



UNION INTERNATIONALE DES TÉLÉCOMMUNICATIONS

**UIT-T**

SECTEUR DE LA NORMALISATION  
DES TÉLÉCOMMUNICATIONS  
DE L'UIT

**G.711**

**Appendice II**  
(02/2000)

SÉRIE G: SYSTÈMES ET SUPPORTS DE  
TRANSMISSION, SYSTÈMES ET RÉSEAUX  
NUMÉRIQUES

Systèmes de transmission numériques – Equipements  
terminaux – Codage des signaux analogiques en  
modulation par impulsions et codage

---

Modulation par impulsions et codage (MIC) des  
fréquences vocales

**Appendice II: Définition d'une charge utile de  
bruit de confort pour utilisation en mode G.711  
dans les systèmes de communication  
multimédias par paquets**

Recommandation UIT-T G.711 – Appendice II

(Antérieurement Recommandation du CCITT)

---

RECOMMANDATIONS UIT-T DE LA SÉRIE G  
**SYSTÈMES ET SUPPORTS DE TRANSMISSION, SYSTÈMES ET RÉSEAUX NUMÉRIQUES**

CONNEXIONS ET CIRCUITS TÉLÉPHONIQUES INTERNATIONAUX	G.100–G.199
<b>SYSTÈMES INTERNATIONAUX ANALOGIQUES À COURANTS PORTEURS</b>	
CARACTÉRISTIQUES GÉNÉRALES COMMUNES À TOUS LES SYSTÈMES ANALOGIQUES À COURANTS PORTEURS	G.200–G.299
CARACTÉRISTIQUES INDIVIDUELLES DES SYSTÈMES TÉLÉPHONIQUES INTERNATIONAUX À COURANTS PORTEURS SUR LIGNES MÉTALLIQUES	G.300–G.399
CARACTÉRISTIQUES GÉNÉRALES DES SYSTÈMES TÉLÉPHONIQUES INTERNATIONAUX HERTZIENS OU À SATELLITES ET INTERCONNEXION AVEC LES SYSTÈMES SUR LIGNES MÉTALLIQUES	G.400–G.449
COORDINATION DE LA RADIODÉLÉPHONIE ET DE LA TÉLÉPHONIE SUR LIGNES	G.450–G.499
<b>EQUIPEMENTS DE TEST</b>	
<b>CARACTÉRISTIQUES DES SUPPORTS DE TRANSMISSION</b>	
<b>SYSTÈMES DE TRANSMISSION NUMÉRIQUES</b>	
EQUIPEMENTS TERMINAUX	G.700–G.799
Généralités	G.700–G.709
<b>Codage des signaux analogiques en modulation par impulsions et codage</b>	<b>G.710–G.719</b>
Codage des signaux analogiques par des méthodes autres que la MIC	G.720–G.729
Principales caractéristiques des équipements de multiplexage primaires	G.730–G.739
Principales caractéristiques des équipements de multiplexage de deuxième ordre	G.740–G.749
Caractéristiques principales des équipements de multiplexage d'ordre plus élevé	G.750–G.759
Caractéristiques principales des équipements de transcodage et de multiplication numérique	G.760–G.769
Fonctionnalités de gestion, d'exploitation et de maintenance des équipements de transmission	G.770–G.779
Caractéristiques principales des équipements de multiplexage en hiérarchie numérique synchrone	G.780–G.789
Autres équipements terminaux	G.790–G.799
RÉSEAUX NUMÉRIQUES	G.800–G.899
SECTION NUMÉRIQUES ET SYSTÈMES DE LIGNES NUMÉRIQUES	G.900–G.999

*Pour plus de détails, voir la Liste des Recommandations de l'UIT-T.*

# **RECOMMANDATION UIT-T G.711**

## **MODULATION PAR IMPULSIONS ET CODAGE (MIC) DES FRÉQUENCES VOCALES**

### **APPENDICE II**

#### **Définition d'une charge utile de bruit de confort pour utilisation en mode G.711 dans les systèmes de communication multimédias par paquets**

##### **Résumé**

Le présent appendice définit un format de charge utile (ou train binaire) de bruit de confort pour utilisation en mode G.711 dans les systèmes de communication multimédias par paquets.

Le format de charge utile est destiné à être utilisé dans les systèmes par paquets dotés d'un grand préfixe d'en-tête dans lesquels le débit de transmission par paquets joue un rôle important dans le débit global du système. Dans cette situation, le recours aux algorithmes VAD/DTX/CNG peut notablement réduire le débit de transmission par paquets et, par conséquent, améliorer l'efficacité de la largeur de bande.

##### **Source**

L'Appendice II à la Recommandation UIT-T G.711, élaboré par la Commission d'études 16 (1997-2000) de l'UIT-T, a été approuvé le 28 février 2000 selon la procédure définie dans la Résolution n° 1 de la CMNT.

## AVANT-PROPOS

L'UIT (Union internationale des télécommunications) est une institution spécialisée des Nations Unies dans le domaine des télécommunications. L'UIT-T (Secteur de la normalisation des télécommunications) est un organe permanent de l'UIT. Il est chargé de l'étude des questions techniques, d'exploitation et de tarification, et émet à ce sujet des Recommandations en vue de la normalisation des télécommunications à l'échelle mondiale.

La Conférence mondiale de normalisation des télécommunications (CMNT), qui se réunit tous les quatre ans, détermine les thèmes d'études à traiter par les Commissions d'études de l'UIT-T, lesquelles élaborent en retour des Recommandations sur ces thèmes.

L'approbation des Recommandations par les Membres de l'UIT-T s'effectue selon la procédure définie dans la Résolution n° 1 de la CMNT.

