

Solution d'allocation de puissance efficace pour de la transmission vidéo scalable en environnement Outdoor réaliste

W. Hamidouche¹

C. Perrine¹

Y. Pousset¹

C. Olivier¹

¹ Laboratoire XLIM, CNRS UMR 2172, Département SIC

Université de Poitiers Téléport 2 Bd Marie et Pierre Curie
BP 30179, 86962 Futuroscope Cedex (France)

{hamidouche, perrine, pousset, olivier}@sic.sp2mi.univ-poitiers.fr

Résumé

Dans ces travaux nous considérons la transmission d'information vidéo au format H.264/SVC dans un contexte de canaux MIMO réalistes tenant compte de techniques de précodage. Ainsi, d'une part, le standard H.264/SVC fournit une qualité de scalabilité spatio-temporelle de l'information vidéo. D'autre part, les précodeurs subdivisent le canal MIMO en sous canaux SISO indépendants associés à des TEB faibles. Ainsi, nous proposons un nouveau schéma exploitant à la fois la hiérarchie intrinsèque du standard H.264/SVC et les qualités de quatre précodeurs (Max-SNR, WF, QoS et $E-d_{min}$). Le schéma proposé est testé sur un canal MIMO réaliste. De plus, nous évaluons ses performances vis-à-vis des erreurs d'estimation du canal selon la norme IEEE802.11n.

Mots clefs

Allocation optimale de la bande passante, H.264/SVC, MIMO, techniques de précodage, précodeur QoS, précodeur $E-d_{min}$, OFDM, tracé de rayons 3D.

1 Introduction

Lors d'une communication sans fil, l'interaction des ondes électromagnétiques avec l'environnement ainsi que la mobilité des utilisateurs engendrent des affaiblissements importants et imprévisibles du signal reçu. Ainsi, la bande passante limitée du canal radio et sa nature instable représentent les éléments principaux à prendre en compte pour maintenir une bonne qualité de service des applications multimédias. Dans cet article nous nous intéressons plus particulièrement à la transmission d'un flux vidéo scalable sur un canal MIMO (Multiple Input Multiple Output). Au niveau de la couche application, nous utilisons le standard H.264/SVC [1] pour compresser la vidéo originale en un flux vidéo scalable et hiérarchisé.

Plusieurs auteurs se sont intéressés à la transmission de vidéos scalables sur un canal MIMO [2]-[5]. Parmi ces solutions, nous nous focalisons sur celles qui considèrent la connaissance de l'information du canal CSI (Channel State Information) au niveau de l'émetteur (Tx-CSI) [3]-[5], no-

tamment appelées systèmes à boucle fermée. Un schéma d'allocation de puissance optimale à été proposé dans [3] minimisant la distorsion du flux vidéo scalable MPEG-4. Cette solution considère la connaissance parfaite de Tx-CSI ainsi que l'information sur le débit-distorsion de la version scalable du codeur vidéo MPEG-4. Un autre schéma de sélection adaptative des sous canaux SISO (Single Input Single Output) pour la transmission de flux vidéo H.264/SVC a été proposé dans [4]. Cette solution, basée sur l'information partielle du canal résumée par le Rapport Signal à Bruit (RSB), affecte chaque flux vidéo, suivant son importance, au sous canal SISO correspondant. Ce schéma offre une stratégie de protection inégale contre les erreurs de transmission aux flux vidéo sans aucune redondance ou une stratégie d'allocation de puissance. Les mêmes auteurs ont proposé par la suite dans [5] un nouveau schéma qui combine les travaux dans [4], le précodeur Water Filing (WF) et une technique de ré-allocation de puissance. Cette combinaison permet d'une part une transmission sans erreur du flux vidéo de qualité de base, et d'autre part, une transmission des flux d'amélioration de qualité avec une modulation de haute efficacité spectrale.

Les travaux présentés précédemment utilisent des techniques d'allocation de puissance standards, telle que le précodeur WF qui est largement surpassé par les nouveaux précodeurs en terme de Taux d'Erreurs Binaire (TEB) et de flexibilité dans le processus d'allocation de puissance. De plus, les stratégies d'allocation de puissances utilisées ne sont pas finement adaptées à l'importance des flux vidéo, tel que le précodeur WF qui plutôt maximise la capacité du canal. Enfin, ces schémas ne sont testés que sur des canaux statistiques, alors que les spécificités environnementales et la corrélation des antennes MIMO ont un grand impact sur la qualité des liens MIMO [9].

