la parole reste compréhensible. Une bande allant des très basses fréquences à quelque kilohertz représente un bon compromis encre fidélité de restitution et ressources en bande-passante. De ce fait, pour transmettre la voix dans le réseau RTC, dont les liaisons inter-centraux sont numériques, on code les fréquences vocales pour les transformer en un signal numérique au débit de base du réseau RTC, représentant l'évolution de l'amplitude du signal vocal.

[0002] On souhaite cependant parfois réduire le débit de la transmission par exemple dans un terminal de synthèse vocale, dont la mémoire de messages doit rester de taille limitée. De même, on peut souhaiter n'utiliser qu'une voie de transmission d'un réseau spécifique, à débit bien inférieur au débit ci-dessus du réseau téléphonique.

[0003] En pareil cas, on tente de reconnaître à la volée les divers phonèmes d'une séquence vocale. On compare ces phonèmes à des références, d'une bibliothèque, associées à des mots de code et on remplace ces phonèmes par les mots de code correspondants, qui décrivent la parole avec une quantité d'informations bien moindre. C'est ainsi que la voix est comprimée.

[0004] En réception, le terminal appelé comporte la même bibliothèque et reconstitue, par synthèse vocale, des signaux analogiques correspondant aux divers mots de code.

[0005] Mais une telle façon de procéder présente l'inconvénient de ne restituer qu'une voix normalisée par la bibliothèque, donc impersonnelle, et il est donc par essence impossible de reconnaître le correspondant pour authentifier un message vocal. Les inflexions ou fluctuations de voix, qui sont aussi de l'information tout comme la signification des paroles elles-mêmes, ne sont donc pas restituées.

[0006] L'article "Speaker normalization algorithms for very low-rate speech coding" de Roucos et al, ICASSP 84, Vol. 1, 19-21 mars 1984, pages 1.11 à 1.14, XP002073267, SAN DIEGO, CA, US, et le brevet DE-A-3 416 238 enseignent d'adapter progressivement une bibliothèque de signaux vocaux à ceux d'un nouveau locuteur. Les différences relevées entre ces signaux servent uniquement à cette adaptation.

[0007] Les signaux de parole sont donc déformés à la restitution tant que la bibliothèque n'a pas atteint un nouvel état d'équilibre.

[0008] WO-A-94 18 668 concerne la transmission de la parole par prédiction linéaire et ne traite pas de la reconnaissance vocale. De surcroît, le procédé de ce document implique des calculs de moyennes, avec une bibliothèque de paramètres auxquels le son a été identifié et non pas de paramètres directement mesurés sur le son capté.

[0009] Patent Abstracts of Japan, Vol. 010, No. 240, JP 61 071 730A enseigne de reconnaître, outre la parole, une caractéristique vocale, sans toutefois indiquer ni sa nature ni, surtout, la façon de déterminer cette caractéristique.

[0010] Une reproduction fidèle de la parole ne peut être garantie.

[0011] La présente invention vis à obtenir un codage de la voix qui permette à la fois une compression de l'information et une restitution personnalisée.

[0012] A cet effet, l'invention concerne tout d'abord un procédé de codage de la parole selon la revendication 1.

[0013] Ainsi, les mots de code primaire vont coder efficacement, de façon compacte, la plus grande partie de l'énergie phonique saisie tandis que les mots de code secondaire vont améliorer la fidélité de restitution sans cependant accroître exagérément le volume de données de code puisqu'ils ne portent que sur une énergie limitée et qu'un faible nombre de bits permet de coder cette énergie marginale modulant l'énergie primaire, standard, correspondant aux mots de code primaire.

[0014] L'invention concerne aussi un terminal de codage de signaux de parole selon la revendication 7.

[0015] L'invention concerne enfin un terminal de décodage de signaux de parole selon la revendication 8.

[0016] Bien que le procédé de l'invention nécessite de disposer au total d'un terminal de codage et d'un terminal de décodage correspondant, chacun de ceux-ci peut être mis séparément sur le marché et la demanderesse entend donc revendiquer l'un et l'autre.

[0017] En particulier, il est avantageux de prévoir un télécopieur comportant des moyens pour insérer les mots de code dans un message de télécopie.

[0018] L'invention sera mieux comprise à l'aide de la description suivante d'un mode de mise en oeuvre préféré du procédé de l'invention, en référence au dessin annexé, sur lequel :

- la figure 1 représente schématiquement un terminal émetteur et un terminal récepteur de signaux vocaux pour la mise en oeuvre du procédé de l'invention,
- la figure 2 représente, en fonction du temps t , l'amplitude A d'un signal de parole,
- la figure 3 représente, en fonction de la fréquence F , l'amplitude K de raies du spectre du signal de la parole 2, et
- la figure 4 est un organigramme illustrant les étapes du procédé.

