



LE COMPROMIS ENTRE BRUIT DE CODAGE ET DEBIT DE TRANSMISSION

M. BONNET^{1,2} O. MACCHI¹

¹LSS, CNRS-ESE, Plateau du Moulon, 91192 Gif-sur-Yvette Cedex, France

²Université Paris VI, 4 Place Jussieu, 75230 Paris Cedex 05, France

RÉSUMÉ

Bien souvent, dans les articles traitant de transmission à débit réduit basée sur la prédiction, la structure utilisée n'est pas justifiée de façon satisfaisante: l'obtention d'une erreur de prédiction de variance plus faible que le signal d'entrée n'est pas suffisante pour obtenir une réduction de débit car suivant la place du quantificateur, le bruit de codage peut alors augmenter considérablement. Ce qu'il faut, c'est fixer le rapport signal à bruit de codage tout en diminuant le débit. Dans cet article, deux classes de codeurs comportant un filtrage et une quantification sont comparées pour lesquelles on calcule explicitement le nombre de bits de quantification en fonction de la puissance du bruit de codage. Le calcul justifie le choix d'une structure de prédiction incluant le quantificateur dans la boucle.

SUMMARY

In most studies dealing with digital transmission at reduced bit rate and based on prediction, the system structure is not well justified: a prediction error with variance smaller than the input signal power does not necessarily yield a reduced bit rate. Depending on the location of the quantizer, the coding noise may even be considerably increased. The real objective is to maintain the power ratio of signal to coding noise while decreasing the bit rate. In this paper we compare two kinds of structure for the coder, including a filter and a quantizer. In both cases, we evaluate the number of transmission bits required to reach a given coding noise power. This confirms the usual choice of including the quantizer in the prediction loop.

I INTRODUCTION

Dans la transmission d'un signal, par exemple acoustique ou d'une image, le codage MIC différentiel [1], basé sur la prédiction, met à profit la redondance du signal pour obtenir une réduction de débit numérique selon le raisonnement intuitif suivant: "l'erreur de prédiction e représente l'innovation; elle comporte moins d'information que le signal s à transmettre; d'ailleurs elle est de variance plus faible donc on peut la transmettre en la quantifiant sur moins de bits".

Le but de cette contribution est de mettre en lumière l'insuffisance de ce raisonnement souvent invoqué dans la littérature [2-3]. Effectivement, il ne suffit pas de prédire, encore faut-il quantifier à bon escient.

Un système de codage étant toujours associé à un système de décodage nous utilisons une mesure de fidélité de reproduction du signal: le *bruit de codage* η , différence entre le signal s à l'entrée du codeur et le signal décodé s' :

$$\eta = s - s' \quad (1)$$

Une autre caractéristique du système de transmission est le débit de transmission qui, pour une fréquence d'échantillonnage fixe est proportionnel au nombre de bits B du quantificateur Q , toujours inclus dans le codeur.

Un système de transmission numérique doit réaliser un compromis entre bruit de codage et débit de transmission.

Nous montrons d'abord qu'il ne suffit pas de réduire la

variance du signal à quantifier pour obtenir une réduction de débit satisfaisante car, suivant la place du quantificateur dans le codeur, le bruit de codage peut augmenter considérablement. Ce qu'il faut, c'est fixer le rapport H/S , où H est la puissance du bruit de codage et S celle du signal, tout en diminuant le débit.

Nous comparons deux classes bien connues de codeurs comportant un filtrage F du signal d'entrée s ainsi qu'une quantification Q . La première classe (figure 1) découple les deux opérations, la quantification étant postérieure. Sous sa forme la plus simple cette classe inclut le codage direct MIC; dans ce cas $F_1 = Id$. La seconde comporte une boucle reliant les deux opérations comme dans le codage MIC différentiel, ce qui rend la prédiction non linéaire, puisque le filtre F_2 (voir figure 2) utilise des entrées quantifiées \tilde{e} . Pour chaque classe nous calculons analytiquement le nombre de bits de quantification en fonction du bruit de codage H , ce qui justifie le choix pratique de la deuxième classe.

Enfin, nous montrons comment la structure couplée peut être justifiée à partir d'un schéma de commande classique du système global $s \rightarrow s'$ incluant codeur et décodeur, commande dotée d'une boucle de contre réaction.

