

Tatouage robuste d'images par Turbo TCQ

Gaëtan LE GUELVUIT¹

¹France Telecom R&D
4, rue du Clos Courtel
35512 Cesson Sévigné cedex, France
gaetan.leguelvuit@orange-ftgroup.com

Résumé – Cet article montre une utilisation de la Turbo TCQ – technique utilisée habituellement pour le codage de source – afin de construire des codes correcteurs adaptés aux canaux avec information adjacente. Nous en faisons une mise en œuvre en l'appliquant au tatouage robuste d'images. Les résultats montrent d'excellentes performances, comparables à l'état de l'art, tout en conservant une complexité mesurée à la fois lors de la phase d'insertion et d'extraction.

Abstract – This article shows how to use Turbo TCQ – an algorithm initially designed for source coding purpose – to design error correcting codes for channels with side information at the encoder. We then apply it for robust image watermarking. Results show good performance, similar to state of the art, but with limited complexity both at the embedding and extraction phases.

Introduction

Depuis quelques années, la recherche en tatouage (*watermarking*) est fortement influencée les principes du schéma de Costa [1], qui permettent – en théorie – de construire un code dont les performances ne dépendent pas du signal hôte à marquer. Les travaux présentés ici montrent une mise en pratique de ces principes, *via* l'utilisation de la Turbo TCQ. En en faisant une utilisation détournée, nous montrons comment construire un code correcteur adapté au problème de communication que représente le marquage de données multimédia.

Dans la première partie, nous revenons sur la notion d'information d'adjacente dans les canaux de communication, puis sur les résultats de Costa. Dans la section suivante, les techniques de quantification pour le tatouage sont abordées et nous décrivons plus particulièrement la Turbo TCQ. Les sections 3 et 4 montrent un exemple de mise en pratique (le tatouage robuste d'images) et présentent les résultats que nous obtenons face à des attaques communes.

1 Information adjacente et schéma de Costa

Nous considérons un canal gaussien avec une contrainte d'émission. Un signal \mathbf{w} est émis avec une énergie limitée définie par

$$\frac{1}{n} \sum_{i=1}^n \mathbf{w}[i]^2 \leq P.$$

Le canal de communication est perturbé par un bruit \mathbf{z} modélisé par une variable aléatoire suivant une loi Normale $Z \sim \mathcal{N}(0, N)$. La capacité du canal – c'est-à-dire le nombre de bits utiles par symbole $\mathbf{w}[i]$ émis – est donnée

par la formule

$$C = \frac{1}{2} \log \left[1 + \frac{P}{N} \right].$$

Si le canal est perturbé par deux sources de bruit indépendantes, naturellement, la variance des deux bruits s'ajoute et diminue d'autant la capacité C . Nous retrouvons ce type de perturbations dans le cadre du tatouage¹ : une marque \mathbf{w} est ajoutée à un signal hôte \mathbf{x} , et le tout est perturbé par une attaque \mathbf{z} . La capacité de la marque serait donc limitée par N (énergie de l'attaque) et par Q (énergie du signal hôte).

Néanmoins, la perturbation \mathbf{x} présente la particularité d'être parfaitement connue au moment de l'émission de \mathbf{w} . Cet état initial du canal est une information adjacente. En 1983, Costa [1] a démontré qu'elle n'avait aucune influence sur la capacité du canal. Cela ouvrit des perspectives importantes pour la problématique du tatouage, car on entrevit la possibilité de construire un schéma dont les performances dépendraient uniquement de l'attaque et pas du tout du signal hôte².

Pour atteindre la capacité, Costa propose dans son article un schéma théorique basé sur la construction d'un dictionnaire surjectif, puis sur une technique d'émission consistant à *diriger* l'information adjacente vers le mot de code, plutôt que de simplement l'ajouter. Il y a donc un codage informé et une émission informée, spécifiques aux canaux avec information adjacente. Dans la suite de cet article, nous proposons un codage adapté s'appuyant sur des techniques de quantification, et une méthode d'émission prenant en compte des critères perceptuels pour minimiser l'impact du tatouage sur la qualité des documents

¹C'est un modèle très simplifié, les signaux rencontrés étant rarement simplement gaussiens.

²Si l'on reste dans une modélisation par canal gaussien classique du problème du tatouage, l'énergie du signal limite très fortement les performances car $P \ll Q$ du fait des contraintes d'invisibilité.

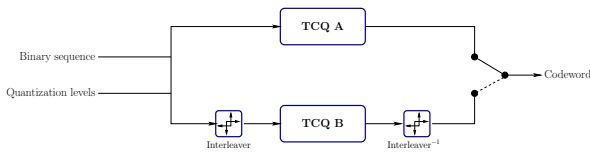


FIG. 1 – Principes de la Turbo TCQ : deux TCQ fonctionnent en parallèle sur des versions entrelacées d'un même signal

marqués.