[0019] Le terminal émetteur référencé 15 sur la figure 1, qui se présente ici sous la forme d'un combiné portable pour réseau de communication radio, comporte un microphone 26 de saisie du signal de parole de son utilisateur, alimentant un convertisseur analogique/numérique 27 relié en sortie à une unité centrale à microprocesseur 28 associée à deux bibliothèques 11 et 12 de séquences sonores ou unités phoniques normalisées,

telles que phonèmes. L'unité centrale 28, qui code la parole, commande un émetteur 29, ici radio, dont les émissions sont reçues par un circuit récepteur 30 d'un terminal 35 de restitution de la parole. La figure 2 représente l'amplitude d'une unité phonique en fonction du temps t et la figure 3 représente l'amplitude K des raies du spectre correspondant à un instant donné.

[0020] Plus précisément, l'unité centrale 28 comporte un comparateur 16 pour comparer les unités phoniques reçues du convertisseur 27 aux unités phoniques en bibliothèque. Comme expliqué plus en détails en regard de la figure 4, le comparateur 16 a pour fonction de sélectionner la référence en bibliothèque qui est la plus semblable au signal alors analysé et a aussi pour fonction de spécifier cette différence, c'est-à-dire de fournir une valeur d'écart pour chacun des critères ayant servi à la sélection. Cette différence est ici transcodée par un circuit de transcodage 17, pour condenser son expression, sous la forme d'un mot de code secondaire qui est mémorisé dans une mémoire 18 sous la commande du comparateur 16. Ce dernier, qui a préalablement mémorisé le mot de code primaire dans la mémoire 18, adresse et commande l'écriture dans celle-ci pour que les deux mots de code, primaire et secondaire, soient associés physiquement comme ils le sont d'un point de vue logique, c'est-à-dire qu'il soit par exemple défini un chaînage entre les deux zones mémoires les contenant.

[0021] Le terminal récepteur 35 comporte une unité centrale 33 effectuant un décodage inverse de la parole pour alimenter un haut-parleur 34. Deux mémoires formant bibliothèques, 31 et 32, ici externes à l'unité centrale 33, sont reliées à celle-ci. Le terminal récepteur 35 est ici un terminal classique pour la réception de messages écrits, appelé pager, agencé en outre pour recevoir des messages vocaux. Divers circuits classiques de saisie de numérotation, d'émission de celle-ci et de visualisation de données n'ont pas été représentés, dans un but de clarté.

[0022] L'unité centrale 33 comporte un circuit 36 d'adressage des bibliothèques 31 et 32, respectivement personnalisée et primaire, à partir des mots de code reçus du circuit récepteur 30. Un circuit tampon 37 reçoit en réponse, de la bibliothèque primaire 32, des spectres d'unités phoniques primaires et les transmet à un circuit 38 de modulation ou composition de ces spectres. Le circuit 38 module ceux-ci en fonction du mot de code secondaire associé au mot de code primaire de lecture de la bibliothèque primaire 32. Le circuit 38 combine ainsi les informations des mots de code primaires et secondaires pour reconstituer le signal de parole capté à l'origine (26). Cette combinaison peut par exemple être une addition ou une multiplication, de raies de fréquences suivie d'une transformation de Fourier inverse ou, encore, elle peut porter directement sur des amplitudes de signal. Dans cet exemple, chaque type d'unité phonique restituée est mémorisée dans la mémoire personnalisée 31, afin d'utiliser directement cette dernière si un même couple de mots de code, primaires et secondaires,

est ultérieurement reçu. En variante, la mémoire 31 pourrait ne contenir que des valeurs de modulation, qu'elle fournirait au circuit 38 après adressage par un mot de code secondaire.

[0023] Les opérations de codage et de décodage vont maintenant être exposées plus en détails, en référence à la figure 4.

[0024] Pour coder la voix, on capte à une étape 1, un signal de parole par le microphone 26 et, ici, on le convertit en un signal numérique dans le convertisseur 27, à une étape 2. Le signal de parole est alors comparé, dans l'unité centrale 28, à une pluralité de signaux de référence de la bibliothèque 11, à une étape 3. La comparaison intervient à la volée, en pratique cycliquement à vitesse élevée par rapport à la vitesse d'évolution du signal de parole analysé. Celui-ci peut être considéré comme étant une séquence d'unités phoniques, spécifiques d'une langue donnée, telles que voyelles, diphthongues ou hiatus, dont une représentation a été initialement mise en bibliothèque 11 et associée à un mot de code, dit primaire, particulier à chacune. Lors de la constitution de la bibliothèque 11 et des bibliothèques 12 et 32 indiquées plus loin, on effectue plusieurs saisies vocales d'un même locuteur ou de plusieurs pour constituer une référence vocale moyenne. De préférence cependant, afin d'améliorer l'efficacité de la reconnaissance future, on mémorise (11, 12), pour chaque unité phonique, plusieurs références pour constituer un domaine de reconnaissance permettant de tolérer des écarts entre locuteurs.

[0025] Chaque unité phonique (fig. 2) correspond à une évolution particulière de l'amplitude A ou énergie du signal de parole et présente une durée susceptible de varier selon la vitesse d'élocution de la personne qui parle.

[0026] L'étape 3 consiste donc à comparer l'évolution des amplitudes des signaux de référence à celle du signal capté. Pour s'affranchir des variations de vitesse d'élocution, on peut par exemple ne considérer que la succession des modulations d'amplitude significatives (variation d'énergie dépassant un seuil), sans y associer une notion de valeur de temps.

