

# Modélisation statistique de l'erreur de quantification dans le codeur AAC. Application à l'algorithme d'allocation de bits.

Olivier DERRIEN<sup>1</sup>, Pierre DUHAMEL<sup>2</sup>, Maurice CHARBIT<sup>1</sup>

<sup>1</sup>ENST TSI / CNRS URA 820  
46, rue Barrault 75634 PARIS Cedex 13, France

<sup>2</sup>CNRS / LSS  
Supélec, Plateau du Moulon 91192 GIF-SUR-YVETTE Cedex, France  
derrien@tsi.enst.fr, pierre.duhamel@lss.supelec.fr  
charbit@tsi.enst.fr

**Résumé** – Dans le codeur MPEG AAC (Advanced Audio Coder), la fonction débit-distorsion dans chaque sous-bande fréquentielle est accessible a priori sous forme paramétrique : Le paramètre de quantification par sous-bande, appelé facteur d'échelle, détermine d'une part la puissance de l'erreur de codage, d'autre part le nombre de bits consommé. Dans cet article, nous proposons une modélisation de l'erreur de codage, paramétrée par le facteur d'échelle, suffisamment simple pour pouvoir être inversée. On se donne une puissance d'erreur par sous-bande, et on est capable de calculer directement la valeur du facteur d'échelle correspondant. La relation liant la distorsion au débit est donc connue sous forme directe.

Ensuite, nous proposons une application à l'algorithme de codage à débit fixe. Contrairement à l'algorithme standard décrit dans l'annexe informative de la norme MPEG-2, il est possible de supprimer la boucle de contrôle de distorsion. Ce nouvel algorithme s'avère donc plus rapide et plus souple vis à vis des critères psychoacoustiques que l'algorithme standard.

**Abstract** – In a MPEG Advanced Audio Coder (AAC), the quantization parameter for each frequency sub-band, called scale-factor, controls the quantization error and the bit rate. Thus, we seem to have only a parametric representation of the rate-distortion function per sub-band. In this paper, we show that an accurate model for the quantization error, depending on the scale-factor value, can lead to a simple expression of the relation between the scale-factor and the error power over the current sub-band. As the reverse relation can be written, we have a direct expression of the distortion-rate function.

Then, we apply these results to the coding algorithm for a fixed output bit-rate. Unlike the standard algorithm described in the informative annex of the MPEG-2 standard, the distortion loop can be removed. This new algorithm happen to be faster and more flexible towards psychoacoustic criterions than the standard one.

## 1 Introduction

Le MPEG AAC (Advanced Audio Coder) est aujourd'hui considéré comme un des standards de compression audio les plus efficaces. Il a été spécifié dans la norme MPEG-2 audio en 1997 [1, 2] puis a été repris et légèrement modifié dans la norme MPEG-4. Il s'agit d'un codeur par transformée spectrale qui partage avec le standard MPEG-1 couche III (dit MP3) le principe de quantification par facteur d'échelle : La compression fait appel à un seul quantificateur scalaire uniforme, l'arrondi à l'entier le plus proche combiné à une non-linéarité, mais comprime ou dilate auparavant le signal au moyen d'un facteur multiplicatif appelé facteur d'échelle. La norme MPEG AAC définit un facteur d'échelle par sous-bande fréquentielle.

Dans chaque sous-bande, le choix de ce paramètre fixe l'erreur de quantification ainsi que le nombre de bits consommé. La fonction débit-distorsion par sous-bande est donc connue a priori sous forme paramétrique, et non sous forme directe. Cette caractéristique a dicté l'architecture de l'algorithme de codage standard, proposé dans l'annexe informative de la norme [1] : Pour choisir les meilleures valeurs des facteurs d'échelle sous contrainte de débit fixe, et si possible en respectant la contrainte de masquage psy-

choacoustique, on fait appel à un algorithme itératif où une boucle de contrôle de la distorsion est imbriquée dans une boucle de contrôle du débit.

Dans cet article, après avoir décrit la quantification et le codage des coefficients transformés, nous proposons un modèle d'erreur de quantification suffisamment simple pour permettre l'inversion de la relation liant le facteur d'échelle à la puissance de l'erreur par sous-bande. Puis nous proposons une application à l'algorithme d'optimisation : La connaissance de la relation distorsion-débit sous forme directe permet de supprimer une boucle et d'obtenir un algorithme particulièrement efficace.