Dans ce papier, nous proposons un nouveau schéma de transmission temps réel de vidéos codées H.264/SVC sur un canal MIMO en utilisant quatre précodeurs, à savoir Max-SNR [6], WF [7], QoS [7] et $E-d_{min}$ [8]. Ces précodeurs décomposent le canal MIMO en sous canaux SISO indépendants et parallèles de différentes impor-

tances, réduisant ainsi la complexité du décodage de Maximum par Vraisemblance (MV) et maintiennent des valeurs de TEB relativement faibles. Les spécificités communes à ces précodeurs leur permettent d'être très adaptées à une transmission temps réel d'un flux vidéo hiérarchisé.

Nous exploitons la flexibilité du précodeur QoS afin d'adapter finement l'allocation de puissance suivant l'importance des flux vidéo et l'état instantané du canal de transmission. Les performances du schéma proposé sont évaluées sur un canal MIMO réaliste. Nous utilisons un modèle de propagation déterministe [10] qui prend en compte toutes les spécificités environnementales (géométrique, électrique) d'un environnement réel fournissant en sortie la Réponse Impulsionnelle Complexe (RIC) correspondante. De plus, nous évaluons la robustesse de ce schéma vis-à-vis des erreurs d'estimation du canal (EC) suivant la norme IEEE802.11n [11].

2 Description du système

2.1 Standard H.264/SVC pour une transmission sans fil

La norme H.264/SVC est la version scalable du standard de compression de vidéos H.264/SVC. La principale attente de cette nouvelle norme est de supporter les scalabilités spatiale, temporelle et en résolution tout en maintenant l'efficacité de compression de la norme H.264/AVC. Ainsi, H.264/AVC conserve la même structure de codage que le standard H.264/AVC, à savoir un codage en deux couches : Video Coding Layer (VCL) et Network Abstraction Layer (NAL). La scalabilité temporelle est efficacement établie en utilisant le concept hiérarchique d'images bidirectionnelles (image-B). La scalabilité spatiale est effectuée en utilisant le mécanisme de prédiction entre les différentes couches. Quant à la scalabilité en résolution, elle est assurée par le concept général d'un codage spatial appelé dans le standard H.264/SVC Medium Grain quality Scalability (MGS). Pour plus détails sur le standard H.264/SVC le lecteur pourra se référer à [1].

Le codec H.264/SVC encode la vidéo originale en plusieurs couches de résolutions temporelles. Le codage de scalabilité en résolution est appliqué sur chacune des couches temporelles afin de fournir les couches de scalabilité en qualité (raffinement ou amélioration). Dans ce papier, le codeur source fournit N_T couches de résolution temporelle et N_Q couches de résolution en qualité. Suivant une stratégie de transmission, ces couches constitueront les N flux vidéo à transmettre, définis dans ce papier par $(\rho_0, \rho_1, \dots, \rho_{(N-1)})$.

Au niveau de la couche physique, les N flux vidéo constituant les paquets NAL seront transmis dans des trames IEEE802.11n de taille constante. Tout paquet vidéo NAL reçu dans une trame erronée sera perdu. Ces pertes de paquets se traduisent soit par des dégradations temporelles ou en résolutions de la vidéo reçue. Afin de faire face aux pertes d'images causées par la perte des couches de

basse résolution, quatre méthodes de robustesses proposées dans [2] ont été implémentées. Ces dernières permettent de remplacer les images perdues par des images correctement reçues appartenant au même GOP (Group Of Pictures). Parmi ces méthodes, on utilise celle qui offre les meilleures performances en terme de qualité des vidéos reçues. La méthode sélectionnée remplace l'image perdue par l'image suivante correctement reçue. Enfin, la distorsion des vidéos reçues est mesurée par le Peak Signal to Noise Ratio (PSNR) moyen de la composante luminance et des deux composantes de chrominance.