2 Codes correcteurs basés sur la quantification

La quantification est une technique populaire pour construire des codes pour canaux avec information adjacente. Le plus connu est le schéma de Costa scalaire (SCS) [2], qui utilise la quantification scalaire de \mathbf{x} pour encoder un message.

2.1 Principes de la Turbo TCQ

La Turbo TCQ (*Trellis Coded Quantization*) [3] est une technique de codage de source reprenant les approches itératives que l'on retrouve appliquées dans les codes correcteurs. Le quantificateur est composé de deux treillis parallèles, comme montré sur la figure 1. Un premier treillis prend en entrée le signal à quantifier, tandis que le second travaille sur une version entrelacée. Les métriques *a posteriori* issues premier quantificateur servent de métriques *a priori* au quantificateur parallèle. Le processus est itéré jusqu'à ce que les sorties convergent.

Nous utilisons cette technique pour construire un code adapté aux canaux avec information adjacente. L'encodage consiste à quantifier \mathbf{x} en forçant les transitions des treillis pour qu'elles correspondent au message à encoder. De cette façon, nous nous assurons que le mot de code obtenu \mathbf{u}^* appartient bien à \mathcal{U}_m , sous-ensemble des mots de codes qui correspondent au message \mathbf{m} . Le décodage reprend le même principe, mais avec en laissant toutes les transitions libres.

2.2 Émission informée

Comme déjà vu dans la section 1, le schéma de Costa comprend en plus d'un dictionnaire adapté une technique d'émission informée. L'information adjacente est dirigée vers le mot de code, par l'introduction d'un paramètre α :

$$\mathbf{w} = \alpha (\mathbf{u}^* - \mathbf{x}). \quad (1)$$

Nous reprenons $\alpha = P/(P + N)$, comme défini par Costa. Les expériences montrent que cette formule donne les meilleurs résultats. Les performances du code montrent un gain important aux techniques plus classiques (SCS, QIM et TCQ simple). Pour un rendement de 1/1 (un bit inséré par échantillon de \mathbf{x}), nous obtenons un gain de 5,5 dB par rapport au SCS. L'utilisation du principe Turbo fait gagner 3,5 dB (voir la figure 2).

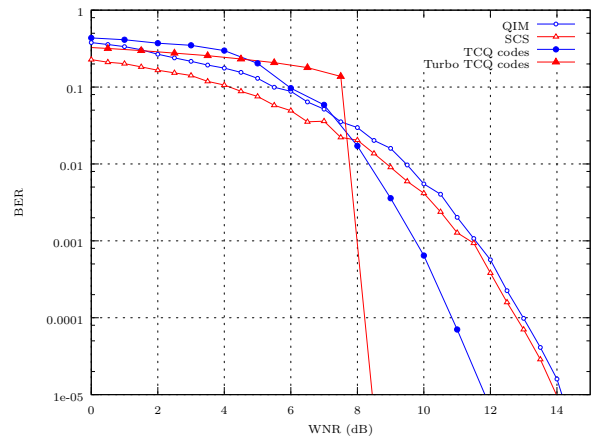


FIG. 2 – Probabilités d'erreur par bit pour un rendement d'un bit par élément hôte marqué

3 Application au tatouage robuste d'images

Dans cette section, nous appliquons le code correcteur basé sur la Turbo TCQ pour construire un schéma de tatouage d'images robuste. Nos choix ont été essentiellement dictés par ceux présentés dans l'article de Miller *et al.* [4], qui décrit une technique de tatouage très performante basée sur les principes informés vus dans la section 1. Nous reprenons ainsi le même domaine d'insertion (et le même choix de coefficients) et la même capacité (un bit utile par bloc marqué). L'image hôte en niveaux de gris est passée dans le domaine DCT en blocs 8×8 . Pour chacun des blocs, douze coefficients sont sélectionnés : le coefficient AC est évité, et nous gardons les douze coefficients suivants dans l'ordre zigzag.

