UNION INTERNATIONALE DES TÉLÉCOMMUNICATIONS

UIT-T

SECTEUR DE LA NORMALISATION
DES TÉLÉCOMMUNICATIONS
DE L'UIT

G.723.1

Annexe A

(11/96)

SÉRIE G: SYSTÈMES ET SUPPORTS DE TRANSMISSION

Systèmes de transmission numériques – Equipements terminaux – Codage des signaux analogiques par des méthodes autres que la MIC

Codeur vocal à double débit pour communications multimédias acheminées à 5,3 et à 6,3 kbit/s

Annexe A: Schéma de compression des silences

Recommandation UIT-T G.723.1 – Annexe A

(Antérieurement Recommandation du CCITT)

RECOMMANDATIONS UIT-T DE LA SÉRIE G

SYSTÈMES ET SUPPORTS DE TRANSMISSION

CONNEXIONS ET CIRCUITS TÉLÉPHONIQUES INTERNATIONAUX	G.100–G.199
SYSTÈMES INTERNATIONAUX ANALOGIQUES À COURANTS PORTEURS	
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 RADIOTÉLÉPHONIE ET DE LA TÉLÉPHONIE SUR LIGNES	G.450–G.499
CARACTÉRISTIQUES DES SUPPORTS DE TRANSMISSION	
SYSTÈMES DE TRANSMISSION NUMÉRIQUES	
EQUIPEMENTS TERMINAUX	G.700–G.799
Généralités	G.700–G.709
Codage des signaux analogiques en modulation par impulsions et codage	G.710–G.719
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
Généralités	G.800–G.809
Objectifs de conception pour les réseaux numériques	G.810–G.819
Objectifs de qualité et de disponibilité	G.820–G.829
Fonctions et capacités du réseau	G.830–G.839
Caractéristiques des réseaux à hiérarchie numérique synchrone	G.840–G.899
SECTIONS NUMÉRIQUES ET SYSTÈMES DE LIGNES NUMÉRIQUES	G.900–G.999
Généralités	G.900–G.909
Paramètres pour les systèmes à câbles optiques	G.910–G.919
Sections numériques à débits hiérarchisés multiples de 2048 kbit/s	G.920–G.929
Systèmes numériques de transmission par ligne à débits non hiérarchisés	G.930–G.939
Systèmes de transmission numérique par ligne à supports MRF	G.940–G.949
Systèmes numériques de transmission par ligne	G.950–G.959
Section numérique et systèmes de transmission numérique pour l'accès usager du RNIS	G.960–G.969
Systèmes sous-marins à câbles optiques	G.970–G.979
Systèmes de transmission par ligne optique pour les réseaux locaux et les réseaux d'accès	G.980–G.999

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

RECOMMANDATION UIT-T G.723.1 – Annexe A

SCHEMA DE COMPRESSION DES SILENCES

Source

L'Annexe A à la Recommandation UIT-T G.723.1, élaborée par la Commission d'études 15 (1993-1996) de l'UIT-T, a été approuvée le 8 novembre 1996 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 (Helsinki, 1er-12 mars 1993).

Dans certains secteurs de la technologie 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.

© UIT 1997

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

		Page
A.1	Introduction	1
A.2	Description du détecteur VAD	2
	A.2.1 Calcul de l'indicateur d'activation de l'adaptation	3
	A.2.2 Filtrage inverse	3
	A.2.3 Calcul de l'énergie filtrée	3
	A.2.4 Calcul du niveau de bruit	4
	A.2.5 Calcul du seuil	4
	A.2.6 Décision VAD	4
	A.2.7 Addition du maintien VAD	4
	A.2.8 Initialisation VAD	4
A.3	Description générale de l'algorithme CNG	5
A.4	Description de l'algorithme CNG du codeur	6
	A.4.1 Calcul de la fonction d'autocorrélation de trame	6
	A.4.2 Calcul du type ($Ftyp_t$) de la trame considérée	7
	A.4.3 Quantification de l'énergie moyenne	8
	A.4.4 Calcul et codage des paramètres SID	9
	A.4.5 Calcul de l'énergie d'excitation CNG	10
	A.4.6 Interpolation des coefficients LSP et mise à jour	11
	A.4.7 Initialisation du bloc COD-CNG	12
A.5	Description du décodeur	12
	A.5.1 Description du bloc DEC-CNG	12
	A.5.2 Transparence de l'effacement des trames pour l'algorithme CNG	13
	A.5.3 Initialisation du bloc DEC-CNG	14
A.6	Compression du flux binaire	14
A.7	Glossaire	15
A.8	Programme source C en virgule fixe et calcul exact	16

Recommandation G.723.1 – Annexe A

SCHEMA DE COMPRESSION DES SILENCES

(Genève, 1996)

A.1 Introduction

La présente annexe décrit le système de compression des silences qui a été conçu pour le codeur vocal de la Recommandation G.723.1. Les techniques de compression des silences servent à réduire le débit transmis pendant les intervalles de silence. Les systèmes assurant une transmission discontinue sont basés sur l'algorithme de détection d'activité vocale (VAD, v*oice activity detection*) et l'algorithme du générateur de bruit de confort (CNG, c*omfort noise generator*) qui permet l'insertion de bruit artificiel pendant les périodes de silence. Cette caractéristique est nécessaire pour éviter de moduler du bruit lorsque la transmission est arrêtée: si le bruit de fond acoustique qui était présent durant les périodes d'activité disparaît brusquement, cette modulation très déplaisante peut même réduire l'intelligibilité du discours.

Le but de l'algorithme VAD est de détecter avec exactitude la présence ou l'absence de parole et de véhiculer cette information vers l'algorithme CNG. Les algorithmes VAD types basent leurs décisions sur plusieurs trames d'information successives de façon à être plus fiables et à éviter de prendre des décisions intermittentes. L'algorithme VAD est contraint de fonctionner sur les mêmes 30 ms de trame d'activité vocale qui seront ultérieurement codées par le codeur vocal ou remplies par du bruit de confort produit par le générateur. Le résultat de l'algorithme VAD est traité par l'algorithme CNG.

La plus grande difficulté de la détection de signal vocal est la présence de tout un éventail de bruits de fond. L'algorithme VAD doit pouvoir détecter un signal vocal même si le rapport signal sur bruit est très faible. Il est impossible de faire une distinction entre le signal vocal et le bruit à l'aide des techniques de détection à niveau unique lorsque le bruit camoufle des bribes de signal. La distinction entre les différents bruits ne peut être faite qu'en prenant en considération les caractéristiques spectrales du signal d'entrée. Pour ce faire, l'algorithme VAD contient un filtre inverse, dont les coefficients sont déterminés par l'algorithme CNG pendant les périodes ne contenant que du bruit. Tout détail complémentaire concernant l'algorithme VAD se trouve au A.2.