[0027] Sur la figure 2, les flèches verticales, ici au nombre de huit, représentent l'amplitude des extréma et forment ainsi une signature, supposée ici représenter une unité phonique déterminée.

[0028] Si, quittant le domaine temporel de la figure 2, on considère maintenant le domaine fréquentiel, la transformée de Fourier de l'amplitude instantanée A du signal en un point quelconque de la courbe de la figure 2 peut être représentée par le spectre de raies de fréquences de la figure 3. En pratique, on considère que l'énergie vocale se trouve essentiellement limitée à trois bandes de fréquences, respectivement situées vers 0,1 kHz et deux bandes entre environ 1 et 3 kHz ainsi que 5 et 7 kHz, respectivement.

[0029] De ce fait, si l'on parcourt la courbe des amplitudes de la figure 2 en fonction du temps t , l'amplitude

K de chaque raie de la figure 3 va se trouver modulée en fonction de l'évolution de l'amplitude A du signal de parole.

[0030] Ainsi, si l'on mémorise la succession des spectres de la figure 3, on peut reconstituer la succession des amplitudes A du signal d'origine, par transformée de Fourier inverse.

[0031] Afin de limiter le nombre de spectres à traiter, on peut n'effectuer que des échantillonnages cycliques, suffisamment rapprochés pour ne pas perdre d'information. On peut encore se limiter à un jeu restreint de spectres des extréma d'amplitude représentés par les huit flèches de la figure 2. Si l'on souhaite encore limiter le nombre de spectres, on peut même ne retenir qu'un seul spectre représentant la moyenne de tous les spectres sur la période de temps considérée de l'unité phonique ou la moyenne des spectres des extréma.

[0032] Comme cela est expliqué plus loin en détail, on va comparer le spectre moyen, ou les spectres, du signal de parole capté à un ou des spectres homologues de signaux de parole de référence en bibliothèque pour, d'une part, sélectionner le signal de parole de référence (unité phonique) le plus semblable au signal capté et, d'autre part, élaborer un signal de différence entre le ou les spectres de ce dernier et le ou les spectres du signal de référence sélectionné. Le signal de différence est mis sous forme de mot de code, dit secondaire, et est associé au mot de code primaire du signal de référence sélectionné (unité phonique reconnue) et constitue ainsi une information additionnelle de modulation ou correction du signal analogique normalisé qui sera restitué à partir du mot de code primaire considéré.

[0033] Les mots de code primaires des unités phoniques successivement sélectionnés, à mesure que se déroule la séquence vocale, sont mémorisés à une étape 4 pour constituer un message, codé selon la norme de la bibliothèque 11.

[0034] Par ailleurs, à une étape 5, on traite plus avant certaines des unités phoniques captées et reconnues, en analysant en détail, ici dans le domaine fréquentiel par transformée de Fourier inverse, leur spectre en fréquence, comme expliqué plus haut, étape 6.

[0035] A une étape 7, on compare le spectre de raies j de l'unité phonique d'identité i concernée, ou les spectres représentant son évolution au cours du temps t, au (x) spectre(s) de l'unité phonique sélectionnée en bibliothèque 11, contenu(s) dans la bibliothèque associée 12. On établit ainsi pour le, ou chaque, spectre, une série de coefficients de pondération Cijt (i = identité du phonème, j = rang fréquentiel de la raie, t = rang temporel) indiquant chacun l'amplitude ou énergie relative de chaque raie j par rapport à son homologue en bibliothèque 12. Exprimé autrement, ces coefficients représentent aussi, bien qu'indirectement, la différence relative (1 - Cijt) entre l'unité phonique reconnue et la référence correspondante en bibliothèque). Les raies dans chacune des trois bandes correspondent en fait à une rangée de mini-bandes de fréquences adjacentes, dans lesquelles

on détecte de l'énergie vocale. L'analyse dans le domaine fréquentiel, ici choisie, fournit ainsi une information plus détaillée que dans le cas d'une analyse dans le domaine temporel de la figure 2, où l'on ne dispose que de l'amplitude instantanée A.

[0036] Ainsi, dans le cas des figures 2 et 3, la série ci-dessus comporte douze coefficients représentant les douze raies dessinées, si bien qu'un tableau de huit telles séries représente l'unité phonique, à travers les huit extréma dessinés. Hormis la réduction du tableau, à une seule série, on peut prévoir de ne retenir qu'un coefficient de pondération moyen pour chacune des trois bandes. Si l'on code chaque coefficient sur simplement 4 bits, l'erreur ne dépasse pas 3 % environ, ce qui est amplement suffisant pour restituer un timbre de voix, d'autant que le signal de correction représente peu d'énergie par rapport au signal normé qu'il corrige, si bien que l'erreur, rapportée au total, est faible.

[0037] On peut ainsi associer ici, au mot de code primaire de l'unité phonique sélectionnée, de l'ordre de la centaine de bits (12 x 8) si l'on retient chaque extrémum, ou seulement 12 bits (4 x 3) pour les trois bandes. Comme le timbre de la voix est surtout fourni par les fréquences hautes de la troisième bande, on peut même ne transmettre que le mot de code secondaire, de correction, relatif à celle-ci.