2.2 Précodeurs dans un contexte de transmission vidéo

Dans ce papier nous étudions la transmission de vidéos H.264/SVC sur un canal MIMO impliquant trois précodeurs diagonaux (Max-SNR, WF, QoS) et un précodeur non diagonal $E-d_{min}$. Ces précodeurs nécessitent la connaissance parfaite du Tx-CSI et permettent d'optimiser un critère propre à chaque précodeur. Par exemple, les précodeurs Max-SNR, WF et $E-d_{min}$ optimisent respectivement le RSB, la capacité du canal et la distance euclidienne. Quant au précodeur QoS il permet de maintenir tout rapport de RSB entre les sous canaux SISO. On définit un système MIMO avec n_T antennes à l'émission et n_R antennes à la réception, noté système MIMO $(n_T \times n_R)$. L'équation correspondante à ce système est donnée par

$$y = GHFs + Gn \quad (1)$$

où s et y représentent respectivement les vecteurs émis et reçus de taille $(b \times 1)$ avec $b = \text{trace}(H) < \min(n_R \times n_T)$, H la matrice du canal de taille $(n_R \times n_T)$, F la matrice de précodage de taille $(n_T \times b)$, G la matrice de décodage de taille $(b \times n_R)$ et n le vecteur de bruit additif de taille $(n_R \times 1)$.

L'étape commune à ces précodeurs est appelée transformation virtuelle. Cette étape permet de diagonaliser le canal et blanchir le bruit. Elle est généralement réalisée par la méthode de décomposition en valeurs singulières. Après la transformation virtuelle du canal, le système MIMO s'écrit sous la forme suivante

$$y = G_d H_v F_d s + G_d n_v \quad (2)$$

où F_d et G_d sont respectivement les matrices de précodage et de décodage, $H_v = G_v H F_v$ est la matrice des valeurs singulières, et $n_v = G_d n$ est le vecteur de bruit de covariance $R_{n_v} = I_b$ (I_b est la matrice identité de taille b).

La puissance du système MIMO doit être limitée à la puissance totale d'émission E_T

$$\|F_d\|_F^2 = E_T \quad (3)$$

avec $\|\cdot\|_F$ est la norme de Frobenius. Les matrices G_v et F_v sont unitaires et sélectionnées afin de blanchir le bruit, diagonaliser le canal et réduire sa dimension à b . Ainsi, le canal MIMO est décomposé en b sous canaux SISO

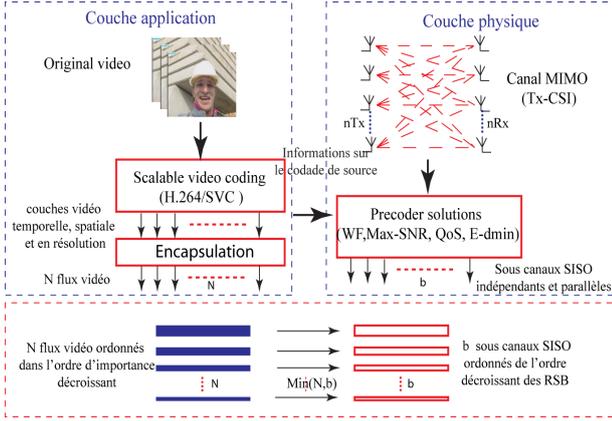


Figure 1 – Stratégie de transmission temps réel du flux vidéo H.264/SVC entre la couche application et la couche physique

indépendants et parallèles représentés par la matrice diagonale H_v

$$H_v = \text{diag}(\sigma_0, \sigma_1, \dots, \sigma_{b-1}) \quad (4)$$

H_v est appelé la matrice du canal virtuel, et la valeur σ_i^2 représente le RSB du sous canal i . Le gain total (σ) d'un système MIMO ($n_T \times n_R$) est calculé par

$$\sigma = \sqrt{\sum_{i=0}^{b-1} \sigma_i^2} \quad (5)$$

Les précodeurs diagonaux sont définis par leurs matrices diagonales F_d et G_d . Quant au précodeur $E-d_{min}$ ses matrices de précodage et décodage ne sont plus diagonales. Ce précodeur offre les meilleures performances en terme de TEB comparé aux précodeurs diagonaux [8]. Le précodeur QoS permet une grande flexibilité dans l'allocation de puissance aux sous canaux SISO, et offre le meilleur compromis entre complexité, performance de TEB et flexibilité [7]-[8]. Ainsi, il nous paraît le plus approprié pour une transmission temps réel d'un flux vidéo scalable.