Dans ce qui suit, nous supposons

- i) que la transmission s'effectue sans erreur,
- ii) que codeur et décodeur sont alignés au sens où le décodeur dispose de la connaissance du filtre F utilisé par le codeur. Ceci peut être obtenu soit par une transmission auxiliaire des paramètres, soit par des algorithmes adaptatifs [4] bien conçus



qui calculent la valeur de F sur la base du signal quantifié \tilde{e} , disponible à la fois au codeur et au décodeur

iii) que tout quantificateur est bien adapté à la puissance de son entrée. Cette dernière hypothèse implique que le rapport signal à bruit de quantification est indépendant de la statistique de l'entrée et uniquement fonction du nombre de niveaux de sortie du quantificateur. Elle implique aussi que le bruit de quantification est blanc (si toutefois le quantificateur possède suffisamment de niveaux).

II STRUCTURE DECOUPLEE

Considérons d'abord le cas de référence où le filtre F_1 est l'identité (donc $s = e$). C'est le cas du codage MIC: le codeur est un simple quantificateur dont la sortie représente à la fois le signal transmis et le signal décodé.

Notons que pour un tel système, bruit de codage et bruit de quantification sont confondus. Soit H_0 la variance du bruit de codage et ρ_0 le rapport signal à bruit de quantification où le quantificateur comporte N_0 niveaux de sortie:

$$H_0 \triangleq E(\eta^2) \quad (2)$$

$$\rho_0 = S/H_0. \quad (3)$$

Dans le cas plus général de la prédiction, l'équation du codeur est, par exemple pour un filtre F_1 transverse

$$e(n) = F_1(s(n)) = s(n) - \sum_{i=1}^L p_i s(n-i). \quad (4)$$

Au décodeur on a (voir fig. 1)

$$\tilde{e}(n) = F_1(s'(n)) = s'(n) - \sum_{i=1}^L p_i s'(n-i). \quad (4')$$

Le bruit de quantification $\varepsilon = e - \tilde{e}$ vaut donc

$$\varepsilon(n) = \eta(n) - \sum_{i=1}^L p_i \eta(n-i), \quad (5)$$

en fonction du bruit de codage η donné en (1). En comparant les équations (4) et (5), on voit que le bruit de quantification découle du bruit de codage par le même filtre de prédiction F_1 qui fait passer de s à e .

D'après le principe d'orthogonalité, on sait [5] que le filtre prédictif \mathcal{P} obtenu en minimisant la variance de e est un filtre blanchisseur. Si l'on suppose que le bruit de quantification est lui-même blanc, ce qui est vrai pour une quantification assez fine, on en déduit que le spectre du bruit de codage η a la même forme que celle de s . L'approximation suivante faisant intervenir le gain de prédiction G est donc justifiée:

$$G \triangleq S / E(\varepsilon^2) = H_1 / E(\varepsilon^2), \quad (6)$$

où H_1 est la puissance du bruit de codage. Ainsi, le rapport signal sur bruit du quantificateur vaut

$$\rho_1 \triangleq \frac{E(\varepsilon^2)}{E(\varepsilon^2)} = \frac{S}{H_1}. \quad (7)$$

Par comparaison avec (3) il vient

$$\rho_1 H_1 = \rho_0 H_0; \quad (8)$$

or ρ dépend uniquement du nombre N de niveaux du quantificateur (lorsque l'hypothèse (iii) est satisfaite) selon

$$\rho = N^2/k, \quad (9)$$

où k est une constante sans importance ici. Donc pour un même nombre de niveaux ($N_0=N_1$, $\rho_0=\rho_1$), il vient $H_1 = H_0$ d'après (8). La structure de prédiction n'a pas fait gagner sur la qualité de la transmission. Si l'on décide néanmoins de réduire le débit de transmission par rapport au codage direct ($N_1 < N_0$), le rapport signal sur bruit de quantification est dégradé ($\rho_1 < \rho_0$). D'après (8) il s'ensuit une augmentation du bruit de codage H_1 .