[0038] A une étape 8, on transforme le signal de différence de spectres en un mot de code secondaire, représentant le tableau ou la série ci-dessus. Lorsque la séquence de parole captée s'achève, les mots de code primaires de l'étape 4 et les mots de code secondaires de l'étape 8 sont associés un à un (étape 9) puis émis sur un réseau de transmission, comme par exemple le réseau téléphonique commuté ou, ici, un réseau de radio messagerie (étape 10).

[0039] Le terminal appelé 35 reçoit le message à une étape 21 et, à une étape 22, un fichier en bibliothèque primaire 32, semblable au fichier de spectres 12, est lu par le circuit 36 pour en extraire les spectres normalisés, primaires, d'après les mots de code primaires. A une étape 23, les mots de code secondaires servent à moduler (38) les amplitudes ou énergies des raies normalisées lues en bibliothèque primaire 32, pour ainsi constituer la bibliothèque personnalisée, 31, d'unités phoniques, c'est-à-dire comportant en particulier le timbre de la voix captée. Les unités phoniques de la bibliothèque personnalisée 31 sont représentées sous forme numérique dans le domaine temporel, après une transformation préalable par transformée de Fourier inverse.

[0040] A une étape 24, les mots de code primaires reçus sont lus successivement pour restituer, par le haut-parleur 34 (étape 25), le signal de parole capté. Pour cela les mots de code primaires lisent la bibliothèque personnalisée 31, qui correspond donc à la bibliothèque 11 mais qui a été personnalisée par les caractéristiques en spectre de la voix captée.

[0041] La constitution de la bibliothèque 31 est, comme évoqué plus haut, facultative et a pour but de mé-

moriser une correction pour chaque mot de code primaire, ce qui évite de répéter l'envoi du mot de code secondaire lorsqu'un même mot de code primaire est émis plusieurs fois. Si, au contraire, on émet systématiquement un mot de code secondaire, celui-ci peut évoluer pour suivre les éventuelles évolutions du timbre. Dans ce cas, la voix restituée est à la fois personnalisée et en outre on restitue l'évolution du timbre au fil du temps.

[0042] Il faut en outre remarquer que, de façon générale, l'analyse et la restitution peuvent porter sur toute la bande de fréquences audibles, allant d'environ 15 Hz à 15 kHz, même si en pratique on peut se limiter à 8 kHz. Les fréquences de la bande allant de 4 à 8 kHz, coupées pour les transmissions classiques par le réseau téléphonique, sont ici analysées et restituées puisqu'on transmet l'information correspondante sous la forme d'une télécommande de la bibliothèque 31, qui contient déjà les raies à ces fréquences, ce qui évite toute transmission explicite de celles-ci.

[0043] Il faut par ailleurs remarquer que, si l'analyse peut ne porter que sur un nombre limité de bandes de fréquences suffisamment caractéristiques en bibliothèque 11, 12, les divers signaux à restituer, en bibliothèque 32, contiennent la totalité des raies initialement saisies, c'est-à-dire couvrent chacun par exemple une bande d'un seul tenant de 15 Hz à 8 kHz.

[0044] Comme exposé au début, l'invention peut s'appliquer en dehors de toute transmission, pour par exemple mémoriser localement un message à restituer ultérieurement, c'est-à-dire qu'il s'agit alors de la fonction d'un magnétophone.

[0045] Dans une autre forme de réalisation, non représentée, les mots de code, primaire et secondaire, sont associés à des données de télécopie pour former un message voix-données. Le message est saisi par le poste téléphonique usuellement associé aux télécopieurs et est restitué par le même moyen au niveau du télécopieur appelé. Les mots de code, émis par un circuit comme 28, sont insérés dans un champ spécifique du message par un microprocesseur gérant le protocole de transmission de télécopies et sont de même extraits en réception pour être traités comme expliqué plus haut. On peut ainsi réaliser une annotation vocale d'un message de télécopie, annotation transmise par exemple comme un en-tête de télécopie.

Revendications

1. Procédé de codage de la parole constituée d'une séquence d'unités phoniques, dans lequel on compare successivement ces unités phoniques à des références en bibliothèque associées à des mots de code primaire spécifiques, et, pour chaque unité phonique, on sélectionne une référence et donc son mot de code spécifique associé, on détermine la différence entre l'unité phonique et la référence, on procède à un transcodage de ladite différence pour

obtenir un mot de code secondaire, on associe le mot de code primaire et le mot de code secondaire et on substitue à l'unité phonique la paire de deux mots de code primaire et secondaire, procédé **caractérisé par le fait que** la comparaison des unités phoniques et des références porte sur l'énergie des raies de fréquences pour lesquelles on détermine des coefficients de pondération, normés par rapport à l'énergie des raies de référence, qui sont transcodés selon les mots de code secondaire.

