

Utilisation de la transformée en ondelettes Quaternionique pour une stratégie de décodage optimisée JPEG 2000 Wireless

DIOWEL TANDIANG¹, PHILIPPE CARRE¹, CLENCY PERRINE¹, CHRISTIAN OLIVIER¹

Laboratoire XLIM-SIC, UMR CNRS 7252

Université de Poitiers - Blvd Curie Teleport 2 BP 30179, 86962 Futuroscope, France

¹ diouwel.tandiang@univ-poitiers.fr, philippe.carre@univ-poitiers.fr, clency.perrine@univ-poitiers.fr, christian.olivier@univ-poitiers.fr

Résumé – Nous proposons dans ce travail une stratégie de décodage à travers la définition d’une métrique de qualité à référence réduite basée sur la décomposition en ondelettes Quaternionique. Afin d’assurer une extraction visant à améliorer la qualité visuelle des images transmises sur un canal sans fil réaliste, nous couplons un processus de classification à la mesure de paramètres dans le domaine des ondelettes Quaternionique afin de sélectionner le nombre de couches du protocole JPEG 2000 Wireless assurant la meilleure reconstruction en terme de qualité visuelle.

Abstract - This paper proposes a decoding strategy aimed at improving the visual quality of transmitted images on realistic wireless channel. In this context, one reduced reference quality assessment based on quaternionic wavelet transform and the information criterions is proposed. The images to be transmitted are hierarchized in quality layers by JPEG2000 Wireless coder. The transmission scheme supporting this work gives good performance in term of quality of service but it shows some limits in terms of visual quality of received images. The quality metric and a classifier proposed in this paper allow obtaining an optimal visual quality.

1 Quelques éléments introductifs

1.1 Contexte applicatif

Dans le cadre de ce travail, nous nous plaçons dans le cadre d’un système de transmission utilisant un codeur JPWL pour la transmission d’images dans des conditions difficiles [1]. Ce système intègre la technologie multi-antennaire MIMO (Multiple Input Multiple Output), c’est à dire qu’il est considéré plusieurs antennes à l’émission et plusieurs antennes à la réception : le canal peut ainsi être décomposé en b sous canaux SISO (Sigle Input Sigle Output), hiérarchisés, indépendants et parallèles, classés par ordre décroissant des valeurs de SNR (Signal Noise Ratio), et correspondant à b couches de qualité [1]. Soit l le nombre de couches que l’on arrive à décoder suivant les conditions du canal. Lorsque le canal est "bon" : $l=b$, et lorsqu’il est "très mauvais" l vaut 0 .

Cette chaîne de transmission présente de nombreuses améliorations en termes de qualité de services (QoS) (minimisation du taux d’erreurs binaires, maximisation du débit...) mais elle montre certaines limites en termes de qualité visuelle sur les images fournies à la réception. Avec le décodeur robuste JPWL, il peut arriver que l’image obtenue en décodant l couches soit de plus mauvaise qualité que l’image obtenue en décodant $(l-1)$ couches. Pourtant le décodeur JPWL décodera systématiquement (sous réserve d’erreurs trop importantes) toutes les couches de qualité.

Nous proposons dans cet article d’utiliser le principe des métriques à références réduites pour choisir la configuration de décodage assurant la meilleure qualité visuelle des images à la réception. En effet, une

métrique à références réduites permet à partir de quelques caractéristiques bien choisies, que l’on fera transiter via le canal de transmission, de modéliser la perception que l’on aura de l’image. Nous proposons alors d’utiliser un système de classification décidant si l’on doit utiliser toutes les couches ou qu’une partie d’entre-elles pour assurer la meilleure qualité visuelle des images à la réception. Ce travail fait suite à une première étude que nous avons menée dans le laboratoire utilisant une partie de l’information associée à la QWT [2].