La première structure étudiée ne permet donc pas l'objectif visé: réduire le débit sans dégrader la qualité de restitution du signal. Avant d'abandonner cette structure, quelques remarques peuvent être faites:

i) la prédiction comportant toujours un gain ($G \geq 1$), on déduit de l'équation (6) que la puissance du bruit de codage est supérieure à celle du bruit de quantification.

ii) d'après (4) et (5) $s(n)$ et $\eta(n)$ ont même allure spectrale (non plate) alors que $e(n)$ et $\varepsilon(n)$ ont des densités spectrales de puissance uniformes (lorsque le prédictif utilisé est proche de l'optimal qui blanchit). Cette remarque sera utilisée dans la suite.

iii) on lit souvent dans la littérature relative au codage MICDA [2-3] que transmettre l'erreur de prédiction e d'un signal, plutôt que le signal lui-même, permet une réduction de débit puisque l'erreur de prédiction, de variance plus faible que le signal d'origine, peut être codée (c'est-à-dire quantifiée) avec moins de bits. Cet argument est insuffisant comme il vient d'être vu. Par exemple, on se persuade facilement qu'une simple division du signal par un facteur k , certes réduit la variance du signal à quantifier mais, augmente en même temps le bruit de codage. Cette réduction forcée de la variance du signal est un cas particulier du système de la figure 1 avec quantificateur hors de la boucle de prédiction, où le prédictif serait réduit à une division.

On voit enfin qu'une augmentation de la variance du signal (par simple multiplication) n'augmente pas le bruit de codage.

III STRUCTURE COUPLEE

La structure 2 est détaillée sur la figure 3 qui fait intervenir un signal reconstitué par le codeur

$$\tilde{s}(n) = \hat{s}(n) + \tilde{e}(n) \quad (10)$$

à partir de la prédiction \hat{s} de s ; $\tilde{s}(n)$ est utilisé à l'entrée du prédictif \mathcal{P} à la place du vrai signal $s(n)$. Ceci est fait dans un but de duplication possible par le décodeur. En d'autres termes le codeur inclut un décodeur local. Dans ce codeur on voit que le quantificateur et le prédictif sont couplés par une boucle. De ce fait, le prédictif du codeur présente une structure non linéaire alors que le décodeur reste linéaire. Selon l'usage, nous conservons cependant le terme "erreur de prédiction" pour désigner ce signal $\tilde{e}(n)$ tout en notant bien qu'il s'agit d'une prédiction non linéaire.

On voit immédiatement sur la figure 3 que, sous les hypothèses (i) et (ii) (alignement), l'on a

$$s'(n) = \tilde{s}(n) \quad (11)$$

Par ailleurs, en rapprochant la définition de l'erreur de prédiction

$$e(n) = s(n) - \hat{s}(n) \quad (12)$$

et l'équation (10) on voit aussitôt que

$$s(n) - \tilde{s}(n) = e(n) - \bar{e}(n). \quad (13)$$

Ainsi

$$s(n) - s'(n) = e(n) - \bar{e}(n), \quad (14)$$

ou encore

$$\eta(n) = \varepsilon(n); \quad (15)$$

le bruit de codage est égal au bruit de quantification. Donc le spectre de η est plat à l'instar de celui de ε .

D'après (15) le rapport signal à bruit de codage vaut

$$S/H_2 = S/E(\varepsilon^2). \quad (16)$$

A l'aide du gain de prédiction G et du rapport signal sur bruit de quantification ρ_2 il vient évidemment

$$S/H_2 = G \rho_2. \quad (17)$$

Par comparaison avec l'équation de référence (3) il vient

$$\rho_2 H_2 = \rho_0 H_0/G. \quad (18)$$

Ainsi pour un même rapport signal sur bruit ρ du quantificateur, donc pour un même nombre de niveaux ($N_2 = N_0$) ou encore pour un même débit dans la transmission, le bruit de codage H est réduit par un facteur égal au gain de prédiction G , lequel est supérieur à 1.

Evidemment, plutôt qu'utiliser la prédiction linéaire pour réduire le bruit de codage on peut aussi l'utiliser pour diminuer le nombre de niveaux du quantificateur, à bruit de codage donné, H . En combinant les équations (9) et (18) il vient alors

$$N_2^2 = N_0^2 / G, \quad (19)$$

et en passant aux débits binaires $D \triangleq \log_2 N$

$$\boxed{D_0 - D_2 = \frac{1}{2} \log_2 G} \quad (20)$$

Cette formule donne, à bruit de codage fixé, la réduction de débit atteinte par la deuxième structure grâce au couplage prédiction quantification, réduction qui est maximale pour le prédicteur qui maximise le gain de prédiction (en minimisant $E(\varepsilon^2)$).