2.3 Stratégie de transmission proposée

Le schéma proposé dans ce papier est composé de deux couches : couche application et couche physique. Les flux vidéo (ρ_i) alimentent la couche physique qui traite chaque flux vidéo d'une manière spécifique selon son importance. Le concept général de notre couche physique est similaire à celui de la couche physique de la norme IEEE802.11n. Les précodeurs présentés dans la section précédente sont implémentés, testés puis intégrés dans la couche physique IEEE802.11n. On utilise les codes correcteurs d'erreurs Low Density Parity Check (LDPC) pour protéger les flux vidéo avec une stratégie EEP (Equal Error Protection). Ainsi, les flux vidéos sont traités par tous les blocs de la chaîne de transmission IEEE802.11n offrant en sortie des trames de taille constante prêtes à la transmission. Au niveau du récepteur, un algorithme basé MV est utilisé pour décoder les symboles reçus, qui sont par la suite,

démodulés, dé-entrelacés et corrigés par le code correcteur d'erreurs LDPC. Après le processus de correction, tous les paquets vidéo NAL transportés sur des trames contenant des erreurs résiduelles seront perdus. Le reste des paquets vidéo, reçus sans erreur, alimentent la couche application pour la reconstruction de la vidéo reçue.

Dans ce papier nous adoptons une stratégie de transmission efficace entre le codeur H.264/SVC et les précodeurs. D'une part, au niveau de la couche application, le codeur H.264/SVC fournit N flux vidéo ($\rho_0, \rho_1, \dots, \rho_{(N-1)}$) ordonnés dans un ordre d'importance décroissant. D'autre part, au niveau de la couche physique, les précodeurs décomposent le canal MIMO en sous canaux SISO ordonnés par ordre de RSB décroissant ($f_0^2 \sigma_0^2, f_1^2 \sigma_1^2, \dots, f_{b-1}^2 \sigma_{b-1}^2$) avec f_i le coefficient de précodage affecté au sous canal SISO i . Ainsi, la stratégie de transmission proposée consiste à affecter les flux vidéos aux sous canaux SISO en considérant une association fine entre l'importance du flux vidéo et la puissance qui lui est allouée par le précodeur. En effet, le flux vidéo (ρ_i) est transmis sur le sous canal SISO de RSB égal à $f_i^2 \sigma_i^2$. La figure 1 illustre la stratégie de transmission adoptée entre la couche application et la couche physique.

3 Résultats de simulation

3.1 Contexte de simulation

Confi.	Zone 1	Zone 2	Zone 3
état du canal	Mauvais	Moyen	Bon
RSB relatives	(0.7, 0.3, 0, 0)	(0.4, 0.3, 0.3, 0)	(0.3, 0.3, 0.25, 0.15)

Tableau 1 – Valeurs relatives de RSB utilisées par le précodeur QoS à chaque état de canal

Dans cette section nous présentons les résultats de simulations afin de montrer les performances du schéma proposé dans un environnement réaliste Outdoor. Nous utilisons un modèle de propagation déterministe 3-D [10], développé au laboratoire XLIM-SIC, pour calculer les RIC d'un trajet de 180 m sur le campus de l'Université de Poitiers. L'émetteur reste statique et le récepteur se déplace sur la trajectoire avec une vitesse de 10 m/s. Les paramètres utilisés au niveau de chaque couche sont présentés ci-dessous :

1. La couche application : on utilise la séquence vidéo de test Foreman au format QCIF 176×144 pixels/image. Les 288 premières images de la vidéo sont codées avec le codeur H.264/SVC à 30 images/seconde. Le codeur de vidéo est paramétré pour fournir 4 couches de résolutions temporelles ($N_T = 4, GOP = 8$) et 4 couches de résolutions en qualité ($N_Q = 4$). Ici, les 4 couches en résolution qualité forment quatre flux vidéo de même taille à transmettre ($N = 4$) ($\rho_0, \rho_1, \rho_2, \rho_3$). Ainsi, le flux ρ_i est exploitable seulement si tous les flux ρ_j d'identifiant $j < i$ sont correctement reçus.
2. La couche physique : les paramètres de transmis-