1.2 Métrique basée QWT

1.2.1 la transformée en ondelette Quaternionique

L’algèbre des Quaternions est une extension des nombres complexes dans laquelle chaque élément est une combinaison linéaire d’un réel et de 3 imaginaires purs i, j and k :

$$\mathbf{H} = \{q = q_0 + iq_1 + jq_2 + kq_3 \mid q_0, q_1, q_2, q_3 \in R\} \quad (1)$$

$$\text{Avec } i^2 = j^2 = k^2 = ijk = -1, \quad ij = -ji = k, \quad k = -kj = i, \quad ki = -ik = j.$$

Il existe une forme polaire définie par un module et 3 phases [3] :

$$q = |q|e^{i\theta}e^{i\psi}e^{i\phi} \quad (2)$$

Nous faisons ainsi apparaitre une notion de module et de 3 phases.

La transformée en ondelettes Quaternionique (QWT) [4] est une évolution de la transformée en ondelette classique basée sur le concept de signal analytique Quaternionique proposé par Bülöw [3] qui généralise le concept de signal analytique aux signaux 2-D avec l’utilisation des quaternions. Cette nouvelle

représentation est intéressante pour la caractérisation des dégradations car elle offre une représentation plus riche de l'information 2-D à travers une analyse Module-Phase invariante par translation [4,5].

Chaque sous-bande de la QWT peut être considérée comme le signal analytique associé avec une bande fréquentielle de l'image. L'information portée par les coefficients de cette représentation reste un sujet de discussion, ceci étant on peut avancer que le module du coefficient QWT, en chaque position, représente une mesure d'énergie alors que les 3 phases (ϕ , θ , ψ) décrivent la structure avec respectivement la notion de translation selon x , y et une information de texture. Nous avons donc proposé d'utiliser ces informations pour analyser les dégradations. Pour cela, nous devons être capable de caractériser la distribution de ces paramètres et ainsi d'observer leurs variations. Ceci a fait l'objet d'un précédent travail [2] que nous rappelons brièvement dans la section suivante.

1.2.2 Modélisation de l'information QWT : critère d'information.

La suite du propos, à la fois en ce qui concerne le choix des modèles [6] mais aussi la définition de la mesure de qualité, repose sur l'information de Kullback-Leibler définie entre deux densités de probabilité (ddp) f et g :

$$KL(f, g) = \int f(x) \ln \frac{f(x)}{g(x)} dx \quad (3)$$

Dans notre utilisation des critères d'information il s'agissait de construire les histogrammes qui résument le mieux la distribution (ddp) des coefficients des différentes sous-bandes de la QWT. Les critères d'information IC permettent de "lisser" les artéfacts aux niveaux des histogrammes des paramètres issus de la QWT facilitant ainsi leur modélisation.

Avec un histogramme à pas égaux, les critères s'écrivent de la manière suivante:

$$IC(k) = \sum_{j=1}^k n_j \log n_j - n \log k - kc_n \quad (4)$$

avec $n_j, j = 1, \dots, n_j$ est le nombre de coefficients dans les réalisations de la classe correspondante j , k est le nombre de classes dans les histogrammes, n est le nombre de coefficient de chaque sous-bande, c_n est la fonction de pénalité. \hat{k} est le nombre de classes de l'histogramme estimé au sens des IC, à savoir celui qui minimise le critère. Le choix des différents IC dépend de la fonction de pénalité, et dans nos études antérieures nous avons choisi le critère d'information bayésien (BIC), consistant, avec $c_n = \log(n)$ [6].

Ainsi, dans le but de trouver les références réduites nous devons modéliser par des lois statistiques les sous-bandes du module et des 3 phases de la QWT afin d'en extraire leurs paramètres qui constitueront nos références réduites. Nous résumons ci-dessous les différents points proposés dans [2].

- Nous utilisons la QWT basée sur l'algorithme Dual tree [7] avec trois niveaux de décompositions.

Avec ces niveaux de décomposition, nous avons une analyse suffisante et fine des détails de l'image. Notons que l'approximation passe-bas n'est pas utilisée car les dégradations observées sont généralement en haute

fréquence. Après l'obtention des coefficients de chaque sous bande de la QWT, un prétraitement est effectué sur les coefficients des sous-bandes du module dans le but de supprimer les coefficients proches de zéro car ces coefficients ne sont pas porteurs d'information.

- Pour modéliser les sous-bandes du module et des 3 phases plusieurs lois statistiques ont été retenues [2].

