HAL
open science

Compression et interactivite : etude de la navigation au recepteur

Aline Roumy, Thomas Maugey

To cite this version:

Aline Roumy, Thomas Maugey. Compression et interactivite : etude de la navigation au recepteur. colloque international francophone de Traitement du Signal et de l'Image, GRETSI, Sep 2015, Lyon, France. hal-01208129

HAL Id: hal-01208129

https://inria.hal.science/hal-01208129

Submitted on 5 Oct 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Compression et interactivité : étude de la navigation au récepteur

Aline ROUMY, Thomas MAUGEY

INRIA, Campus Beaulieu, 35042 Rennes, France
aline.roumy@inria.fr, thomas.maugey@inria.fr

Résumé – Nous étudions l'interactivité pour la compression de sources multiples dans un réseau de capteurs, où l'utilisateur peut choisir librement la source qu'il souhaite observer. Ce problème, que nous appelons compression avec navigation interactive, doit (i) exploiter la corrélation entre les sources, (ii) compresser les sources hors ligne et donc indépendamment de la navigation d'un utilisateur et (iii) transmettre sans délai à réception d'une requête, les données relatives à la source demandée. Nous dérivons les bornes théoriques optimales de compression et transmission et montrons que, contrairement aux résultats de Draper et al [1], le coût de l'interactivité ne se mesure que dans le coût de stockage et non dans le débit transmis. Enfin, appliqués au cas de la TV avec point de vue libre, nos résultats montrent qu'il est encore possible d'améliorer les schémas de compression pratiques existants.

Abstract – We consider the problem of interactive source coding in a sensor network, where the user can freely choose the source (sensor) he wants to observe. Such an interactive navigation scheme must (i) take into account the possible correlation between the sources (ii) encode the data off-line and therefore independently of any navigation and finally (iii) transmit without any delay upon request of a source. We derive the information theoretical limits in terms of rate and storage of such interactive source coding scheme. Moreover we show a counterintuitive result: freely choosing its source at the receiver incurs a loss in terms of storage only and not in the transmission rate. This corrects a result of Draper et al.[1] which claims otherwise. Finally, as an application, we consider the problem of video compression with free viewpoint interactivity and derive the gap between the optimal scheme proposed here and interactive schemes proposed so far.

La multiplication du nombre de capteurs dans notre environnement soulève de nombreuses questions quant à la conservation (compression) et la transmission des données collectées qui, par nature, sont de grande taille et avec des dépendances complexes. Traiter la totalité de ces données n'étant pas toujours possible (l'interprétation nécessitant des coûts calculatoires prohibitifs, capacité limitée de stockage du récepteur) ou pas souhaité (intérêt de l'utilisateur pour un seul capteur à la fois), il est constaté que, dans la plupart des scénarios multi-capteurs, le traitement des différentes sources ne s'effectue pas simultanément. Ainsi, de plus en plus de schémas offrent une *interactivité* à l'utilisateur, lui permettant de choisir la source qu'il souhaite recevoir, et donc d'effectuer une libre *navigation* parmi les sources. Cette liberté, si elle apporte une nouvelle fonctionnalité à l'utilisateur, introduit, au contraire, une nouvelle contrainte pour la compression des données en vue d'un stockage intelligent pouvant fournir à la demande, et donc avec le minimum de délai, le choix de l'utilisateur. C'est cette question de compression avec interactivité que nous étudions ici.

La notion d'interactivité a déjà été abordée en théorie de l'information, mais sous un angle différent. L'interactivité peut signifier une communication bi-directionnelle, où deux sources souhaitent s'échanger leurs données avec [2] ou sans pertes [3]. Ou bien, l'interactivité signifie une aide apportée à une communication mono-directionnelle, sous la forme d'une voie de retour, qui ne permet pas de transmettre plus d'informations mais permet de rendre le code plus simple (canal binaire à effacement [4]) ou avec une plus grande décroissance de la probabilité d'erreur ([5]). Dans ces deux cas, l'interactivité implique des communications symétriques (avec des volumes de données échangées du même ordre de grandeur). En revanche, nous nous intéressons au cas où l'interactivité implique une communication asymétrique : à la communication des données de la source, est ajouté une transmission relative au choix de la source et donc de volume négligeable. Dans la suite, un tel schéma est appelé schéma *avec navigation interactive*.