2. Procédé selon la revendication 1, dans lequel on détermine ladite différence à partir d'une succession de spectres correspondant à une succession d'amplitudes de l'unité phonique considérée.
3. Procédé selon la revendication 2, dans lequel on ne considère que les amplitudes correspondant à des extrema.
4. Procédé selon la revendication 1, dans lequel on détermine ladite différence à partir d'un spectre unique moyen de l'unité phonique considérée.
5. Procédé selon l'une des revendications 1 à 4, dans lequel on limite la comparaison fréquentielle à trois bandes de fréquences.
6. Procédé selon la revendication 5, dans lequel on exprime le coefficient de pondération des raies de chaque bande par un unique coefficient.
7. Terminal de codage de signaux de parole, comportant des moyens (26, 27) pour saisir une séquence d'unités phoniques et la transmettre à des moyens comparateurs (16) agencés pour successivement comparer les unités phoniques à des références en bibliothèque (11, 12), et ainsi y sélectionner à chaque fois un mot de code primaire spécifique de l'une des références, la comparaison des unités phoniques et des références posant sur l'énergie des raies de fréquences, les moyens comparateurs (16) étant agencés pour déterminer une différence entre l'unité phonique saisie considérée et la référence correspondant au mot de code sélectionné, pour en déduire un coefficient de pondération normé par rapport à l'énergie des raies de référence et pour transmettre le coefficient à des moyens de transcodage (17), prévus pour fournir en réponse un mot de code secondaire mémorisé dans des moyens mémoires (18) à les mots de code, respectivement primaire et secondaire, sont associés.
8. Terminal de décodage de signaux de parole, comportant des moyens (30) de réception de signaux représentant des mots de code primaires de références d'unités phoniques en bibliothèque (32), et des moyens de décodage (33) agencés pour sélec-

tionner certaines des références en bibliothèque (32) d'après les mots de code primaires reçus et pour commander en conséquence un transducteur (34) de restitution des signaux de parole, la comparaison des unités phoniques et des références portant sur l'énergie des raies de fréquences, les moyens de décodage (33) étant agencés pour décoder en outre des mots de code secondaire, de correction, associés aux mots de code primaire et correspondant à des coefficients de pondération normés par rapport à l'énergie des raies de référence, et pour corriger en conséquence (38) les références sélectionnées par modulation de ces références par les mots de code secondaire.

9. Terminal selon l'une des revendications 7 et 8, constitué par un télécopieur comportant des moyens pour insérer les mots de code dans un message de télécopie.

Patentansprüche

1. Verfahren zur Codierung von aus einer Folge von phonischen Einheiten bestehender Sprache, wobei man diese phonischen Einheiten nacheinander mit Referenzen in einer Bibliothek vergleicht, die zu speziellen primären Codewörtern gehören, und wobei man für jede phonische Einheit eine Referenz und folglich deren spezielles zugehöriges Codewort auswählt, die Differenz zwischen der phonischen Einheit und der Referenz bestimmt, eine Codeumsetzung der Differenz durchführt, um ein sekundäres Codewort zu erhalten, das primäre Codewort und das sekundäre Codewort einander zuordnet und die phonische Einheit durch das Paar aus primärem und sekundärem Codewort ersetzt, wobei das Verfahren **dadurch gekennzeichnet ist, daß** der Vergleich der phonischen Einheiten und der Referenzen die Energie von Frequenzlinien betrifft, für welche man bezogen auf die Energie von den Referenzlinien normierte Wichtungskoeffizienten bestimmt, die entsprechend den sekundären Codewörtern umcodiert sind.
2. Verfahren nach Anspruch 1, wobei man die Differenz ausgehend von einer Folge von Spektren bestimmt, die einer Folge von Amplituden der betrachteten phonischen Einheit entspricht.
3. Verfahren nach Anspruch 2, wobei man nur die Amplituden betrachtet, die Extrema entsprechen.
4. Verfahren nach Anspruch 1, wobei man die Differenz ausgehend von einem einzigen mittleren Spektrum der betrachteten phonischen Einheit bestimmt.
5. Verfahren nach einem der Ansprüche 1 bis 4, wobei man den Frequenzvergleich auf drei Frequenzbänder beschränkt.
6. Verfahren nach Anspruch 5, wobei man den Wichtungskoeffizienten der Linien jedes Bandes durch einen einzigen Koeffizienten ausdrückt.
7. Endgerät zur Codierung von Sprachsignalen, mit Mitteln (26, 27) zur Erfassung einer Sequenz von phonischen Einheiten und zur Übertragung dieser an Komparatormittel (16), die ausgebildet sind, um die phonischen Einheiten mit Referenzen in einer Bibliothek (11, 12) zu vergleichen, und so dort jedesmal ein spezielles primäres Codewort von einer der Referenzen auszuwählen, wobei der Vergleich der phonischen Einheiten und der Referenzen die Energie von Frequenzlinien betrifft, wobei die Komparatormittel (16) ausgebildet sind, um eine Differenz zwischen der betrachteten erfaßten phonischen Einheit und der dem ausgewählten Codewort entsprechenden Referenz zu bestimmen, daraus einen bezogen auf die Energie der Referenzlinien normierten Wichtungskoeffizienten abzuleiten und den Koeffizienten an Mittel zur Codeumsetzung (17) zu übertragen, die zur Lieferung eines sekundären Codewortes als Antwort vorgesehen sind, das in Speichermitteln (18) gespeichert wird, wo jeweils die primären und sekundären Codewörter zugeordnet sind.
8. Endgerät zur Decodierung von Sprachsignalen, mit Mitteln (30) zum Empfangen von Signalen, die primäre Codewörter von Referenzen phonischer Einheiten in einer Bibliothek (32) darstellen, und mit Mitteln zur Decodierung (33), die ausgebildet sind, um bestimmte Referenzen in der Bibliothek (32) gemäß den empfangenen primären Codewörtern auszuwählen und um entsprechend einen Wandler (34) zur Sprachsignalwiedergabe zu steuern, wobei der Vergleich der phonischen Einheiten und der Referenzen die Energie von Frequenzlinien betrifft, wobei die Decodiermittel (33) ausgebildet sind, um außerdem sekundäre Codewörter zur Korrektur zu decodieren, die den primären Codewörtern zugeordnet sind und bezogen auf die Energie der Referenzlinien normierten Wichtungskoeffizienten entsprechen, und um entsprechend (38) die ausgewählten Referenzen durch Modulation dieser Referenzen mit Hilfe der sekundären Codewörter zu korrigieren.
9. Endgerät nach Anspruch 7 oder 8, gebildet durch ein Telefaxgerät, das Mittel zum Einfügen der Codewörter in eine Telefaxnachricht aufweist.