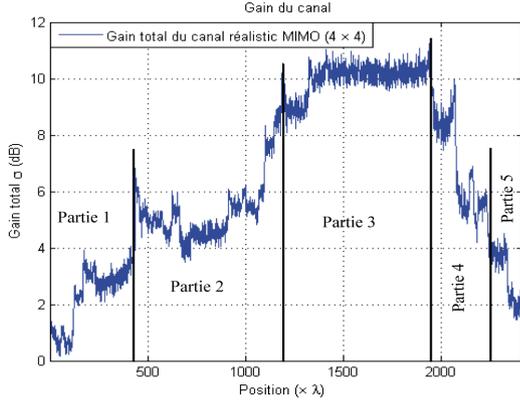


Figure 2 – Variation du gain en fonction de la position sur le trajet

sion (puissance d'émission, fréquence porteuse, système OFDM) sont fixés suivant la norme IEEE802.11n. Nous considérons un canal MIMO (4×4). La distance entre les antennes est fixée à 0.4λ , où λ représente la fréquence porteuse. Cela correspond généralement à la corrélation minimale entre les antennes [9]. Nous considérons la modulation 4-QAM pour tous les flux vidéo, qui sont protégés par le code correcteur LDPC de rendement $1/2$. Cette configuration de couche physique nous permet de transmettre à un débit utile de $12 \times b$ Mb/s.

Suivant l'état du canal de transmission, nous considérons trois configurations du précodeur QoS. Ces configurations tiennent en compte, d'une part, de l'état instantané du canal en sélectionnant uniquement un sous ensemble des canaux SISO, et d'autre part, de l'importance des flux vidéo en adoptant une stratégie UEP. Les valeurs relatives de RSB utilisées à chaque état de canal sont données dans le Tableau 1, suivant trois zones sur lesquelles nous revenons dans la section suivante. Nous voulons notifier que cette configuration est choisie expérimentalement parmi d'autres configurations moins performantes, et qu'elle ne représente en aucun cas la solution optimale.

3.2 Résultats et discussions

Le gain total σ du canal réaliste MIMO (4×4) en fonction de la position du récepteur est illustré sur la figure 2. Suivant les valeurs du gain total, on peut classer la trajectoire du récepteur en trois zones d'études. La première zone représente un canal de mauvaise qualité et couvre les parties 1 et 5 de la trajectoire. La zone 2 représente un canal de qualité moyenne et couvre les parties 2 et 4 de la trajectoire. Quant à la partie 3 de la trajectoire, elle représente un canal de bonne qualité et correspond à la zone 3. La figure 3 illustre les performances du schéma proposé le long de la trajectoire du récepteur. Ces courbes montrent la variation du PSNR des vidéos reçues en fonction du déplacement du récepteur sur la trajectoire pour les différents précodeurs utilisés. Nous constatons les hautes performances du précodeur QoS en utilisant la configura-