IV APPROCHE PAR L'AUTOMATIQUE

Nous montrons maintenant comment le choix de la structure couplée précédente peut être justifié à partir d'un schéma de commande possédant une boucle de contre réaction. La commande agit sur les paramètres du filtre de boucle \mathcal{P} afin d'identifier la sortie s' d'un filtre inconnu \mathcal{F} avec un signal de référence d (voir figure 4).

Dans le cas qui nous occupe nous considérons le système de transmission lui-même avec son entrée s , sa sortie s' , le signal de référence d étant le signal s afin d'obtenir une erreur $\eta = s - s'$ la plus faible possible.

Le critère à réaliser est la minimisation du bruit de codage η . Le bruit minimal que l'on peut espérer atteindre est le bruit

dû au quantificateur. On cherche donc comment doivent être les filtres \mathcal{F} et \mathcal{P} pour que le bruit de codage soit égal au bruit de quantification.

En notant $F(z)$ et $P(z)$ les fonctions de transfert des filtres \mathcal{F} et \mathcal{P} , les équations du système de la figure 4 s'écrivent:

$$e(z) = s(z) - P(z) s'(z), \quad (21)$$

$$s'(z) = F(z) \bar{e}(z). \quad (22)$$

Le bruit de codage s'exprime par

$$\eta(z) = s(z) - s'(z) = e(z) + [P(z) - 1] F(z) \bar{e}(z), \quad (23)$$

ou, en faisant intervenir le bruit de quantification

$$\eta(z) = e(z) - \bar{e}(z) + \bar{e}(z) [1 + P(z) F(z) - F(z)]. \quad (24)$$

Pour que les bruits de codage et de quantification soient égaux, il suffit donc que le crochet soit nul:

$$F(z) = 1 / [1 - P(z)]. \quad (25)$$

Nous appliquons la propriété (9) en considérant le même quantificateur que celui utilisé dans le codage direct. Si ce quantificateur reçoit en entrée le signal s , il a pour sortie s' ; s'il reçoit e , la sortie est \bar{e} et nous écrivons l'égalité des rapports signaux à bruits:

$$\frac{S}{H_0} = \frac{E(\varepsilon^2)}{E[(e - \bar{e}^2)]}. \quad (26)$$

Lorsque la condition (25) est satisfaite, l'expression (26) donne

$$E[(s - s')^2] = \frac{E(\varepsilon^2)}{S} H_0. \quad (27)$$

Les quantités H_0 et S étant fixées cette formule montre que le filtre \mathcal{P} doit minimiser la variance de e pour minimiser celle du bruit de codage η .

Toujours sous la condition (25), la figure 4 peut se mettre sous la forme 5 qui montre que le codeur dispose d'une référence de la sortie s' du décodeur. Cette référence est le signal reconstitué \tilde{s} qui, en l'absence d'erreur de transmission est égal au signal de sortie s' .

Nous voyons donc que, au codeur, la prédiction s'effectue sur le signal reconstitué \tilde{s} et non sur le signal d'entrée s . Le codeur inclut un décodeur. La structure optimale ainsi trouvée correspond bien à celle du schéma de la figure 3.

CONCLUSION

Pour justifier la structure utilisée dans le codage MIC différentiel nous tenons compte de deux mesures de performances: le débit de transmission et le bruit de codage. Nous évaluons alors ces performances pour deux structures; la première comporte un filtre suivi d'un quantificateur, la seconde, structure couplée, présente le quantificateur dans sa boucle de prédiction. Les performances de chaque structure sont comparées à celles d'un système de codage de référence ne comportant qu'un quantificateur (comme c'est le cas du codage MIC). Par rapport à ce codage, nous montrons que seule la structure couplée permet une réduction de débit, sans dégradation du bruit de codage.



