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

(11) EP 0 734 013 A2

(12) DEMANDE DE BREVET EUROPEEN

(43) Date de publication:
25.09.1996 Bulletin 1996/39

(51) Int Cl.[6]: **G10L 9/14**

(21) Numéro de dépôt: **96410028.3**

(22) Date de dépôt: **19.03.1996**

(84) Etats contractants désignés:
DE FR GB IT

(30) Priorité: **24.03.1995 FR 9503735**

(71) Demandeur: **SGS-THOMSON MICROELECTRONICS S.A.**
F-94250 Gentilly (FR)

(72) Inventeur: **Bouraoui, Mustapha**
38000 Grenoble (FR)

(74) Mandataire: **de Beaumont, Michel**
1bis, rue Champollion
38000 Grenoble (FR)

(54) Determination d'un vecteur d'excitation dans un codeur CELP

(57) La présente invention concerne un procédé de détermination d'un vecteur d'excitation (Copt) dans un codeur de signal de parole CELP, ledit vecteur appartenant à un sous-ensemble associé à un ensemble plus grand de vecteurs d'excitation susceptibles de maximiser un critère (m). Le procédé comprend les étapes consistant à présélectionner un vecteur d'excitation (C_O) ayant pour composantes celles de même signe que les échantillons correspondants à un vecteur cible (T), et, si le vecteur d'excitation (C_O) présélectionné n'appartient pas audit sous-ensemble, à sélectionner comme vecteur d'excitation celui qui maximise ledit critère (m) parmi les vecteurs (C'_O... C'_N) du sous-ensemble qui sont associés respectivement au vecteur présélectionné et aux vecteurs (C_1... C_N) les plus proches de celui-ci dans l'ensemble plus grand.

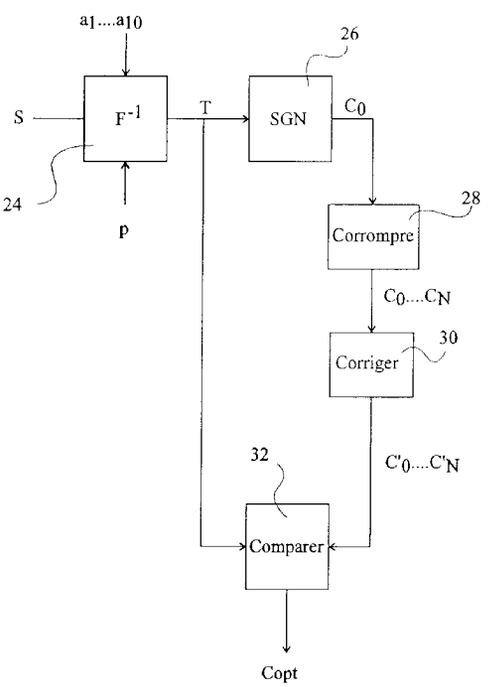


Fig 3

EP 0 734 013 A2

Description

La présente invention concerne la compression des signaux de parole à transmettre sur une ligne téléphonique et plus particulièrement la détermination d'un vecteur d'excitation dans le cadre d'une compression selon la méthode de la Prédiction Linéaire Excitée par Code (CELP).

La figure 1 représente très schématiquement un circuit de compression CELP. Un tel circuit est basé sur une modélisation des cordes vocales et de la chambre de résonance constituée par les cavités buccale, de la gorge et du larynx. Une telle méthode de compression est donc optimisée pour le traitement des signaux de parole.

Les cavités buccale, de la gorge et du larynx sont modélisées par un filtre de "prédiction linéaire" 10 dont la fonction de transfert comporte en général dix pôles. Les cordes vocales sont modélisées par une excitation E traitée par un filtre en peigne 12.