Le but de l'algorithme CNG est de créer un bruit qui correspond au bruit de fond effectif avec un coût de transmission aussi faible que possible. A l'extrémité de transmission, l'algorithme CNG utilise les informations d'activité vocale que l'algorithme VAD donne pour chaque trame, puis calcule les paramètres codés nécessaires pour synthétiser le bruit artificiel à l'extrémité de réception. Ces paramètres forment les trames de descripteur d'insertion de silence (SID, s*ilence insertion descriptor*), qui requièrent moins de bits que les trames d'activité vocale et qui sont transmises durant les périodes d'inactivité.

La principale caractéristique de l'algorithme CNG est que la transmission des trames SID n'est pas périodique: pour chaque trame inactive, l'algorithme prend la décision d'envoyer ou non une trame SID, en fonction du résultat d'une comparaison entre la trame d'inactivité en question et la trame SID précédente. La transmission des trames SID est ainsi limitée aux trames dont le spectre en puissance de bruit a changé.

Pendant la transmission des trames d'inactivité, le bruit de confort est synthétisé au niveau du décodeur en introduisant une excitation pseudo-blanche dans le filtre de synthèse à court terme. Les paramètres utilisés pour caractériser le bruit de confort sont les coefficients LPC (codage prédictif linéaire, *linear prediction coding*) du filtre de synthèse et l'énergie du signal d'excitation.

L'algorithme calcule, au niveau du codeur et pour chaque trame SID, un ensemble de paramètres LPC et quantifie sur 24 bits les coefficients LSP (raies spectrales doubles) correspondants à l'aide du quantificateur LSP. Il évalue également l'énergie d'excitation et la quantifie sur 6 bits. Cette méthode conduit à des trames SID codées sur 4 bits composés des deux bits de débit et d'information DTX.

Une caractéristique importante de l'algorithme CNG est sa méthode d'évaluation du spectre de bruit ambiant de chaque trame SID. L'algorithme prend en compte la stationnarité locale du signal d'entrée ou sa non-stationnarité.

L'excitation correspond finalement à l'excitation au débit supérieur du codec G.723.1. Puisque le spectre de l'excitation fixe donne des résultats moins probants, l'excitation à long terme est aussi utilisée afin d'obtenir un meilleur bruit blanc. L'algorithme choisit au hasard les codes des paramètres caractéristiques de l'excitation à long terme (retard et gains) et des paramètres du répertoire de séquences codées fixes (peigne de Dirac, positions et signes des impulsions). Pour chaque couple de sous-trames, il calcule le gain de l'excitation fixe pour obtenir une énergie globale à partir de l'énergie SID transmise.

Il est nécessaire d'effectuer le calcul de l'excitation au niveau du codeur et du décodeur afin d'assurer la synchronisation des deux systèmes.

Afin de simplifier la procédure au niveau du récepteur, le postfiltre harmonique est éteint pendant la génération de bruit de confort, celui-ci n'étant pas un signal vocal.

Les résultats des tests réalisés dans le cadre du schéma VAD/DTX/CNG et décrits dans la présente annexe seront publiés ultérieurement sous forme d'appendice à l'Annexe A de la Recommandation G.723.1.[1]

A.2 Description du détecteur VAD

Ce paragraphe décrit le détecteur d'activité vocale (VAD, *voice activity detector*) utilisé dans le codeur vocal G.723.1. La fonction de ce détecteur est d'indiquer si chaque trame de 30 ms produite par le codeur vocal contient ou ne contient pas de signaux vocaux. La décision VAD prise au niveau de la trame *t*, notée Vad_t, est l'entrée du bloc COD-CNG, qui calcule le type $Ftyp_t$, conformément à la description du A.3 et à l'illustration de la Figure A.1. La performance de l'algorithme VAD est caractérisée par la quantité de signal vocal audible mutilé et le pourcentage d'activité vocale que cet algorithme indique.

Le détecteur VAD est à la base un détecteur d'énergie. L'énergie du signal filtré inverse est comparé à un seuil. Si celui-ci est dépassé, il s'agit de signal vocal. Ce seuil est calculé à l'aide d'une procédure à deux étapes. D'abord, le niveau de bruit est mis à jour en fonction de sa valeur précédente et de l'énergie du signal filtré. Puis, le seuil est calculé à partir du niveau de bruit à l'aide d'une approximation logarithmique.

Le maintien permet de déclarer quelques-unes des premières trames de silence qui suivent une rafale vocale comme étant encore du signal vocal. Il sert à éliminer la mutilation des signaux vocaux de faible niveau. Il n'est rajouté qu'aux rafales vocales qui dépassent un certain laps de temps afin d'éliminer les pics de bruits élargis.

---

[1] Il convient de noter que les conditions de test n'étaient pas suffisamment sévères dans le cas des mobiles.

A.2.1 Calcul de l'indicateur d'activation de l'adaptation

Un indicateur d'activation de l'adaptation, noté Aen_t pour la trame considérée *t*, sert à s'assurer que le niveau de bruit du détecteur VAD n'est adapté qu'en l'absence de signaux vocaux. Il est basé sur le fait que le bruit de fond et les silences ne sont ni des signaux vocaux ni des ondes sinusoïdales:

– détection de signal vocal/ non vocal:

les délais tonals en boucle ouverte de la trame précédente et de la trame considérée sont utilisés pour déterminer si les signaux sont des signaux vocaux. Notons $L_{OL}^{j}, j = 0,1,2,3$ ces 4 valeurs. On calcule d'abord le délai minimal $L_{OL}^{\min} = \mathrm{Min}\left(L_{OL}^{j}, j = 0,1,2,3\right)$. On évalue ensuite la valeur du compteur $pc \in [1,2,3,4]$ indiquant le nombre de délais L_{OL}^{j} qui se trouvent au voisinage d'un multiple de $L_{OL}^{\min}$ (±3). Si *pc* est égal à 4, le signal est considéré comme étant un signal vocal;

– détection d'onde sinusoïdale: (figure déjà dans la Contribution COM 15-255)

le détecteur d'onde sinusoïdale suivant fait partie de la fonction d'analyse LPC du codeur G723.1:

soit $k_i^t[2]$ le deuxième coefficient de réflexion calculé à l'aide de la récurrence de Durbin pour chaque sous-trame $i = 0,...,3$ de la trame *t*.

Si $k_i^t[2] \geq 0{,}95$ pour au moins 14 des 15 dernières valeurs, alors une onde sinusoïdale est détectée (*SinD* = 1). Sinon, *SinD* = 0;

– calcul de l'indicateur d'activation de l'adaptation:

$$\begin{cases} Aen_t = Aen_{t-1} + 2 & \text{si } pc = 4 \text{ ou } SinD = 1 \\ Aen_t = Aen_{t-1} - 1 & \text{sinon} \end{cases}$$

Aen_t appartient à l'intervalle $[0,6]$.

A.2.2 Filtrage inverse

La trame du signal d'entrée: $\{s[n]\}_{n=60..239}$, est traité par filtrage inverse par un filtre à réponse impulsionnelle finie (FIR) $A_{no}(z)$ dont les coefficients sont $\{a_{no}[j]\}_{j=1...10}$. Ce filtre est calculé par le bloc CNG et donne une estimation du filtre LPC associé au bruit de fond du moment.

$$e_t'[n] = s[n] + \sum_{j=1}^{10} a_{no}[j] \cdot s[n-j] \quad n = 60 \rightarrow 239 \tag{A-1}$$

où $e_t'[n]$ est le signal filtré inverse.

