

UNE NOUVELLE STRUCTURE NUMERIQUE DE PREDICTION ADAPTATIVE
POUR LE TRANSCODAGE MIC/MICDA

Madeleine BONNET^{*,**} et Odile MACCHI^{*}

^{*} Laboratoire des Signaux et Systèmes, CNRS-ESE,
Plateau du Moulon, 91190 Gif-sur-Yvette

^{**} Université Pierre et Marie Curie, LIF
4, Place Jussieu, 75230 - Paris Cedex 05

Résumé

L'algorithme de transmission numérique à 32 kb/s normalisé par le CCITT n'assure pas en régime transitoire la propriété de non accumulation des dégradations, propriété souhaitable lors de transcodages successifs.

Une nouvelle structure numérique est proposée où tous les filtres considérés traitent des signaux quantifiés et où l'adaptation des prédicteurs tout en cherchant encore à maximiser le gain de prédiction (pour réduire le débit) vise à éviter l'accumulation des dégradations en forçant le signal (MIC ou MICDA) à être le plus proche possible du niveau quantifié de référence.

Grâce à ce nouveau système l'accumulation des dégradations est évitée tout en assurant, pour un seul transcodage, un bruit de quantification de l'ordre de celui de l'algorithme normalisé.

I. Introduction

Pour augmenter la capacité des réseaux le CCITT a normalisé récemment [1] un algorithme permettant une réduction du débit de 64 kb/s (8KHzx8 bits) à 32 kb/s (8KHzx4 bits) selon le codage MICDA (MIC différentiel adaptatif), qui utilise à la fois la prédiction adaptative et la quantification adaptative. L'adaptativité est nécessaire pour tenir compte d'une part des différents types de signaux transisant sur les lignes (parole, données, tonalités) et d'autre part des variations statistiques à court terme du signal de parole (puissance, spectre).

Pour être acceptables dans un réseau entièrement numérisé, les transcodages successifs MIC-MICDA et MICDA-MIC (voir Fig.1) ne doivent pas accumuler les dégradations. Celles-ci proviennent principalement des diverses opérations de quantification; le bruit de quantification doit donc être indépendant du nombre de transcodages. Dans la suite nous désignerons par (P) cette propriété. Sur la figure 1, la différence ($S^1 - \bar{s}^2$) entre l'entrée et la sortie de la chaîne a) doit être égale à celle ($S^1 - \bar{s}^3$) de la chaîne b), où S^1 est le signal analogique d'entrée, \bar{s}^2 (resp. \bar{s}^3) le signal linéarisé MIC après le premier (resp. deuxième) transcodage. L'opération E_{MIC}^{oQ} représente l'opération de quantification MIC. Il est facile de comprendre que si les codes MICDA I^1 et I^2 sont égaux et si les paramètres du décodeur D_2 sont égaux à ceux de D_1 , on a $\bar{s}^2 = \bar{s}^3$ car l'évolution du décodeur dépend seulement de son entrée I. Donc la non-accumulation des bruits de quantification est liée à l'alignement asymptotique des paramètres du deuxième décodeur sur ceux du premier, ainsi qu'à l'égalité des erreurs de prédiction.

Avec l'algorithme normalisé cette propriété (P) n'est pas obtenue en présence de signaux à bande étroite [2]. Initialisé différemment (par exemple à la suite d'erreurs de transmission ou de microcoupures) D_2 n'arrive jamais à s'aligner sur D_1 et les bruits de quantification se cumulent à chaque transcodeur. Les

Abstract

After transmission errors and/or microbreaks, the 32 kbit/s ADPCM coder accumulates the quantizing distortions when tandem connections of ADPCM-codecs with a 64 kbit/s PCM are interfaced.

A new digital structure is proposed, where all the filters are processing quantized signals. The predictors updating uses a trade off criterion which aims at maximizing the prediction gain (in order to reduce the transmission rate) but also tries to avoid the accumulation of distortion by constraining the signal to be as close as possible to its quantized level.

While it avoids the accumulation of distortion, this new system, when there is a single transcoder, achieves a signal to quantization noise ratio in the same range as the standardized algorithm.

causes de ce phénomène ont été analysées dans [3] et [4].