Un signal de parole numérisé S est analysé par trames par un circuit d'analyse 14. A chaque trame, le circuit d'analyse 14 détermine des coefficients a_1 à a_{10} de la fonction de transfert du filtre 10, le pas p du filtre en peigne 12, et un gain G appliqué à l'excitation E en 16 à l'entrée du filtre 12. Les valeurs a_i, p et G sont calculées à chaque trame pour tenir compte respectivement des variations de la cavité buccale, du spectre de fréquence des cordes vocales et de l'amplitude du son. On tente de cette manière d'obtenir que la sortie du filtre 10 soit égale au signal S. Alors, au lieu de transmettre les échantillons du signal S, on transmet les coefficients a_i, p, et G afin qu'un décodeur qui reçoit ces coefficients reconstitue les trames correspondantes du signal S.

Bien entendu, le décodeur doit également connaître l'excitation E à utiliser. La détermination des coefficients ai, p et G ne pose pas de problème particulier. Toutefois, la procédure de recherche de l'excitation optimale reste la plus lourde en termes de charge de calcul et il est toujours d'un grand intérêt de la simplifier même au prix d'une sensible réduction de la qualité de la compression.

A l'origine du codage CELP, l'excitation E était sélectionnée dans une table 18 (appelée "codebook") contenant plusieurs excitations possibles qui représentaient, en fait, des morceaux de bruit blanc. Dans ce cas, un circuit de commande 20 balaye la table 18 jusqu'à ce que la différence e, formée en 22, entre la trame courante du signal S et la trame correspondante en sortie du filtre 10 soit minimale. (Bien entendu, au lieu de comparer le signal S à la sortie du filtre 10, on peut également comparer l'excitation E à la trame du signal S ayant subi le traitement inverse des filtres 10 et 12.)

Avec cette technique, outre les coefficients a_i, p et G, on fournit l'adresse C sélectionnant la meilleure excitation E dans la table 18 à un décodeur muni d'une table homologue.

Chaque excitation contenue dans la table 18 est une séquence d'échantillons numériques correspondant respectivement aux échantillons de chacune des trames du signal à comprimer. Pour que la compression soit de qualité acceptable, il est nécessaire de stocker un nombre relativement important, de l'ordre de 1000, de séquences d'excitation.

Pour limiter la complexité de la procédure de recherche, on a proposé que chaque échantillon d'une séquence d'excitation ne puisse prendre que trois valeurs, à savoir, 0, 1 ou -1 (séquence d'excitation ternaire). On s'est aperçu que cela ne modifiait pas la qualité de la compression de manière perceptible.

La figure 2 représente un exemple de séquence d'excitation E que l'on a proposé pour réduire davantage la complexité de la recherche. Cette séquence d'excitation est dite binaire. Elle comprend plusieurs échantillons non-nuls de valeurs 1 et -1, deux échantillons non-nuls, ou impulsions, étant séparés par un nombre constant d'échantillons nuls, ici 3. Une telle séquence d'excitation peut être représentée par un nombre binaire (ou code d'excitation) C dont les bits sont associés aux impulsions et correspondent à la polarité de celles-ci. En procédant ainsi, le code C fourni par le circuit de commande 20 correspond directement à une séquence d'excitation et la table 18 est supprimée. La complexité est par ailleurs réduite du fait que les échantillons à prendre en compte se réduisent aux impulsions, dont le nombre est, dans l'exemple de la figure 2, quatre fois inférieur au nombre total d'échantillons d'une séquence. En outre, la structure des filtres 10 et 12 est simplifiée.

Cette technique dégrade très légèrement la qualité de la compression, mais cette dégradation est compensée de manière simple par un traitement destiné à supprimer les effets de la régularité de l'espacement des échantillons non-nuls.

A chaque code C est associé un vecteur d'excitation C ayant pour composantes les valeurs 1 et -1 correspondant aux bits à 0 et 1 du code C. Les termes "vecteur" et "code" sont confondus dans la suite de la description.

Afin de réduire davantage le nombre d'essais pour minimiser l'erreur, on a proposé de limiter le nombre de vecteurs ou codes d'excitation possibles à un sous-ensemble représentatif d'un ensemble plus grand. L'article intitulé "A Comparison of some Algebraic Structures for CELP Coding of Speech" de J.P. Adoul et C. Lamblin dans Proc. ICASSP, 1987, décrit une telle méthode. Pour créer un sous-ensemble représentatif de tous les codes C de N bits, on forme l'ensemble des valeurs de n bits (n<N), dont chacune est complétée par N-n bits de correction d'erreur.