Les caractéristiques inhérentes aux schémas avec navigation interactive sont multiples. La première est que les sources peuvent être corrélées entre elles (ce qui est le cas dans la plupart des scénarios multi-capteurs). Cela implique que le débit de compression d'une source donnée peut être diminué grâce à la connaissance d'une autre source, gardée en mémoire par l'observateur durant sa navigation. On en vient au deuxième aspect. L'interactivité étant le choix d'un utilisateur, la diffusion de l'information se fait pour cet utilisateur uniquement et donc en streaming. Ainsi, les sources sont compressées hors ligne sans connaître le schéma de navigation puis stockées sur un serveur. L'utilisateur interagit avec celui-ci (et non avec les encodeurs directement) et se voit communiquer un flux particulier extrait de l'information stockée. Enfin, le dernier aspect concerne la réactivité du serveur de données à une requête. Cela empêche tout simplement l'encodage *en-ligne* ou de manière similaire *le ré-encodage*, qui consiste à encoder une source juste après la requête de l'observateur, en fonction de sa navigation passée (et donc des sources qu'il a encore en mémoire). Ce type de schéma permettrait en effet d'optimiser les débits, mais est, en pratique, inconcevable, surtout lorsque le nombre d'utilisateurs est important.

Pour résumer, les difficultés des schémas avec navigation interactive viennent du fait que l'encodeur doit pouvoir exploiter

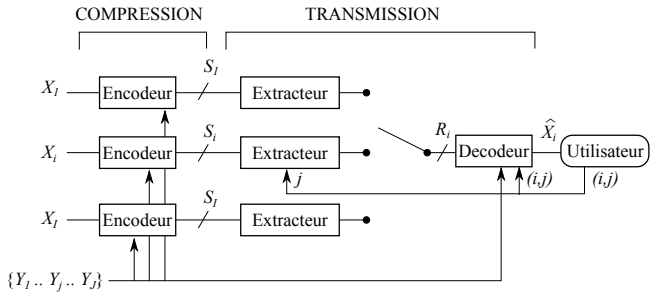

FIG. 1: Codage de sources avec navigation interactive.

la corrélation entre les sources, tout en ne sachant pas quelle sera l'information adjacente disponible au décodeur. Un tel schéma vise à *stocker* et *transmettre* un minimum d'information. Dans ce papier, nous calculons les bornes de débit et stockage atteignables pour un schéma avec navigation interactive et montrons que, contrairement aux résultats de [1], le coût de l'interactivité se mesure uniquement sur le stockage mais pas sur le débit transmis. Ce gain vient du fait que lorsqu'un utilisateur fait une requête auprès du serveur, le serveur a accès aux précédentes requêtes de ce même utilisateur. Enfin, nous étudions le cas de la TV avec point de vue libre et montrons qu'il est encore possible d'améliorer les schémas de compression pratiques existants.

1 Compression avec navigation interactive

Soit I sources, notées $\{X_i\}_{i\in[1,I]}$. L'utilisateur choisit la source qu'il souhaite recevoir parmi l'ensemble $[1, I]$. Comme l'utilisateur peut changer de source à tout moment et que les sources sont corrélées, le décodeur peut réaliser une estimation de la nouvelle source souhaitée X_i à partir des sources précédemment demandées qui constituent le *chemin de navigation* noté j. Cette estimée est modélisée par une information adjacente (IA) Y_j parfaitement synchronisée avec la source nouvellement demandée. Cette IA dépendant du chemin navigation, il existe autant d'informations adjacentes Y_j que de chemins de navigation, nombre que nous notons J.

Le problème de la navigation interactive consiste à stocker pour transmettre l'information à moindre coût sous hypothèse de grande réactivité du serveur. Aussi, pour satisfaire cette contrainte de délai, il n'est pas possible de décompresser/compresser la source demandée en fonction de la requête. Au contraire, on souhaite stocker (i.e. compresser) la source demandée X_i en tenant compte d'une incertitude quant aux requêtes passées. La Figure 1 présente le schéma de transmission. L'encodeur comprend deux phases. Tout d'abord, la phase de compression, où le chemin de navigation est inconnu. Cette incertitude est montrée sous la forme d'un ensemble d'IA $\{Y_1 \dots Y_J\}$ disponible à l'encodeur sans savoir laquelle de ces informations sera réellement disponible au décodeur. Suit une phase de transmission, où l'émetteur connaît parfaitement quelles vues (passées et présentes) ont été demandées. Pour satisfaire la contrainte de faible délai, l'émetteur est seulement autorisé à extraire des données stockées, l'information à transmettre.