Dans certains secteurs des technologies de l'information qui correspondent à la sphère de compétence de l'UIT-T, les normes nécessaires se préparent en collaboration avec l'ISO et la CEI.

## NOTE

Dans la présente Recommandation, l'expression "Administration" est utilisée pour désigner de façon abrégée aussi bien une administration de télécommunications qu'une exploitation reconnue.

## DROITS DE PROPRIÉTÉ INTELLECTUELLE

L'UIT attire l'attention sur la possibilité que l'application ou la mise en œuvre de la présente Recommandation puisse donner lieu à l'utilisation d'un droit de propriété intellectuelle. L'UIT ne prend pas position en ce qui concerne l'existence, la validité ou l'applicabilité des droits de propriété intellectuelle, qu'ils soient revendiqués par un Membre de l'UIT ou par une tierce partie étrangère à la procédure d'élaboration des Recommandations.

A la date d'approbation de la présente Recommandation, l'UIT n'avait pas été avisée de l'existence d'une propriété intellectuelle protégée par des brevets à acquérir pour mettre en œuvre la présente Recommandation. Toutefois, comme il ne s'agit peut-être pas de renseignements les plus récents, il est vivement recommandé aux responsables de la mise en œuvre de consulter la base de données des brevets du TSB.

© UIT 2000

Droits de reproduction réservés. Aucune partie de cette publication ne peut être reproduite ni utilisée sous quelque forme que ce soit et par aucun procédé, électronique ou mécanique, y compris la photocopie et les microfilms, sans l'accord écrit de l'UIT.

## TABLE DES MATIÈRES

	<b>Page</b>
Appendice II – Définition d'une charge utile de bruit de confort pour utilisation en mode G.711 dans les systèmes de communication multimédias par paquets .....	1
II.1 Domaine d'application .....	1
II.2 Définition de la charge utile du bruit de confort.....	1
II.2.1 Niveau de bruit .....	1
II.2.2 Coefficients de réflexion .....	2
II.2.3 Compression de la charge utile.....	2
II.3 Directives concernant l'utilisation.....	3
II.3.1 Facteurs affectant la qualité de fonctionnement du système .....	3
II.3.2 Illustration des économies de largeur de bande effectuées dans les applications de réseau par paquets.....	4
II.4 Résultats.....	5
II.5 Exemple de solution.....	7
II.5.1 Description de l'algorithme.....	7
II.5.2 Configuration testée.....	10



## Recommandation G.711

### MODULATION PAR IMPULSIONS ET CODAGE (MIC) DES FRÉQUENCES VOCALES

#### APPENDICE II

#### Définition d'une charge utile de bruit de confort pour utilisation en mode G.711 dans les systèmes de communication multimédias par paquets

(Genève, 2000)

##### II.1 Domaine d'application

Le présent appendice définit un format de charge utile (ou train binaire) de bruit de confort pour utilisation dans les systèmes de communication multimédias par paquets selon la Recommandation UIT-T G.711. Le format de la charge utile est générique et peut également être utilisé avec d'autres codecs vocaux dépourvus de fonction de transmission discontinue (DTX, *discontinuous transmission*) intégrée comme ceux des Recommandations UIT-T G.726 [1], G.727 [2], G.728 [3] et G.722 [4]. Le format indique une spécification minimale en matière d'interfonctionnement pour la communication des paramètres relatifs au bruit de confort. L'analyse et la synthèse du bruit de confort ainsi que les algorithmes de détection d'activité vocale (VAD, *voice activity detection*) et DTX ne sont pas spécifiés et sont propres à la mise en œuvre. Toutefois, un exemple de solution a été testé et est décrit dans le présent document. Il fait appel aux algorithmes VAD et DTX décrits à l'Annexe B de la Recommandation G.729 [5] et à un algorithme de génération de bruit de confort (CNG, *comfort noise generation*), donné à titre indicatif.

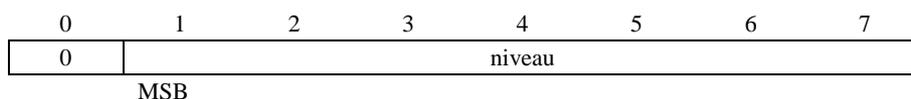
Le format de la charge utile est destiné à être utilisé dans les systèmes par paquets dotés d'un grand préfixe d'en-tête dans lesquels le débit de transmission par paquets joue un rôle important dans le débit global du système. Dans cette situation, le recours aux algorithmes VAD/DTX/CNG peut notablement réduire le débit de transmission par paquets et, par conséquent, améliorer l'efficacité de la largeur de bande.

##### II.2 Définition de la charge utile du bruit de confort

La charge utile du bruit de confort comprend une description du niveau de bruit et des informations relatives au spectre sous forme de coefficients de réflexion. L'utilisation des informations concernant le spectre est facultative et l'ordre du modèle tous pôles n'est pas spécifié. Le codeur peut déterminer l'ordre approprié du modèle en fonction de facteurs tels que qualité, complexité, bruit environnemental prévu et largeur de bande du signal. L'ordre du modèle n'est pas explicitement transmis étant donné qu'il peut être déterminé à partir de la longueur de la charge utile au niveau du récepteur. En raison de la complexité ou pour d'autres motifs, le décodeur peut réduire l'ordre du modèle en mettant à zéro les coefficients de réflexion d'ordre supérieur.

##### II.2.1 Niveau de bruit

Le niveau de bruit est exprimé en  $-dBov$ , les valeurs comprises entre 0 et 127 représentant une gamme comprise entre 0 et  $-127$  dBov. dBov est le niveau par rapport à la surcharge du système. Le niveau de bruit est comprimé avec le bit de plus fort poids (MSB, *most significant bit*) en premier lieu, le bit non utilisé étant toujours mis à 0 conformément à la Figure II.1.