A.2.3 Calcul de l'énergie filtrée

L'énergie, Enr_t, est calculée à partir du signal filtré inverse de la trame en question à l'aide de l'équation:

$$Enr_t = \frac{1}{80} \sum_{n=60}^{239} e_t'^2[n] \tag{A-2}$$

A.2.4 Calcul du niveau de bruit

Le niveau de bruit associé à la trame *t*, $Nlev_t$, est mis à jour en fonction de sa valeur précédente et de la valeur de l'énergie précédente Enr_{t-1} ainsi que de l'indicateur d'activation de l'adaptation Aen_t. Cette mise à jour se caractérise par un front avant en pente douce et un amortissement rapide. L'intervalle dynamique du niveau de bruit pour la trame *t* se limite à l'intervalle $\left[Nlev_{\min}, Nlev_{\max}\right]$.

1) Si $Nlev_{t-1} > Enr_{t-1}$, le niveau de bruit est d'abord tronqué:

$$Nlev_t = \begin{cases} 0{,}25 \cdot Nlev_{t-1} + 0{,}75 \cdot Enr_{t-1} & \text{si } Nlev_{t-1} > Enr_{t-1} \\ Nlev_{t-1} & \text{sinon} \end{cases} \quad \text{(A-3)}$$

2) Ensuite, si l'adaptation est activée, le niveau $Nlev_t$ est augmenté; dans le cas contraire, le niveau est diminué d'une faible quantité:

$$Nlev_t = \begin{cases} 1{,}03125 \times Nlev_t & \text{si } Aen_t = 0 \\ 0{,}9995 \times Nlev_t & \text{sinon} \end{cases} \quad \text{(A-4)}$$

$$\text{avec} \begin{cases} Nlev_{\min} = 128 \\ Nlev_{\max} = 131071 \end{cases}$$

A.2.5 Calcul du seuil

La relation entre le niveau de bruit de la trame *t*, $Nlev_t$, et le seuil, *Thr*, est définie par l'équation suivante qui est une approximation logarithmique:

$$Thr = \begin{cases} 5{,}012 & \text{si } Nlev = 128, \\ 10^{0{,}7 - 0{,}05 \log_2 \frac{Nlev}{128}} & \text{si } 128 < Nlev < 16384 \\ 2{,}239 & \text{si } Nlev \geq 16384 \end{cases} \quad \text{(A-5)}$$

A.2.6 Décision VAD

La décision VAD dépend de la comparaison entre le seuil, *Thr*, et l'énergie Enr_t considérée.

$$Vad_t = \begin{cases} 1 & Enr_t \geq Thr \\ 0 & Enr_t < Thr \end{cases} \quad \text{(A-6)}$$

A.2.7 Addition du maintien VAD

Un maintien de 6 trames n'est ajouté qu'en cas de rafales vocales (Vad_t = 1) de 2 trames ou plus .

A.2.8 Initialisation VAD

Toutes les variables statiques de l'algorithme VAD sont initialisées à 0 à l'exception des variables suivantes:

$$\begin{aligned} Nlev_{-1} &= 1024 \\ Enr_{-1} &= 1024 \\ L_{OL}^{j} &= 1 \qquad j = 0{,}1 \\ L_{Ol}^{j} &= 60 \qquad j = 2{,}3 \end{aligned} \quad \text{(A-7)}$$

A.3 Description générale de l'algorithme CNG

L'algorithme est divisé en 2 blocs, appelés COD-CNG et DEC-CNG, qui sont respectivement situés au niveau du codeur et du décodeur. Au niveau du codeur (voir la Figure A.1), le bloc COD-CNG utilise la fonction d'autocorrélation du signal vocal qui est calculée pour chaque sous-trame de 60 échantillons, les anciens échantillons d'excitation et les coefficients LSP de la trame précédente.

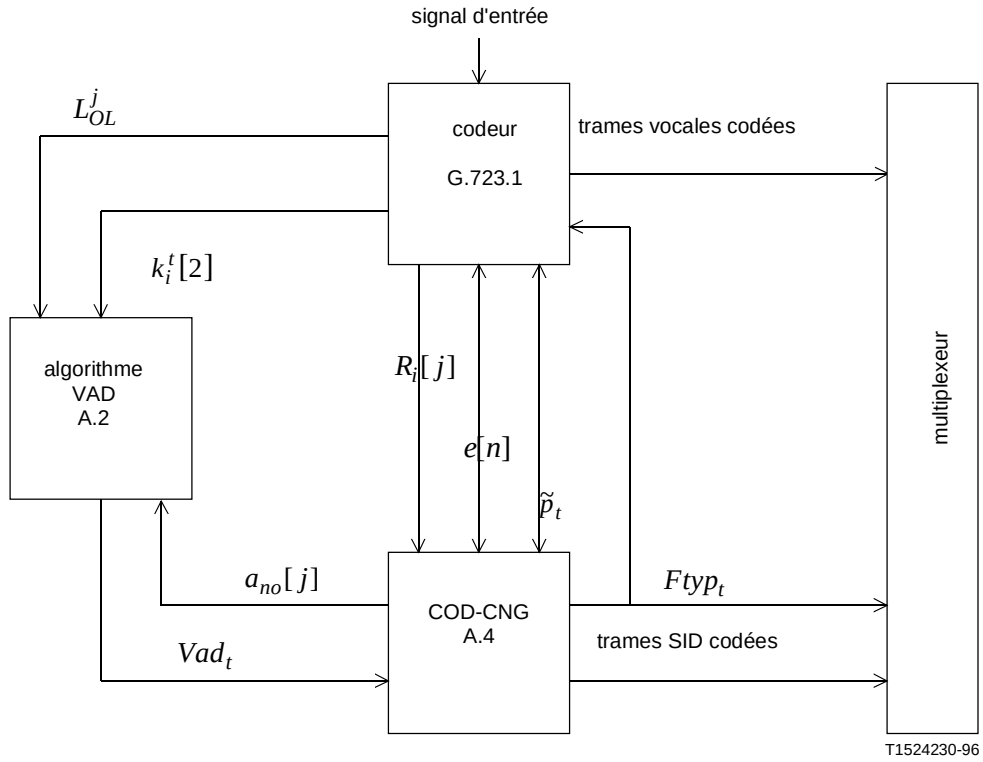

FIGURE A.1/G.723.1

Schéma de principe du codeur avec algorithmes VAD/CNG

Pour les trames d'inactivité vocale, le bloc COD-CNG calcule les échantillons d'excitation CNG afin de synchroniser le décodeur local du codeur sur le décodeur distant.

En raison du codage prédictif des coefficients LSP du dispositif de la Recommandation G.723.1, une entrée/sortie similaire avec mise à jour est réalisée pour les paramètres LSP pendant la transmission des trames d'inactivité vocale.