Sans perte de généralité et pour simplifier les notations, nous fixons la vue courante demandée et la notons X. Le schéma de

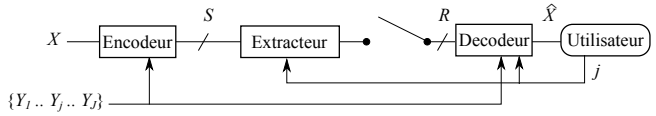

FIG. 2: Compression incrémentale et universelle avec IA partielle à l'encodeur et parfaite au décodeur.

transmission se simplifie et devient la Figure 2. Ce schéma de transmission présente des similarités avec des schémas classiques de codage de source. Par exemple, dans le cas de la compression conditionnelle, une source X doit être compressée sous l'hypothèse qu'une IA Y, corrélée à la source X, est disponible à l'encodeur et au décodeur. La source peut alors être compressée sans perte au taux de $H(X|Y)$. Dans notre schéma, si la réalisation de l'IA est disponible au décodeur, en revanche elle ne l'est pas à l'encodeur (en effet, l'encodeur connaît l'ensemble des réalisations possibles mais pas celle qui sera choisie au niveau du décodeur). Notre schéma est donc plus proche d'un codage de source distribué, où la réalisation de l'IA Y est connue du récepteur uniquement. Si la loi de la corrélation est connue à l'encodeur, alors la source peut être compressée au taux de $H(X|Y)$ [6]. Ce même débit serait atteignable dans notre schéma, si la source pouvait être réencodée à chaque requête. En revanche, la compression doit être réalisée sans connaissance de l'IA et l'incertitude sur la réalisation de l'IA à l'encodeur engendre une incertitude quant à la loi jointe (la corrélation) des sources. Dans ce cas, le débit atteignable est supérieur à $H(X|Y)$ et sa valeur dépend du modèle de corrélation [7]. Enfin, les schémas classiques ne tiennent pas compte de la possibilité d'extraire une partie du flux compressé. La compression avec navigation interactive (Fig. 1 et 2) est donc un nouveau problème, pour lequel nous calculons les bornes théoriques de compression dans la section suivante.

2 Modèle de source

Les taux de compression dépendent des caractéristiques des sources. Aussi, nous précisons dans cette section la source considérée et définissons formellement le problème traité.

Définition 1 (Source statique [7, Def. 3])**.** *Une* source statique (X, Y) *est une suite de variables aléatoires i.i.d. de distribution jointe* $\mathbb{P}(X, Y)$ *choisie aléatoirement dans l'ensemble fini de distributions* $\{\mathbb{P}(X, Y_j)\}_{j\in[1,J]}$*. Cet ensemble de distributions est muni d'une loi* $\mathbb{P}(Y = Y_j) = p_j$*, qui correspond au choix de la distribution* $\mathbb{P}(X, Y_j)$*. Aussi, la source statique est complètement déterminée par* $\mathcal{P} = \{\mathbb{P}(X, Y_j), p_j\}_{j\in[1,J]}$.

Avant de traiter le problème général de la compression avec navigation interactive (Fig. 1), nous considérons le sous- problème (Fig. 2) défini ci-dessous :

Problème 2. *Soit* (X, Y) *une source statique définie par* $\mathcal{P} = \{\mathbb{P}(X, Y_j), p_j\}_{j\in[1,J]}$. La compression incrémentale et universelle avec IA partielle à l'encodeur et parfaite au décodeur *(voir Figure 2) consiste à compresser la source* X*, sachant qu'une IA* $Y_j \in \{Y_1 \dots Y_J\}$ *sera disponible au décodeur. De plus, l'encodeur connaît la réalisation de chaque IA* Y_j *pour chaque* j *(sans connaître la valeur de* j*) et doit fournir une*

représentation incrémentale de la source X. Ainsi, lorsque la valeur de j est révélée à l'émetteur, celui-ci fournit une représentation de X en extrayant une partie de la représentation préalablement compressée. E, la loi $\mathcal{P}$ n'est connue ni de l'émetteur ni du récepteur, rendant le schéma universel.

