

Une méthode d'allocation de débit basse-résolution pour des données parcimonieuses

Aurélia FRAYSSE¹, Béatrice PESQUET-POPESCU², Jean-Christophe PESQUET³

¹Laboratoire des Signaux et Systèmes
CNRS, Université Paris Sud, Supélec, 3 rue Joliot Curie, 91192 Gif Sur Yvette, France

²Laboratoire Traitement des Signaux et Images
Telecom ParisTech, 46 Rue Barrault, 75634 Paris Cedex 13, France

³Laboratoire d'Informatique Gaspard Monge
Université Paris-Est, 77454 Champs sur Marne, France
fraysse@lss.supelec.fr, beatrice.pesquet@telecom-paristech.fr
jean-christophe.pesquet@univ-paris-est.fr

Résumé – Cet article présente une méthode d'allocation de débit pour le codage de signaux ou d'images basée sur un critère d'optimalité débit-distorsion. Notre objectif est de construire un algorithme d'optimisation par sous-bandes comme celui utilisé dans des normes telles que JPEG2000 mais en adoptant une approche analytique. Pour cela, nous nous plaçons dans le cadre d'une quantification scalaire uniforme de données dont les coefficients de décomposition en sous-bandes sont modélisés par des distributions de probabilité de type mélanges Bernoulli-Gaussiens Généralisés (BGG). Ce modèle se révèle particulièrement bien adapté pour des données dont la représentation est parcimonieuse dans le domaine ondelettes. Grâce à des approximations récentes de l'entropie et de la distorsion pour de tels modèles, nous proposons un algorithme efficace d'optimisation débit-distorsion, permettant de calculer les pas de quantification optimaux.

Abstract – This paper deals with a bit allocation method for image or signal coding based on a rate-distortion optimality criterion. Our objective is to design a subband optimization algorithm such as that used in the JPEG2000 standard, but by adopting an analytical approach. For doing this, we consider the context of a uniform scalar quantization of the data whose subband decomposition coefficients are modelled by Bernoulli-Generalized Gaussian (BGG) mixture probability distributions. This model appears to be particularly well-adapted for data with sparse representation in the wavelet domain. Thanks to recent approximations of the entropy and the distortion for such models, we propose an efficient rate-distortion optimization algorithm to calculate the optimal quantization steps.

1 Introduction

Le but de cet article est de proposer une méthode d'allocation de débit optimale dans le cadre d'une compression avec pertes de données issues d'une analyse en sous-bandes. Dans ce contexte, les méthodes analytiques classiques (où la solution prend une forme explicite) s'appuient sur une optimisation à haute résolution de critères débit-distorsion, en utilisant généralement une quantification scalaire uniforme. Bien que sous-optimale, cette quantification reste largement utilisée de par sa simplicité. L'inconvénient majeur de cette approche est qu'elle repose sur des approximations haut débit et est, par conséquent, limitée dans son champ d'applications. Diverses méthodes sont apparues pour passer outre cette restriction. Ces méthodes sont principalement de deux types. On trouve des méthodes analytiques, telle que celles dans [1], qui proposent des critères basés sur l'entropie diffé-

rentielle. D'autres méthodes sont purement numériques, en liaison notamment avec les résultats exposés dans [8], et s'appuient sur des techniques itératives. Ainsi, dans [8] est effectuée une recherche itérative des points critiques du Lagrangien. Cette méthode a été ensuite étendue dans [9] à une analyse en sous-bandes. Les méthodes de [6, 5] proposent aussi de trouver la meilleure représentation des images en paquets d'ondelettes et de coder les jeux de coefficients grâce à des critères débit-distorsion.

Dans le cas présent, nous proposons un algorithme applicable à des signaux présentant une représentation creuse lors d'une analyse en sous-bandes (par exemple, une décomposition sur une base d'ondelettes). Cette méthode repose sur une optimisation débit-distorsion permettant d'obtenir une forme analytique du pas de quantification optimal dans chaque sous-bande, ainsi que des distorsions correspondantes. Nous nous proposons pour cela d'utiliser

une approximation de l'entropie, plus fine que celle donnée par la formule de Bennett [3], grâce à laquelle il nous est possible de sortir du cadre haut débit, tout en obtenant une solution quasi explicite du problème d'allocation.