C'est pourquoi nous proposons une nouvelle structure et une nouvelle adaptation du codeur (et du décodeur) que nous comparons à l'algorithme normalisé. Les simulations montrent que la propriété (P) est alors obtenue et en outre que la dégradation ($S^1 - \bar{s}^2$) due au premier transcodage est du même ordre que celle obtenue avec l'algorithme normalisé.

II. Structure et algorithme normalisés

L'étude de la non accumulation des dégradations peut être effectuée avec seulement deux transcodages comme sur la figure 2, où les indices 2 et ' entre () correspondent respectivement au codeur 2 et au décodeur 2. Cette figure représente la structure prédicteur/quantificateur utilisée dans l'algorithme normalisé (appelé (N)), en l'absence d'erreurs de transmission. La réduction de débit est obtenue par prédiction récursive adaptative du signal MIC d'entrée téléphonique \bar{s}^1 et par quantification adaptative de l'erreur de prédiction e^1 . Au codeur, le prédicteur de type ARMA (2,6) reçoit un signal "reconstitué"

$$\bar{s}_n^1 = \bar{e}_n^1 + \hat{s}_n^1 \quad (1)$$

où \hat{s}^1 est le signal prédit selon

$$\hat{s}_n^1 = \sum_{i=1}^2 a_i^1(n) \bar{s}_{n-i}^1 + \sum_{j=1}^6 b_j^1(n) \bar{e}_{n-j}^1 \triangleq A_n^1 \bar{S}_n^1 + B_n^1 \bar{E}_n^1, \quad (2)$$

où A_n^1 et B_n^1 sont les vecteurs paramètres AR et MA du prédicteur à l'instant n et $\bar{S}_n^1 = (\bar{s}_{n-1}^1, \bar{s}_{n-2}^1)^T$ et $\bar{E}_n^1 = (\bar{e}_{n-1}^1, \dots, \bar{e}_{n-6}^1)^T$ les échantillons passés des signaux. L'erreur de prédiction à l'instant n

$$e_n^1 = \bar{s}_n^1 - \hat{s}_n^1, \quad (3)$$

est transformée en \bar{e}_n^1 par le quantificateur adaptatif MICDA E_{DA}^{oQ} . Seule l'erreur I^1 (codée sur 4 bits) est transmise au décodeur; ceci implique que le codeur simule le décodeur. C'est pourquoi le signal reconstitué \bar{s}^1 est à l'entrée du prédicteur et non le



signal \bar{s}^1 . En supposant acquis l'alignement de E_{DA} sur E_{DA}^1 (démontré en [3]) on peut écrire

$$\bar{e} = e^1; \bar{e}^1 = e^2. \quad (4)$$

Dans (N), les adaptations des filtres transverse A et récursif B du prédicteur sont découplées; ces deux filtres sont mis en cascade et adaptés respectivement par leur sortie. L'adaptation, de type gradient stochastique (LMS) est, pour la partie MA du codeur, et du décodeur

$$B_{n+1}^1 = B_n^1 + \beta \bar{e}_n^1 \bar{E}_n^1; B_{n+1} = B_n + \beta \bar{e}_n^1 \bar{E}_n^1 \quad (5)$$

où β est un pas d'incrémentent positif.

Dans ce contexte, pour obtenir la propriété (P), on montre que l'alignement de tous les organes successifs doit être atteint. L'algorithme normalisé s'avère incapable d'obtenir l'alignement asymptotique à partir de conditions initiales différentes. En conséquence il accumule les bruits de quantification.

III. Nouvelle structure, nouvel algorithme

III.1. Structure

La nouvelle structure proposée (fig. 3) considère le prédicteur/quantificateur dans son ensemble. La représentation linéaire \bar{s}^1 du signal MIC entre directement dans le filtre transverse A^1 au codeur et au décodeur le signal reconstitué \bar{s} subit la quantification MIC avant d'être injecté dans le filtre A.