Définition 3 (Paire débit-stockage atteignable)**.** *Soit (X, Y) une source statique définie par $\mathcal{P} = \{\mathbb{P}(X, Y_j), p_j\}_{j\in[1,J]}$ qui fournit une suite de n symboles. La* paire débit-stockage (R, S) *est* atteignable *pour le* problème 2 *si, quelque soit la requête $j \in [1, J]$, il existe une suite de (codeur/ extracteur de bits/ décodeur) indexée par n, la longueur de la séquence traitée, qui peut estimer au décodeur la source X sans erreur (i.e. $\hat{X} = X$) lorsque n tend vers l'infini.*

Définition 4 (La région débit-stockage)**.** *La* région débit-stockage *atteignable pour le* problème 2 *est la fermeture de l'ensemble des paires débit-stockage atteignables (R, S).*

3 Coût de l'interactivité

Cette section établit la région débit–stockage pour la compression interactive de données tel que définie dans le problème 2.

Théorème 5 (Région débit–stockage)**.** *Soit (X, Y) une source statique définie par $\mathcal{P} = \{\mathbb{P}(X, Y_j), p_j\}_{j\in[1,J]}$, qui doit être compressée selon les hypothèses du problème 2 (voir Figure 2). Soit S le stockage nécessaire pour que la source X soit reconstruite quelque soit l'IA Y disponible au décodeur, où $Y \in \{Y_j\}_{j\in[1,J]}$. Soit R le débit nécessaire à transmettre lorsque l'émetteur reçoit la requête j, (qui signifie que Y_j est disponible au décodeur). La région débit-stockage est l'ensemble des (R, S) qui vérifient*

$$R \geq H(X|Y_j) \quad (1)$$
$$S \geq \max_{k:\, p_k>0} H(X|Y_k) \quad (2)$$

où $H(X|Y_j)$ est l'entropie conditionnelle de X sachant Y_j. Cette région correspond à la zone blanche de la Figure 3.

Discussion: Draper *et al.*[1] établit que la région débit stockage est définie par

$$R \geq \max_{k:\, p_k>0} H(X|Y_k) \quad (3)$$
$$S \geq \max_{k:\, p_k>0} H(X|Y_k) \quad (4)$$

Or, lorsqu'un utilisateur fait une requête auprès du serveur, ce dernier a accès aux précédentes requêtes de ce même utilisateur. Autrement dit, au moment de la transmission, l'émetteur connaît j et donc l'information adjacente Y_j disponible au décodeur.

Preuve: ***Atteignabilité.*** [8] propose un schéma de codage incrémental pour le problème de codage de source distribué universel. Une source X est compressée sachant qu'une IA est disponible au décodeur uniquement. L'encodeur transmet un mot de code incrémental et dès que le décodeur peut décoder il en informe l'encodeur (ACK). [8] montre que le débit $H(X|Y_j)$ est atteignable si Y_j est disponible au décodeur. Dans notre contexte, le mot de code incrémental est stocké jusqu'à la taille maximale nécessaire soit $\max_{k:\, p_k>0} H(X|Y_k)$. Cette longueur est connue du serveur car il dispose de l'ensemble des réalisations de toutes les IAS. Puis, à réception de la requête j, il délivre le code jusqu'à la longueur $H(X|Y_j)$. Enfin, le code [8] atteint ces débits sous hypothèse universelle i.e. l'ensemble de lois $\mathcal{P}$ n'est connu ni de l'émetteur ni du récepteur. On en déduit que $(R, S) = (H(X|Y_j), \max_{k:\, p_k>0} H(X|Y_k))$ est atteignable. Ce point porte l'étiquette (4) dans la Figure 3.

Réciproque: borne externe. Nous utilisons ici les réciproques pour des schémas de codage disposant de plus d'information que dans notre schéma.

• $R \geq H(X|Y_j)$. Preuve par l'absurde. Soit le schéma de codage, où l'IA Y_j et les lois jointes sont connues de l'encodeur et du décodeur. Si $R < H(X|Y_j)$ la source X ne peut être reconstruite sans erreur d'après la réciproque du théorème de Slepian Wolf [6]. Aussi, la source ne peut être reproduite pour notre schéma qui dispose de moins d'informations.