Pour trouver le meilleur vecteur d'excitation C, on cherche généralement à maximiser un critère de sélection défini par :

$$m = scal^2(T, C_i)/mod^2(FC_i),$$

où C_i est le vecteur d'excitation essayé ; où T est un vecteur cible formé par des échantillons de la trame en

cours d'analyse du signal S ayant subi le traitement inverse des filtres 10 et 12, ces échantillons étant ceux correspondant aux valeurs 1 et -1 du vecteur C_i et où F désigne la matrice représentant la fonction de transfert des filtres 10 et 12, dans laquelle on n'a conservé que les rangées correspondant aux valeurs 1 et -1 du vecteur C_i. Les notations scal(.,.) et mod(.,.) désignent respectivement le produit scalaire et le module.

L'essai de tous les vecteurs d'excitation C_i selon ce critère représente un nombre de calculs considérable à effectuer entre les arrivées de deux trames du signal S.

Il a été constaté que le dénominateur du critère m est approximativement constant, quel que soit le vecteur d'excitation C_i. Ainsi, le critère m est approximativement maximisé en maximisant le numérateur. Ce numérateur est maximisé lorsque chaque composante du vecteur d'excitation C_i est celle de même signe que l'échantillon correspondant du vecteur cible T. En d'autres termes, un code d'excitation optimal approximatif est obtenu directement en prenant pour ses bits les bits de signe des échantillons du vecteur cible (ou les inverses des bits de signe).

Cette solution n'est pas applicable dans le cas où les codes d'excitation utilisables sont limités à un sous-ensemble représentatif d'un ensemble plus grand obtenu, par exemple, à l'aide d'un code correcteur d'erreurs.

Un objet de la présente invention est de prévoir un procédé permettant de limiter le nombre de calculs nécessaires pour maximiser le critère m susmentionné dans le cas où les codes d'excitation utilisables appartiennent à un sous-ensemble représentatif d'un ensemble plus grand.

Pour atteindre cet objet, la présente invention prévoit un procédé de détermination d'un vecteur d'excitation associé à une trame d'un signal de parole à comprimer, ledit vecteur appartenant à un sous-ensemble associé à un ensemble plus grand de vecteurs d'excitation susceptibles de maximiser un critère, et ayant pour composantes des valeurs 1 et -1 correspondant à une séquence d'échantillons d'excitation d'un filtre de prédiction linéaire. Le critère est égal au carré du rapport entre, d'une part, le produit scalaire du vecteur d'excitation par un vecteur cible formé par des échantillons de la trame ayant subi un filtrage de prédiction linéaire inverse et, d'autre part, le module du vecteur d'excitation ayant subi un filtrage de prédiction linéaire direct. Le procédé comprend les étapes consistant à présélectionner un vecteur d'excitation ayant pour composantes celles de même signe que les échantillons correspondants du vecteur cible, ou celles de signe inverse, et, si le vecteur d'excitation présélectionné n'appartient pas audit sous-ensemble, à sélectionner comme vecteur d'excitation celui qui maximise ledit critère parmi les vecteurs du sous-ensemble qui sont associés respectivement au vecteur présélectionné et aux vecteurs les plus proches de celui-ci dans l'ensemble plus grand.

Selon un mode de réalisation de la présente invention, les vecteurs d'excitation sont associés à des codes d'excitation dont les bits correspondent aux signes des composantes des vecteurs d'excitation, le sous-ensemble de codes d'excitation associé audit sous-ensemble de vecteurs étant formé par des valeurs binaires complétées par des bits de correction d'erreurs, tout code d'excitation étant associé à un code d'excitation du sous-ensemble par une fonction de correction d'erreurs. Le procédé comprend les étapes consistant à former le groupe comprenant le code présélectionné associé au vecteur présélectionné et les codes les plus proches de celui-ci, de manière que chacun de ces codes les plus proches diffère du code présélectionné par un seul bit, à soumettre les codes de ce groupe à la fonction de correction d'erreur pour obtenir un groupe de codes corrigés appartenant au sous-ensemble, et à sélectionner comme code d'excitation, parmi les codes corrigés, celui associé au vecteur qui maximise ledit critère.