2 Modèle utilisé

Dans le cadre de la représentation par ondelettes d'une image, des distributions de probabilité Gaussiennes Généralisées (GG) sont classiquement utilisées pour modéliser les coefficients de chaque sous-bande [4]. En partant de ce modèle statistique, nous nous proposons de considérer des signaux ayant une représentation parcimonieuse. C'est pourquoi dans la suite, nous supposons que si le signal est décomposé sur J sous-bandes, les distributions des coefficients dans chaque sous-bande peuvent être caractérisées par un modèle plus général, de type mélange Bernoulli-Gaussien Généralisé (BGG). Pour chaque sous-bande $j \in \{1, \dots, J\}$, il y a donc n_j (tels que $\sum_{j=1}^J n_j = n$) coefficients dont la loi est donnée par :

$$\forall \xi \in \mathbb{R}, \quad g_j(\xi) = (1 - \epsilon_j)\delta(\xi) + \epsilon_j f_j(\xi) \quad (1)$$

où $\epsilon_j \in [0, 1]$ est un paramètre de mélange, δ est la masse de Dirac et f_j est une densité de probabilité GG d'exposant β_j et de facteur d'échelle ω_j :

$$\forall \xi \in \mathbb{R}, \quad f_j(\xi) = \frac{\beta\omega^{1/\beta_j}}{2\Gamma(1/\beta_j)} e^{-\omega|\xi|^{\beta_j}}. \quad (2)$$

Remarquons que nous nous restreignons dans cette classe aux fonctions log-concaves, telles que $\beta_j \in [1, 2]$ pour $j \in \{1, \dots, J\}$. Les données ainsi modélisées sont quantifiées à l'aide d'un quantificateur scalaire uniforme de pas de quantification $q_j > 0$. Si X_j désigne la variable en entrée du quantificateur, la variable obtenue en sortie \bar{X}_j vaut alors 0 si $|X_j| < \frac{q_j}{2}$ et est égale à r_i dès que $(i - \frac{1}{2})q_j \leq X_j < (i + \frac{1}{2})q_j$ avec

$$\forall i \geq 1, \quad r_i = -r_{-i} = (i + \zeta_j)q_j \quad (3)$$

où $\zeta_j \in]-\frac{1}{2}, \frac{1}{2}[$. La condition classique de reconstruction centrée correspondrait ici à $\zeta_j = 0$. Notons que, dans nos travaux récents, il a été démontré que la quantification uniforme permettait d'atteindre asymptotiquement les bornes de Shannon à bas débit, sous certaines hypothèses [2].

Dans cet article, nous quantifions le débit à l'aide de l'entropie d'ordre 0. De plus, nous autorisons des fonctions de distorsions différentes dans chaque sous-bande, correspondant à des normes L^{p_j} de l'erreur de quantification, où $p_j \geq 1$ peut varier en fonction de j . Ceci pourrait permettre, par exemple, la prise en compte de critères plus proches du système psycho-visuel humain.

3 Principe de l'algorithme proposé

La méthode analytique classique d'allocation de débit utilise la formule de Bennett pour approcher l'entropie. Cette formule exprime l'entropie à haute-résolution comme une application affine de \mathbf{l} , où $\mathbf{l} = (l_1, \dots, l_J)$, avec $\forall j \in \{1, \dots, J\}$, $l_j = \log_2(q_j)$. Plus précisément, l'entropie de chaque coefficient de la j -ème sous-bande vaut

$$H_{g_j}(l_j) = H_{\epsilon_j} + \epsilon_j(h_{\beta_j}(\omega_j) - l_j) + O(l_j 2^{l_j}), \quad (4)$$

où $H_{\epsilon_j} = -\epsilon_j \log_2 \epsilon_j - (1 - \epsilon_j) \log_2(1 - \epsilon_j)$ est l'entropie d'une variable aléatoire de Bernoulli de paramètre ϵ_j et $h_{\beta_j}(\omega_j)$ est l'entropie différentielle de la loi GG. Cette formule correspond néanmoins à un développement asymptotique, pour des pas de quantification tendant vers zéro. Son utilisation n'est donc parfaitement valide que dans des conditions haut débit. Dans notre étude, nous remplaçons cette approximation affine par une fonction de cette même variable \mathbf{l} , donnée par $n^{-1} \sum_{j=1}^J n_j g_j(l_j)$ où, pour tout $j \in \{1, \dots, J\}$ et $l_j < l_j^2$,

$$g_j(l_j) = \begin{cases} a_j^1(l_j - l_j^1) + a_j^2(l_j^1 - l_j^2) & \text{si } l_j \leq l_j^1 \\ a_j^2(l_j - l_j^2) & \text{si } l_j > l_j^1, \end{cases} \quad (5)$$

avec $l_j^1 < l_j^2$. En prenant a_j^1 et a_j^2 négatifs, la fonction g_j est continue et décroissante. Pour s'assurer de la positivité de l'entropie, on pose aussi $g_j(l_j) = 0$ pour tout $l_j \geq l_j^2$.