Cette structure est non linéaire tant au décodeur qu'au codeur, alors que le décodeur normalisé demeurait linéaire. Elle présente donc un caractère symétrique puisque codeur et décodeur traitent tous les deux des niveaux quantifiés dans les parties transverse et récursive du filtrage contrairement à la structure normalisée. Ceci évite la boucle qui, au codeur, simule le décodeur suivant et dont il a été montré [3] qu'elle rend le codeur instable avec l'adaptation normalisée. D'après (4) les relations entre les signaux de deux transcodeurs successifs (cf. fig. 3) s'écrivent

$$I = I^1, \bar{s}^2 = \bar{s}, I^1 = I^2, \bar{s}^3 = \bar{s}^1. \quad (6)$$

Il vient

$$e_n^i = \bar{s}_n^i - [A_n^{iT} \bar{S}_n^i + B_n^{iT} \bar{E}_n^i] \quad \text{au codeur } i \quad (7)$$

$$e_n^i - \bar{e}_n^i = \epsilon_{DA}^i(n) \quad (\text{erreur de quantification MICDA}) \quad (8)$$

$$\tilde{s}_n = \bar{e}_n^1 + [A_n^T \bar{S}_n^2 + B_n^T \bar{E}_n^1] \quad \text{au décodeur } 1 \quad (9)$$

$$\tilde{s}_n - \bar{s}_n^2 = \epsilon_{MIC}^1(n) \quad (\text{erreur de quantification MIC}) \quad (10)$$

$$\tilde{s}_n^1 = \bar{e}_n^2 + [A_n^{1T} \bar{S}_n^3 + B_n^{1T} \bar{E}_n^2] \quad \text{au décodeur } 2 \quad (9')$$

$$\tilde{s}_n^1 - \bar{s}_n^3 = \epsilon_{MIC}^2(n). \quad (10')$$

III.2. Adaptation du prédicteur

Le nouveau mode d'adaptation des prédicteurs tend à la fois à maximiser le gain de prédiction pour réduire le débit et à minimiser les bruits de quantification pour éviter l'accumulation des dégradations. Pour cela le critère utilisé au codeur est la minimisation de la quantité $J^1(P) + \lambda J^1(Q)$, avec

$$J^1(P) = E\{(e_n^1)^2\}, J^1(Q) = E\{(e_n^1 - \bar{e}_n^1)^2\}. \quad (11)$$

En pratique on omet les moyennes et (11) devient $J_n^1(C) = (e_n^1)^2 + \lambda (e_n^1 - \bar{e}_n^1)^2$, dont on écrira le gradient. Le gradient de $(e_n^1 - \bar{e}_n^1)^2$ est obtenu en supposant que les variations de A^1 et B^1 sont suffisamment faibles pour que les niveaux \bar{E}_n^1 ne changent pas donc pour que

$$\nabla(e_n^1 - \bar{e}_n^1) = \nabla e_n^1. \quad (12)$$

Pour la même raison $\nabla \bar{E}_n^1$ est supposé nul. On obtient donc d'après (7)

$$\nabla_{A^1} J_n^1(C) = -2[e_n^1 + \lambda(e_n^1 - \bar{e}_n^1)] \bar{S}_n^1, \quad (13)$$

$$\nabla_{B^1} J_n^1(C) = -2[e_n^1 + \lambda(e_n^1 - \bar{e}_n^1)] \bar{E}_n^1.$$

En faisant intervenir l'erreur quantifiée e^1 et le bruit de quantification (8), l'adaptation au codeur peut s'écrire sous la forme

$$A_{n+1}^1 = A_n^1 + \alpha[\bar{e}_n^1 + (\lambda+1) e_{DA}^1(n)] \bar{S}_n^1, \quad (AC)$$

$$B_{n+1}^1 = B_n^1 + \beta[\bar{e}_n^1 + (\lambda+1) e_{DA}^1(n)] \bar{E}_n^1. \quad (BC)$$

