

TRANSMISSION DE PAROLE A 16 KBIT/S SUR LIAISON

A EVASION DE FREQUENCE FORTEMENT BRUITEE

O. BOSCH - R. GIRAULT

THOMSON-CSF
66, rue du Fossé Blanc / B.P. 156
92231 GENNEVILLIERS CEDEX

RESUME

La nécessité d'établir des liaisons en phonie dans un milieu très perturbé est de plus en plus grande. Une solution classique repose sur le codage de la parole par un codeur Delta et une transmission en évasion de fréquence à 16 kbit/s. Nous présentons une solution originale fondée sur l'utilisation de vocodeurs à bas débit de haute qualité et sur la mise en oeuvre de codes correcteurs d'erreurs performants. Les résultats obtenus à partir de simulations montrent la très bonne qualité de la parole synthétique malgré une liaison fortement bruitée.

ABSTRACT

Transmission of speech on highly perturbed radio links is often needed. A classical solution lies on speech coding with a Delta coder and frequency hopping transmission at 16 kbit/s rate. The solution we propose here is based on high quality low rate multipulse linear predictive coders (MP - LPC) and efficient error correcting codes. The results of simulations show the high quality of the synthetic voice after transmission through a highly perturbed channel.

I - INTRODUCTION

Dans le contexte des télécommunications protégées, c'est-à-dire devant résister au brouillage volontaire ou non, à l'écoute et à l'intrusion, on voit apparaître des systèmes à base d'étalement de spectre, où la bande de fréquences occupée est largement supérieure à ce qui serait strictement nécessaire pour acheminer le débit d'information utile souhaité. Parmi les différentes techniques d'étalement de spectre possible, l'évasion de fréquence (ou E.V.F.) est la plus utilisée pour les fréquences inférieures à 16 GHz. Cette technique consiste à émettre pendant une durée constante (appelée aussi "palier") à fréquence fixe et bande étroite, puis à passer rapidement sur une autre fréquence, à un rythme qui, dans les produits actuellement sur le marché, s'échelonne entre quelques dizaines et quelques centaines de sauts par seconde.

Dans beaucoup d'applications, la transmission en bande étroite de la phonie conduit à l'utilisation de codeurs de parole. Une technique classique de codage de la parole est le codage delta adaptatif à 16 kb/s. La résistance de ce type de codage aux erreurs de transmission est essentiellement liée à la redondance propre de la parole et au pouvoir correcteur de l'appareil auditif.

L'évolution récente des techniques de codage de la parole par prédiction linéaire et codage multi-impulsionnel de l'excitation (MP-LPC) permet de proposer une nouvelle solution au problème posé. En effet, pour un débit en ligne donné, typiquement 16 kb/s, l'utilisation de vocodeurs d'un débit largement inférieur permet de protéger l'information utile par des codes correcteurs-détecteurs d'erreurs efficaces dont l'effet est de remplacer une partie de la redondance naturelle de la parole par de la redondance artificielle capable de corriger à elle seule la plupart des erreurs de transmission.

Ces techniques, purement numériques, permettent alors de garantir un confort d'écoute uniforme quelque soit le taux de paliers brouillés sous réserve qu'il soit inférieur à un seuil défini par la capacité de correction des codes. Dans le cadre du présent exposé on présente les résultats de différentes simulations qui

permettent, en fonction du débit alloué à la parole codée, d'évaluer l'efficacité des stratégies de protection utilisées.

II - CODAGE DE LA PAROLE

Dans les simulations mises en oeuvre pour cette étude, le codage de la parole est réalisé à partir d'une analyse par prédiction linéaire et codage multi-impulsionnel de l'excitation.

La prédiction linéaire, technique bien connue (1), permet d'extraire périodiquement du signal de parole, d'une part les coefficients d'un filtre auto-régressif modélisant le conduit vocal, d'autre part un signal d'excitation représentant l'erreur de prédiction ainsi réalisée.

L'intervalle de temps sur lequel ce filtre est supposé invariant est de l'ordre de 20 ms. Le codage multi-impulsionnel consiste alors à modéliser sur ces intervalles le signal d'excitation par un train d'impulsions dont l'amplitude et la position sont déterminées afin de satisfaire à un critère d'optimalité au sens de la perception auditive (2). L'algorithme de recherche des impulsions est fondé sur une méthode d'analyse-par-synthèse dont le principe est représenté sur la figure 1.