L'algorithme COD-CNG code les trames SID et rend la décision finale $Ftyp_t$ (type de la trame t) sous forme d'une des trois valeurs 0, 1 ou 2; ces valeurs correspondent respectivement à une trame non transmise, une trame d'activité vocale ou une trame SID.

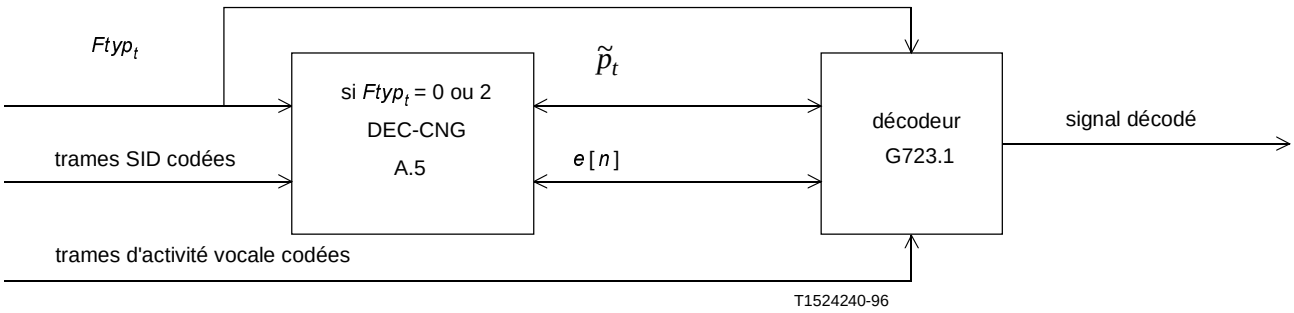

FIGURE A.2/G.723.1

Schéma de principe d'un décodeur avec algorithmes VAD/DTX

Au niveau du récepteur (voir la Figure A.2), le bloc DEC-CNG ne traite que les trames d'inactivité vocale, pour lesquelles l'information d'entrée $Ftyp_t$ est égale à 0 ou 2 (trame non transmise ou trame SID). Le bloc DEC-CNG décode les trames SID et calcule – à la fois pour les trames SID et les trames non transmises – les coefficients LSP du moment et l'excitation à l'aide de la même méthode que celle du bloc COD-CNG.

Puis le décodeur G.723.1 synthétise le bruit de confort à l'aide de l'excitation CNG et des coefficients LSP.

A.4 Description de l'algorithme CNG du codeur

Pour chaque trame de 240 échantillons (d'activité ou d'inactivité vocale), le bloc COD-CNG traite les données venant du détecteur VAD et du codeur, produit l'information $Ftyp_t$ et la trame SID codée selon la procédure décrite par la Figure A.3 et détaillée dans A.4.1 à A.4.7.

A.4.1 Calcul de la fonction d'autocorrélation de trame

Fichier: COD_CNG.C	Procédure: Update_Acf()	Mise à jour de la fonction d'autocorrélation

Pour chaque trame *t* (d'activité ou d'inactivité vocale), les coefficients d'autocorrélation (calculés dans le codeur comme cela est décrit au 2.4 de la Recommandation G.723.1) $R_i[j]$, j = 0 à 10 des 4 sous-trames numérotées de i = 0 à 3 sont ajoutés. La fonction d'autocorrélation cumulée de la trame *t* examinée, est donnée par:

$$R^t[j] = \sum_{i=10}^{3} R_i[j], \text{ pour } j = 0 \text{ à } 10 \qquad \text{(A-8)}$$

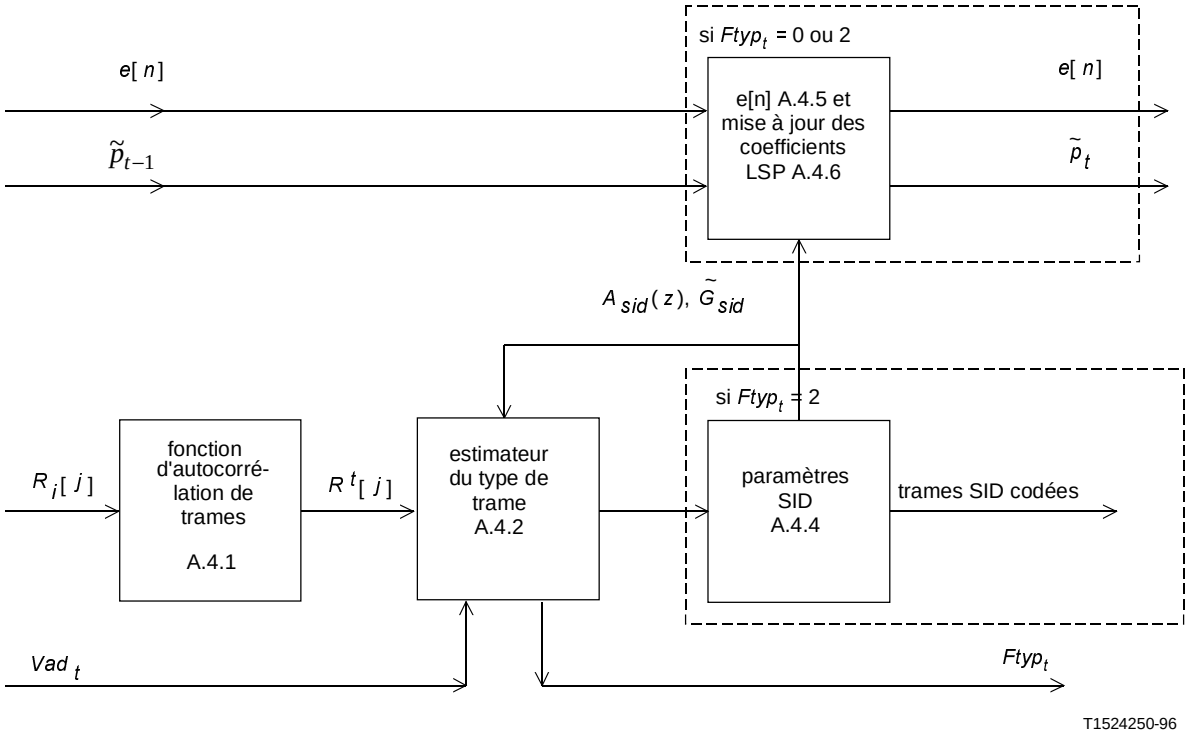

FIGURE A.3/G.723.1

Schéma de principe de l'algorithme CNG du codeur

A.4.2 Calcul du type ($Ftyp_t$) de la trame considérée

Fichier: COD_CNG.C	Procédure: Cod_Cng()	Partie principale: COD_CNG
Fichier: COD_CNG.C	Procédure: LpcDiff()	Comparaison avec la distance d'Itakura
Fichier: LPC.C	Procédure: Durbin()	Récurrence Levinson-Durbin

Si la trame t considérée est une trame d'activité vocale (Vad_t = 1), alors $Ftyp_t$ = 1 et aucun autre traitement n'est effectué.