## 2 Quantification et codage dans le codeur MPEG-2 AAC

Le schéma fonctionnel d'un codeur AAC est représenté en figure 1. Le signal est transformé dans un domaine spectral au moyen d'une transformée en cosinus appelée MDCT (Modified Discrete Cosine Transform). On note  $X(k)$  les coefficients transformés (réels) dans la sous-bande  $s$ . Le signal d'entrée est compris entre -32768 et +32767.

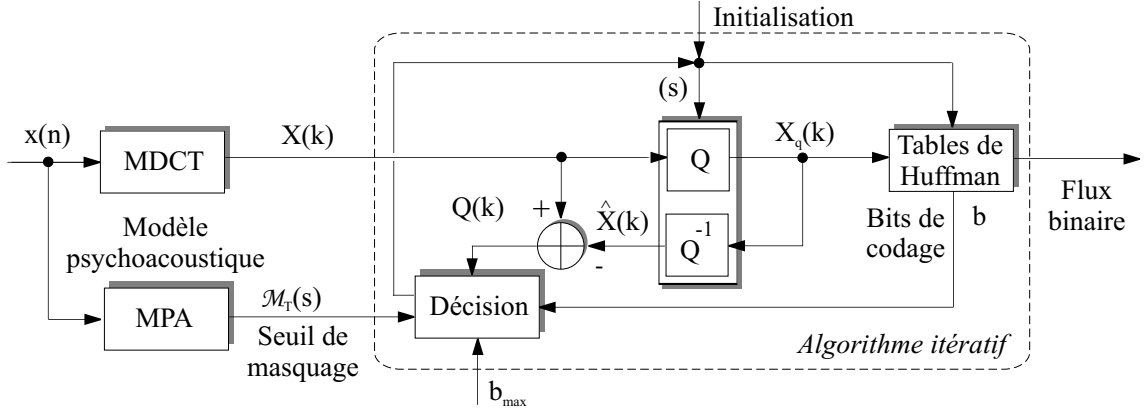


FIG. 1 – Schéma de fonctionnement d'un codeur AAC classique.

Le module des coefficients spectraux est donc généralement grand devant 1. L'opération de quantification, décrite par l'équation (1), est paramétrée par le facteur d'échelle  $A(s)$ .

$$X_q(k) = \text{Round} \left[ |X(k)|^{\frac{3}{4}} A(s) \right] \quad (1)$$

Les valeurs du facteur d'échelle dépendent d'un paramètre  $\gamma(s)$  entier :  $A(s) = 2^{\frac{3}{16}\gamma(s)}$ . L'opération de quantification inverse, décrite par l'équation (2), permet d'obtenir les coefficients spectraux reconstruits.

$$\hat{X}(k) = \text{Signe} [X(k)] (X_q(k) A(s))^{\frac{4}{3}} \quad (2)$$

On note  $Q(k)$  l'erreur de quantification dans le domaine transformé. Sa puissance sur la sous-bande  $s$  s'écrit :

$$\begin{aligned} \mathcal{P}_Q(s) &= \sum_{k=k_{\min}(s)}^{k_{\max}(s)} Q(k)^2 = \\ &= \sum_{k=k_{\min}(s)}^{k_{\max}(s)} \left( \left( \text{Round} \left[ \frac{|X(k)|^{\frac{3}{4}}}{A(s)} \right] A(s) \right)^{\frac{4}{3}} - |X(k)| \right)^2 \end{aligned} \quad (3)$$

Ensuite, les coefficients  $X_q(k)$  et  $\gamma(s)$  sont exprimés en binaire au moyen de tables de Huffman. Dans la sous-bande  $s$ , pour un signal  $X(k)$  donné, le choix de  $A(s)$  fixe donc la puissance de l'erreur et le nombre de bits consommé. Mais l'expression (3) reste complexe, et n'est pas aisément inversible. Nous allons donc chercher une expression approchée plus simple.