Grace aux IC, nous construisons les histogrammes des 9 ddp avec un nombre de classes optimal facilitant ainsi leur modélisation. Ensuite, les modèles sont choisis suivant le critère de minimisation de l'information de Kullback-Leibler. Certaines sous-bandes de la phase ϕ et θ ne sont pas retenues (les ddp des coefficients sont des lois uniformes) et 3 sous-bandes sont seulement retenues ?. Pour le module et la phase Ψ les 9 sous-bandes d'analyses sont retenues. Pour résumer les modèles et paramètres associés sont :

1. Pour le module, le modèle de Gaussienne Inverse a été choisi associé avec une Moyenne (μ) et un Paramètre de forme (α) ;
2. Pour la phase θ le Mélange de Gaussienne Généralisé a été retenu associé à plusieurs Moyennes (μ), et Paramètres de forme(α), et écart-types (σ)
3. Enfin les phases Ψ , ϕ sont modélisées par une Gaussienne Généralisée associée à une Moyenne (μ), un Paramètre de forme(α), et l'écart-type (σ).

Après cette étapes de modélisation, les paramètres de chaque sous bande vont constituer nos références réduites (63 au total). Ce sont ces paramètres et l'image originale qui seront envoyés sur le canal de transmission. A la réception de ces paramètres, les ddp sont reconstruites à partir de ceux-ci (la nature des lois est supposée connue du récepteur), et nous construisons également les distributions des coefficients des sous bandes de l'image reçue.

1.2.3 Caractérisation de la dégradation

La mesure de la dégradation peut être une fonction représentant la « différence » entre les distributions des sous-bandes de l'image originale (ddp connues par leurs références réduites) et l'image dégradée [4], comme nous l'avons fait dans notre précédente proposition à savoir :

$$D_{QWT}^{total} = \log_2 \left(1 + \sum_n^N \bar{KL} + \sum_m^M \bar{KL} + \sum_s^S \bar{KL} + \sum_t^T \bar{KL} \right) \quad (5)$$

Avec N , M , S et T les nombres respectifs de sous-bandes du module et des phases ϕ , θ et Ψ .

Notons que dans notre travail précédent, dans le cadre de certaines attaques que l'on peut par exemple trouver dans la base Live [8], le module et la phase suffisaient pour détecter les dégradations. Dans un contexte plus réaliste, face aux attaques canal qui ont pour incidences des décalages horizontaux et verticaux, toutes les phases ont une certaine importance.

Cependant, on peut noter dans l'équation (5) qu'il n'existe pas de pondération entre les différentes bandes et entre les différentes caractéristiques. Ces

pondérations vont elles aussi avoir un lien fort avec la modélisation psychovisuelle que l'on peut avoir des dégradations. C'est pourquoi elles ne peuvent être issues que d'un processus d'apprentissage. C'est l'objet de la dernière partie de notre article.

2 Stratégie de décodage basée QWT

2.1 Mesure directe : métrique de qualité

Pour illustrer la stratégie de décodage que l'on propose, on utilise un parcours en milieu urbain où l'émetteur est fixe et le récepteur se déplace à une vitesse de 5m/s. Nous avons alors une alternance successivement de mauvaise (zones notées 1), moyenne (zones notées 2) ou bonne condition (zones notées 3).

Pour mesurer les dégradations du canal réaliste nous avons considéré des images reçues à des positions variées du récepteur. Nous décodons à chaque fois les $(l-1)$ couches de l'image et les l couches de l'image envoyée suivant les conditions du canal. Sur la figure 1 nous présentons ces différentes images avec la mesure DQWT selon l'équation (5).