**Figure II.1/G.711 – Compression binaire du niveau de bruit**

## II.2.2 Coefficients de réflexion

Les informations spectrales sont transmises au moyen des coefficients de réflexion [6]. A partir du polynôme

$$A(z) = 1 - \sum_{j=1}^M \alpha_j z^{-j}$$

obtenu par prédiction linéaire, on peut obtenir l'ensemble des coefficients de réflexion à partir de l'ensemble des coefficients de codage prédictif linéaire (LPC) au moyen d'une récurrence vers l'arrière de type

$$k_i = -a_i^{(i)}$$

$$a_j^{(i-1)} = \frac{a_j^{(i)} + a_i^{(i)} a_{i-j}^{(i)}}{1 - k_1^2} \quad 1 \leq j \leq i-1$$

où  $i$  va de  $M$  à 1 avec la condition initiale

$$a_j^{(M)} = \alpha_j \quad 1 \leq j \leq M$$

A noter que les formules ci-dessus donnent comme solution:

$$k_1 = -\frac{r_i}{r_0}$$

où  $r_i$  est le  $i^{\text{e}}$  coefficient d'autocorrélation du signal d'entrée.

Chaque coefficient de réflexion peut avoir des valeurs comprises entre  $-1$  et  $1$  et est quantifié uniformément sur 8 bits. La valeur quantifiée est représentée par l'indice  $N$  à 8 bits, dans lequel  $N = 0, \dots, 254$ , l'indice  $N = 255$  étant réservé pour un usage ultérieur. Chaque indice  $N$  est comprimé dans un octet séparé, le bit de plus fort poids venant en premier. La valeur quantifiée de chaque coefficient de réflexion peut être obtenue à partir de son indice correspondant comme suit:

$$\hat{k}_i(N_i) = \frac{258}{32768} \cdot (N_i - 127) \quad \text{pour } N_i = 0, \dots, 254; -1 < \hat{k}_i(N_i) < 1$$

## II.2.3 Compression de la charge utile

Le premier octet de la charge utile doit contenir le niveau de bruit comme indiqué à la Figure II.1. Les coefficients de réflexion quantifiés sont comprimés dans les octets ultérieurs dans un ordre croissant comme indiqué à la Figure II.2,  $M$  étant l'ordre du modèle.



**Figure II.2/G.711 – Format de compression de la charge utile du bruit de confort**

La longueur totale de la charge utile est de  $M + 1$  octets. Il convient de noter qu'avec un modèle d'ordre 0 (c'est-à-dire sans informations concernant l'enveloppe spectrale), seul le niveau d'énergie est transmis.

### II.3 Directives concernant l'utilisation

On trouvera à la Figure II.3 un schéma fonctionnel représentant le système de communication vocale doté des fonctions VAD/DTX/CNG. La fonction VAD a pour rôle de différencier les segments vocaux actifs des segments vocaux inactifs dans le signal d'entrée. Pendant les segments vocaux inactifs, le rôle de la fonction CNG est de décrire suffisamment le bruit ambiant tout en réduisant au minimum le débit de transmission. Une trame de descripteur d'insertion de silence (SID, *silence insertion descriptor*) contenant une description du bruit est comprimée dans la charge utile du bruit de confort (CN) et envoyée au récepteur. L'algorithme DTX détermine quand une trame SID est transmise. La trame SID peut être envoyée périodiquement ou seulement en cas de modification importante des caractéristiques du bruit de fond. Au niveau du récepteur, l'algorithme CNG utilise les informations contenues dans la trame SID pour mettre à jour le modèle de génération de bruit et produire ensuite la quantité voulue de bruit de confort.

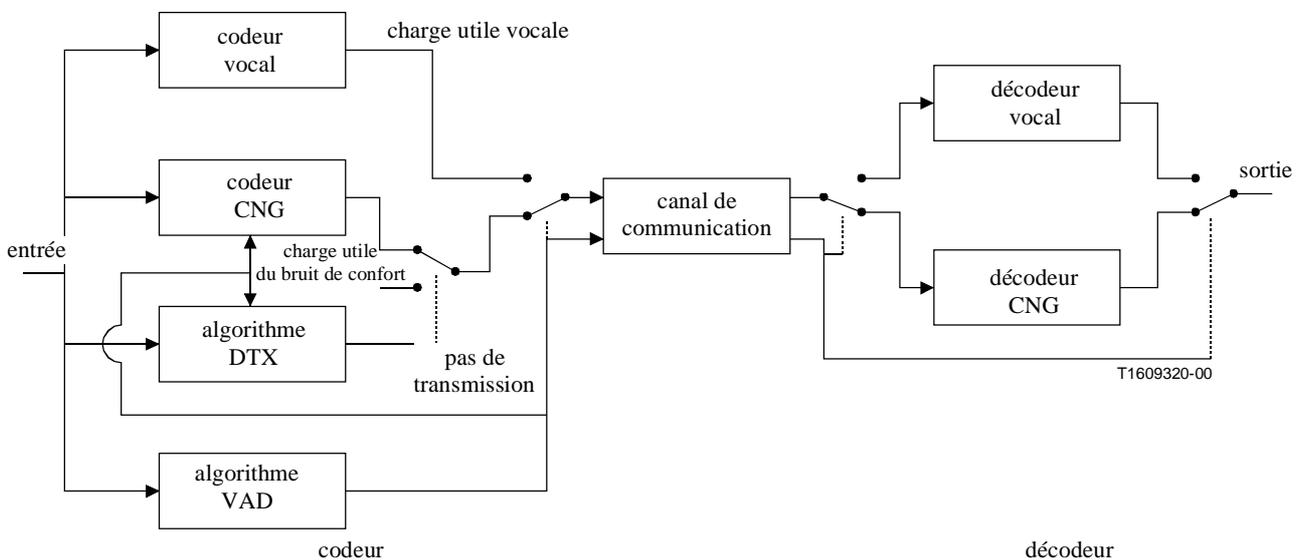


Figure II.3/G.711 – Système de communication vocale doté de la fonction DTX

#### II.3.1 Facteurs affectant la qualité de fonctionnement du système

Les éléments VAD/DTX/CNG ont pour objet de réduire le débit de transmission pendant les périodes d'inactivité vocale tout en maintenant un niveau acceptable de qualité de sortie. La qualité et l'efficacité sont affectées par la qualité de fonctionnement de chacun des éléments susmentionnés. Il faut veiller à prendre conjointement en compte les caractéristiques des algorithmes VAD, DTX et CNG, sinon il peut en résulter une mauvaise qualité de fonctionnement du système.