Le rendement de notre code est de 1/1 (un bit utile par échantillon du signal hôte). Pour réduire ce rendement, nous utilisons une projection : les douze échantillons DCT sont projetés sur un vecteur pseudo-aléatoire³ (composé de +1 et de -1) afin de former une valeur unique. L'ensemble des valeurs obtenues par projection (une valeur par bloc DCT) forme le signal \mathbf{x} :

$$\mathbf{x}[i] = \sum_{j=1}^{12} \mathbf{X}[i, j] \times \mathbf{G}[i, j],$$

avec \mathbf{X} les coefficients DCT et \mathbf{G} l'ensemble des vecteurs de projection. Nous pouvons ensuite calculer \mathbf{u}^* (quantification de \mathbf{x} sous contraintes, comme vu en section 2.1), puis \mathbf{w} par la formule 1. Les coefficients DCT marqués \mathbf{Y} sont donnés par la projection inverse :

$$\mathbf{Y}[i, j] = \mathbf{X}[i, j] + \frac{\mathbf{w}[i]}{12} \times \mathbf{G}[i, j].$$

Comme le montre cette équation, la marque \mathbf{w} est répartie uniformément sur les coefficients à marquer. Or, rien ne nous empêche de choisir une autre répartition. Il est ainsi possible de prendre en compte une mesure perceptuelle. La

³La génération de ces vecteurs est paramétrée par une clef secrète.

marque \mathbf{w} – initialement calculée dans le domaine projeté – est alors distribuée sur les coefficients en fonction de leur importance psychovisuelle. Nous utilisons pour cela un algorithme classique de *water filling*. La pondération perceptuelle choisie est le modèle de Watson, également utilisé dans l'article que nous prenons comme référence [4].

L'extraction de la marque reprend les mêmes étapes : transformation de l'image considérée dans le domaine DCT, application de la projection puis décodage de la suite de valeurs obtenues. La technique est très rapide, car sur un ordinateur portable classique (type Intel Core Duo à 2 GHz), l'insertion d'une marque prend 400 ms (ce qui comprend la lecture de l'image depuis le disque, la DCT, le calcul du modèle de Watson, la Turbo TCQ, l'insertion de la marque, la DCT inverse et l'enregistrement du fichier) et l'extraction 200 ms.

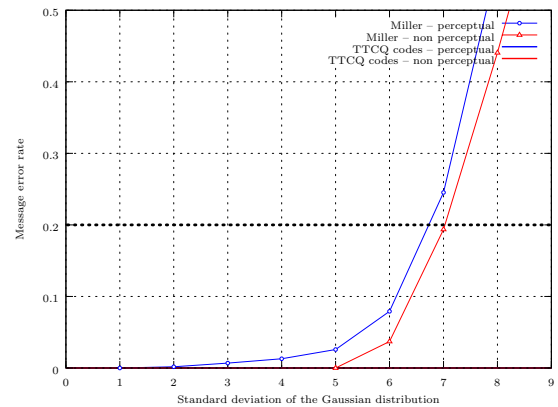
4 Résultats des tests de robustesse

Le schéma est testé sur un ensemble de 1200 images en niveaux de gris de 368×240 pixels. Nous insérons une marque de 1380 bits utiles par image (un bit par bloc DCT). La distorsion d'insertion (paramétrée par le pas de quantification Δ utilisé dans la Turbo TCQ) est fixée telle que la distance de Watson moyenne entre image d'origine et image marquée soit proche de celle de l'article de référence. Sans prise en compte du modèle de Watson à l'insertion, nous obtenons une distance moyenne de 96,84 (contre la moyenne de 101,52 pour [4]), et 31,04 avec prise en compte (contre 31,6).

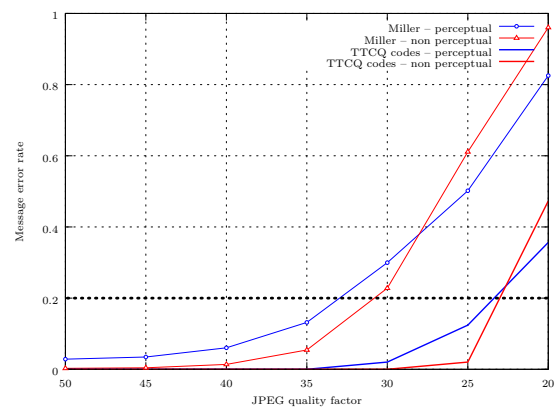
La figure 3(a) montre la robustesse de notre schéma face à l'ajout de bruit gaussien. Sur les 1200 images, nous n'avons noté aucune erreur pour un bruit d'écart type jusqu'à 9. Le schéma de référence n'est plus considéré comme robuste (plus d'un message extrait sur cinq présente au moins une erreur) à partir d'un écart type de 7.

La figure 3(b) présente les résultats face à la compression JPEG. L'abscisse montre la qualité de compression (plus la valeur est basse, et plus l'image est dégradée). Notre schéma reste robuste jusqu'à un facteur de qualité de 23 (celui de référence résiste jusqu'à 30 dans le cas le plus favorable).