Sinon, la décision *trame non transmise* ou *trame SID* est prise en fonction du résultat de la procédure suivante:

le filtre LPC $A_t(z)$ de la trame t considérée est calculé à l'aide de la récurrence de Durbin (voir 2.4 de la Recommandation G.723.1) avec en entrée les valeurs $R^t[\,j\,]$. Les coefficients de $A_t(z)$ s'écrivent: $a_t[j]$, j=1 à 10. La récurrence de Durbin donne également l'énergie résiduelle E_t, qui sera utilisée comme estimation de l'énergie de la trame d'excitation.

Puis le type $Ftyp_t$ de la trame considérée est déterminé de la manière suivante:

- si cette trame est la première de la période d'inactivité, on décide que c'est une trame SID; la variable $\overline{E}$, qui est la somme des énergies de la trame, est prise égale à E_t, et le nombre, k_E, de sous-trames considéré dans le calcul est initialisé à 1:

$$\left(Vad_{t-1} = 1\right) \Rightarrow \begin{cases} Ftyp_t = 2 \\ \overline{E} = E \\ k_E = 1 \end{cases} \qquad \text{(A-9)}$$

- si le filtre considéré diffère de manière significative du filtre SID précédent, ou si l'énergie d'excitation en question diffère sensiblement de l'énergie SID précédente, alors on décide que c'est une trame SID ($Ftyp_t$ = 2);
- sinon, si cette trame n'est pas la première de la période d'inactivité et si le filtre LPC considéré et l'énergie d'excitation sont similaires à ceux de la trame SID, alors elle est non transmise ($Ftyp_t$ = 0).

Les filtres LPC et les énergies sont comparés à l'aide des méthodes suivantes:

comparaison des filtres LPC

Les filtres LPC et SID étudiés sont considérés comme sensiblement différents si la distance Itakura entre les deux filtres dépasse le seuil fixé, celui-ci s'exprimant de la manière suivante:

$$\sum_{j=0}^{10} R_a[i] \times R^t[i] \geq E_t \times thr1 \qquad \text{(A-10)}$$

où $R_a[j]$, j = 0 à 10 est une fonction déterminée à partir de l'autocorrélation des coefficients du filtre SID et est donnée par:

$$\begin{cases} R_a[j] = 2\sum_{k=0}^{10-j} a_{sid}[k] \times a_{sid}[k+j] & \text{si } j \neq 0 \\ R_a(0) = \sum_{k=0}^{10} a_{sid}[k]^2 \end{cases} \qquad \text{(A-11)}$$

avec $a_{sid}[0] = 1$

La valeur de *thr*1 est fixée à 1,2136.

comparaison des énergies

En incrémentant k_E jusqu'à sa valeur maximale de 3, on calcule d'abord la somme des énergies de trame $\overline{E} = \sum_{i=t-k_E+1}^{t} E_t$.

Puis $\overline{E}$ est quantifiée sur 6 bits à l'aide du quantificateur pseudo-logarithmique décrit au A.4.3. L'indice de gain codé, $GInd_t$, est comparé à l'indice de gain codé SID précédent, $GInd_{sid}$. Si la différence dépasse le seuil *thr2*=3, les deux énergies seront considérées comme sensiblement différentes.

A.4.3 Quantification de l'énergie moyenne

Fichier: UTIL_CNG.C	Procédure: Qua_SidGain()	Quantifie le gain SID
Fichier: UTIL_CNG.C	Procédure: Dec_SidGain()	Décode le gain SID

La procédure de quantification fonctionne sur la somme des énergies $\overline{E}$, et le décodage donne un gain qui correspond à la valeur décodée de la racine de l'énergie moyenne.

Un facteur d'échelle α_w = 2,70375 est introduit pour prendre en compte l'effet dû à l'élargissement de fenêtre et de bande passante qui est présent dans les fonctions d'autocorrélation des sous-trames $R_i[j]$.

La valeur utilisée à l'entrée du quantificateur de gain est:

$$G = \alpha_w \times \sqrt{\frac{1}{k_E \times 240} \overline{E}}$$, limitée à l'intervalle [0, 352]

Le quantificateur pseudo-logarithmique divise l'intervalle [0, 352] en 3 segments d'indice *isg* = 0 à 2 et de longueur *N*[*isg*] = 16, 16, 32, les résolutions associées étant 2, 4 et 8.

Soit $G_{isg}[j]$, $j = 0$ à $N[isg]$-1 les valeurs décodées du segment *isg*. Ces valeurs sont données par:

$$G_{isg}[j] = G_{isg}[0] + j \times 2^{(isg+1)} \quad \text{(A-12)}$$

La procédure utilise G^2 pour calculer l'indice *isg* du segment qui contient *G* et l'indice i_s du gain $G_{isg}(i_s)$ le plus proche de *G*.

L'indice de quantification considéré est donné par:

$$GInd_t = 16 \times isg + i_s \quad \text{(A-13)}$$

Le décodage est réalisé par la formule suivante:

$$Q^{-1}(GInd_t) = G_{isg}[0] + (GInd_t - \lfloor isg / 16 \rfloor) \times 2^{isg+1} \quad \text{(A-14)}$$

où $\lfloor x \rfloor$ représente le plus grand entier $\leq$ x.

A.4.4 Calcul et codage des paramètres SID

Fichier: COD_CNG.C	Procédure: Cod_Cng()	Partie principale: COD_CNG
Fichier: COD_CNG.C	Procédure: ComputePastAvFilter()	Calcule les moyennes de l'ancien filtre
Fichier: COD_CNG.C	Procédure: LpcDiff()	Comparaison avec la distance Itakura
Fichier: COD_CNG.C	Procédure: CalcRc()	Calcule la fonction RC du codage LPC
Fichier: LPC.C	Procédure: Durbin()	Récurrence Levinson-Durbin
Fichier: LSP.C	Procédure: AtoLsp()	Convertit les coefficients LPC en coefficients LSP
Fichier: LSP.C	Procédure: Lsp_Qnt()	Quantification LSP
Fichier: LSP.C	Procédure: Lsp_Inq()	Quantification LSP inverse

Lorsque la trame considérée est une trame SID, les paramètres SID sont calculés et quantifiés. Il est à noter que ces paramètres serviront à décider si les prochaines trames d'inactivité vocale seront des trames SID jusqu'à l'émission de la prochaine trame SID.

Calcul du filtre LPC SID [$A_{sid}(z)$] et mise à jour du filtre LPC VAD [$A_{no}(z)$]

D'abord les valeurs moyennes de l'ancien filtre LPC $\overline{A}_p(z)$, réalisé à partir des 3 trames précédant la trame considérée, sont estimées à l'aide de la récurrence de Durbin avec en entrée la fonction d'autocorrélation suivante:

$$\overline{R}_p[j] = \sum_{k=t-3}^{t-1} R^k[j], \text{ pour } j = 0 \text{ à } 10 \quad \text{(A-15)}$$

les fonctions d'autocorrélation $R^k[j]$ étant les fonctions cumulées calculées à l'aide de l'équation (A-8).