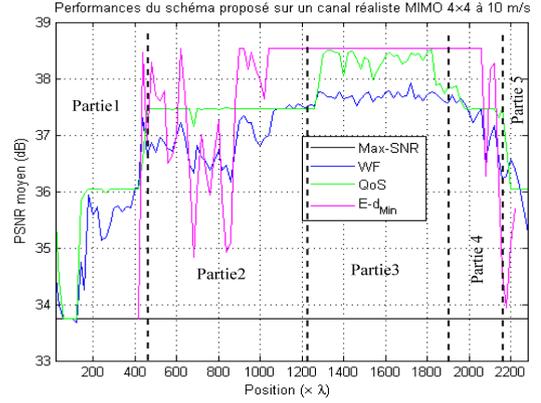


Figure 3 – Performance du schéma proposé sur un canal MIMO (4×4) réaliste

tion appropriée à chaque zone du canal (Tableau 1). Le précodeur Max-SNR quant à lui assure une réception sans erreur du flux vidéo de qualité de base et cela quelque soient les conditions du canal. Nous pouvons aussi remarquer l'adaptation des précodeurs WF et QoS aux variations du canal de transmission. En effet, la qualité des vidéos reçues en utilisant ces deux précodeurs est améliorée suivant l'augmentation de la puissance du signal reçue. Dans de mauvaises conditions du canal (zone 1), le précodeur $E-d_{min}$ n'arrive même pas à assurer une bonne réception du flux vidéo de qualité de base, alors qu'en zone 3 ce précodeur offre les meilleures performances en recevant correctement les quatre flux vidéos. Nous constatons aussi que les précodeurs adoptants une stratégie UEP, tels que QoS et WF assurent en moyenne les meilleures performances. De plus, une considération plus fine du contenu de l'information transmise améliore davantage les performances du précodeur QoS par rapport à celles du précodeur WF. Le gain en terme de qualité visuelle apportée par le précodeur QoS par rapport au précodeur WF est illustré sur les figures 4(a) et 4(b). Cette figure montre clairement que la qualité visuelle de la vidéo reçue avec le précodeur QoS est nettement meilleure à celle reçue avec le précodeur WF.

Nous pouvons aussi constater sur la figure 3 que les performances du précodeur $E-d_{min}$ ne sont pas calculées dans la zone 1. Pour cette raison, ses performances moyennes sur la trajectoire ne peuvent être objectivement comparées avec les trois autres précodeurs (diagonaux). Ainsi, sur les Figures 5(a) et 5(b) nous illustrons séparément à la fois les performances moyennes du schéma proposé sur la trajectoire ainsi que sa robustesse vis-à-vis des erreurs d'EC pour respectivement les précodeurs diagonaux et le précodeur non-diagonal $E-d_{min}$. La figure 5(a) montre que le précodeur QoS assure en moyenne les meilleures performances comparées aux autres précodeurs diagonaux. De plus, on remarque que tous les précodeurs diagonaux restent très robustes aux erreurs d'EC et maintiennent les mêmes performances avec et sans erreurs d'estimation. Cependant, sur la 5(b) nous remarquons que le précodeur E-

d_{min} devient très sensible aux erreurs d'EC. En effet, les erreurs d'EC à une vitesse de 10 m/s entraînent une perte entre 1 et 2 dB en termes de PSNR sur les vidéos reçues.



(a) PSNR = 35.67 dB



(b) PSNR = 34.11 dB

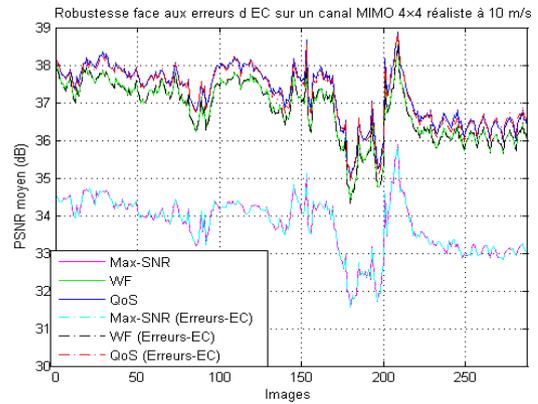
Figure 4 – Illustration de la qualité visuelle des vidéos reçues par les précodeurs QoS (haut) et WF (bas) sur un canal MIMO réaliste (zone 1)

4 Conclusion

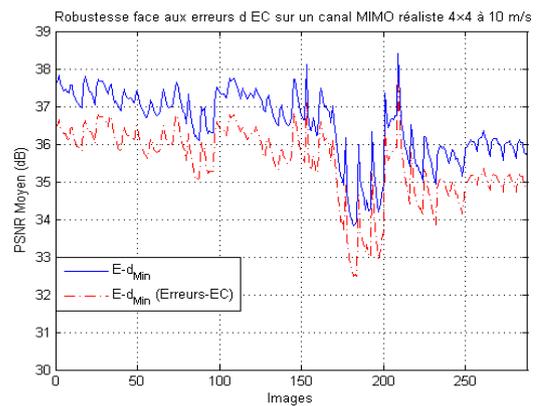
Dans ce papier nous avons développé une solution prometteuse pour la transmission temps réel d'un flux vidéo H.264/SVC sur un canal MIMO. La hiérarchie fournie par le codeur vidéo H.264/SVC est efficacement exploitée par quatre précodeurs. Nous avons montré que les précodeurs les plus performants en terme de TEB ($E-d_{min}$) ne sont pas nécessairement les plus appropriés pour une transmission de vidéo. Cependant, une meilleure prise en compte du contenu des flux vidéo (QoS) permet de garantir la meilleure qualité de service. De plus, contrairement au précodeur $E-d_{min}$, le précodeur QoS reste très robuste aux erreurs d'EC même avec des vitesses de déplacement relativement élevées de l'ordre de 10 m/s.

