



UNION INTERNATIONALE DES TÉLÉCOMMUNICATIONS

UIT-T

SECTEUR DE LA NORMALISATION
DES TÉLÉCOMMUNICATIONS
DE L'UIT

Appendice II
Rec. G.728

(11/95)

**ASPECTS GÉNÉRAUX DES SYSTÈMES
DE TRANSMISSION NUMÉRIQUES**

**CARACTÉRISTIQUES DE TRANSMISSION
VOCALE**

Appendice II à la
Recommandation UIT-T G.728

(Antérieurement «Recommandation du CCITT»)

AVANT-PROPOS

L'UIT-T (Secteur de la normalisation des télécommunications) est un organe permanent de l'Union internationale des télécommunications (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, 1^{er}-12 mars 1993).

L'Appendice II à la Recommandation UIT-T G.728, que l'on doit à la Commission d'études 15 (1993-1996) de l'UIT-T, a été approuvé le 13 novembre 1995 selon la procédure définie dans la Résolution n° 1 de la CMNT.

NOTE

Dans le présent Appendice, l'expression «Administration» est utilisée pour désigner de façon abrégée aussi bien une administration de télécommunications qu'une exploitation reconnue de télécommunications.

© UIT 1996

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

	<i>Page</i>
II.1 Introduction	1
II.2 Caractéristiques de la transmission vocale.....	1
II.3 Caractéristiques de transmission des signaux non vocaux.....	2
II.4 Signaux vocaux artificiels.....	2

CARACTÉRISTIQUES DE TRANSMISSION VOCALE

(Genève, 1995)

(Cet appendice ne fait pas partie intégrante de la présente Recommandation)

II.1 Introduction

Le présent appendice a pour objectif d'indiquer dans les grandes lignes quelles sont les caractéristiques de transmission vocale de l'algorithme de prédiction linéaire à faible retard avec excitation par code (LD-CELP) à 16 kbit/s lorsqu'il est relié à d'autres éléments du réseau. Des orientations générales sont également données pour les signaux vocaux et non vocaux.

Dans cet appendice, l'usage de l'algorithme LD-CELP à 16 kbit/s fait référence à l'algorithme spécifié dans le corps de la Recommandation G.728; l'algorithme MICDA à 32 kbit/s est utilisé en référence à l'algorithme à 32 kbit/s spécifié dans le corps de la Recommandation G.726; et l'algorithme MIC à 64 kbit/s est utilisé en référence à l'algorithme spécifié dans le corps de la Recommandation G.711.

II.2 Caractéristiques de la transmission vocale

II.2.1 Codage unique

Dans des conditions de transmission sans erreur, la qualité perçue d'un codec avec prédiction linéaire à faible retard avec excitation par code (LD-CELP) à 16 kbit/s est inférieure à celle d'un codec MIC à 64 kbit/s, mais est équivalente à celle d'un codec conforme au codage MICDA à 32 kbit/s.

On a trouvé que les caractéristiques du codec avec prédiction linéaire à faible retard avec excitation par code (LD-CELP) à 16 kbit/s n'étaient pas affectées de façon importante par le bruit (distribution gaussienne) pour un taux d'erreur sur les bits (BER) (*bit error ratio*) jusqu'à 10^{-3} , et avait des caractéristiques équivalentes à celles d'un codec conforme au codage MICDA à 32 kbit/s pour un taux d'erreur sur les bits de 10^{-2} .

La valeur de planification attribuée au codec avec prédiction linéaire à faible retard avec excitation par code (LD-CELP) à 16 kbit/s se trouve dans la Recommandation G.113.

II.2.2 Caractéristiques de transmission vocale en cas d'interconnexion avec des systèmes de codage de type analogique

II.2.2.1 Mises en cascade successives de codecs LD-CELP à 16 kbit/s

Quand un codec LD-CELP à 16 kbit/s est mis en cascade avec des dispositifs de codage multiples de la parole, ses caractéristiques sont considérées comme étant équivalentes à celles d'un codec MICDA à 32 kbit/s pour un nombre de dispositifs en cascade allant jusqu'à trois. On trouvera des règles précises pour la mise en cascade de codecs LD-CELP à 16 kbit/s dans la Recommandation G.113.

II.2.2.2 Caractéristiques pour une association avec un codec MICDA à 32 kbit/s

Les caractéristiques de parole quand un codec LD-CELP à 16 kbit/s est relié en cascade à un ou plusieurs systèmes de codage conformes à l'algorithme MICDA à 32 kbit/s sont équivalentes dans les deux configurations suivantes:

$$\begin{aligned} &G.728 + G.726 + G.728, \\ &G.726 + G.728 + G.726, \end{aligned}$$

le symbole «+» dans la formule ci-dessus désignant une interconnexion.

II.2.3 Caractéristiques de transmission vocale du codec LD-CELP à 16 kbit/s pour la mise en cascade synchrone

Des essais subjectifs ont montré que les caractéristiques de transmission vocale lors de la mise en cascade synchrone (c'est-à-dire l'interconnexion de deux codecs LD-CELP ou davantage par interfaces de codage MIC à 64 kbit/s) sont équivalentes aux caractéristiques de la transmission vocale pour la mise en cascade asynchrone (voir II.2.2.1). De ce fait, le codec LD-CELP n'a pas de propriétés de mise en cascade synchrone équivalentes à celles du codage MICDA à 32 kbit/s (voir la Recommandation G.726).