##### II.3.1.1 Algorithme VAD

L'algorithme VAD a pour rôle de classer le signal d'entrée comme signal vocal actif, signal vocal inactif ou bruit de fond. Le fait de classer un signal vocal inactif comme signal vocal actif affecte l'efficacité du système en ce sens que cela entraîne une augmentation inutile du débit de transmission. Dans ce cas, la qualité vocale n'est pas affectée. Toutefois, lorsqu'un signal vocal actif est classé par erreur comme signal vocal inactif, il est écrété et la qualité vocale se détériore. La plupart des algorithmes DTX appliquent une période de maintien lors du passage d'une période

d'activité vocale à une période d'inactivité vocale pour éviter un écrêtage de la fin du signal vocal. Pendant la période de maintien, les trames vocales inactives sont reclassées comme trames actives. La période de maintien est également importante en ce sens qu'elle permet au codeur CNG d'obtenir une estimation précise du bruit ambiant.

### II.3.1.2 Algorithme DTX

L'algorithme DTX détermine la fréquence de transmission des trames SID pendant les périodes d'inactivité vocale. Les algorithmes simples envoient une mise à jour périodique (par exemple de 5 à 30 Hz). Les algorithmes plus complexes analysent le signal d'entrée et ne procèdent à une transmission qu'en cas de détection d'une modification importante du bruit ambiant [5].

### II.3.1.3 Algorithme CNG

L'algorithme CNG a pour fonction de décrire et de reproduire le bruit ambiant. Celui-ci peut être décrit de manière appropriée par son énergie et son contenu spectral. Pour éviter de brusques changements des caractéristiques du bruit de confort, il est important d'établir la moyenne des estimations de paramètres pendant une période de temps. La quantité appropriée de calculs de valeurs moyennes dépend du bruit ambiant, de la qualité de fonctionnement et du temps de maintien de l'algorithme VAD, ainsi que de la fréquence de mise à jour de l'algorithme DTX.

L'ordre du modèle utilisé est un facteur qui contribue à la précision de l'estimation des valeurs spectrales. L'ordre optimal dépend du bruit ambiant présent et de la largeur de bande du signal. Il est également primordial de faire correspondre les caractéristiques spectrales de bruit produit par l'algorithme CNG avec celles du codec de parole. En conséquence, il est proposé que tout prétraitement du signal d'entrée avant analyse dans le codeur de parole soit également effectué dans le codeur de bruit de confort.

## II.3.2 Illustration des économies de largeur de bande effectuées dans les applications de réseau par paquets

Le Tableau II.1 montre comment le recours à la transmission discontinue dans un système de communication par paquets peut notablement réduire le débit de transmission et, en conséquence, améliorer l'efficacité de la largeur de bande. L'exemple donné est fondé sur l'utilisation d'un préfixe de paquet de 40 octets, d'une activité vocale de 60% et d'une fréquence de mise à jour de l'algorithme DTX de 10 Hz.

**Tableau II.1/G.711 – Economies effectuées en matière de largeur de bande**

Codec	Débit (bit/s)	Taille des paquets (ms)	Débit IP (bit/s)	Charge utile du bruit de confort 1 octet		Charge utile du bruit de confort 11 octets	
				Débit IP (débit moyen en bit/s)	Economies effectuées (%)	Débit IP (débit moyen en bit/s)	Economies effectuées (%)
G.711	64 000	5 ms	128 000	78 112	39,0	78 432	38,7
G.711	64 000	10 ms	96 000	58 912	38,6	59 232	38,3
G.711	64 000	20 ms	80 000	49 312	38,4	49 632	38,0
G.726	32 000	5 ms	96 000	58 912	38,6	59 232	38,3
G.726	32 000	10 ms	64 000	39 712	38,0	40 032	37,5
G.726	32 000	20 ms	48 000	30 112	37,3	30 432	36,6
G.728	16 000	5 ms	80 000	49 312	38,4	49 632	38,0
G.728	16 000	10 ms	48 000	30 112	37,3	30 432	36,6
G.728	16 000	20 ms	32 000	20 512	35,9	20 832	34,9

Avec, par exemple, un en-tête RTP/UDP/IP de 40 octets, une activité vocale de 60% et une fréquence de mise à jour DTX de 10 Hz, le débit IP moyen selon la Recommandation G.711 et avec une charge utile de bruit de confort de 11 octets est donné par la formule suivante:  $((64\ 000\ \text{bit/s}) + (40\ \text{octets} \times 8\ \text{bit/octet} \times (1,0/0,005\ \text{s}))) \times (0,6) + ((40 + 11)\ \text{octets} \times 8\ \text{bit/octet} \times 10/\text{s}) \times (0,4) = 78\ 432\ \text{bit/s}$ .