• $S \geq \max_{k:\, p_k>0} H(X|Y_k)$. Preuve par l'absurde, comme ci-dessus. □

Une conséquence directe du Th. 5 est l'évaluation du coût de l'interactivité :

Corollaire 6. *(*Coût de l'interactivité*). Le coût de l'interactivité qui permet à un utilisateur de choisir librement sa source se mesure en stockage et non en débit transmis et vaut (seulement) $\Delta S = \max_{k:\, p_k>0} H(X|Y_k) - H(X|Y_j)$.*

Preuve: Sans interactivité, l'IA est connue de l'encodeur et du décodeur. Le schéma est conditionnel et $S = R = H(X|Y_j)$ est atteignable. Ce point porte l'étiquette (1) dans la Fig. 3. □

Nous généralisons maintenant au cas de sources multiples (Fig. 1)

Théorème 7 (Région débit–stockage: cas de sources multiples en temps partagé (time sharing))**.** *Soit une multisource (X, Y) définie par $\mathcal{P} = \{\mathbb{P}(X_i, Y_j), p_{ij}\}_{j\in[1,J]}$ où $p_{ij} = \mathbb{P}(X = X_i, Y = Y_j)$. Supposons une émission en temps partagé de (X, Y) : pour n symboles, np_{ij} symboles consécutifs de la source X_i sont émis et np_{ij} symboles consécutifs de l'IA Y_j sont disponibles au décodeur de manière synchronisée avec la source X_i. De plus,ces np_{ij} symboles sont générés de manière i.i.d. selon la loi $\mathbb{P}(X_i, Y_j)$. Ces symboles sont traités (i.e. compressés puis transmis selon la requête du décodeur) avec les hypothèses du problème 2. Soit S le stockage nécessaire pour que X soit reconstruite quelque soit les requêtes (i, j) de l'utilisateur. Soit R le débit nécessaire à transmettre lorsque les requêtes (i, j) sont disponibles à l'encodeur, en extrayant le flux nécessaire du flux stocké. La région débit–stockage atteignable est:*

$$R \geq \sum_{i,j} p_{ij} H(X_i|Y_j) \quad (5)$$
$$S \geq \sum_{i} \max_{k:\, p_k>0} H(X_i|Y_k) \quad (6)$$

Le débit d'une navigation est donc la moyenne des débits des sources demandées (5).

Preuve: Pour chaque source i, nous utilisons le code incrémental défini dans la preuve du théorème 5 □

4 Exemple de schéma à navigation interactive: la TV à point de vue libre

Lors de la restitution d'une vidéo multi-vues (provenant de la capture d'une scène 3D par un réseau de caméras) à un util-

isateur, la visualisation se fait sur un écran 2D, stéréoscopique, ou autostéréoscopique (affichant plusieurs points de vue simultanément). Les systèmes de rendu existant limitent donc le nombre de vues qui peuvent être vues simultanément et l'utilisateur sélectionne en direct les vues qu'il souhaite observer. Ce scénario s'appelle la *TV à point de vue libre* [9] et constitue un domaine de recherche et développement populaire dans la communauté de codage vidéo.

Le schéma de la Fig. 1 peut modéliser un système de TV à point de vue libre. En effet, une source X_i correspond à une vue d'une scène 3D. Cette source délivre des suites de symboles, où chaque symbole correspond à un pixel et une suite de symboles à une image. L'utilisateur en changeant de vue, réalise un partage de temps entre les vues, comme défini dans le Th. 7. Les images successives d'une vidéo ainsi que les images capturant le même instant sous différents angles étant fortement corrélées, cette dépendance spatio-temporelle est modélisée par une information adjacente notée Y [10]. Plus précisément, Y est la meilleure estimation de la source courante demandée,

Si l'on rapproche les résultats théoriques obtenus en Sec. 3 et les méthodes de transmission existantes pour la TV à point de vue libre, nous pouvons faire l'analyse suivante. La Figure 3 compare la paire débit–stockage atteignable à des systèmes de codage existants. Tout d'abord (1) correspond au cas, où l'information adjacente est parfaitement connue de l'encodeur et du décodeur i.e. au moment du stockage la navigation entre les vues est figée et il n'y a donc plus d'interactivité. C'est donc le point de référence pour calculer le coût de l'interactivité. Le point (2) correspond aux schémas [11, 12, 13] qui proposent de stocker a source X pour tous les chemins de navigation possibles afin de limiter le débit transmis. Bien qu'ils parviennent à atteindre le débit de transmission de (1), cette stratégie nécessite une forte capacité de stockage. Afin de limiter le stockage, [14] propose de stocker et transmettre le "pire cas", i.e. le nombre de bits correspondant à l'information adjacente la moins corrélée (voir (3) dans la Fig. 3). Nous avons montré dans ce papier, qu'un code incrémental serait en mesure d'améliorer les performances de débit de [14, 1] pour atteindre (4) et ainsi avoir un coût d'interactivité impactant uniquement le stockage. Cette amélioration est possible car, lorsqu'un utilisateur fait une requête auprès du serveur, le serveur a accès aux précédentes requêtes de ce même utilisateur. La transmission peut donc être faite avec connaissance de l'information adjacente disponible au niveau du récepteur.