Pour $l_j^1 = -\infty$, $a_j^1 = 0$, $a_j^2 = -\epsilon_j$ et $\epsilon_j l_j^2 = H_{\epsilon_j} + \epsilon_j h_{\beta_j}(\omega_j)$, on se ramène à la formule de Bennett. Le fait de remplacer l'approximation haute résolution dans chaque sous-bande par une forme affine par morceaux offre cependant plus de flexibilité en termes d'approximation et permet ainsi d'avoir des résultats plus concluants pour de plus grandes valeurs des l_j . Par exemple, pour obtenir une approximation plus fine de l'entropie, on peut prendre :

$$a_j^1 = -\epsilon_j \quad (6)$$

$$a_j^1 l_j^1 + a_j^2 l_j^2 = -H_{\epsilon_j} - \epsilon_j h_{\beta_j}(\omega_j) \quad (7)$$

et, fixer (l_j^2, a_j^2) de façon à ce que le graphe de g_j soit tangent à celui de l'approximation de l'entropie donnée dans [2] en un point quelconque $l_j > l_j^1$. De façon similaire, nous considérons l'approximation suivante de la distorsion d'ordre $p_j \geq 1$ pour la j -ème sous-bande :

$$d_j(l_j) = \begin{cases} \frac{\epsilon_j \eta_j (p_j + 1)}{\nu_j} (\alpha_j^1 2^{l_j} \gamma_j^1 + \delta_j^1) & \text{si } l_j \leq l_j^1 \\ \frac{\epsilon_j \eta_j (p_j + 1)}{\nu_j} (\alpha_j^2 2^{l_j} \gamma_j^2 + \delta_j^2) & \text{si } l_j > l_j^1, \end{cases} \quad (8)$$

où $\nu_j = \left(\frac{1}{2} + \zeta_j\right)^{p_j+1} + \left(\frac{1}{2} - \zeta_j\right)^{p_j+1}$, $\eta_j > 0$ est un coefficient de pondération éventuelle et $\gamma_j^1 > 0$, $\gamma_j^2 > 0$, $\alpha_j^1 > 0$, $\alpha_j^2 > 0$, δ_j^1 et δ_j^2 sont des paramètres d'ajustements. En particulier, en prenant $\gamma_j^1 = p_j$, $\eta_j = \alpha_j^1 = \nu_j / (p_j + 1)$ et $\delta_j^1 = 0$, l'expression obtenue est cohérente avec l'approximation basse résolution $d_j(l_j)(1 + O(2^{l_j}))$ de

$E[|X_j - \bar{X}_j|^{p_j}]$. Pour assurer la continuité de la fonction d_j , on impose alors $\delta_j^2 = \alpha_j^1 2^{l_j p_j} - \alpha_j^2 2^{l_j \gamma_j^2}$.

La distorsion globale D est alors donnée par la moyenne des distorsions d_j . Grâce aux approximations précédentes illustrées par la figure 1, le problème d'allocation de débit se formule comme un problème d'optimisation, qui est de minimiser la distorsion globale sous la contrainte que l'entropie globale reste inférieure ou égale au débit total R_{\max} (en bpp).

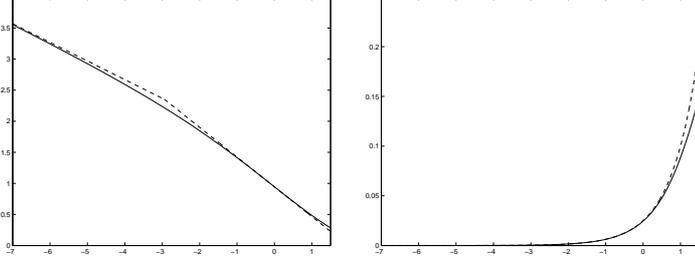


FIG. 1 – Approximations (en tirets) de l'entropie (à gauche) et de la distorsion (à droite) dans une sous-bande en fonction de l_j pour $\epsilon_j = 0.3$, $\beta_j = 4/3$, $\omega_j = 1$ et $p_j = 2$.