Du à la complexité des algorithmes à mettre en oeuvre et au temps de calcul qui en résulterait, il est difficile d'obtenir le train d'impulsions optimal en une seule itération. Il a donc été proposé de calculer séquentiellement les impulsions et d'obtenir ainsi une séquence sous-optimale (2). On peut, en outre, réduire la quantité de calculs nécessaire à la détermination de cette séquence en remplaçant la méthode d'analyse-par-synthèse par une recherche de maxima d'intercorrélation (3).

Sur la figure 2 on a représenté le schéma synoptique du procédé de codage de la parole décrit ci-avant.

A la synthèse, la reconstitution de la parole est réalisée par un procédé symétrique du précédent (figure 3).

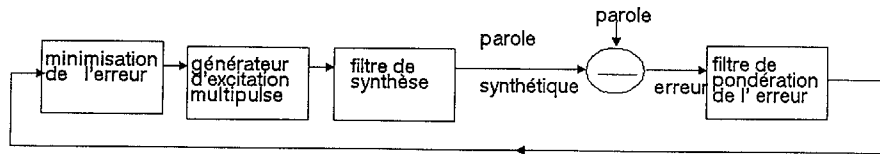
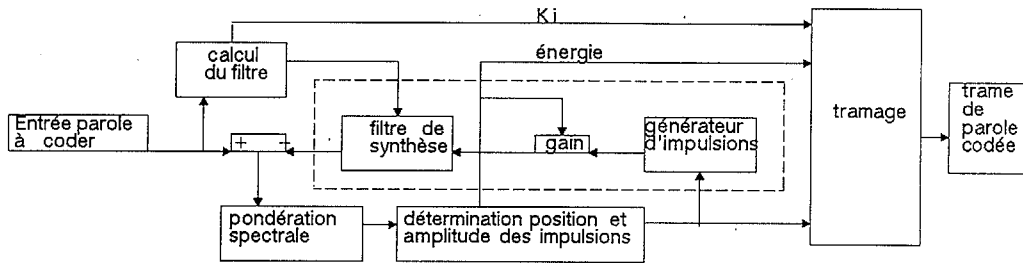
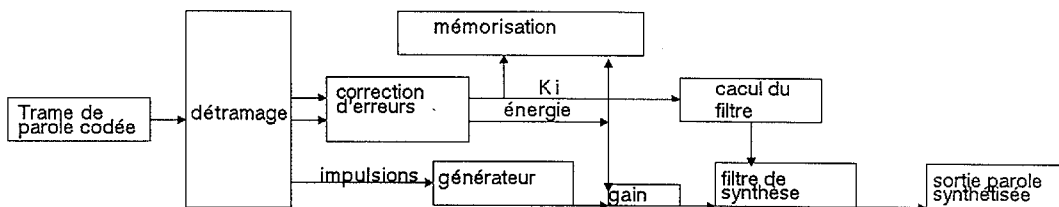


FIGURE 1

ANALYSE
FIGURE 2SYNTHÈSE
FIGURE 3

Après la phase d'analyse, réalisée sur un signal de parole préalablement filtré entre 300 et 3400 Hz et échantillonné à 8 kHz, on dispose pour chaque fenêtre de 25 ms (soit 200 échantillons) des paramètres suivants : les dix coefficients de réflexion K_i du filtre de prédiction, l'amplitude et la position des N impulsions d'excitation.

Avant de constituer la trame à transmettre, il faut quantifier ces paramètres. Pour cela, les dix K_i sont codés globalement sur 36 bits. On quantifie ensuite sur cinq bits l'amplitude de l'impulsion maximale (on a ainsi une mesure de l'énergie du signal dans la fenêtre analysée). Il reste alors à coder l'amplitude des autres impulsions par rapport à ce maximum et leur position respective dans la fenêtre : ceci est réalisé globalement sur huit bits pour chacune des impulsions.