Les coefficients moyens de l'ancien filtre LPC s'écrivent $\overline{a}_p[j]$, j = 1 à 10.

Le filtre LPC VAD associé au bruit et utilisé dans A.2 est ensuite mis à jour avec les coefficients $\bar{a}_p[j]$ mais uniquement si l'indicateur d'activation de l'adaptation Aen_t le permet:

$$\text{si } Aen_t = 0 \text{ alors } a_{no}[j] = \bar{a}_p[j], j = 1,2,...10 \tag{A-16}$$

puis

$$A_{sid}(z) = \begin{cases} A_t(z) \text{ si la distance entre } A_t(z) \text{ et } \bar{A}_p(z) \text{ est} \geq thr1 \\ \bar{A}_p(z) \text{ sinon} \end{cases} \quad \text{voir l'éq. (A-10)}$$

La distance entre le filtre LPC considéré $A_t(z)$ et les valeurs moyennes de l'ancien filtre LPC $\bar{A}_p(z)$ est calculée de la même manière qu'au A.4.2.

Les coefficients $a_{sid}[j], j = 1 \rightarrow 10$ du nouveau filtre LPC SID sont convertis en coefficients LSP et ceux-ci sont quantifiés à l'aide de séquences codées sur 24 bits (voir 2.5 de la Recommandation G.723.1). La valeur décodée sera notée $\tilde{p}_{sid}$.

Gain SID

La valeur quantifiée du gain SID est donnée par:

$$GInd_{sid} = GInd_t \tag{A-17}$$

et la valeur décodée est notée $\tilde{\tilde{G}}_{sid}$.

A.4.5 Calcul de l'énergie d'excitation CNG

Fichier: UTIL_CNG.C	Procédure: Calc_Exc_Rand()	Calcul de l'excitation
Fichier: UTIL_CNG.C	Procédure: random_number()	Génération de nombre aléatoire
Fichier: UTIL_CNG.C	Procédure: distG()	Utilisé pour choisir le gain de l'excitation
Fichier: UTIL_LBC.C	Procédure: Sqrt_lbc()	Racine carrée
Fichier: UTIL_LBC.C	Procédure: Rand_lbc()	Séquence pseudo-aléatoire

La mise à jour du signal d'excitation est effectuée pour les trames SID et celles qui sont non transmises.

Soit $\tilde{G}_t$ le gain d'excitation cible exprimé sous forme de la racine carrée de l'énergie moyenne qui doit être obtenue pour la trame *t* d'excitation synthétisée considérée. $\tilde{G}_t$ est calculé à l'aide de méthode de lissage suivante:

$$\tilde{\tilde{G}}_t = \begin{cases} \tilde{\tilde{G}}_{sid} & \text{si } Vad_{t-1} = 1 \\ \frac{7}{8}\tilde{\tilde{G}}_{t-1} + \frac{1}{8}\tilde{\tilde{G}}_{sid} & \text{sinon} \end{cases} \tag{A-18}$$

Les 240 échantillons de la trame sont divisés en 2 blocs de 120, chaque bloc comprenant deux sous-trames de 60 échantillons.

Pour chaque bloc, les échantillons d'excitation CNG sont synthétisés à l'aide de l'algorithme suivant:

Les paramètres LPT (prédicteur à long terme) des 2 sous-trames sont d'abord choisis:

- le délai tonal de la 1ère sous-trame est choisi aléatoirement dans l'intervalle [123, 143];
- les indices des vecteurs gain des 2 sous-trames sont choisis aléatoirement dans l'intervalle [0,49], cet intervalle correspondant aux 50 premiers vecteurs des 170 entrées du répertoire de séquences codées de gain;
- le décalage du délai tonal associé à la seconde sous-trame est pris égal à 0 pour le premier bloc et à 3 pour le second bloc.

Les vecteurs fixes du répertoire de séquences codées des deux sous-trames sont ensuite construits en choisissant de façon aléatoire le peigne de Dirac, les signes des impulsions et les positions, correspondant au diagramme associé à l'excitation fixe au débit supérieur.

Puis un seul gain d'excitation fixe est calculé pour les deux sous-trames du bloc.

Le vecteur d'excitation adaptative du bloc considéré est noté *u[n]*, *n* = 0 à 119 et le vecteur d'excitation fixe *v[n]*, *n* = 0 à 119.

Le gain de l'excitation fixe est obtenu en calculant la valeur *Gf* qui conduit à l'énergie moyenne du bloc la plus proche de l'énergie cible $\tilde{G}_t^2$:

$$\text{Choisir le gain } Gf \text{ pour lequel } \left| \frac{1}{20} \sum_{n-0}^{119} \left(u[n] + Gf \times v[n]\right)^2 - \tilde{G}_t^2 \right| \text{ est minimal} \qquad \text{(A-19)}$$

A noter que *Gf* peut être négatif.

Posons $C(X) = aX^2 + 2bX + c$ avec:

$$a = \left(\sum_{n=0}^{119} v[n]^2 \right),\ b = \left(\sum_{n=0}^{119} u[n]v[n] \right),\ c = \left(\sum_{n=0}^{119} u[n]^2 - 120\tilde{G}_t^2 \right)$$

Étudions l'équation *C(X)* = 0:

si le discriminant est ≤ 0 alors on choisit $Gf = -\frac{b}{a}$, sinon on calcule les deux racines et on choisit celle dont la valeur absolue est la plus petite.

Puis on limite *Gf*: G*f* ≤ 5000

Finalement, on construit l'excitation CNG du bloc à l'aide de:

$$e[n] = u[n] + Gf \times v[n], n = 0 \text{ à } 119 \qquad \text{(A-20)}$$

A.4.6 Interpolation des coefficients LSP et mise à jour

Fichier: COD_CNG.C	Procédure: Cod_Cng()	Partie principale: COD_CNG
Fichier: LSP.C	Procédure: Lsp_Int()	Interpolateur LSP

Les ensembles de coefficients LPC interpolés des trames SID et des trames non transmises sont calculés à l'aide de la valeur $\tilde{p}_{sid}$ et du vecteur LSP précédent $\tilde{p}_{t-1}$ fourni au codeur COD-CNG.

La mise à jour LSP est également effectuée: $\tilde{p}_t = \tilde{p}_{sid}$.

A.4.7 Initialisation du bloc COD-CNG

L'initialisation suivante doit être faite sur les fonctions d'autocorrélation de trames, le gain de l'excitation cible, l'information VAD et la valeur de départ (*rseed*) utilisée par le générateur aléatoire pour calculer l'excitation CNG:

$$\begin{cases} R^k[j] = 0 \text{ pour } j = 0,\ldots,10 \text{ et } k = -1,-2,-3 \\ \tilde{G}_{-1} = 0 \\ Vad_{-1} = 1 \\ rseed = 12345 \end{cases}$$

Il n'est pas nécessaire d'initialiser les autres variables statiques du bloc COD-CNG.