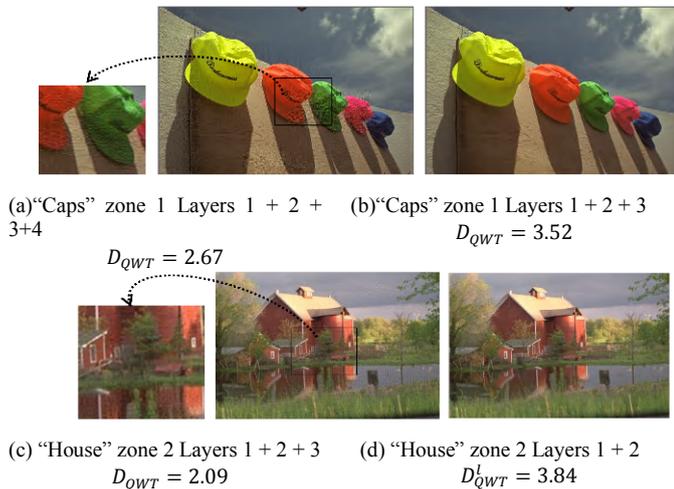


Figure 1: (a), (b), (c), (d), exemples de décodage avec la mesure basée QWT

Notre mesure QWT est une fonction de la somme des mesures de l'information de Kullback-Leibler entre les sous-bandes de l'image originale et celles de l'image dégradée et donc plus l'image est dégradée plus la valeur de la métrique doit être grande. Nous constatons que ce n'est pas la réalité dans les exemples exposés ci-dessus : l'image de gauche est plus dégradée que celle de droite avec une mesure QWT indiquant l'inverse.

Comme nous l'avons abordé, le problème de la métrique DQWT selon l'équation (5) réside dans la pondération des différentes mesures entre les bandes et les paramètres. Nous sommes donc face à un problème de fusion de mesures qui doit faire émerger une décision. C'est pourquoi nous proposons de nous écarter de la définition d'une métrique évaluant un indice de dégradation, et nous proposons de nous placer dans une stratégie de classification qui doit répondre à la question suivante : étant données les mesures basées QWT quelle est la meilleure qualité visuelle d'image entre celle issue du décodage partiel et celle issue du décodage total ?

2.2 Fusion des mesures et prise de décision

Nous appliquons une classification suivant la méthode des k plus proches voisins (KppV) car nous possédons une base d'apprentissage restreinte. Pour mettre en place le système de prise de décision, nous avons dû créer une base de données d'apprentissage associée à un test subjectif. Cette base permet d'avoir des « exemples » associant différents types de décodage et donc différentes mesures de dégradation basées QWT et différentes perceptions de qualité. A ce jour, notre système d'apprentissage repose sur 4 images très différentes par leur contenu issues de la base Live, dégradées par différents niveaux d'attaque canal et donc décodées à partir de toutes les couches ou d'une manière partielle.

Nous obtenons alors 350 images (50% décodage total et 50% décodage partiel) que nous avons présentées à des observateurs afin qu'ils choisissent selon eux la meilleure image entre celle issue du décodage partiel et celle issue du décodage total. Cette opération labélise la base d'image suivant deux classes : image nécessitant un décodage complet et image nécessitant un décodage partiel pour obtenir la meilleure qualité visuelle.

En parallèle, pour chacun de ces couples d'images, un vecteur de mesures des différences entre les modélisations des paramètres QWT de ces deux images et des paramètres fournis dans le flux binaire résultant de la modélisation de l'image originale sont construits. Ils constitueront les coordonnées des images dans l'espace des paramètres.

Ensuite, lors de l'utilisation de notre stratégie, pour une image transmise pouvant être décodée d'une manière complète ou partielle, nous prenons la décision suivant le protocole :

- nous effectuerons la mesure des différences entre les modélisations des paramètres QWT de ces deux images et des paramètres résultant de la modélisation de l'image originale
- nous étudions les k voisins de la base d'apprentissage dont les coordonnées (donc le vecteur de différences) sont les plus proches du vecteur de mesures du nouveau couple d'images arrivant.
- Nous sélectionnons le meilleur décodage suivant la classe majoritairement présente dans le proche voisinage.

Après différents tests, nous avons choisi un nombre de voisins $k = 5$.

2.3 Applications de la stratégie de décodage basée QWT

Afin d'évaluer la qualité de notre stratégie, nous construisons différents groupes d'images tests (non utilisées lors de l'apprentissage) par tirage aléatoire avec la contrainte d'avoir des paires d'images.