Claims

1. Speech coding method consisting of a sequence of phonic units, in which these phonic units are successively compared with library references associated with specific primary code words and, for each phonic unit, a reference and therefore its associated specific code word is selected, the difference between the phonic unit and the reference is determined, code conversion of the said difference is carried out in order to obtain a secondary code word, the primary code word and the secondary code word are associated and the pair of two primary and secondary code words replace the phonic unit, **characterised in that** the comparison of the phonic units and the references is concentrated on the energy of the frequency rays for which weighting coefficients are determined, normed in relation to the reference rays, which are code-converted according to the secondary code words.

5
10
15
20
2. Method according to Claim 1, whereby the said difference is determined from a succession of spectra corresponding to a succession of amplitudes of the phonic unit in question.

25
3. Method according to Claim 2, whereby only the amplitudes corresponding to extremes are considered.
4. Method according to Claim 2, whereby the said difference is determined from a single mean spectrum of the phonic unit in question.

30
5. Method according to one of the Claims 1 to 4, whereby the frequential comparison is limited to three frequency bands.

35
6. Method according to Claim 5, whereby the weighting coefficient of the lines of each band are expressed by a single coefficient.

40
7. Speech signal coding terminal consisting of means (26, 27) for catching a sequence of phonic units and transmitting it to comparative means (16) arranged for successively comparing the phonic units with library references (11, 12) and thus each time selecting from them a primary code word specific to one of the references, the comparison of the phonic units and the references concentrating on the energy of the frequency rays, the comparative means (16) being arranged to determine a difference between the caught phonic unit in question and the reference corresponding to the code word selected, in order to deduce from it a weighting coefficient normed in relation to the energy of the reference rays and to transmit the coefficient to code converting means (17), designed to supply a secondary code word in response, memorised in memory

45
50
55
8. Speech signal decoding terminal consisting of means (30) for receiving signals representing primary code words of references of phonic units in the library (32), and means of decoding (33) arranged for selecting some of the library references (32) from the primary code words received, and consequently to control a transducer (34) for return of the speech signals, the comparison of the phonic units and of the references concentrating on the energy of the frequency rays, the decoding means (33) being arranged for decoding in addition secondary correction code words, associated with the primary code words and corresponding to weighting coefficients normed in relation to the energy of the reference rays, and consequently to correct (38) the references selected by modulation of these references by the secondary code words.
9. Terminal according to one of the Claims 7 and 8, consisting of a fax machine with means for inserting the code words into a fax message.

means (18), with which the code words, primary and secondary respectively, are associated.