### 3 Modélisation de l'erreur de quantification

On constate que  $\mathcal{P}_Q(s)$  est une fonction globalement croissante de  $A(s)$ , même si elle peut être localement non-monotone. En codage, la contrainte de qualité perçue s'exprime sous forme d'un seuil sur  $\mathcal{P}_Q(s)$ , noté  $T(s)$ , qui se traduit donc par un seuil sur  $A(s)$ . Comme la consommation de bits est une fonction globalement décroissante de  $A(s)$ , la solution qui respecte la contrainte de puissance maximale et qui minimise la consommation de bits est celle qui produit une puissance d'erreur la plus proche du seuil par valeur inférieure. On pose donc le problème ainsi : Etant donné un jeu de coefficients spectraux  $X(k)$ , on cherche une borne supérieure sur  $A(s)$  qui approche au mieux  $\mathcal{P}_Q$  de  $T(s)$  par valeur inférieure.

#### 3.1 Approche déterministe

On note  $\varepsilon(k)$  l'erreur d'arrondi dans l'équation (1). Il vient :

$$\hat{X}(k) = \text{Signe} [X(k)] \left( |X(k)|^{\frac{3}{4}} + A(s)\varepsilon(k) \right)^{\frac{4}{3}} \quad (4)$$

En hypothèse haute-résolution, lorsque  $|X(k)|^{\frac{3}{4}} \gg A(s)\varepsilon(k)$ , on peut développer au premier ordre :

$$\hat{X}(k) \approx \text{Signe} [X(k)] \left( |X(k)| + \frac{4}{3} A(s)\varepsilon(k) |X(k)|^{\frac{1}{4}} \right) \quad (5)$$

D'où :

$$Q(k)^2 \approx \frac{16}{9} A(s)^2 |\varepsilon(k)|^2 |X(k)|^{\frac{1}{2}} \quad (6)$$

On sait que  $|\varepsilon(k)| \leq \frac{1}{2}$ , donc la puissance de l'erreur sur l'ensemble de la sous-bande peut être majorée :

$$\mathcal{P}_Q(s) \leq \frac{4}{9} A(s)^2 \sum_{k=k_{\min}(s)}^{k_{\max}(s)} |X(k)|^{\frac{1}{2}} \quad (7)$$

On en déduit donc une borne sur  $A(s)$  qui assure  $\mathcal{P}_Q(s) \leq T(s)$  :

$$A(s)^2 \leq \frac{9}{4} \frac{T}{\sum_{k=k_{\min}(s)}^{k_{\max}(s)} |X(k)|^{\frac{1}{2}}} \quad (8)$$

Cette borne est atteignable lorsque  $|\varepsilon(k)| = \frac{1}{2}$  pour toutes les valeurs de  $k$ . Comme cette situation est très rare lorsqu'on traite des signaux de musique, il est préférable d'adopter une approche probabiliste.

#### 3.2 Approche probabiliste

Dans cette section, on supprime l'indice  $s$ . On modélise les  $X(k)$  comme une suite de variables aléatoires. L'équation (3) définit donc une transformation déterministe, paramétrée par  $A$ .  $\mathcal{P}_Q$  est donc aussi une variable aléatoire qu'on autorise à dépasser  $T$  mais en contrôlant le pourcentage de temps où ce dépassement se produit, c'est à dire en fixant la probabilité de dépassement du seuil. La contrainte de qualité s'écrit alors :

$$\text{Proba} \{ \mathcal{P}_Q \leq T \} \geq \alpha \quad (9)$$

Où  $\alpha$  est un niveau de confiance compris entre 0 et 1. Bien qu'il soit possible d'écrire la loi de probabilité de  $\mathcal{P}_Q$  en

fonction de celles des  $X(k)$ , cette formulation s'avère inutilisable. Il est toutefois possible de simplifier l'approche : On suppose que les variables aléatoires  $Q(k)$  sont indépendantes et identiquement distribuées (iid). Ce qui signifie que l'erreur est supposée blanche par sous-bande. Si on connaît les deux premiers moments de la variable  $Q^2(k)$ , on peut appliquer le théorème de la limite centrale [3] :

$$\sqrt{K} \left( \frac{1}{K} \mathcal{P}_Q - m_{Q^2} \right) \xrightarrow{K \rightarrow \infty} \mathcal{N} (0, \sigma_{Q^2}^2) \quad (10)$$

où on note  $K = k_{max} - k_{min} + 1$ . On a donc convergence vers une loi normale centrée. On en déduit, pour tout  $M$  :

$$\text{Proba} \left( \sqrt{K} \left( \frac{1}{K} \mathcal{P}_Q - m_{Q^2} \right) \leq M \right) \xrightarrow{K \rightarrow \infty} \frac{1}{2} \left[ 1 + \text{erf} \left( \frac{M}{\sqrt{2} \sigma_{Q^2}} \right) \right] \quad (11)$$