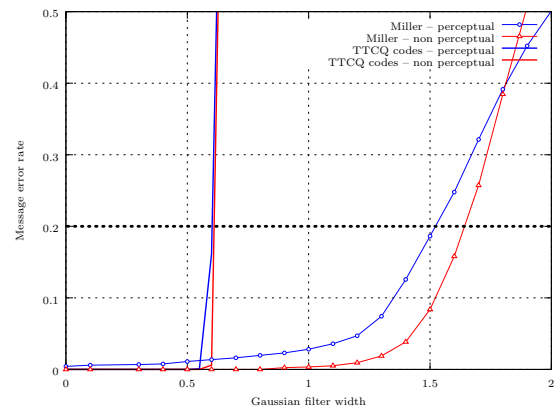
La figure 3(c) montre les performances du schéma pour des attaques par floutage de l'image marquée. On voit alors les limites des codes par quantification. En effet, le filtre de flou atténue les coefficients DCT, ce qui désynchronise les pas de quantification lors du décodage. Le code décroche rapidement (robuste jusqu'à 0,6) alors que le schéma de référence se comporte bien (robuste jusqu'à 1,55). Cette faiblesse face aux changements d'échelle est confirmée par les résultats de la figure 3(d), où les pixels des images marquées sont multipliés par un scalaire. De part sa conception, le schéma de Miller y est quasi insensible, contrairement à celui proposé ici.



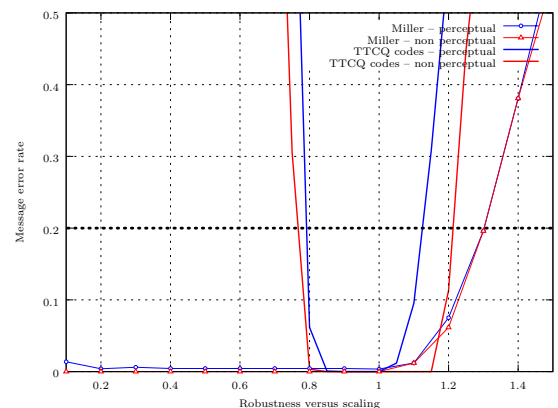
(a) Contre l'ajout de bruit gaussien



(b) Contre la compression JPEG



(c) Contre un flou gaussien



(d) Contre le facteur d'échelle

FIG. 3 – Résultats des tests de robustesse pour 1380 bits utiles dans des images de dimensions 368×240 pixels

Conclusion et perspectives

Cet article a présenté une mise en pratique simple des principes décrits dans l'article de Costa sur les canaux avec information adjacente. Nous avons montré que l'utilisation d'un code correcteur adapté au canal du tatouage et d'une technique d'insertion informée permet d'obtenir d'excellentes performances, à la fois en termes de capacité (plus de 1300 bits dans des imagerie de 368×240 pixels) et de robustesse. La technique de codage par treillis que nous utilisons présente en plus des propriétés intéressantes pour la stéganographie [5].

De plus, l'utilisation d'une fonction de projection avant le code correcteur permet à la fois de moduler le rendement global du code pour fixer le compromis robustesse/capacité, et de prendre en compte des pondérations externes pour adapter la marque au document hôte (poids perceptuels, paramètres issus de la théorie des jeux, marque de type PSC, etc.).

Néanmoins, l'utilisation de quantificateurs pour le tatouage montre ses limites face aux changements d'échelle. Ce point a été étudié pour les schémas de type QIM ou SCS, en proposant des techniques pour estimer le changement d'échelle avant l'extraction [6]. Les solutions issues de ces travaux pourraient être directement intégrées dans notre schéma.

Références

- [1] M. H. M. Costa. Writing on dirty paper, IEEE. Trans. on Information Theory, 29(3) : 439–441, Mai 1983.
- [2] J. J. Eggers, R. Baüml, R. Tzchoppe et B. Girod. Scalar Costa scheme for information embedding, IEEE Trans. on Signal Processing, Avr. 2003.
- [3] V. Chappelier, C. Guillemot et S. Marinkovic. Turbo trellis coded quantization, in *Proc. of Int. Symp. on Turbo Codes*, Brest, France, Sep. 2003.
- [4] M. L. Miller, G. J. Doërr et I. J. Cox. Applying informed coding and informed embedding to design a robust, high capacity watermark, IEEE Trans. on Image Processing, 3(6) : 792–807, 2004.
- [5] J. Bourgeois, G. Le Guelvouit, A. Ould Bouya, C. Delpha et R. Boyer. Analyse stéganographique du schéma scalaire de Costa, in *GRETSI Symp. on Image and Signal Processing*, Sep. 2007.
- [6] I. D. Shterev, R. L. Lagendijk. Amplitude scale estimation for quantization-based watermarking, IEEE Trans. on Signal Processing, 11(54) : 4146–4154, Nov. 2006.