Naturellement le décodeur ne dispose pas de l'erreur de quantification e_{DA}^1 . Il faut la reconstituer. Pour exposer l'adaptation utilisée au décodeur nous évaluons le bruit dû à un transcodage en définissant l'hypothèse d'alignement des transcodeurs qui suppose que les signaux et les paramètres sont dans le même état jusqu'à l'instant $n-1$; en particulier

$$A_n^1 = A_n^2, B_n^1 = B_n^2, \bar{S}_n^3 = \bar{S}_n^2. \quad (14)$$

A l'étape n , la dégradation due au deuxième transcodage s'exprime à l'aide de (7), (9'), (10'),

$$\begin{aligned} \bar{s}_n^2 - \bar{s}_n^3 &= e_n^2 + (A_n^{2T} \bar{S}_n^2 + B_n^{2T} \bar{E}_n^2) \\ &+ \epsilon_{MIC}^2(n) - (e_n^2 + A_n^{1T} \bar{S}_n^3 + B_n^{1T} \bar{E}_n^2) \end{aligned} \quad (15)$$

soit encore, d'après (14):

$$\bar{s}_n^2 - \bar{s}_n^3 = \epsilon_{DA}^2(n) + \epsilon_{MIC}^2(n). \quad (16)$$

Avec l'hypothèse d'alignement précédente, la propriété (P) implique que $\bar{s}_n^2 = \bar{s}_n^3$ ou

$$e_n^2 - \bar{e}_n^2 = -(\bar{s}_n^1 - \bar{s}_n^3). \quad (17)$$

D'après (17) à partir du deuxième transcodeur, l'erreur MIC doit compenser l'erreur MICDA. Au décodeur, l'erreur de quantification MIC est connue. Il y a donc un moyen d'estimer l'erreur ϵ_{DA}^1 non disponible. Bien que (17) concerne le décodeur 2 nous l'appliquons au décodeur 1, dont l'adaptation se calque sur celle du codeur selon

$$A_{n+1} = A_n + \alpha[\bar{e}_n^1 + (\lambda+1) (\bar{s}_n^2 - \bar{s}_n^3)] \bar{S}_n^2 \quad (AD)$$

$$B_{n+1} = B_n + \beta[\bar{e}_n^1 + (\lambda+1) (\bar{s}_n^2 - \bar{s}_n^3)] \bar{E}_n^1 \quad (BD)$$

Justification : A quel critère $J(D)$ correspond cette adaptation du décodeur? Puisque

$$\tilde{s}_n = \bar{e}_n^1 + A_n^T \bar{S}_n^2 + B_n^T \bar{E}_n^1, \quad (18)$$

la minimisation du bruit de quantification $J_n(Q) = (\tilde{s}_n - \bar{s}_n^2)^2$ donne, pour le gradient stochastique et en supposant les variations de A et B assez faibles pour que les niveaux \bar{s}_n^2 ne changent pas:

$$\nabla_A J_n(Q) = -2(\bar{s}_n^2 - \tilde{s}_n) \bar{S}_n^2, \quad (19)$$

$$\nabla_B J_n(Q) = -2(\bar{s}_n^2 - \tilde{s}_n) \bar{E}_n^1, \quad (20)$$

si bien que (AD), par exemple, s'écrit:

$$A_{n+1} = A_n + \alpha[\bar{e}_n^1 + \bar{s}_n^2 - \tilde{s}_n] \bar{S}_n^2 - \frac{\alpha}{2} \lambda \nabla_A J_n(Q) \quad (21)$$

Le second terme à droite de (21) s'écrit aussi $\alpha(\bar{s}_n^2 - \tilde{s}_n) \bar{S}_n^2$. C'est le gradient stochastique de la quantité $\frac{\alpha}{2} J(P)$; $J(P) = E(\bar{s}_n^2 - \tilde{s}_n)^2$. Sa présence dans l'algorithme tend à minimiser l'erreur entre le niveau MIC \bar{s}_n^2 et sa prédiction \tilde{s}_n . On remarque que lorsque $\tilde{s}_n = \bar{s}_n^2$ (égalité qui correspond à l'alignement du codeur 2 et du décodeur 1) alors $J(P) = E\{(e_n^2)^2\}$, e_n^2 étant l'erreur de prédiction au deuxième codeur. Ainsi, comme le codeur, le décodeur utilise pour s'adapter une erreur de prédiction non quantifiée bien que celle-ci ne lui soit pas disponible, selon

$J_n(D) = (\bar{s}_n^2 - \hat{s}_n)^2 + \lambda(\bar{s}_n^2 - \bar{s}_n)^2$. Le nouvel algorithme baptisé (NUM) s'écrira {(AC),(BC),(AD),(BD)}.