Problème 1 Trouver $\tilde{\mathbf{l}}$ minimisant

$$\mathbf{l} = (l_1, \dots, l_J) \in \mathbb{R}^J \mapsto D(\mathbf{l}) = \sum_{j=1}^J \frac{n_j}{n} d_j(l_j)$$

sur l'ensemble C défini par

$$C := \{\mathbf{l} = (l_1, \dots, l_J) \in \mathbb{R}^J \mid \sum_{j=1}^J \frac{n_j}{n} g_j(l_j) \leq R_{\max}\}. \quad (9)$$

La principale difficulté dans ce problème réside dans la non-différentiabilité des différentes fonctions g_j . Notons $l_j^0 = \max(l_j^1, \frac{\alpha_j^1 l_j^1 + \alpha_j^2 l_j^2 + n n_j^{-1} R_{\max}}{\alpha_j^1})$, pour tout $j \in \{1, \dots, J\}$. On constate alors qu'il suffit de rechercher le minimiseur de la distorsion sur le pavé $[l_1^0, l_1^2] \times \dots \times [l_J^0, l_J^2]$.

Pour lever le problème de la non-différentiabilité des g_j en l_j^1 , nous découpons ce pavé en sous-pavés de la forme $[l_1^{b_1}, l_1^{b_1+1}] \times \dots \times [l_J^{b_J}, l_J^{b_J+1}]$, pour $\mathbf{b} = (b_1, \dots, b_J) \in \{0, 1\}^J$. Sur chacun de ces sous-pavés, l'entropie et la distorsion globales sont des fonctions convexes. Nous obtenons ainsi 2^J domaines sur lesquels nous sommes ramenés à des problèmes d'optimisation convexe.

4 Solution explicite

Considérons un pavé $\mathbf{P}_{\mathbf{b}} := [l_1^{b_1}, l_1^{b_1+1}] \times \dots \times [l_J^{b_J}, l_J^{b_J+1}]$ donné, et notons $(\mathcal{P}_{\mathbf{b}})$ le problème de minimisation convexe sur ce pavé.

Pour simplifier les expressions, introduisons les notations suivantes. Pour $j \in \{1, \dots, J\}$, on note :

$$a_j = a_j^{b_j}, \gamma_j = \gamma_j^{b_j}, \alpha_j = \alpha_j^{b_j} \quad (10)$$

$$c_j = \begin{cases} -a_j^1 l_j^1 - a_j^2 l_j^2 & \text{si } b_j = 0 \\ -a_j^2 l_j^2 & \text{si } b_j = 1 \end{cases} \quad (11)$$

$$N_j = -\frac{n_j a_j}{\gamma_j} \quad (12)$$

$$\underline{\lambda}_j = \kappa_j 2^{\gamma_j l_j^{b_j}} \quad \text{et} \quad \bar{\lambda}_j = \kappa_j 2^{\gamma_j l_j^{b_j+1}}, \quad (13)$$

où $\kappa_j = -\epsilon_j \eta_j \alpha_j \gamma_j \ln 2 / (a_j \alpha_j^1)$.

La solution du problème $(\mathcal{P}_{\mathbf{b}})$ est alors donnée par :

- Proposition 1**
1. Si $\sum_{j=1}^J \frac{n_j}{n} g_j(l_j^{b_j+1}) > R_{\max}$, alors $(\mathcal{P}_{\mathbf{b}})$ n'admet pas de solution.
 2. Si $\sum_{j=1}^J \frac{n_j}{n} g_j(l_j^{b_j}) < R_{\max}$ alors la solution de $(\mathcal{P}_{\mathbf{b}})$ est donnée par $\tilde{\mathbf{l}}_{\mathbf{b}} = (l_1^{b_1}, \dots, l_J^{b_J})$.
 3. Sinon, il existe trois sous-ensembles \mathbb{I} , \mathbb{J} et \mathbb{K} de $\{1, \dots, J\}$ et un unique vecteur solution $\tilde{\mathbf{l}}_{\mathbf{b}}$ dont les composantes sont données par : $\forall j \in \{1, \dots, J\}$,

$$\tilde{l}_{j,\mathbf{b}} = \begin{cases} l_j^{b_j} & \text{si } j \in \mathbb{I} \\ \frac{1}{\gamma_j} \log_2 \left(\frac{\bar{\lambda}_j}{\kappa_j} \right) & \text{si } j \in \mathbb{J} \\ l_j^{b_j+1} & \text{si } j \in \mathbb{K} \end{cases} \quad (14)$$

où $\tilde{\lambda} \in]\underline{\lambda}_j, \bar{\lambda}_j[$ si $j \in \mathbb{J}$ et dépend de R_{\max} , de \mathbf{b} ainsi que de \mathbb{I} , \mathbb{J} , \mathbb{K} .