On peut alors constituer la trame de parole codée comme suit :

- $8 \times N$ bits pour les N impulsions
- 36 bits pour les K_i
- 5 bits pour l'amplitude maximale.

Soit, 201 bits par trame pour 20 impulsions par fenêtre ou 161 bits par trame pour 15 impulsions par fenêtre ; ce qui conduit à des débits, sans protection contre les erreurs de transmission, de 8040 et 6440 b/s.

III - PROTECTION DE LA LIAISON

Afin de minimiser l'effet des aléas de transmission (erreurs isolées ou paliers brouillés) sur la qualité de la liaison deux types de protection sont appliqués. D'abord une protection dite externe qui vise à protéger les paramètres sensibles de la trame, puis une protection interne mettant en oeuvre des codes correcteurs-détecteurs d'erreurs et d'effacements associés à des procédés d'entrelacement.

1 - Protection externe de la trame

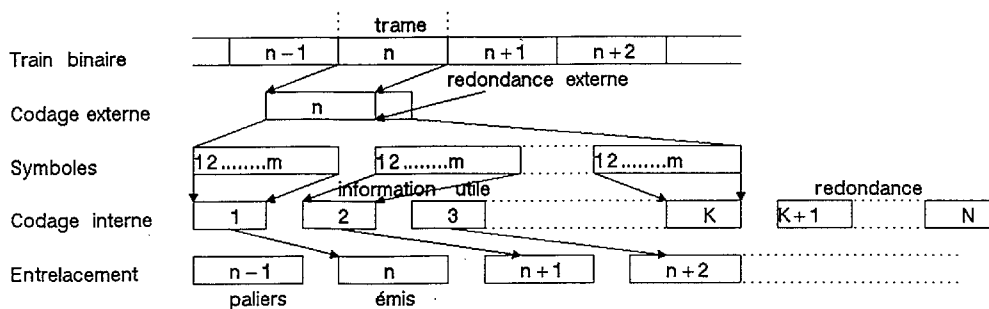
Les coefficients du filtre (K_i) contiennent l'essentiel de l'information spectrale de la trame. Une erreur sur les bits les plus significatifs conduit à une distorsion très perceptible de la parole synthétisée. Les 16 bits les plus significatifs des K_i sont protégés par un code de Hamming modifié (22,16) permettant de détecter deux erreurs et d'en corriger une. Les 5 bits codant l'énergie sont eux aussi sensibles et donc protégés par un code de Hamming modifié (10,5) de mêmes caractéristiques. Lors du décodage, si le code détecte des erreurs sur les K_i , les coefficients du filtre de la trame précédente sont utilisés pour réaliser la synthèse. Si des erreurs sont détectées sur les bits codant l'énergie, la valeur de l'énergie de la trame précédente diminuée de 2 dB est utilisée.

2 - Protection interne de la trame

a - Code correcteur détecteur d'erreurs et d'effacements

Le code de Reed Solomon a été choisi en raison de sa très grande efficacité comparée à celles d'autres codes reposant sur les mêmes principes, code BCH par exemple, et parce qu'il est particulièrement performant en présence d'erreurs ou d'effacements par paquets. En effet, une transmission EVF perturbée se traduit par une perte de paliers, synonyme d'erreurs par paquets.

Chaque trame constituée d'un train binaire est segmentée en K blocs (dits symboles) de m bits d'information utile. Le débit en ligne étant fixé, le nombre maximum N de symboles formant une trame est limité par la contrainte $N < 2^m - 1$. Un mot de code est donc formé de K symboles d'information et $N - K$ symboles de redondance (figure 4).



FORMATION ET COMPOSITION DES PALIERS

FIGURE 4

Un code de Reed Solomon (N, K) est capable de corriger N_{err} erreurs et N_{eff} effacements simultanément tant que la relation $2xN_{err} + N_{eff} \leq N-K$ est vérifiée. Pour notre application, un symbole est déclaré effacé si il appartient à un palier brouillé. (On considère, en effet, que le récepteur est capable de fournir une information de qualité à chaque palier reçu).

La simulation permet de faire varier les paramètres suivants :

- débit utile du vocodeur,
- nombre de bits par symbole,
- paramètres du code interne,
- nombre de symboles par palier,
- statistiques d'apparition des erreurs.