Selon un mode de réalisation de la présente invention, les bits de correction d'erreurs sont des bits d'un code correcteur de Hamming.

Ces objets, caractéristiques et avantages ainsi que d'autres de la présente invention seront exposés en détail dans la description suivante de modes de réalisation particuliers, faite à titre non-limitatif à l'aide des figures jointes parmi lesquelles :

la figure 1, précédemment décrite, illustre une méthode de compression CELP ;
la figure 2, précédemment décrite, représente un exemple de séquence d'excitation et du code correspondant ; et
la figure 3 illustre des traitements à effectuer selon la présente invention pour sélectionner un vecteur d'excitation optimal dans le cas où celui-ci appartient à un sous-ensemble obtenu en utilisant un code correcteur d'erreurs.

Pour maximiser le critère m susmentionné, on a constaté que le dénominateur de ce critère, le carré du module du vecteur FC_i, est approximativement constant, quel que soit le vecteur d'excitation C_i. Cette approximation est relativement bonne puisque le module du vecteur C_i est constant. Ainsi, pour maximiser le critère m de manière approchée, il suffit de maximiser un critère simplifié qui est le produit scalaire du vecteur cible T par le vecteur d'excitation C_i. Ce produit scalaire est maximal lorsque chaque composante (1 ou -1) du vecteur C_i est celle de même signe que l'échantillon correspondant du vecteur cible T. On obtient ainsi directement un vecteur d'excitation optimal approché Copt à partir du vecteur cible T.

Ce procédé n'est pas directement applicable dans le cas où les codes d'excitation possibles appartiennent à un sous-ensemble représentatif d'un ensemble plus grand, par exemple lorsque ce sous-ensemble est formé à partir de valeurs de n bits auxquels on ajoute N-n bits d'un code correcteur d'erreurs. En effet, il est alors très probable que le code d'excitation trouvé n'appar-

tienne pas au sous-ensemble. Dans ce cas, on pourrait envisager de ramener le code d'excitation trouvé à un code d'excitation appartenant au sous-ensemble en lui appliquant une fonction de correction d'erreur associée au code correcteur. On tombe alors sur le code d'excitation du sous-ensemble qui est le plus proche du code d'excitation trouvé. Cette "correction d'erreur" a pour conséquence de modifier au moins un bit du code d'excitation, ce bit pouvant avoir dans certains cas une grande influence sur la valeur du critère m, de manière que le code d'excitation final fournisse de mauvais résultats.

A titre d'exemple, on utilise ci-après un code correcteur de Hamming noté H(N, n, 3), où 3 désigne la distance minimale de Hamming séparant deux éléments appartenant au sous-ensemble représentatif. La distance de Hamming entre deux valeurs est définie comme étant le nombre de différences bit à bit entre ces deux valeurs. Avec cette solution, on crée un sous-ensemble de 2^n vecteurs d'excitation de N bits.

Selon l'invention, on forme un groupe de codes d'excitation comprenant un code initial trouvé en maximisant le critère m simplifié ainsi que tous les autres codes obtenus à partir de celui-ci en modifiant un seul bit. Ceci a pour conséquence, en utilisant un code correcteur de Hamming permettant de corriger un bit unique (distance de Hamming minimale 3), que chacun des codes d'excitation du groupe est proche d'un code distinct du sous-ensemble utilisable. Ensuite, on applique à chacun des codes du groupe la fonction de correction d'erreur de Hamming, ce qui ramène chacun des codes du groupe au code le plus proche du sous-ensemble. On obtient un groupe de codes "corrigés" appartenant au sous-ensemble et qui "entoure" le code initialement trouvé. Parmi les codes corrigés, on retient comme code optimal approché celui qui maximise le critère m complet, c'est-à-dire en calculant son numérateur et son dénominateur.