Ces travaux montrent clairement qu'une adaptation adéquate de l'allocation de puissance en fonction de l'importance du flux vidéo et l'état instantané du canal améliore

la qualité de service. Ainsi, dans la suite de nos travaux nous développerons un précodeur QoS adaptatif qui minimise la distorsion de la vidéo reçue. En se basant sur un critère pertinent, cette solution considèrera conjointement les couches application et physique et permettra d'atteindre la solution d'allocation de puissance optimale quelque soient les conditions du canal.



(a) Précodeurs diagonaux



(b) Précodeur non-diagonal

Figure 5 – Robustesse du schéma proposé vis-à-vis les erreurs d'EC sur un canal réaliste MIMO (4×4) à 10 m/s

Remerciement

Ce travail entre dans le cadre du projet ANR CAIMAN.

Références

- [1] T. Wiegand, G. Sullivan, J. Reichel, H. Schwarz, M. Wien, *Joint Draft of SVC Amendment*, Joint Video Team (JVT), Doc. JVT-W201, San Jose, CA, 2007.
- [2] M. K. Jubran, M. Bansal, L. P. Kondi and R. Grover, *Accurate Distortion Estimation and Optimal Bandwidth Allocation for Scalable H.264 video transmission Over MIMO systems*, IEEE Trans. Image Processing, Vol. 10, no. 1 pp. 106-116, Jan. 2009.
- [3] Z. Ji, Q. Zhang, W. Zhu, Z. Guo, J. Lu, *Power efficient MPEG-4 FGS video transmission over MIMO-OFDM*

systems, Proc. IEEE ICC, Vol. 5, pp. 3398-3402, May 2003.

- [4] D. Song, C. W. Chen, *Scalable H.264/AVC Video Transmission over MIMO Wireless Systems with adaptive channel selection based on partial channel information*, IEEE Trans. Circuits Syst. Video Technol., Vol. 17, no. 9, pp. 1218-1226, Sep. 2007.
- [5] D. Song, C. W. Chen, *Maximum-throughput delivery of SVC-based video over MIMO systems with time-varying channel capacity*, Elsevier Journal of Visual Commun. And Image Represent Vol. 19, no. 8, pp. 520-528, Dec. 2008.
- [6] P. Stoica and G. Ganesan, *Maximum-SNR spatial-temporal formatting designs for MIMO channels*, IEEE Trans. Signal Processing, vol. 50, no. 12, pp. 3036-3042, Dec. 2002.
- [7] H. Sampath, P. Stoica, and A. Paulraj, *Generalized Linear Precoder and Decoder Design for MIMO Channels Using the Weighted MMSE Criterion*, IEEE Trans. Commun. Vol. 49, no. 12, pp. 2198-2206, Dec. 2001.
- [8] B. Vrigneau, J. Letessier, P. Rostaing, L. Collin and G. Burel, *Extension of the MIMO Precoder based on the Minimum Euclidean Distance : a cross-form matrix*, IEEE Signal Processing Vol. 2, no. 2, pp. 135-146, May 2008.
- [9] C. Pereira, Y. Pousset, R. Vauzelle, P. Combeau, *Sensitivity of the MIMO Channel Characterization to the Modeling of the Environment*, IEEE Transactions on Antennas and Propagation, Vol. 57, no. 4, pp. 1218-1227, Mar. 2009.
- [10] F. Mora, L. Aveneau, *Optimized scanning of a visibility graph data structure for efficient ray-tracing*, ECWT, Paris, Oct. 2005.
- [11] IEEE Standard for Information Technology-Part 11 : *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications Amendment : Enhancements for Higher Throughput (802.11n)*, 2009.