Preuve : Les deux premiers cas se déduisent trivialement des propriétés de croissance des fonctions d_j et de décroissance des fonctions g_j , pour tout $j \in \{1, \dots, J\}$.

Supposons que les hypothèses des points 1 et 2 ne soient pas satisfaites. Dans ce cas, le minimum est obtenu par la méthode de dualité de Fenchel (voir [7]).

Dans le cadre de la proposition 1, introduisons g la fonction indicatrice du convexe fermé

$$C_{\mathbf{b}} := \{\mathbf{l} = (l_1, \dots, l_J) \in \mathbb{R}^J \mid \sum_{j=1}^J \frac{n_j}{n} (a_j l_j + c_j) \leq R_{\max}\} \quad (15)$$

et la fonction définie pour tout $\mathbf{l} \in \mathbb{R}^J$ par $f(\mathbf{l}) = D(\mathbf{l}) + \iota_{\mathbf{P}_{\mathbf{b}}}(\mathbf{l})$, où $\iota_{\mathbf{P}_{\mathbf{b}}}$ est la fonction indicatrice du pavé $\mathbf{P}_{\mathbf{b}}$. En calculant les conjuguées de ces fonctions, le problème dual de $(\mathcal{P}_{\mathbf{b}})$ se ramène à une maximisation de la fonction

$$\forall \lambda \in \mathbb{R}_+ \quad \Phi(\lambda) = \lambda \left(\sum_{j=1}^J \frac{n_j}{n} c_j - R_{\max} \right) - \sum_{j=1}^J \varphi_j(\lambda), \quad (16)$$

où $\forall j \in \{1, \dots, J\}$ et $\forall \lambda \in \mathbb{R}_+$,

$$\varphi_j(\lambda) = \begin{cases} \frac{N_j}{n} (\gamma_j l_j^{b_j} \lambda - \frac{\lambda}{\ln 2}) & \text{si } \lambda \leq \underline{\lambda}_j \\ \frac{N_j \lambda}{n \ln 2} \left(\ln \left(\frac{\lambda}{\kappa_j} \right) - 1 \right) & \text{si } \underline{\lambda}_j < \lambda < \bar{\lambda}_j \\ \frac{N_j}{n} (\gamma_j l_j^{b_j+1} \lambda - \frac{\bar{\lambda}_j}{\ln 2}) & \text{si } \lambda \geq \bar{\lambda}_j. \end{cases} \quad (17)$$

La fonction Φ est concave sur \mathbb{R}^+ , et admet un unique maximum en $\tilde{\lambda}$ tel que $\Phi'(\tilde{\lambda}) = 0$. Notons

$$\mathbb{I} = \{j \in \{1, \dots, J\} \mid \Phi'(\underline{\lambda}_j) \leq 0\} \quad (18)$$

$$\mathbb{K} = \{j \in \{1, \dots, J\} \mid \Phi'(\bar{\lambda}_j) > 0\} \quad (19)$$

$$\mathbb{J} = \{1, \dots, J\} \setminus (\mathbb{I} \cup \mathbb{K}). \quad (20)$$

La solution $\tilde{\lambda}$ est alors donnée par

$$\tilde{\lambda}^{N_{\mathbb{J}}} = 2^{-nR_{\max} + \sum_{j=1}^J n_j c_j - \sum_{j \in \mathbb{I}} N_j \gamma_j l_j^{b_j} - \sum_{j \in \mathbb{K}} N_j \gamma_j l_j^{b_j+1}} \prod_{j \in \mathbb{J}} \kappa_j^{N_j} \quad (21)$$

où $N_{\mathbb{J}} = \sum_{j \in \mathbb{J}} N_j$. Remarquons que si $j \in \mathbb{I}$ ou $j \in \mathbb{K}$, la valeur trouvée dans (21) n'appartient pas à l'intervalle $]\underline{\lambda}_j, \bar{\lambda}_j[$.