A.5 Description du décodeur

A l'extrémité réceptrice, le bloc DEC-CNG traite les trames SID et les trames non transmises afin de synthétiser le bruit de confort.

Les méthodes développées pour traiter les effacements de trames sont décrites ci-dessous.

A.5.1 Description du bloc DEC-CNG

Fichier: DEC_CNG.C	Procédure: Dec_Cng()	Partie principale: DEC_CNG
Fichier: UTIL_CNG.C	Procédure: Calc_Exc_Rand()	Calcul de l'excitation
Fichier: UTIL_CNG.C	Procédure: random_number()	Génération de nombre aléatoire
Fichier: UTIL_CNG.C	Procédure: distG()	Utilisé pour choisir le gain de l'excitation
Fichier: UTIL_LBC.C	Procédure: Sqrt_lbc()	Racine carrée
Fichier: UTIL_LBC.C	Procédure: Rand_lbc()	Séquence pseudo-aléatoire
Fichier: LSP.C	Procédure: Lsp-Inq()	Quantification LSP inverse
Fichier: LSP.C	Procédure: Lsp_Int()	Interpolateur LSP
Fichier: UTIL_CNG.C	Procédure: Qua_SidGain()	Quantification du gain SID
Fichier: UTIL_CNG.C	Procédure: Dec_SidGain()	Décodage du gain SID

La Figure A.4 donne une description générale sur la génération du bruit de confort au niveau du décodeur.

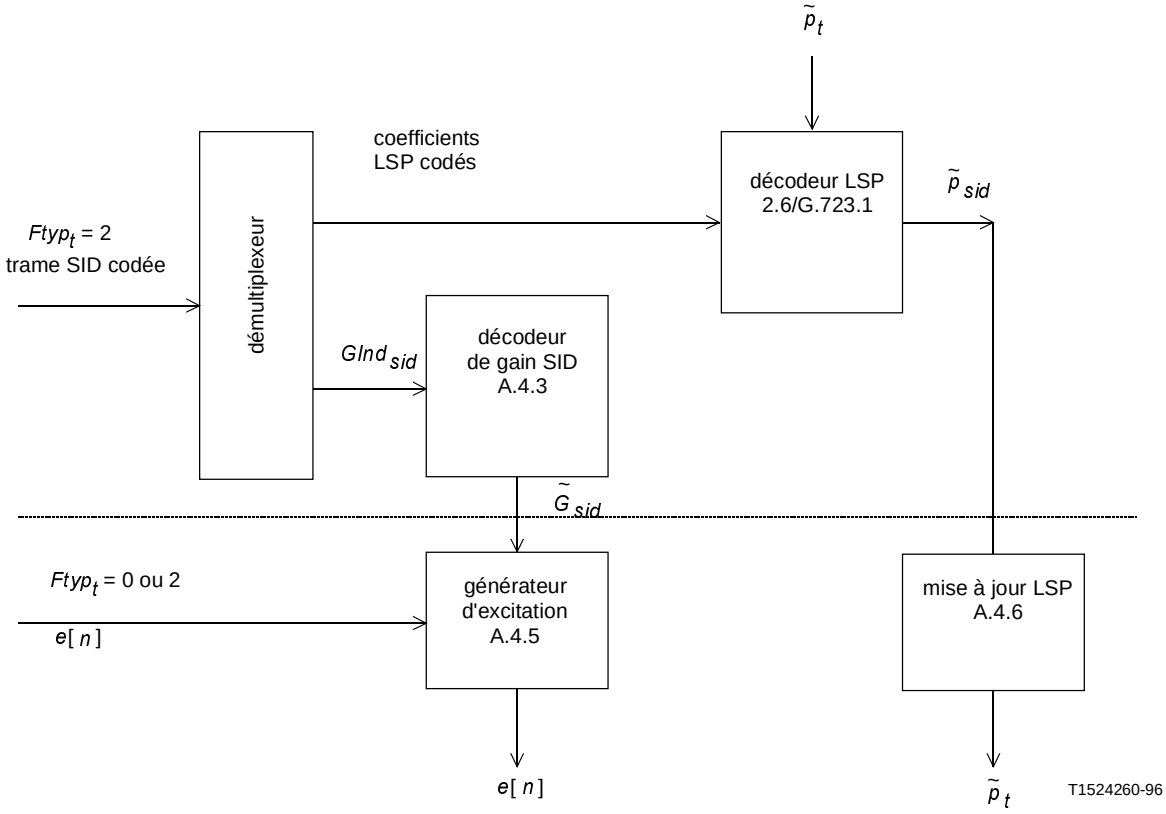

FIGURE A.4/G.723.1

Schéma de principe de l'algorithme CNG du décodeur

Quand le décodeur reçoit une trame SID, le bloc DEC-CNG décode les paramètres SID.

En présence de trames SID et de trames non transmises, le module DEC-CNG utilise les paramètres SID décodés pour calculer les coefficients LSP et l'excitation du bruit de confort, qui sera synthétisée par le module de synthèse du décodeur.

L'information type de trame CNG ($Ftyp_t$) (pour la trame t) fournie au niveau du récepteur est la même que celle calculée par le bloc COD-CNG au niveau du codeur.

- Lorsque $Ftyp_t = 2$, les paramètres de la trame SID sont décodés: $\tilde{p}_{sid}$ pour les coefficients LSP et $\tilde{G}_{sid}$ pour le gain décodé.
- Lorsque $Ftyp_t = 0$, $Ftyp_{t-1}$ est testé pour vérifier que l'information SID n'a pas été effacée (voir A.5.2). Si $Ftyp_{t-1} = 1$, l'énergie *Enr*, qui a été calculée par le décodeur G723.1 durant le traitement de la dernière trame valide, est quantifiée et décodée à l'aide de la même méthode que celle utilisée au 3.1 de la Recommandation G.723.1 pour l'énergie moyenne, si ce n'est qu'il n'y a pas de facteur d'échelle α_w. La valeur décodée est le gain reconstitué $\tilde{G}_{sid}$.

Puis, dans les deux cas, l'excitation CNG est calculée conformément à la procédure utilisée pour le bloc COD-CNG et décrite dans A.4.5. Le nouveau vecteur LSP, $\tilde{p}_{sid}$, sert à calculer les coefficients LPC interpolés; la mise à jour LSP est effectuée: $\tilde{p}_t = \tilde{p}_{sid}$.

A.5.2 Transparence de l'effacement des trames pour l'algorithme CNG

Fichier: DECOD.C	Procédure: Decod()	Décodage de trames

Lorsqu'un effacement de trame est détecté par le décodeur, son type dépend du type de la trame précédente:

– si la trame précédente était une trame d'activité vocale, alors la trame effacée en question est considérée comme étant une trame d'activité vocale;

– sinon si la trame précédente était une trame SID ou une trame non transmise, la trame effacée en question est considérée comme étant une trame non transmise:

$$\begin{cases} Ftyp_{t-1} = 1 & \Rightarrow \quad Ftyp_t = 1 \\ Ftyp_{t-1} = 0 \text{ ou } 2 & \Rightarrow \quad Ftyp_t = 0 \end{cases} \tag{A-21}$$

Si une trame non transmise a été effacée, aucune erreur n'est introduite.