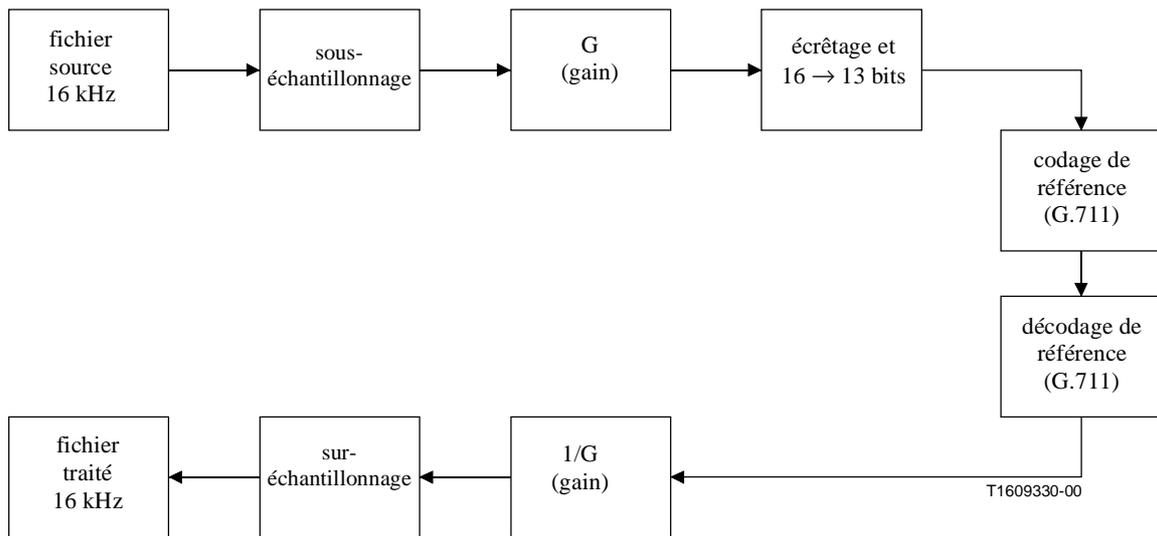
## II.4 Résultats

On a procédé à une évaluation subjective d'un exemple d'implémentation de l'algorithme CNG utilisant la charge utile du bruit de confort. La méthode d'évaluation utilisée est la méthode d'évaluation par catégories absolues (ACR, *absolute category rating*) définie dans la Recommandation UIT-T P.800. Le matériel vocal utilisé au cours de l'expérience consistait en des phrases simples, courtes et significatives prononcées en anglais nord-américain. Les séquences vocales sources ont été filtrées selon le système IRS modifié (Recommandation UIT-T P.830, Annexe D) et disposées par paires. Les phrases duraient de 7 à 8 secondes et étaient séparées par un intervalle de 1 seconde environ. L'évaluation a été effectuée à la fois dans des conditions non bruyantes et dans des conditions bruyantes (bruits de bavardage, de rue, de bureau et de voiture).

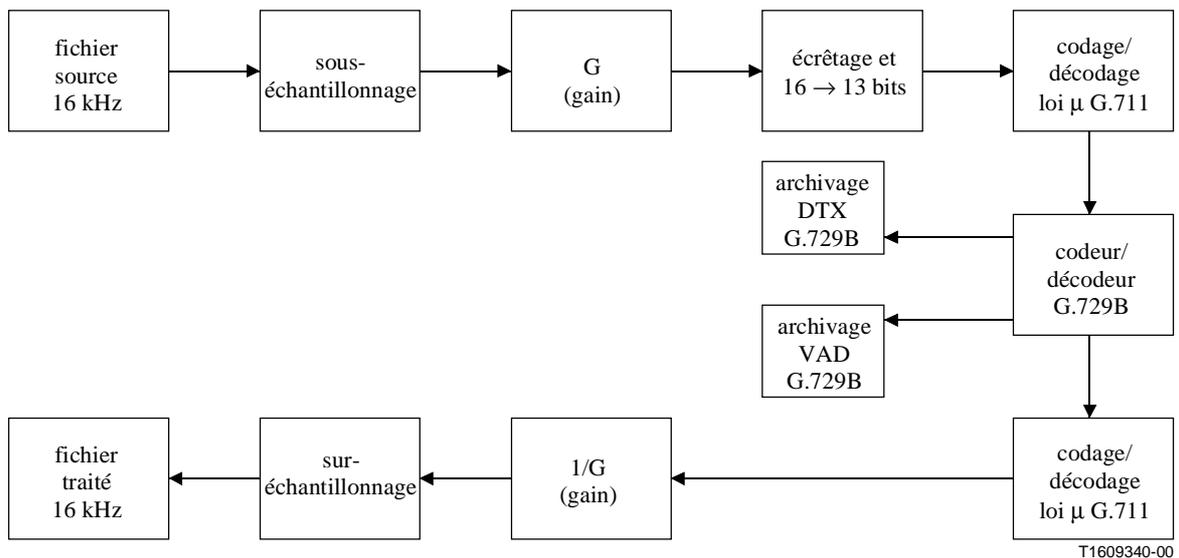
Le codec vocal utilisé lors de l'expérience était un codec G.711, le traitement appliqué étant celui de la procédure indiquée à la Figure II.4. Seul l'algorithme CNG a été mis en œuvre. On a utilisé les algorithmes VAD et DTX décrits à l'Annexe B de la Recommandation G.729 [5]. Les fichiers d'archivage contenant les décisions VAD et DTX ont été obtenus selon la procédure indiquée à la Figure II.5, l'indicateur "SYNC" étant validé pour aligner le fichier de sortie sur le fichier d'entrée.