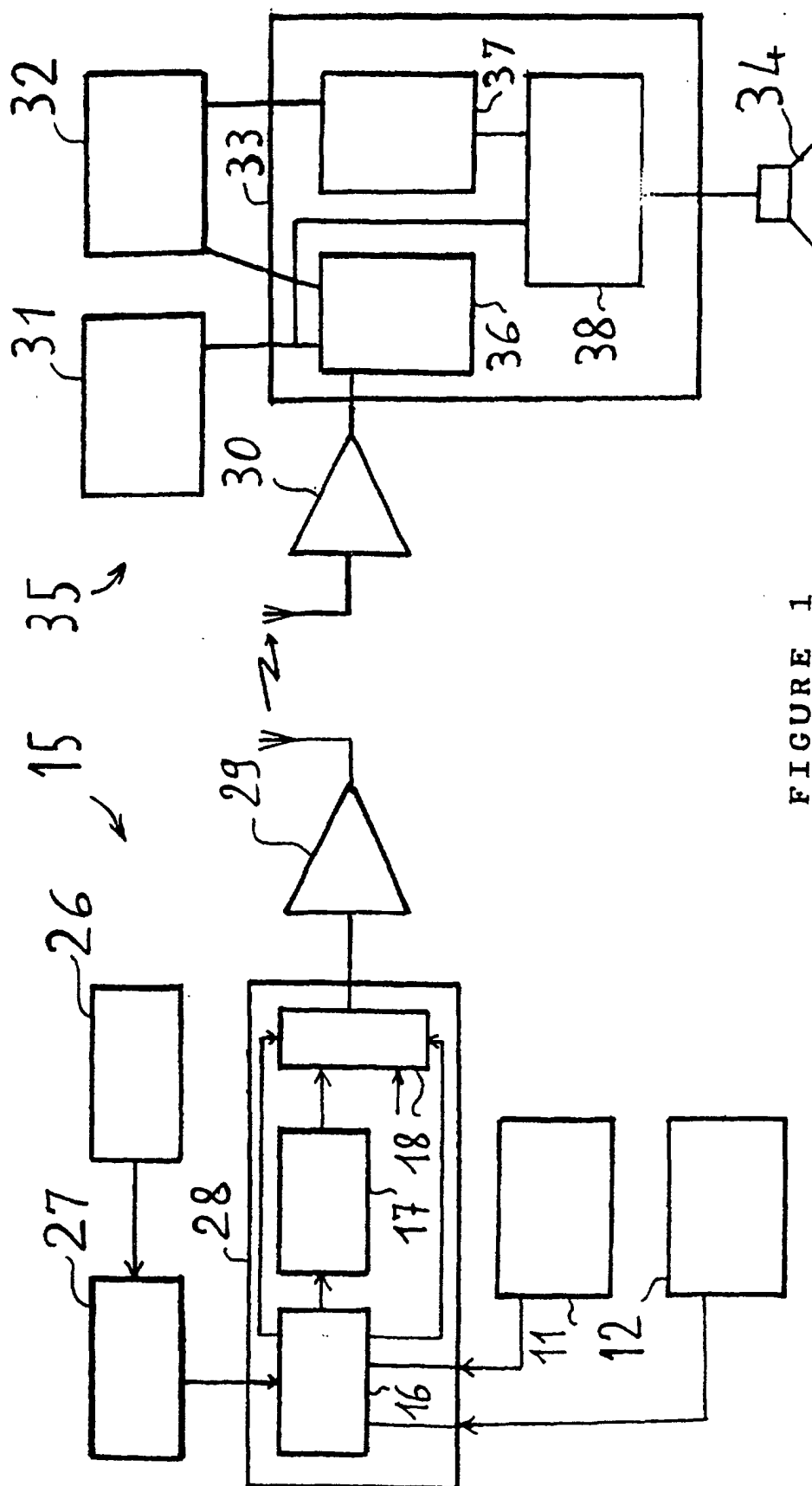


FIGURE 1

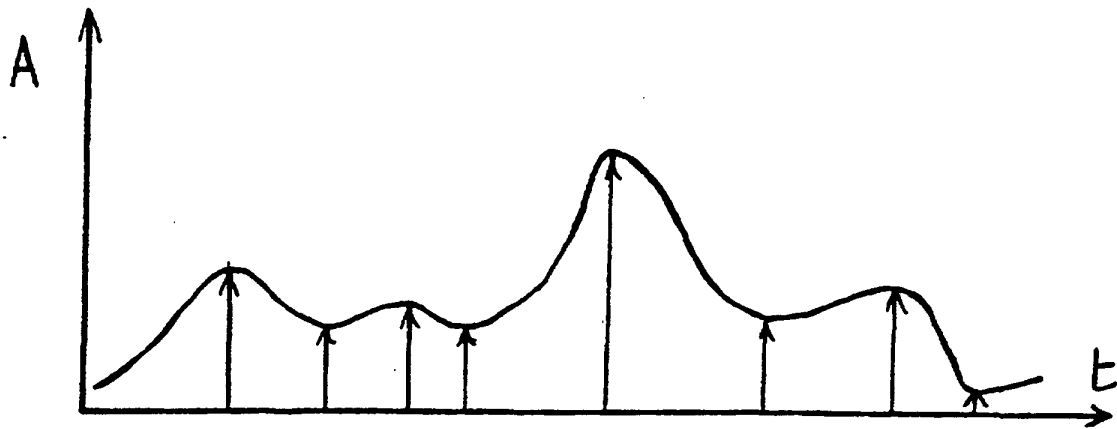


FIGURE 2