Un cas limite intéressant correspond à $\lambda=0$. Au décodeur, l'erreur entre crochets qui pilote l'incrément devient égale à $\bar{s}_n^2 - \hat{s}_n$, erreur de prédiction au deuxième codeur (lorsque $\hat{s}=\hat{s}^2$). Au codeur l'adaptation utilise la sortie e avant quantification contrairement à (N) qui utilise \bar{e} et perd ainsi toute une part d'information. Ainsi (NUM) (resp. (N)) cherche à décorréliser e_n (resp. \bar{e}_n) de $\bar{e}_{n-1}, \dots, \bar{e}_{n-6}$. En utilisant toute l'information de e , (NUM) met en oeuvre une stratégie supérieure.

IV. Résultats de simulation

Pour l'algorithme (NUM), le quantificateur MICDA est soit normalisé [1], soit amélioré par lissage selon [3]. Comme pour l'algorithme (N) nous introduisons un paramètre d'oubli δ dont le rôle est de favoriser l'alignement des prédicteurs et des quantificateurs. Ainsi, par exemple pour le codeur, on utilise (NUM) sous la forme

$$A_{n+1}^1 = (1-\delta_A) A_n^1 + \alpha[\bar{e}_n^1 + (\lambda+1) e_{DA}^1(n)] \bar{S}_n^1 \quad (AC)$$

Les paramètres ont été ajustés par simulations: $\alpha=2 \cdot 10^{-9}$, $\beta=4 \cdot 10^{-9}$, $\delta_A=2^{-7}$, $\delta_B=2^{-8}$. A l'initialisation, les quantificateurs sont alignés et les prédicteurs sont tels que $A^1=A^2=A^3=A^4=(0.5, -0.5)^T$; $b_i^1=b_i^2=b_i^3=b_i^4=0, i=1, \dots, 6$.

La propriété (P) sera atteinte si $\bar{s}^2=\bar{s}^3$. Mais elle n'est pas suffisante pour un bon fonctionnement. Encore faut-il que le bruit de quantification dû au premier transcodage soit tolérable. Pour comparer les bruits de quantification de (N) et de (NUM) nous mesurons la dégradation due au i -ème transcocodeur par $(S/B)_i = 10 \log_{10} (E(\bar{s}_i^2) / E(\bar{s}_i^2 - \bar{s}_i^1)^2)$ ($i=1,2$). Sur le tableau qui suit on notera que (N) accuse des dégradations $(S/B)_i$ voisines à chaque transcodage. Ainsi le rapport signal sur bruit final se dégrade d'environ 3 dB à chaque nouveau transcodage comme il est dit dans [6].

Le tableau concerne d'abord une entrée \bar{s}^1 sinusoïdale, d'amplitude 3001 de fréquence 815 Hz, ayant subi la quantification MIC selon la loi A. Pour $\lambda=0$ et 1 les performances de (NUM) au premier transcocodeur sont égales ou supérieures à celles de (N); au deuxième transcocodeur la différence est d'au moins 10 dB en faveur de (NUM). Les courbes présentées correspondent à $\lambda=0$ avec un quantificateur

normalisé. La figure 4 est relative à $(S/B)_0=49$ dB. La propriété (P) est satisfaite en moins d'une seconde. Au moyen de l'écart entre les erreurs quantifiées, la figure 5 caractérise l'alignement des quantificateurs aux premier et deuxième transcocodeurs. Comme il a été vu dans l'introduction la propriété (P) et l'alignement asymptotique des erreurs de prédiction sont liés.