Le traitement G.711 avec bruit de confort a été obtenu selon la procédure indiquée à la Figure II.6. Le fichier source a été sous-échantillonné et ajusté en ce qui concerne le niveau au moyen d'un gain G, puis codé au moyen du codage G.711 combiné à l'algorithme CNG. Les données d'entrée ont été mises en tampon dans des trames de 10 ms. Le codage de trame de l'algorithme CNG a été aligné en fonction du début du fichier de paroles en vue d'une "synchronisation" avec le tramage correspondant aux fichiers d'archivage VAD et DTX. Sur la base d'une trame de 10 ms, les fichiers d'archivage VAD et DTX ont permis de contrôler le fonctionnement de l'algorithme CNG. Pour les trames vocales actives, le processus G.711 a été appliqué pour le traitement des trames de données d'entrée. Pour les trames inactives, c'est l'algorithme CNG qui a été utilisé. L'indicateur DTX a commandé la mise à jour des paramètres CNG. Au niveau du décodeur, l'indicateur VAD a permis d'indiquer si la trame courante était une trame vocale active ou inactive. Un gain complémentaire 1/G (destiné à produire un niveau d'écoute constant) a ensuite été appliqué et le résultat a été suréchantillonné et mémorisé comme "fichier traité".

Les résultats de cette expérience ACR effectuée avec bruit ambiant ont montré que, dans tous les cas intéressants, le traitement G.711 combiné à l'algorithme d'essai CNG donne des résultats équivalents au traitement G.711 sans algorithme VAD/CNG. Ces résultats concernent les cas sans bruit de fond comme les cas avec bruit de fond (bavardage, voiture, bureau et rue).

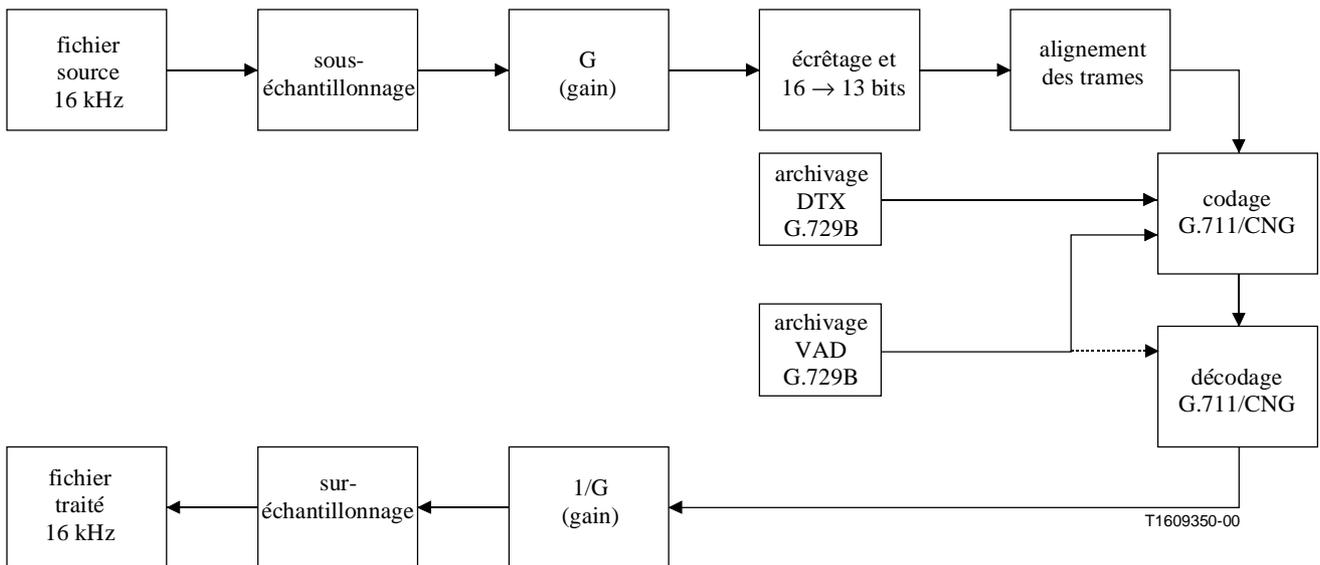


**Figure II.4/G.711 – Traitement G.711 sans algorithme CNG**



G.729B est l'Annexe B à la Recommandation G.729

**Figure II.5/G.711 – Traitement G.729B en vue de l'obtention des fichiers d'archivage VAD/DTX**



G.729B est l'Annexe B à la Recommandation G.729

**Figure II.6/G.711 – Traitement G.711 avec algorithme CNG**

## II.5 Exemple de solution

Le présent sous-paragraphe décrit un système de génération de bruit de confort utilisant le format de charge utile de bruit de confort décrit dans le présent appendice et auquel il a été fait appel pour l'évaluation décrite au II.4.

### II.5.1 Description de l'algorithme

#### II.5.1.1 Codeur

Le programme appelant doit solliciter le codeur à chaque trame. Pour les trames vocales actives, le signal d'entrée est prétraité et les tampons internes sont mis à jour avant renvoi. Pour les trames inactives, les estimées de l'énergie et du contenu spectral du bruit de fond sont mises à jour. Dans le cas d'une trame SID, les paramètres estimés sont quantifiés et insérés dans le tampon de canal pour transmission vers le décodeur. La fréquence de mise à jour des trames SID a été déterminée par l'algorithme DTX conformément à la Recommandation G.729, Annexe B [5]. On trouvera dans les sous-paragraphe suivants les caractéristiques du codeur CNG.