5 Conclusion

Ce papier établit les bornes théoriques d'un schéma de compression multi-sources avec libre choix de la source par l'utilisateur. Ce nouveau problème, que nous appelons codage de source avec navigation interactive, trouve des applications dans les réseaux de capteurs mais aussi la TV à point de vue libre. Nous avons ainsi montré que le coût de l'interactivité se mesure en stockage et non en débit transmis, comme ce qui avait été établi précédemment dans [1].

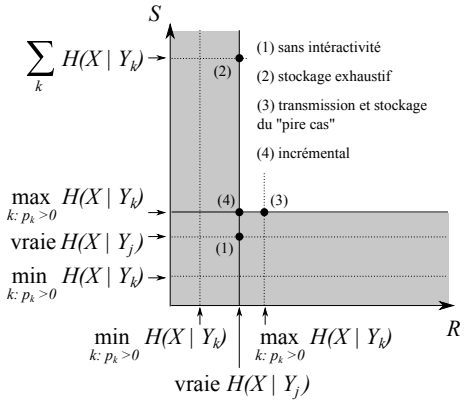

FIG. 3: Région débit-stockage (R, S) atteignable.

Bibliographie

[1] S. C. Draper and E. Martinian, "Compound conditional source coding, Slepian-Wolf list decoding, and applications to media coding," in *IEEE Int. Symp. Inform. Theory*, 2007.

[2] A. H. Kaspi, "Two-way source coding with a fidelity criterion," *IEEE Trans. on Inf. Th.*, vol. 31, no. 6, pp. 735 – 740, 1985.

[3] A. El Gamal and A. Orlitsky, "Interactive data compression," in *Symposium on Foundations of Computer Science*, 1984.

[4] C. Shannon, "The zero error capacity of a noisy channel," *IRE Trans. Inf. Theory*, vol. 2, no. 6, pp. 8–19, Nov 1956.

[5] J. P. M. Schalkwijk and T. Kailath, "A coding scheme for additive noise channels with feedbac – I: No bandwidth constraint," *IEEE Trans. on Inf. Th.*, vol. 12, pp. 172 – 182, Apr 1966.

[6] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Transactions on Information Theory*, vol. 19, no. 4, pp. 471–480, July 1973.

[7] E. Dupraz, A. Roumy, and M. Kieffer, "Source Coding with Side Information at the Decoder and Uncertain Knowledge of the Correlation," *IEEE Transactions on Communications*, vol. 62, no. 1, pp. 269 – 279, Jan. 2014.

[8] S. C. Draper, "Universal Incremental Slepian-Wolf Coding," in *Allerton Conf. on Communication, contro and computing*, 2004.

[9] M. Tanimoto, "FTV: Free-viewpoint television," *IEEE Signal Processing Magazine*, vol. 27, no. 6, pp. 555–570, Jul. 2012.

[10] L. Toni, T. Maugey, and P. Frossard, "Correlation-aware packet scheduling in multi-camera networks," *IEEE Trans. on Multimedia*, vol. 16, no. 2, pp. 496–509, 2014.

[11] Y. Liu, Q. Huang, S. Ma, D. Zhao, and W. Gao, "RD-optimized interactive streaming of multiview video with multiple encodings," *Journal on Visual Commun. and Image Repr.*, 2010.

[12] H. Kimata, M. Kitahara, K. Kamikura, and Y. Yashima, "Free-viewpoint video communication using multi-view video coding," *NTT Technical Review*, vol. 2, no. 8, pp. 21–26, Aug. 2004.

[13] S. Shimizu, M. Kitahara, H. Kimata, K. Kamikura, and Y. Yashima, "View Scalable Multiview Video Coding Using 3-D Warping With Depth Map," *IEEE Trans. on Circ. and Syst. for Video Technology*, vol. 17, no. 11, pp. 1485–1495, Nov. 2007.

[14] G. Cheung, A. Ortega, and NM. Cheung, "Interactive streaming of stored multiview video using redundant frame structures," *IEEE Trans. on Image Proc.*, vol. 3, no. 3, pp. 744–761, Mar. 2011.