Le tableau donne aussi les résultats concernant la phrase "Annie s'ennuie loin de mes parents" issue de [5] représentée figure 6, après quantification MIC. Avec $\lambda=0$ la perte au premier transcocodeur est de 3 dB par rapport à (N) mais le gain au deuxième transcocodeur est de plus de 25 dB avec le quantificateur lissé.

Conclusion

Nous préconisons l'utilisation de (NUM) avec $\lambda=0$ et le quantificateur lissé introduit dans [3]. Ce nouveau système supprime presque totalement la dégradation au deuxième transcocodeur au prix - pour la parole seulement - d'une perte de 3 dB sur le rapport signal/bruit de quantification au premier transcocodeur. Avec (N) le rapport signal à bruit se dégradait d'environ 3 dB à chaque transcodage. Donc après deux transcodages, (NUM) et (N) donnent la même dégradation. Dans les transcodages ultérieurs avec (NUM), les valeurs très élevées de $(S/B)_i$, pour $i \geq 2$, n'entraînent aucune dégradation supplémentaire tandis que (N) perd 3 dB à chaque transcodage. Ainsi (NUM) est très supérieur à (N).

Alors que (N) conserve un aspect de transmission analogique en cherchant à ajuster le signal reconstitué \bar{s} sur le signal d'entrée \bar{s} au moyen d'une parfaite prédiction ($\hat{s}=\bar{s}$), l'algorithme (NUM) procède d'une approche globalement numérique tendant à ajuster l'entrée \bar{s}^2 du transcocodeur 2 sur celle, \bar{s}^1 , du transcocodeur 1.

References

- [1] CCITT, Red Book, Recommendation G721, Tome III-3, Oct. 1984.
- [2] M. TAKA, R. MARUTA, A. LE GUYADER, "Synchronous tandem algorithm for 32 kb/s ADPCM", Session 23, GLOBECOM 84, Atlanta.
- [3] M. BONNET, O. MACCHI, M. JAÏDANE, "A multicriterion view of jointly adaptive prediction/quantization with application to ADPCM transmission. Parts I, II, October 1986.
- [4] M. BONNET et al. "Anatomy of 32 kb/s ADPCM CCITT algorithm: misadjustment of tandem connections in transient state", GLOBECOM 86, Dec. 1986, Houston, Texas.
- [5] P. COMBESURE, "20 listes de 10 phrases en français, phonétiquement équilibrées", Note technique, CNET, Lannion, Oct. 1980.
- [6] JIN-ICK LEE, CHONG KWAN UN, "Performance analysis of PCM/ADPCM transcoding systems", IEEE Trans. on Comm., vol.33, n°12, pp.1303-1308, Dec. 1985.

Remerciements

Ce travail supporté financièrement par le CNET Lannion a bénéficié des conseils de Messieurs Alain Charbonnier et Xavier Maitre que nous remercions vivement.

		Sinusoïde		Parole	
		(S/B) ₁	(S/B) ₂	(S/B) ₁	(S/B) ₂
(N)		27.6	31	20.7	25.8
(NUM)	λ				
quantifi- cateur normalisé	0	27	49	17.6	30.6
	1	25.5	44.5	15.7	79.6
	2	25.7	41.6		
quantifi- cateur lissé	0	28.4	40.7	17.7	51.6
	1	28	76	15.4	38.8
	2	27	40.7	13.5	30.4

Tableau: Performances comparées de (NUM) et (N).

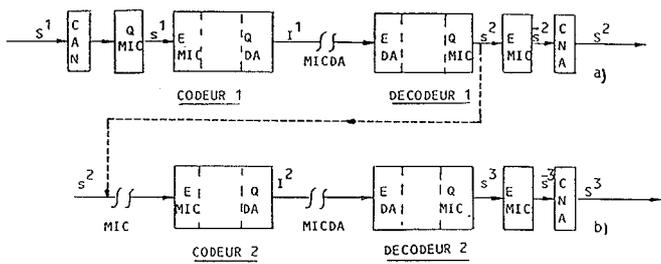


Figure 1: Codage MICDA inséré dans une chaîne de transmission MIC.