##### II.5.1.1.1 Prétraitement

Un filtre passe-haut IIR (réponse impulsionnelle infinie) du premier ordre soumet le signal d'entrée à un prétraitement pour supprimer tout élément basse fréquence non souhaité. Le filtre passe-haut est donné par:

$$H(z) = \frac{1 - z^{-1}}{1 - (127/128)z^{-1}}$$

### II.5.1.1.2 Analyse d'autocorrélation

Les coefficients d'autocorrélation normalisés  $r_m$  et l'énergie de trame  $E$  sont calculés à partir du signal prétraité fenêtré au moyen d'une fenêtre asymétrique de 25 ms. Pour une fréquence d'échantillonnage de 8,0 kHz, la fenêtre est donnée par:

$$w(n) = \begin{cases} 0,54 - 0,46 \cos\left(\frac{2\pi n}{339}\right) & n = 0, 1, \dots, 169 \\ \cos\left(\frac{2\pi(n-170)}{119}\right) & n = 170, 171, \dots, 199 \end{cases}$$

Les moyennes glissantes des coefficients d'autocorrélation normalisés et de l'énergie de trame sont ensuite calculées pour la  $n^{\text{ie}}$  trame comme suit:

$$\begin{aligned} \bar{r}_m(i) &= \bar{r}_m(i-1) \cdot \beta_1 + \bar{r}_m(i) \cdot (1,0 - \beta_1) & m = 1, 2, \dots, M \\ \overline{LE}(i) &= \overline{LE}(i-1) \cdot \beta_2 + \overline{LE}(i) \cdot (1,0 - \beta_2) \end{aligned}$$

où  $LE$  est le logarithme de base 2 de l'énergie de trame et  $M$  l'ordre du modèle.  $\beta_1$  et  $\beta_2$  sont des constantes qui dépendent de la taille de trame. Si celle-ci est inférieure ou égale à 7,5 ms,  $\beta_1$  et  $\beta_2$  sont mis à 0,8; dans le cas contraire, ils sont mis à 0,6. Les moyennes sont réinitialisées aux valeurs de trames courantes si la trame précédente était une trame vocale active.

### II.5.1.1.3 Calcul du coefficient de réflexion

L'erreur quadratique moyenne entre les coefficients d'autocorrélation normalisés instantanés et moyens est calculée selon l'équation suivante:

$$d = \frac{1}{M} \sum_{m=1}^M (\bar{r}_m(i) - r_m(i))^2$$

Si  $d$  est inférieur à un seuil adaptatif  $Th$  et que la dernière trame est inactive, les coefficients moyens  $\bar{r}_m(i)$  sont utilisés pour le calcul des coefficients de réflexion, sinon ce sont les coefficients instantanés  $r_m(i)$  qui sont utilisés. Le seuil  $Th$  est déterminé à chaque trame selon l'algorithme suivant:

```
if (PrevVad == 1)
    Th = 0.0
else
    Th += 0.2857 * (FRAME_SIZE / SAMPLING_RATE)
    if (Th > 0.06)
        Th = 0.06
    end
end
```

Les coefficients de réflexion  $k_m(i)$  sont calculés à partir des coefficients d'autocorrélation choisis à l'aide de l'algorithme de Levinson-Durbin.

### II.5.1.1.4 Quantification

Pour les trames de descripteur d'insertion de silence (SID, *silence insertion descriptor*), l'énergie  $\overline{LE}(i)$  et les coefficients de réflexion  $k_m(i)$  sont quantifiés et comprimés selon le format spécifié pour la charge utile.

### II.5.1.2 Décodeur

Le décodeur produit un bruit de confort en faisant passer un signal d'excitation de bruit blanc mis à l'échelle à travers un filtre de synthèse de prédiction linéaire. On trouvera les détails du processus dans les sous-paragraphes suivants.

#### II.5.1.2.1 Mise à jour des paramètres

Les coefficients de réflexion provenant de la dernière trame SID reçue sont utilisés dans la trame courante. Soit  $LE_{SID}$  les derniers paramètres de bruit de confort reçus, l'énergie ayant été convertie d'unités dBov en logarithme de base 2. L'énergie utilisée dans la trame courante est donnée par l'équation suivante:

$$LE(i) = LE(i-1) \cdot \alpha + LE_{SID} \cdot (1,0 - \alpha)$$

où  $\alpha = 0,9$ . Cette procédure de lissage est effectuée pour éviter toute modification brutale de l'énergie du signal dans le bruit de confort.

#### II.5.1.2.2 Génération du signal d'excitation

Un générateur de nombres aléatoires à distribution Gaussienne est utilisé pour produire la séquence  $Rn$  qui est mise à l'échelle au moyen du facteur  $\eta$  par rapport à l'énergie correcte selon l'équation suivante:

$$\eta = \sqrt{\frac{E(i) \cdot \prod_{m=1}^M (1,0 - \hat{k}(N_m)^2)}{\frac{1}{L} \cdot \sum_{j=0}^{L-1} Rn(j)^2}}$$

où  $L$  est la longueur du signal d'excitation et  $E(i)$  l'énergie de trame.

Une approximation constante du dénominateur de l'équation ci-dessus permet d'éviter l'opération produit scalaire et de réduire la complexité.

#### II.5.1.2.3 Synthèse de prédiction linéaire

Les coefficients de réflexion sont convertis en coefficients de prédiction linéaire pour être utilisés dans le filtre de synthèse de prédiction linéaire (LP, *linear prediction*) selon la récurrence suivante [6]:

$$\begin{aligned} a_i^{(i)} &= -\hat{k}_i(N_i) \\ a_j^{(i)} &= a_j^{(i-1)} + \hat{k}_i(N_i) a_{i-j}^{(i-1)} \quad 1 \leq j \leq i-1 \end{aligned}$$

ces équations étant résolues pour  $i = 1, 2, \dots, M$  et l'ensemble final étant défini comme suit:

$$\alpha_j = a_j^{(p)} \quad 1 \leq j \leq M$$

Le filtre de synthèse de prédiction linéaire est défini comme suit:

$$\frac{1}{A(z)} = \frac{1}{1 - \sum_{j=1}^M \alpha_j z^{-j}}$$

Le signal d'excitation est transféré à travers le filtre en vue de produire le bruit de confort final. La longueur du signal d'excitation  $L$  est généralement égale à la longueur de trame. Cependant, pour la première trame inactive suivant une trame active,  $L$  est égal à la longueur de trame majorée de

l'ordre ( $M$ , *model order*) du modèle. Dans ce cas, il n'est pas tenu compte des premiers échantillons de sortie  $M$  provenant du filtre de synthèse.