Donc, lorsque  $K$  est suffisamment grand, on a :

$$\text{Proba} (\mathcal{P}_Q \leq T) \simeq \frac{1}{2} \left[ 1 + \text{erf} \left( \frac{T - K m_{Q^2}}{\sqrt{2K} \sigma_{Q^2}} \right) \right] \quad (12)$$

Pour déterminer les moments de  $Q^2(k)$ , nous avons adopté une approche expérimentale : En supposant que les  $X(k)$  sont iid gaussiens centrés, et que le quantificateur fonctionne avec suffisamment de niveaux (hypothèse haute-résolution), on mesure asymptotiquement :

$$\begin{cases} m_{Q^2} &= c_1 A^2 \sqrt{\sigma_X} \\ \sigma_{Q^2}^2 &= c_2^2 A^4 \sigma_X \end{cases} \quad (13)$$

avec  $c_1 \simeq 0.12$  et  $c_2 \simeq 0.13$ . En supposant que ces expressions restent valables pour les signaux de musique, on obtient finalement la borne :

$$\text{Proba} (\mathcal{P}_Q \leq T) \geq \alpha \Leftrightarrow A^2 \leq C(K, \alpha) \frac{T}{\sqrt{\sigma_X}} \quad (14)$$

où la constante s'écrit :

$$C(K, \alpha) = \frac{1}{\sqrt{2K} c_2 \text{erfinv}(2\alpha - 1) + K c_1} \quad (15)$$

Pour estimer  $\sigma_X$  dans chaque sous-bande, on utilise la variance empirique et on obtient l'expression suivante, équivalente à la condition (9) :

$$A^2 \leq \frac{C(K, \alpha) T}{\left( \frac{1}{K} \sum_{k=k_{min}}^{k_{max}} X(k)^2 \right)^{\frac{1}{4}}} \quad (16)$$

Cette expression est d'autant plus juste que  $K$  est grand. En pratique,  $K$  est compris entre 4 et 32. Le modèle est donc meilleur dans les sous-bandes en hautes fréquences. De plus, cette condition n'a de sens que tant que le seuil d'erreur visé  $T$  est inférieur à la puissance du signal dans la sous-bande :  $\mathcal{P}_X = \sum_{k=k_{min}}^{k_{max}} X(k)^2$ . Dans le cas contraire, nous utilisons un modèle dégénéré du fonctionnement du quantificateur, où  $Q(k) = X(k)$ .

### 3.3 Validation expérimentale

Partant du niveau de confiance  $\alpha$ , nous avons vérifié sur des signaux réels la validité de la borne décrite par l'équation 16 en mesurant le pourcentage du temps où le seuil

de puissance est respecté, appelé taux de respect du seuil. Pour diverses valeurs de  $\alpha$ , on indique dans le tableau 1 la valeur mesurée du taux, en moyenne sur toutes les sous-bandes. Dans le codeur AAC, on a pris pour valeur de  $T$  le seuil de masquage, l'algorithme de codage consistant simplement à choisir comme valeur du facteur d'échelle la borne définie par l'équation 16 (arrondie à la valeur possible la plus proche). On ne comptabilise pas les sous-bandes masquées ( $\mathcal{P}_X > T$ ). Le fichier audio de référence est un extrait de "Tom's Diner" par Suzanne Vega. L'estimation du taux a été effectuée sur 300 blocs, soit 15300 tests de masquage. On observe une bonne correspondance des valeurs.

$\alpha$	0.9	0.8	0.7	0.6	0.5
Taux	0.89	0.80	0.72	0.63	0.54

TAB. 1 – Niveau de confiance et taux moyen de respect du seuil de puissance.