- a) Sortie d'un transcodage MIC/MICDA/MIC
- b) Sortie du 2^{ème} transcodage MIC/MICDA/MIC

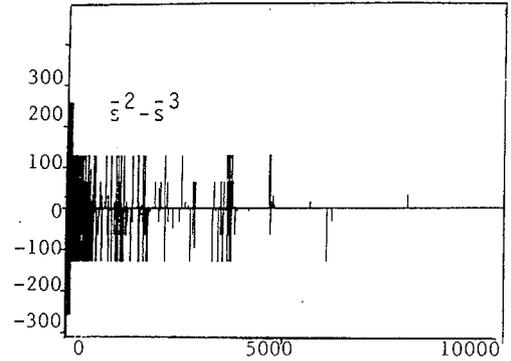


Figure 4: Dégradation au deuxième transcodeur.

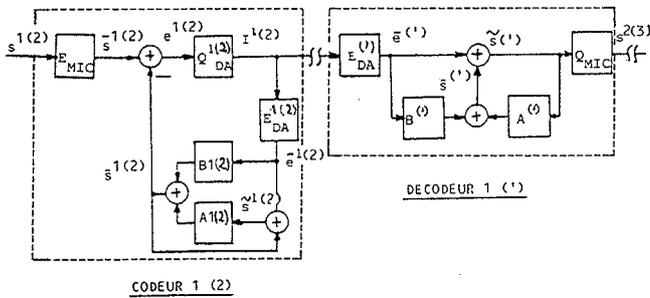


Figure 2: Structure normalisée, deux transcodages successifs.

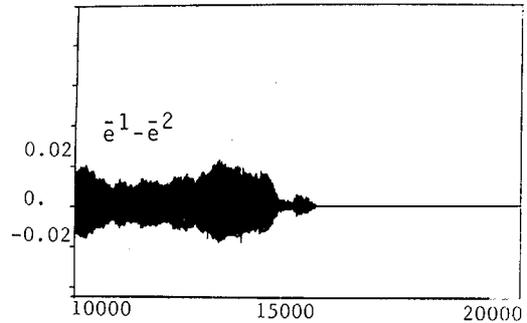


Figure 5: Alignement des quantificateurs 1 et 2 (égalité des erreurs MICDA).

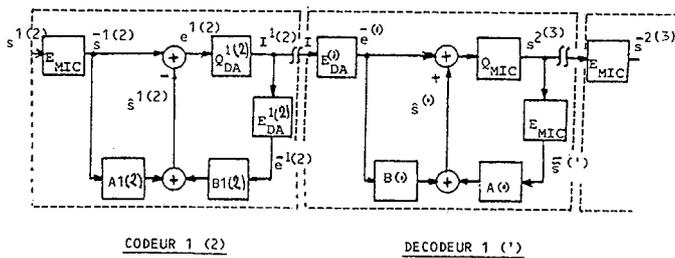


Figure 3: Nouvelle structure symétrique, deux transcodages successifs.

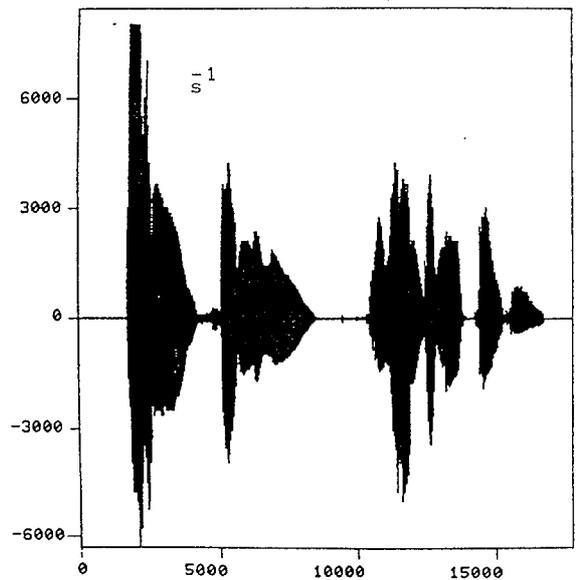


Figure 6: Signal de parole, quantifié MIC: "Annie s'ennuie loin de mes parents".