### II.5.1.3 Délai

L'algorithme relatif au bruit de confort ne comprend intrinsèquement aucun délai.

### II.5.1.4 Complexité

L'algorithme a été mis en œuvre dans un système de 16 bits à virgule fixe au moyen de la bibliothèque des opérateurs de base de l'UIT-T. On trouvera au Tableau II.2 un résumé de l'utilisation de la mémoire et des ressources pour différentes tailles de trame avec une fréquence d'échantillonnage de 8,0 kHz et pour un modèle tous pôles du dixième ordre. Le nombre WMOPS (millions d'opérations pondérées par seconde) est obtenu au moyen du compteur d'opérations de la bibliothèque et représente le cas le plus défavorable. La ROM est de la taille estimée sur un processeur DSP type à virgule fixe.

**Tableau II.2/G.711 – Ressources CNG nécessaires pour un modèle du dixième ordre**

Taille de trame	RAM (mots)	ROM (mots)	WMOPS
5 ms	650	1300	1,1
10 ms	690	1300	0,66
20 ms	760	1300	0,47

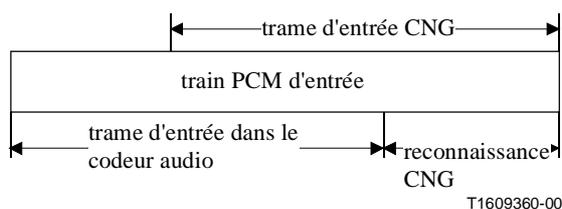
## II.5.2 Configuration testée

L'algorithme testé est spécifié au Tableau II.3.

**Tableau II.3/G.711 – Configuration CNG testée**

Paramètre	Valeur d'essai
Fréquence d'échantillonnage	8,0 kHz
Taille de trame	10 ms
Ordre du modèle	10
Temps de reconnaissance	5 ms

Au cours de l'essai, on a ajouté un temps de reconnaissance de 5 ms en retardant l'arrivée du signal d'entrée dans le codec de parole associé (G.711) comme indiqué à la Figure II.7. On a introduit le temps de reconnaissance pour adapter de manière appropriée l'utilisation de l'algorithme VAD spécifié dans la Recommandation G.729, Annexe B, à l'exemple de solution CNG. Dans la pratique, il est possible d'éviter d'introduire un temps de reconnaissance en ajoutant une trame de maintien supplémentaire à l'algorithme VAD spécifié par la Recommandation G.729, Annexe B.



**Figure II.7/G.711 – Temps de reconnaissance CNG ajouté pendant l'essai**

## Références

- [1] Recommandation CCITT G.726 (1990), *Modulation par impulsions et codage différentiel adaptatif (MICDA) à 40, 32, 24, 16 kbit/s.*
- [2] Recommandation CCITT G.727 (1990), *Modulation par impulsions et codage différentiel adaptatif (MICDA) imbriqué à 5, 4, 3 et 2 bits par échantillon.*
- [3] Recommandation CCITT G.728 (1992), *Codage de la parole à 16 kbit/s en utilisant la prédiction linéaire à faible délai avec excitation par code.*
- [4] Recommandation CCITT G.722 (1988), *Codage audiofréquence à 7 kHz à un débit inférieur ou égal à 64 kbit/s.*
- [5] Recommandation UIT-T G.729 Annexe B (1996), *Schéma de compression des silences pour la Recommandation G.729 optimisé pour les terminaux conformes à la Recommandation V.70.*
- [6] RABINER (L.R.), SCHAFER (R.W.): *Digital processing of speech signals*, Prentice-Hall, 1978.
- [7] Recommandation UIT-T G.191 (1996), *Outils logiciels pour la normalisation du codage des signaux vocaux et audiofréquence.*





## SERIES DES RECOMMANDATIONS UIT-T

Série A	Organisation du travail de l'UIT-T
Série B	Moyens d'expression: définitions, symboles, classification
Série C	Statistiques générales des télécommunications
Série D	Principes généraux de tarification
Série E	Exploitation générale du réseau, service téléphonique, exploitation des services et facteurs humains
Série F	Services de télécommunication non téléphoniques
<b>Série G</b>	<b>Systèmes et supports de transmission, systèmes et réseaux numériques</b>
Série H	Systèmes audiovisuels et multimédias
Série I	Réseau numérique à intégration de services
Série J	Transmission des signaux radiophoniques, télévisuels et autres signaux multimédias
Série K	Protection contre les perturbations
Série L	Construction, installation et protection des câbles et autres éléments des installations extérieures
Série M	RGT et maintenance des réseaux: systèmes de transmission, de télégraphie, de télécopie, circuits téléphoniques et circuits loués internationaux
Série N	Maintenance: circuits internationaux de transmission radiophonique et télévisuelle
Série O	Spécifications des appareils de mesure
Série P	Qualité de transmission téléphonique, installations téléphoniques et réseaux locaux
Série Q	Commutation et signalisation
Série R	Transmission télégraphique
Série S	Equipements terminaux de télégraphie
Série T	Terminaux des services télématiques
Série U	Commutation télégraphique
Série V	Communications de données sur le réseau téléphonique
Série X	Réseaux pour données et communication entre systèmes ouverts
Série Y	Infrastructure mondiale de l'information et protocole Internet
Série Z	Langages et aspects informatiques généraux des systèmes de télécommunication