Notre modélisation de la relation liant  $\mathcal{P}_Q(s)$  à  $A(s)$  fournit donc une expression analytique simple de la valeur de  $A(s)$  pour laquelle  $\mathcal{P}_Q(s)$  est le plus proche par valeur inférieure d'un seuil  $T(s)$  quelconque. Cette expression existe en deux versions : Une déterministe, qui garantit un strict respect du seuil, mais qui risque d'être très coûteuse en débit, les réglages étant effectués pour tenir compte du pire cas atteignable. Une autre probabiliste, qui autorise des dépassements contrôlés du seuil, mais qui correspond à des réglages plus réalistes. Comme on dispose aussi de la relation liant  $A(s)$  au nombre de bits consommé, on connaît la relation liant un seuil d'erreur quelconque au nombre de bits nécessaire, les facteurs d'échelle n'étant plus que des intermédiaires de calcul.

## 4 Application à l'algorithme d'allocation de bits

Nous proposons maintenant une application des résultats obtenus dans la section précédente. La procédure de codage selon la norme AAC consiste à choisir les valeurs des facteurs d'échelle. La contrainte principale concerne le nombre de bits consommé qui doit respecter le débit fixé. Ensuite, vient une contrainte de qualité. La psychoacoustique donne une condition d'inaudibilité de l'erreur de codage sous la forme d'un seuil de puissance d'erreur par sous-bande, appelé seuil de masquage et calculé par le modèle psychoacoustique. Généralement, la contrainte de débit ne permet pas de respecter ce seuil. Il faut donc mettre en ouvre un algorithme de gestion de la sous-optimalité.

La norme MPEG-2 propose, dans une annexe informative, un tel algorithme qui comporte deux boucles imbriquées : Une boucle de contrôle de distorsion, visant le respect du seuil de masquage, et une boucle de contrôle de débit, la solution finale consistant en un compromis.

Nous proposons alors une simplification de ce principe en ne conservant que la boucle de contrôle de débit : Connaissant le seuil de masquage  $\mathcal{M}_T(s)$  et les puissances du signal par sous-bandes  $\mathcal{P}_X(s)$ , chaque seuil de puissance d'erreur  $T(s)$  intermédiaire entre ceux derniers cor-

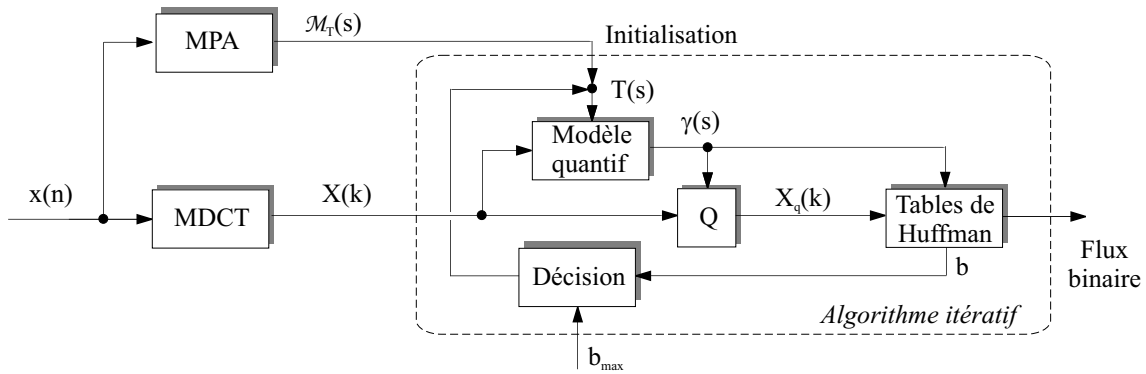


FIG. 2 – Schéma de fonctionnement du nouveau codeur AAC.

respondra à une solution de codage, donc un jeu de facteurs d'échelle et un nombre de bits consommé. On se donne une famille de seuils  $T(s)$  indicée par un paramètre réel  $\delta \in [0, 1]$  :

$$T(s) = \mathcal{P}_X(s)^\delta \mathcal{M}_T(s)^{1-\delta} \quad (17)$$

$\delta = 0$  suppose  $T = \mathcal{M}_T$  et donc une qualité maximale.  $\delta = 1$  suppose  $T = \mathcal{P}_X$  et donc une qualité minimale. On y ajoute une contrainte de distance maximale entre  $T(s)$  et  $\mathcal{M}_T(s)$  dans les basses fréquences. Expérimentalement, il s'est avéré qu'un écart maximal de 5 dB pour les sous bandes 1 à 10 en fenêtres longues, et 1 à 4 en fenêtres courtes, donne de bons résultats.