Dans l'expérimentation menée, nous avons construit 8 groupes d'images composés chacun de 25 paires d'images. Nous avons mené une étude comparative entre les résultats de notre stratégie basée classification

QWT et ceux issus d'une métrique à références réduites RRIQA [9] et d'une métrique de qualité à référence complète SSIM [10]. A chaque test, le nombre de décisions erronées commises est relevé. Le nombre de décision erronée est le nombre de fois que la mesure où notre système a choisi l'image issue du décodage partiel comme étant de meilleure qualité alors que les utilisateurs avaient choisi majoritairement l'image issue du décodage total (et vice-versa).

Dans le tableau ci-dessous, nous indiquons le nombre d'erreurs pour toutes les métriques considérées.

	SSIM	RRIQA	Notre stratégie
Test 1	7	8	3
Test 2	6	3	3
Test 3	9	17	2
Test 4	3	7	1
Test 5	5	5	2
Test 6	6	6	2
Test 7	8	10	1
Test 8	5	5	1
Mean	6.12	8.25	1.86
Standard-deviation	1.76	4.71	0.78

Notre système obtient en moyenne de 92.56% de réponses correctes contre 67% pour RRIQA et 75.52% pour SSIM. Nous remarquons également une variabilité selon les différentes bases d'images. Cette observation confirme que le contenu de l'image présente un impact sur l'efficacité de la stratégie de décodage. Concernant notre stratégie, nous avons étudié les décisions erronées, et nous avons constaté qu'elles correspondaient aux couples d'images que les observateurs avaient classées comme étant les couples d'images visuellement très proches en termes de dégradations.

Nous avons proposé une stratégie de prise de décision de décodage dans le cadre d'une transmission sans fil d'image dans un environnement difficile de transmission caractérisé par le phénomène de multi-trajets, de mobilité et de bruit radioélectrique.

La méthode utilise des références réduites basées sur la décomposition QWT, les IC et les modélisations des

coefficients. Les premiers résultats obtenus sont encourageants avec en moyenne 92.56% de prise de décision correcte lors du décodage.

Deux points d'études sont actuellement poursuivis :

- L'augmentation de la base de données afin d'incorporer d'autres dégradations canal.
- La définition d'un autre système de classification ne nécessitant par la conservation des exemples issus de la base d'apprentissage.

3 References

- [1] Abot M., Olivier C., Perrine C., Pousset Y, A link adaptation scheme optimized for wireless JPEG2000 transmission over realistic MIMO systems, Signal Processing : Image Communication, Elsevier, 93 (5), 1064-1078, May 2013.
- [2] Traore A., Carré P., Olivier C., "Reduced-reference metric based on the quaternionic wavelet coefficient modelling by Information Criteria", IEEE-ICIP, Paris (France), Oct. 2014.
- [3] T. Bülow, "Hypercomplex spectral signal representation for the processing and analysis of images". Thesis, Aug. 1999.
- [4] W L Chan, H Choi, R Baraniuk, Coherent multiscale image processing using dual-tree quaternion wavelets, IEEE Trans. Image process. 2008, 17(7), 1069-1082
- [5] R Soulard, P Carré, Quaternionic wavelets for texture classification," Pat. Recognition Letters 32, 1669-1678(2011)
- [6] Alata O., Olivier C., Pousset Y., Law recognitions by Information Criteria for the statistical modelling of small. scale fading of the radio mobile channel, Signal Processing Journal, Elsevier, 93 (5), 1064 -1078, May 2013.
- [7] N G Kingsbury, The dual-tree complex wavelet transform: a new technique for shift invariance and directional filters. IEEE Digital Signal Proc. Workshop on DSP (Bryce Canyon, USA , 1998), pp. 2543-2560
- [8] "Laboratory for Image and Video Engineering" in collaboration with Center for Perceptual Systems at The University of Texas at Austin
- [9] Z. Wang and E. P. Simoncelli, "Reduced-reference image quality assessment using a wavelet-domain natural image statistic model," Human Vision and Electronic Imaging X, Proc. SPIE, vol. 5666, San Jose, CA, Jan. 2005.
- [10] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity", IEEE Transactions on Image Processing, vol. 13, no. 4, pp. 600-612, Apr. 2004.