II.2.4 Caractéristiques de transmission avec un codage différent de celui de l'algorithme LD-CELP à 16 kbit/s et du codage MICDA à 32 kbit/s

En général, l'interconnexion du codec LD-CELP à 16 kbit/s avec un nombre de dispositifs inférieur ou égal à deux semble être équivalente à l'interconnexion du codeur MICDA à 32 kbit/s avec les mêmes dispositifs. Cependant, au moment où le présent appendice était rédigé, seule une quantité limitée de données était disponible sur ce sujet. Cependant, des précautions doivent être prises quand l'interconnexion est effectuée vers des codecs avec un codage différent de celui de l'algorithme LD-CELP à 16 kbit/s et du codage MICDA à 32 kbit/s.

II.3 Caractéristiques de transmission des signaux non vocaux

Il convient de noter que le codec LD-CELP à 16 kbit/s est un système adaptatif qui a été optimisé pour la parole. Il convient de prendre des précautions particulières en effectuant les mesures avec les signaux non vocaux parce que les hypothèses normales de l'invariance dans le temps et de la linéarité ne devraient pas être faites (par exemple en réalisant les essais de maintenance du réseau en utilisant des tonalités de mesure).

II.3.1 Caractéristiques de transmission avec des tonalités d'information

Les essais ont montré que les tonalités d'information créées par le réseau, conformes aux Recommandations Q.35, pouvaient être facilement reconnues après le passage dans un codec LD-CELP à 16 kbit/s (un codage).

II.3.2 Caractéristiques de transmission pour la musique

Les essais ont montré qu'il n'y avait pas d'effets gênants pour une grande diversité de musiques (voir également II.4).

II.3.3 Caractéristiques de transmission de la signalisation multifréquence à double tonalité (DTMF) (*dual-tone multi-frequency*)

En général, les caractéristiques du codec LD-CELP à 16 kbit/s en codage unique sont perçues comme étant équivalentes à celles du codage MICDA à 32 kbit/s et du codage MIC à 64 kbit/s.

II.3.4 Caractéristiques de transmission avec le système de signalisation entre enregistreurs du système de signalisation 5

En général, les caractéristiques du codec LD-CELP à 16 kbit/s en codage unique sont perçues comme étant équivalentes à celles du codage MICDA à 32 kbit/s et du codage MIC à 64 kbit/s.

II.3.5 Caractéristiques de transmission pour les données de bande vocale

En général, les caractéristiques de fonctionnement du codec LD-CELP à 16 kbit/s (même pour un codage simple) sont perçues comme étant nettement inférieures à celles du codage MICDA à 32 kbit/s et du codage MIC à 64 kbit/s. Il convient cependant de noter que le codage LD-CELP à 16 kbit/s peut recevoir la plupart (mais pas tous) des modems à données dans la bande vocale fonctionnant à un débit de 2400 bit/s ou à un débit inférieur, dans des conditions réalistes de circuits de transmission, pour autant que le filtre et le postfiltre de pondération de perception soient tous deux déconnectés.

II.4 Signaux vocaux artificiels

Au cours des essais du codec LD-CELP à 16 kbit/s, on a découvert qu'un certain type de données introduites pouvait entraîner la divergence des étages du codeur et du décodeur, provoquant une forte distorsion du signal de sortie. Ce comportement a été observé pour la première fois pour certains signaux musicaux et pour la parole artificielle (voir la Recommandation P.50). Il a été ensuite observé pour certains sons voisés soutenus lorsqu'ils sont prononcés par certains locuteurs. Il n'a pas été observé pendant le discours réel.

Afin de permettre l'apparition de ce phénomène, les états du codeur et du décodeur doivent être légèrement différents au moment où commence l'introduction des données. La différence pourrait être causée par des différences légères de calcul, comme cela serait normal avec un codeur et un décodeur provenant de constructeurs différents. Même dans les codeurs et les décodeurs fonctionnant de façon absolument identique, de telles différences pourraient être provoquées par des erreurs de circuits de transmission ou en raison de la réinitialisation à des instants différents dans le codeur et le décodeur. Il a été observé que le niveau de distorsion du signal de sortie augmentait si la précision des calculs était réduite dans les réalisations du codeur et du décodeur.

Une seconde condition nécessaire à la divergence est que le signal d'entrée réponde à certaines caractéristiques. Tout d'abord, le signal doit contenir dans son spectre un certain nombre de pics très fins (le plus petit nombre de pics observé comme étant à l'origine de ce phénomène est 14). Ensuite, ces pics doivent avoir des fréquences relativement stationnaires sur plusieurs centaines de millisecondes pour permettre un accroissement des écarts dans le temps).

Pendant la période de divergence, le signal de sortie du décodeur ne devient pas instable au sens où il deviendrait illimité. Il reste un signal fini mais le contenu de son spectre diffère considérablement de celui du signal d'entrée. Si le signal d'entrée change, le codeur et le décodeur convergent de nouveau en donnant au signal de sortie une haute qualité.

Bibliographie

Speech communication, Volume 21, n° 2, juin 1993 (numéro spécial sur la norme CCITT à propos du codage de la parole à 16 kbit/s).