L'algorithme de codage fonctionne alors comme suit :

- Initialisation :  $\delta = 0$
- Boucle sur  $\delta$ 
  - Calcul de  $T(s)$
  - Calcul des valeurs-seuils des facteurs d'échelle selon l'équation (16)
  - Calcul du nombre de bits de codage nécessaires  $b$
  - Si  $b > b_{max}$ , augmenter  $\delta$

En effet, la relation liant  $\delta$  à  $b$  est globalement décroissante, et très localement non-monotone, comme le montre la figure 3. Un tel algorithme va donc permettre d'atteindre rapidement une solution satisfaisante. La structure du codeur AAC ainsi obtenu est représentée en figure 2.

L'avantage de cette architecture par rapport à celle de l'algorithme standard est d'une part sa plus grande rapidité : Nous avons mesuré une division du temps d'exécution moyen de l'algorithme d'un facteur 7, toutes choses égales par ailleurs dans le codeur. D'autre part, il est caractérisé par une grande flexibilité vis à vis des critères psychoacoustiques : Il peut s'adapter à d'éventuels modèles donnant, pour un niveau de qualité perçue donnée, la répartition spectrale idéale de l'erreur de quantification. Nous avons, à défaut, proposé un critère qui a essentiellement valeur d'exemple.

## 5 Conclusion

Dans cet article, nous présentons une modélisation de la puissance de l'erreur de quantification paramétrée par le facteur d'échelle dans le codeur MPEG AAC. Il s'agit d'une simplification importante de la relation exacte, mais

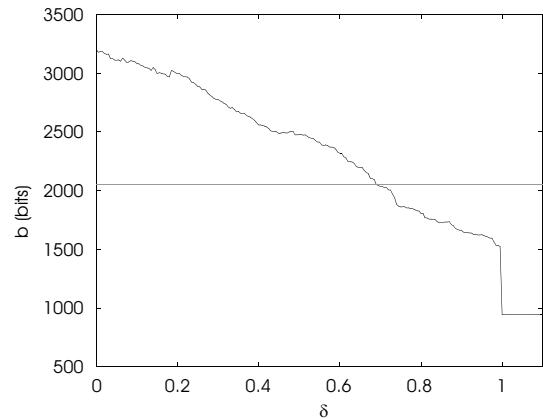


FIG. 3 – Exemple de relation liant  $\delta$  à  $b$  sur une fenêtre de 1024 coefficients spectraux. Le trait horizontal représente le nombre de bits disponible (2048 bits, pour un débit de 64 kbps).

qui reste fiable, comme le montrent les expériences. La simplicité de cette relation permet son inversion. Il est alors possible de connaître la fonction distorsion-débit du codeur sous-bande par sous-bande, sous forme directe. Nous avons ensuite appliqué cet outil à l'algorithme d'allocation de bits du codeur. Cette nouvelle méthode d'optimisation du codage est plus simple que celle proposée dans le document normatif. Sa convergence est assurée, et la charge de calcul est fortement réduite.

Cette étude pourra être prolongée en cherchant des critères de sous-optimalité issus de la psychoacoustique donnant les meilleurs résultats perceptuels. En effet, nous disposons maintenant des outils de réglage du codeur permettant de contrôler efficacement le niveau de bruit par sous-bande.

## Références

- [1] International Organization for Standardization, *ISO/IEC 13818-7 (MPEG-2 Advanced Audio Coding, AAC)*, 1997.
- [2] O. Derrien, S. Larbi, M. Perreau-Guimares, and N. Moreau, "Le codeur MPEG-2 AAC expliqué aux traités de signaux," *Annales des Télécommunications*, 2000.
- [3] A. Borovkov, *Statistique mathématique*. MIR, 1997.