Si une trame SID est effacée, il existe deux possibilités:

• s'il ne s'agit pas de la première trame SID de la période d'inactivité considérée, alors les paramètres SID précédents sont conservés;

• s'il s'agit de la première trame d'une période d'inactivité, une protection spéciale a été prise:

comme mentionné dans A.5.1, ce cas est détecté quand $Ftyp_{t-1} = 1$ et $Ftyp_t = 0$.

Cette combinaison d'événements n'implique pas que la trame précédente était une trame d'activité vocale correcte: plusieurs trames jusqu'à la précédente peuvent avoir été effacées. Ce qui est certain c'est que la dernière trame correcte était une trame d'activité vocale, que la trame présente n'était pas effacée et que la trame SID supposée fournir l'information de la trame non transmise considérée est perdue.

Pour rétablir l'information SID, le bloc DEC-CNG utilise les paramètres fournis par la partie principale du décodeur G.723.1:

• les coefficients LSP de la dernière trame d'activité vocale sont utilisés pour déterminer $\tilde{p}_{sid}$;

• l'énergie *Enr* calculée par le décodeur durant la procédure d'interpolation résiduelle (voir 3.10.2 de la Recommandation G.723.1) appliquée aux 120 derniers échantillons d'excitation de la dernière trame d'activité vocale est utilisée pour rétablir $\tilde{G}_{sid}$ conformément à la méthode décrite au A.5.1.

Finalement, pour éviter de désynchroniser le générateur aléatoire utilisé pour calculer l'excitation, la séquence pseudo-aléatoire est réinitialisée pour chaque trame d'activité vocale au niveau du codeur et la partie codage: *rseed* = 12345.

A.5.3 Initialisation du bloc DEC-CNG

Seules les variables suivantes doivent être initialisées:

$$\begin{cases} \tilde{G}_{sid} = 0 \\ \tilde{P}_{sid} = \text{vecteur } LSP \quad \text{constant} \quad P_{DC} \\ Vad_{-1} = 1 \\ rseed = 12345 \end{cases}$$

A.6 Compression du flux binaire

Le Tableau A.1 montre le flux de données des trames SID en fonction des notations utilisées dans l'article 4 de la Recommandation G.723.1.

TABLEAU A.1/G.723.1

Compression binaire des trames SID

Octets transmis	PARx_By, ...
1	LPC_B5 ... LPC_B0, VADFLAG_B0, RATEFLAG_B0
2	LPC_B13 ... LPC_B6
3	LPC_B21 ... LPC_B14
4	GAIN_B5 ... GAIN_B0, LPC_B23, LPC_B22

A.7 Glossaire

$a_{no}[j]$	coefficients du filtre LPC associé au bruit
L_{OL}^{j}	délais tonals en boucle ouverte de la trame précédente et de la trame considérée
pc	compteur de délais tonals pour la détermination des signaux vocaux
Aen_t	indicateur d'activation de l'adaptation
$e_t^{'}[n]$	signal de bruit inverse filtré à l'entrée pour la trame t
Enr_t	énergie du signal de bruit inverse filtré à l'entrée pour la trame t
$Nlev_t$	niveau de bruit au niveau de la trame t
$Nlev_{min}$	borne inférieure du niveau de bruit $Nlev_t$
$Nlev_{max}$	borne supérieure du niveau de bruit $Nlev_t$
Thr	seuil adapté de la décision VAD
$k_i^t[2]$	deuxième coefficient de réflexion de la sous-trame i de la trame t
SinD	indicateur de détection d'onde sinusoïdale (1: sinusoïde détectée, 0: sinon)
e[n]	vecteur d'excitation combiné décodé
$R_t[j]$	fonction d'autocorrélation des sous-trames i, j = 0,1,...,10
thr_2	seuil correspondant à la distance associée aux énergies
$R_a[j]$	autocorrélation modifiée des coefficients LPC
$Gind_{sid}$	indice de gain SID
$\tilde{G}_{sid}$	gain SID décodé
G	gain d'excitation utilisé à l'entrée du quantificateur SID
$Gind_t$	indice de gain de la trame t
isg	indice du segment du quantificateur de gain SID
N[*isg*]	longueur du segment du quantificateur de gain SID
$G_{isg}[j]$	valeurs de gain décodées du segment isg, j=0,1, ..., N[isg]-1
i_s	indice de gain relatif au segment
α_w	facteur d'échelle en énergie
a_{sid}	coefficient LPC du vecteur SID

$\overline{a}_p$	moyenne des coefficients de l'ancien filtre LPC
$\overline{R}_p[j]$	somme des anciennes fonctions d'autocorrélation
$\tilde{\tilde{p}}_{sid}$	vecteur LPC SID décodé
$u[n]$	vecteur d'excitation du répertoire de séquences codées adaptatives
$v[n]$	vecteur d'excitation du répertoire de séquences codées fixes
$\tilde{\tilde{G}}_t$	gain d'excitation cible pour la synthèse de l'excitation
a, b, c	coefficients de l'équation de minimisation de l'énergie
C(X)	équation de minimisation de l'énergie
Gf	gain du répertoire de séquences codées fixes pour la synthèse de l'excitation CNG
rseed	valeur de départ du générateur aléatoire

A.8 Programme source C en virgule fixe et calcul exact

Tous les détails de l'algorithme de compression des silences sont inclus dans le programme en code C ANSI en virgule fixe et calcul exact. En cas de différence entre les descriptions figurant ci-dessus et le programme en C, celui-ci est présumé correct. Ce programme fait partie du logiciel distribué par l'UIT-T avec la Recommandation G.723.1.

SÉRIES DES RECOMMANDATIONS UIT-T

Série A	Organisation du travail de l'UIT-T
Série B	Moyens d'expression
Série C	Statistiques générales des télécommunications
Série D	Principes généraux de tarification
Série E	Réseau téléphonique et RNIS
Série F	Services de télécommunication non téléphoniques
Série G	**Systèmes et supports de transmission**
Série H	Transmission des signaux autres que téléphoniques
Série I	Réseau numérique à intégration de services
Série J	Transmission des signaux radiophoniques et télévisuels
Série K	Protection contre les perturbations
Série L	Construction, installation et protection des câbles et autres éléments des installations extérieures
Série M	Maintenance: systèmes de transmission, de télégraphie, de télécopie, circuits téléphoniques, et circuits loués internationaux
Série N	Maintenance: circuits internationaux de transmission radiophoniques et télévisuels
Série O	Spécifications des appareils de mesure
Série P	Qualité de transmission téléphonique
Série Q	Commutation et signalisation
Série R	Transmission télégraphique
Série S	Equipements terminaux de télégraphie
Série T	Equipements terminaux et protocoles 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 Z	Langages de programmation