La figure 3 représente sous forme de schéma le procédé selon l'invention qui vient d'être décrit. La trame analysée du signal S subit en 24 le traitement inverse de celui des filtres 10 et 12 de la figure 1. On obtient ainsi un vecteur cible T, dont on ne garde que les échantillons correspondant aux impulsions des séquences d'excitation. En 26, on ne retient des échantillons du vecteur T que les bits de signe (ou leurs inverses) pour fournir un code d'excitation initial C_0. Ce code C_0 est "corrompu" en 28 pour former un groupe de codes comprenant le code C_0 et tous les autres codes C_1 à C_N obtenus en modifiant un seul bit du code C_0. Chaque code C_0 à C_N subit en 30 une "correction d'erreur" pour fournir un groupe de codes corrigés C'_0 à C'_N. En 32, chacun des vecteurs associés aux codes corrigés est comparé au vecteur cible T, et on retient comme code d'excitation optimal approché Copt celui associé au vecteur qui maximise le critère m complet.

Généralement, pour obtenir de meilleurs résultats, la position de la première impulsion des séquences d'excitation E est variable. Dans l'exemple de la figure 2, cette position peut être l'une des quatre premières, ce qui est déterminé par deux bits supplémentaires à transmettre au décodeur et qui multiplie par quatre le nombre de vecteurs d'excitation à essayer. Dans ce cas, on commence par former, pour chacune des quatre positions possibles, un vecteur cible et un vecteur d'excitation dont les composantes sont celles de même signe que celles du vecteur cible correspondant. On retient comme vecteur d'excitation optimal approché, parmi les quatre vecteurs ainsi obtenus, celui qui maximise le critère m complet.

Revendications

1. Procédé de détermination d'un vecteur d'excitation (Copt) associé à une trame d'un signal de parole (S) à comprimer, ledit vecteur appartenant à un sous-ensemble associé à un ensemble plus grand de vecteurs d'excitation susceptibles de maximiser un critère (m), et ayant pour composantes des valeurs 1 et -1 correspondant à une séquence d'échantillons d'excitation (E) d'un filtre de prédiction linéaire (10), ledit critère étant égal au carré du rapport entre, d'une part, le produit scalaire du vecteur d'excitation par un vecteur cible (T) formé par des échantillons de la trame ayant subi un filtrage de prédiction linéaire inverse et, d'autre part, le module du vecteur d'excitation ayant subi un filtrage de prédiction linéaire direct, caractérisé en ce qu'il comprend les étapes suivantes :

   - présélectionner un vecteur d'excitation (C_0) ayant pour composantes celles de même signe que les échantillons correspondants du vecteur cible, ou celles de signe inverse ; et
   - si le vecteur d'excitation (C_0) présélectionné n'appartient pas audit sous-ensemble, sélectionner comme vecteur d'excitation celui qui maximise ledit critère (m) parmi les vecteurs (C'_0... C'_N) du sous-ensemble qui sont associés respectivement au vecteur présélectionné et aux vecteurs (C_1... C_N) les plus proches de celui-ci dans l'ensemble plus grand.

2. Procédé selon la revendication 1, dans lequel les vecteurs d'excitation sont associés à des codes d'excitation dont les bits correspondent aux signes des composantes des vecteurs d'excitation, le sous-ensemble de codes d'excitation associé audit sous-ensemble de vecteurs étant formé par des valeurs binaires complétées par des bits de correction d'erreurs, tout code d'excitation étant associé à un code d'excitation du sous-ensemble par une fonction de correction d'erreurs ; caractérisé en ce qu'il comprend les étapes suivantes :

   - former le groupe comprenant le code présélec-

tionné associé au vecteur présélectionné (C_0) et les codes (C_1... C_N) les plus proches de celui-ci, de manière que chacun de ces codes les plus proches diffère du code présélectionné par un seul bit ;

- soumettre les codes de ce groupe à la fonction de correction d'erreur pour obtenir un groupe de codes corrigés (C'_0... C'_N) appartenant au sous-ensemble ; et
- sélectionner comme code d'excitation (Copt), parmi les codes corrigés, celui associé au vecteur qui maximise ledit critère (m).