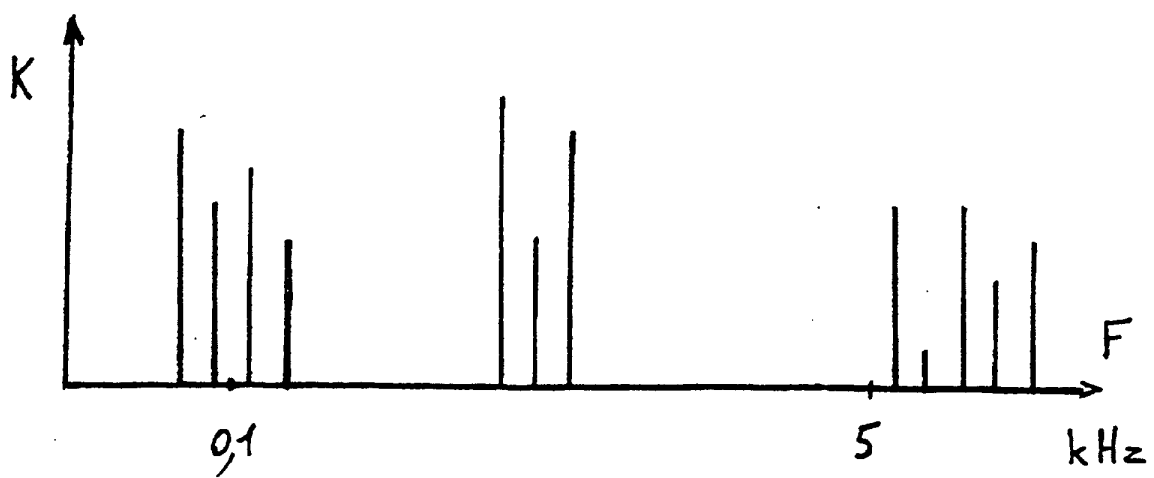


FIGURE 3

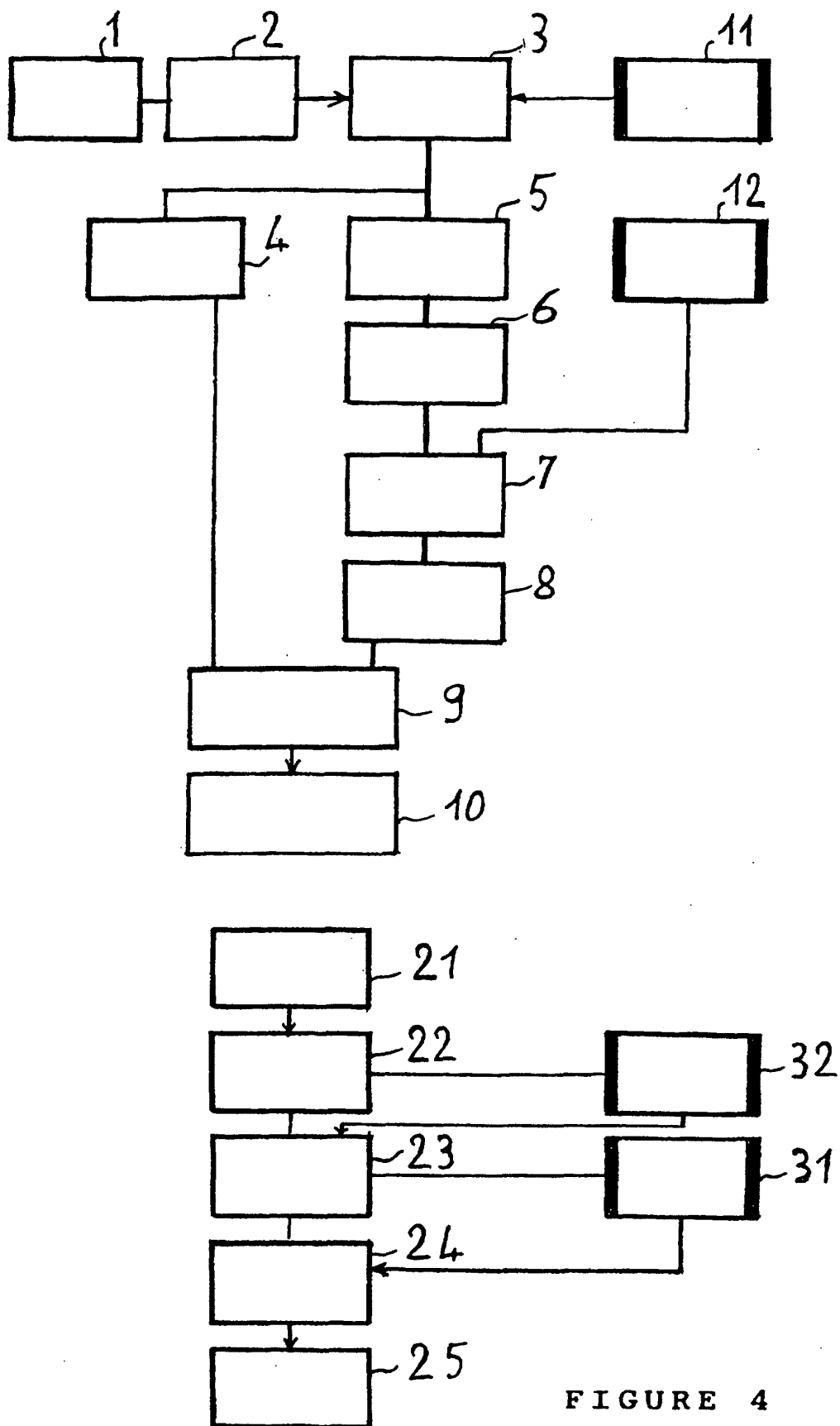


FIGURE 4