D'autre part le logiciel offre la possibilité d'utiliser deux codes Reed Solomon pour répartir la capacité de correction en fonction de la sensibilité aux erreurs des différents champs d'une trame vocodeur (impulsions ou autres paramètres).

Avec un vocodeur utilisant 20 impulsions par trame, soit un débit utile de 8 kb/s l'utilisation de deux codes RS(32,23) et RS(23,8) donne pour un taux de paliers brouillés de 25% des résultats satisfaisants.

L'augmentation de la redondance disponible peut être obtenue avec un vocodeur générant 15 impulsions par trame (soit 6,4 kb/s de débit utile). Un code RS (56,25) donne alors des résultats remarquables pour un taux de paliers brouillés de 30% et 1% d'erreurs isolées.

L'entrelacement introduit un retard de $2xLxN$ symboles. Pour des paliers portant $L=5$ symboles de 7 bits et une trame formée de $N=56$ symboles, le retard total est de 245 ms. Ce retard est certes important mais les simulations montrent que sans entrelacement, le taux de paliers brouillés ne saurait dépasser dans la configuration précédente une valeur de 15% pour une qualité équivalente.

V - CONCLUSION

L'utilisation conjointe de semi-vocodeurs à faible débit de qualité satisfaisante et de codes correcteurs performants permet de trouver un bon compromis entre la suppression de la redondance intrinsèque de la parole et l'adjonction de redondance artificielle. En effet, la bonne qualité de la parole synthétique obtenue permet de réduire le débit utile alloué à la parole, donc de dégager une redondance importante pour un débit en ligne fixé. Cette redondance permet d'effectuer des transmissions de phonie numérisée en EVF dans un milieu très perturbé et ce, sans dégradation notable des performances tant que la capacité de correction n'est pas dépassée. L'évolution de la technologie permet de proposer dès à présent l'utilisation de ces techniques pour des systèmes de télécommunications avancés tant militaires (tactique) que civils (radiotéléphone cellulaire).

Cette étude a été en partie financée par la Direction des Recherches, Etudes et Techniques du ministère de la défense.

b - Entrelacement

Du fait de la nature impulsive des errata, les erreurs sont corrélées d'un bit à l'autre et donc d'un symbole à l'autre sur un même palier. Afin de perturber indépendamment les différents symboles d'un même mot de code, on émet chaque symbole sur un palier différent. Cette répartition est obtenue à partir d'un entrelacement matriciel. Ce type d'entrelacement entraîne un retard à l'émission et à la réception fonction du nombre de symboles par palier. Si L est le nombre de symboles par palier et N le nombre de symboles par trame le retard total est égal à $2xNxL$ symboles dont une moitié est imputable à l'émission et l'autre à la réception.

IV - SIMULATION

La structure de la simulation, réalisée sur VAX 785, est présentée ci-dessous (figure 5) :

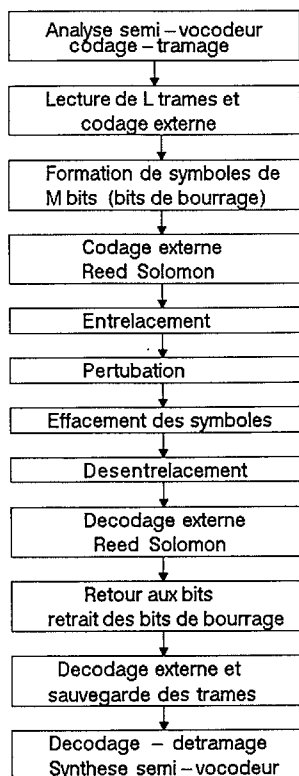


FIGURE 5

REFERENCES

- (1) J.D. MARKEL and A.H. GRAV
*Linear Prediction of Speech. Springer Verlag,
Berlin 1976.*
- (2) B.S. ATAL and J.R. REMDE
*A new model of LPC excitation for producing
natural sounding speech at low bit rates.
ICASSP 1982.*
- (3) T. ARASEKI, K. OZAWA, S. ONO and K. OCHIAI
*Multi-pulse excited coder based on a maximum
cross-correlation search algorithm.
GLOBECOM 1983.*