3. Procédé selon la revendication 2, caractérisé en ce que les bits de correction d'erreurs sont des bits d'un code correcteur de Hamming.

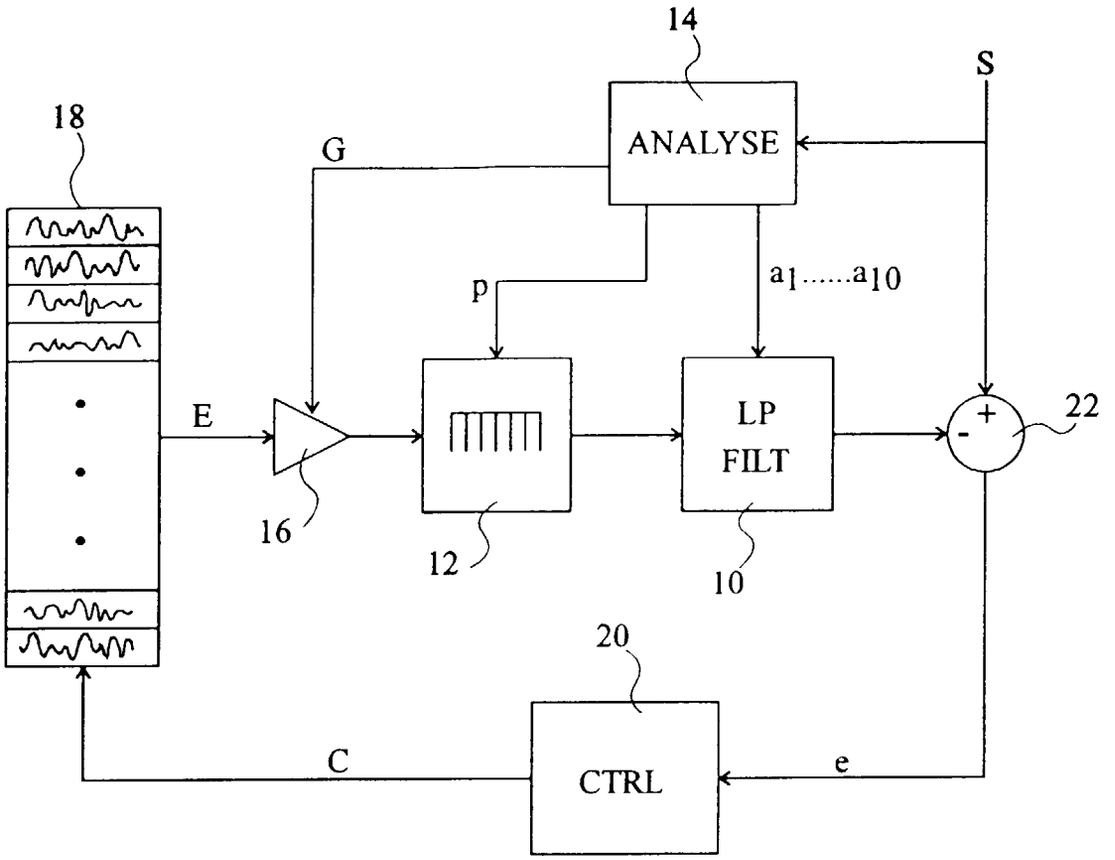

14
S
18
ANALYSE
G
p
a1.....a10
E
LP
FILT
22
16
12