A partir de cette valeur $\tilde{\lambda}$, on retrouve les valeurs optimales de \mathbf{l} en considérant les points critiques de f , ce qui nous donne comme résultat (14). \square

Cette forme des pas optimaux sur un sous-pavé donné nous permet de déduire un nombre fini de valeurs candidates pour la distorsion minimale. Notons que ces valeurs peuvent être calculées indépendamment les unes des autres. Une fois ce calcul effectué, il ne nous reste qu'à déterminer sur quel sous-pavé la distorsion minimale est effectivement atteinte. Nous avons alors au plus 2^J candidats possibles. Les hypothèses des points 1 et 2 permettent de réduire le nombre de tests à effectuer.

5 Exemple illustratif

Appliquons la méthode proposée à la compression d'une image satellitaire de Marseille, décomposée sur une base d'ondelettes orthonormales (symlets de longueur 8 sur 3 niveaux de résolution). Les distributions des coefficients de chaque sous-bandes ont été modélisées à l'aide de lois BGG en utilisant une méthode de maximum de vraisemblance. La courbe distorsion-débit (avec $p_j = 2$ pour tout j) prédite par la proposition 1 sont indiquées sur la figure 2 (tracé noir) et comparée à celle obtenue par quantification scalaire de l'image avec les pas optimaux que nous avons calculés (tracé en bleu). On constate que ces courbes sont en assez bonne adéquation et que l'approche proposée fournit des résultats cohérents à bas débit. Nous n'avons pas représenté les résultats obtenus à un débit supérieur à 2 bpp car la différence entre les deux courbes devient alors imperceptible. Notons, en revanche, que pour un débit en dessous de 1 bpp, la formule d'allocation classique basée sur les seules approximation haute résolution fournit des résultats peu exploitables, même si l'on y intègre la contrainte de positivité de l'entropie.

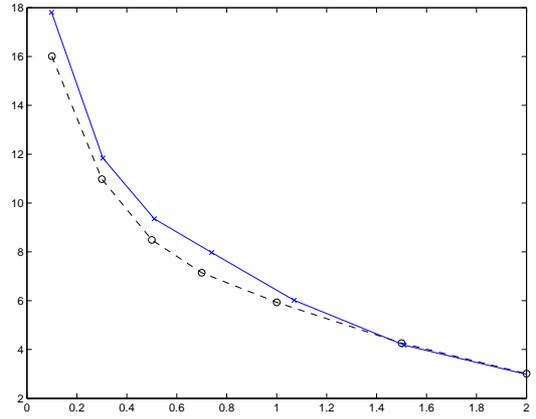


FIG. 2 – Evolution de $D^{1/2}$ en fonction du débit global en bpp.

Références

- [1] T. André, M. Antonini, M. Barlaud, and R.M. Gray. Entropy-based distortion measure and bit allocation for wavelet image compression. *IEEE Trans. on Image Processing*, 16(12) :3058–3064, 2007.
- [2] A. Fraysse, B. Pesquet-Popescu, and J.-C. Pesquet. On the uniform quantization of a class of sparse sources. accepted, *IEEE Trans. on Information Theory*, 2009.
- [3] H. Gish and J. Pierce. Asymptotically efficient quantizing. *IEEE Trans. on Information Theory*, 14(5) :676–683, 1968.
- [4] S. Mallat. A theory for multiresolution signal decomposition : The wavelet representation. *IEEE Trans. Pattern Anal. Machine Intell.*, 11 :674–693, 1989.
- [5] K. Ramchandran, A. Ortega, and M. Vetterli. Bit allocation for dependent quantization with applications to multiresolution and mpeg video coders. *IEEE Trans. on Image Processing*, 3(5) :533–545, 1994.
- [6] K. Ramchandran and M. Vetterli. Best wavelet packet bases in a rate-distortion sense. *IEEE Trans. on Image Processing*, 2(2) :160–175, 1993.
- [7] R.T. Rockafellar. *Convex analysis*. Princeton Landmarks in Mathematics, 1997.
- [8] Y. Shoham and A. Gersho. Efficient codebook allocation for an arbitrary set of vector quantizers. *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '85.*, 10 :1696–1699, 1985.
- [9] P.H. Westerink, J. Biemond, and D.E. Boekee. An optimal bit allocation algorithm for sub-band coding. *IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP*, pages 757–760 vol.2, 1988.