REFERENCES

- [1] J.L. FLANAGAN & al., "Speech Coding", IEEE Trans. on Com., vol. 27, n° 4, April 1979, pp.710-736.
- [2] N.S. JAYANT, P. NOLL, "Digital Coding of Waveforms. Principles and Applications to Speech and Video", Prentice-Hall 1984.
- [3] P. FLEISCHER, "Sufficient conditions for achieving minimum distortion in a quantizer", IEEE Int. Conv. Rec. April 72, pp.104-111.
- [4] M. BONNET, O. MACCHI, M. JAIDANE, "Theoretical analysis of the ADPCM CCITT algorithm", article accepté à IEEE Trans. on Communications, 1989.
- [5] A. BLANC-LAPIERRE, B. PICINBONO, "Fonctions aléatoires", 1981.
- [6] A. GERSHO, "Principles of quantization", IEEE Trans. on Circuits and Systems, vol. 25, n° 7, July 78, pp. 427- 436.

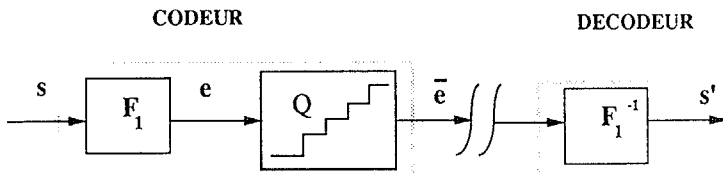


Figure 1: Structure découplée.

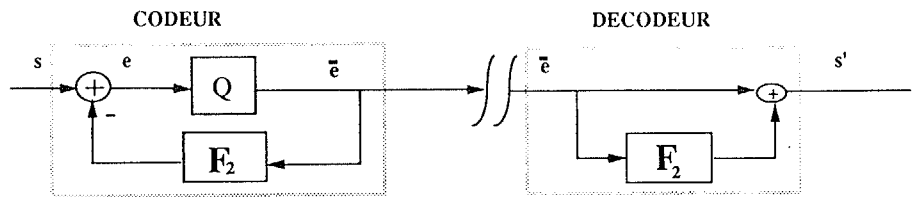


Figure 2: Structure couplée.

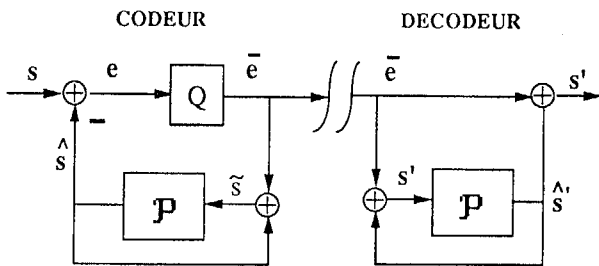


Figure 3: Structure couplée montrant le signal reconstitué.

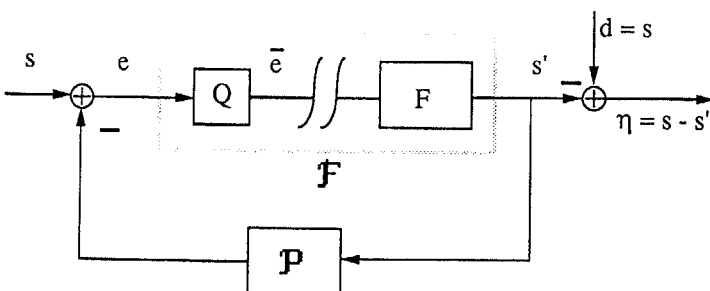


Figure 4: Schéma de commande classique avec boucle fermée.

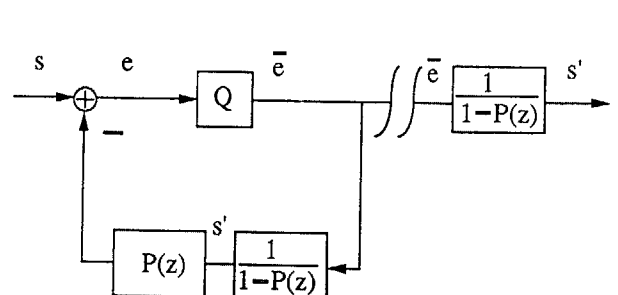


Figure 5: Système montrant la nécessité du signal reconstitué au codeur.