10
20
C
CTRL
e

Fig 1

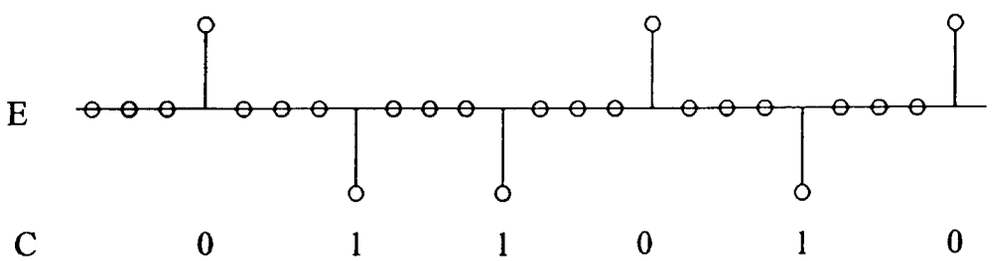

E
C
0
1
1
0
1
0

Fig 2

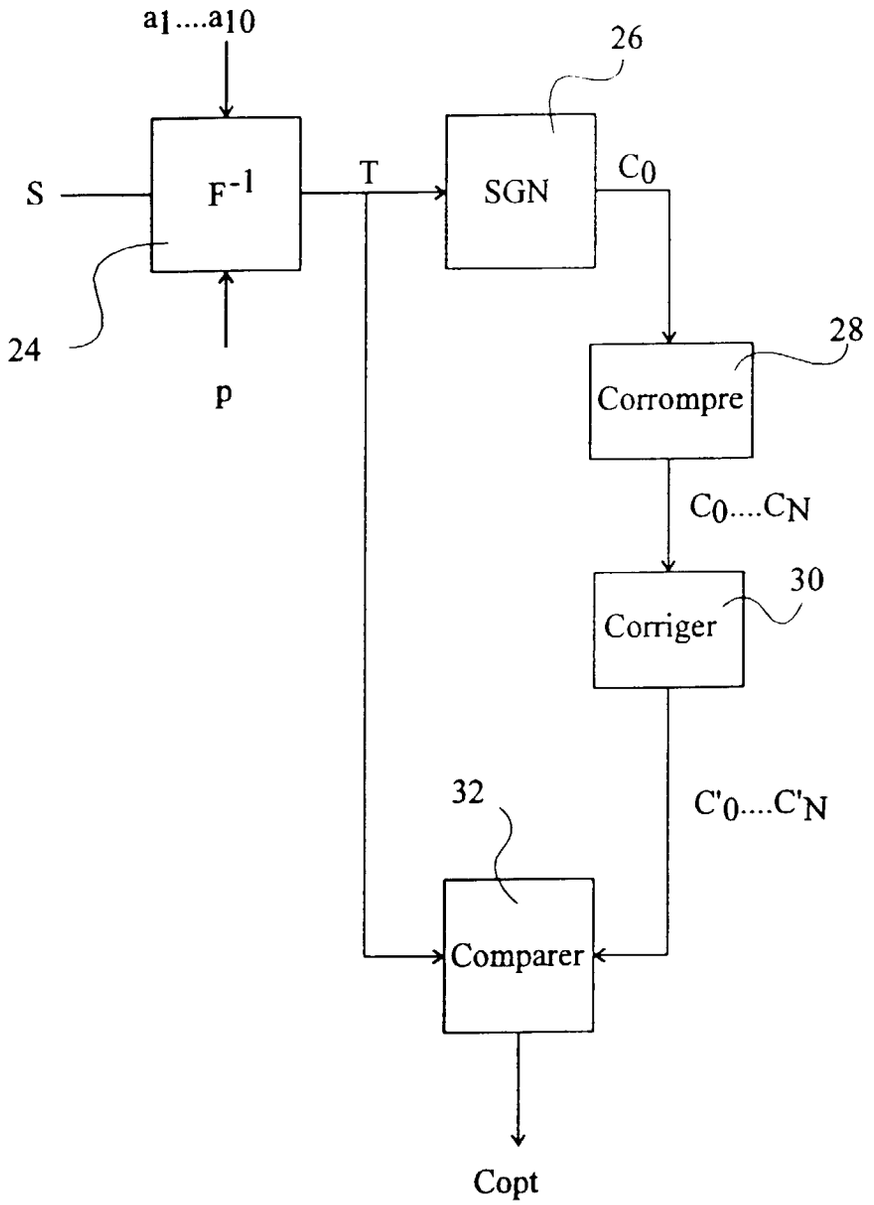

a1....a10
26
S
F-1
T
SGN
C0
24
p
28
Corrompre
C0....CN
30
Corriger
32
C'0....C'N
Comparer
Copt

Fig 3