

Description multiple par transformation de Radon discrète exacte

Benoit Parrein, Nicolas Normand, Dominique Barba

Institut de Recherche en Communication et Cybernétique de Nantes, UMR 6597
1 rue de la Noé, BP 92101, 44321 Nantes¹
benoit.parrein@polytech.univ-nantes.fr, nicolas.normand@polytech.univ-nantes.fr
dominique.barba@polytech.univ-nantes.fr

Résumé – On propose dans ce papier un schéma de codage conjoint source-canal avec une mise en œuvre de la description multiple via un opérateur de projection. Le choix des angles et de la topologie du support est précisé pour le cas de source non scalable et scalable. Une optimisation est formulée pour spécifier les niveaux de protections à apporter à chaque couche d’une source en fonction de son importance et de la probabilité de pertes de projections dans le canal. Ce système peut facilement supporter des standards actuels de codage scalable (JPEG, JPEG2000, MPEG...).

Abstract – We propose in this paper a joint source-channel coding with an implementation of multiple description via projection operator. Angles choice and support topology are precised for a non scalable and scalable source. An optimization is formulated to specify protection levels for each layer depending to pertinence and packet loss probability. This system can easily protect standards with scalable coding (JPEG, JP2K, MPEG...).

1 Introduction

Les systèmes de communication ou de stockage distribuent de plus en plus les informations à transmettre ou à conserver. Les canaux multiples empruntés assure une robustesse dans la délivrance d’un volume d’information. Considérons un système à deux canaux : si l’information envoyée sur les deux canaux est identique et que les deux canaux fonctionnent, la moitié de l’information est inutile. Le codage à description multiple évite ce coût élevé de redondance en produisant des canaux équivalents et complémentaires. La réception d’une seule description parmi les deux fournit une reconstruction avec un critère minimum de fidélité. Si les deux descriptions sont reçues, on retrouve une l’information originale. Aucune distinction n’est faite quant à l’effacement des descriptions. Aussi cette technique est fortement adaptée aux canaux de type *best-effort* que nous empruntons au quotidien sur un réseaux de paquets tel que l’Internet. Le codage à description multiple a été inventé dans le laboratoire Bell. Son histoire décrit une boucle des plus particulières. Elle débute par une motivation pratique telle que la transmission de la parole sur le réseaux téléphonique, devient un problème majeur en théorie de l’information pour enfin revenir vers une application. Nous proposons dans ce papier une généralisation du problème de la description multiple à N canaux supportant la protection scalable d’une source. Les descriptions équivalentes et complémentaires sont produites via une transformation de Radon discrète exacte que l’on appelle couramment la transformation Mojette. Aussi pour la suite, on parlera indifféremment de descriptions, de pro-

jections et de paquets.

2 Description multiple Mojette

2.1 Rappels sur la transformation Mojette

La transformation Mojette a déjà été présentée dans les précédentes éditions du GRETSI [1]. Aussi nous rappellerons brièvement les principaux résultats obtenus.

La transformation Mojette permet la représentation d’une image $f[k, l]$ en une série finie de projections $proj_{p,q}[m]$. L’angle de projection θ est défini par le couple d’entiers premiers entre eux (p, q) avec $\tan(\theta) = \frac{q}{p}$. Chaque élément de la projection appelé bin résulte de la somme des pixels situés sur la droite discrète $m = -qk + pl$. Ces pixels sont appelés antécédents du bin considéré. L’expression de la transformation Mojette directe s’écrit :

$$\mathcal{M}_{p,q}f[k, l] = \sum_k \sum_l f[k, l] \Delta[m + qk - pl] \quad (1)$$

où Δ est le symbole de Krönecker.

$\mathcal{M}f$ représente un ensemble de N projections $\mathcal{M}_{p_i,q_i}f[k, l]$ tel que

$$\mathcal{M}f = \{proj_{p_i,q_i}, i = 1, \dots, N\}. \quad (2)$$

Le nombre de bins dans une projection (p, q) est fonction de l’angle et des dimensions du support à projeter. En effet, on a :

$$\#bins = |p|(Q - 1) + |q|(P - 1) + 1. \quad (3)$$

On trouve dans Katz [2] les conditions de restructibilité d’une image de taille (P, Q) avec un ensemble de N projections :

$$P \leq P_N = \prod_{i=1}^N |p_i| \text{ ou } Q \leq Q_N = \prod_{i=1}^N |q_i|. \quad (4)$$

¹. École Polytechnique de l’Université de Nantes, rue Christian Pauc, La Chantrerie, BP60601, 44306 Nantes cedex 3.

La morphologie mathématique étend les conditions de reconstruction à des formes convexes. Une forme est reconstituable par un ensemble donné de projections si et seulement si le résultat d'érosions successives par les éléments structurants à 2 pixels (ÉS2P) associés aux vecteurs de projections est l'ensemble vide. Le nombre d'éléments dans la nouvelle représentation est supérieur au nombre de pixels dans l'image. La redondance naturelle dans le cas d'une transformation non dégénérée peut-être calculée par *Red* avec :

$$Red = \frac{\text{Nombre de Bins}}{\text{Nombre de Pixels}} - 1. \quad (5)$$

La description multiple Mojette se base sur la topologie du support à projeter. L'allocation spatiale d'un élément d'information (ixel) sur ce support est fonction de son niveau de priorité dans la représentation scalable de la source. Le support se comporte ainsi comme un buffer que l'on qualifie de géométrique. La section suivante présente brièvement le choix des angles et des formes du buffer pour assurer un seul niveau de protection.

2.2 Construction d'un buffer géométrique

On se propose de générer un ensemble de N projections tel qu'un sous-ensemble de M projections reconstruisent exactement le buffer géométrique. La redondance se traduira par $N - M$ projections supplémentaires équivalentes. Chaque élément du buffer géométrique nécessite au plus M projections pour être reconstruit avec les C_M^N cas possibles de recevoir M projections parmi N . De cette manière, les N descriptions générées sont équivalentes en termes de reconstruction. On définit l'indice de protection comme étant la proportion du nombre de projections nécessaires sur le nombre de projections envoyées soit $\frac{M}{N}$. La figure 2(a) présente l'allocation d'un vecteur source $\{X_k\}_{k=1}^N$ sur un buffer géométrique avec $M = 3$ et $N = 5$ *i.e.* reconstituable par 3 projections parmi 5 pour l'ensemble de projections considéré.

Sans restrictions, on choisit de satisfaire la contrainte verticale du critère de Katz à savoir $Q \leq \prod_{i=1}^P |q_i|$. Les choix à venir seront valables pour la contrainte horizontale. Si un nombre M de projections suffisent à reconstruire le support rectangulaire, la contrainte verticale peut être restreinte à l'égalité soit $Q = \prod_{i=1}^P |q_i|$. L'équivalence de toutes les combinaisons de M projections est donc assurée par des valeurs de q_i constantes pour tout $i = 1, 2, \dots, N$. Les angles de projections sont donc du type $S = \{(p_i, q) \forall i = 1, 2, \dots, N\}$. Des supports rectangulaires avec une hauteur Q de Mq ixels sont reconstituables par tout les sous-ensembles de S de M projections. Initialement la valeur q a été choisie sous réserve que $PGCD(p_i, q) = 1 \forall i = 1, 2, \dots, N$. En outre, de par la π -périodicité, les projections (p, q) et $(-p, -q)$ sont identiques. On a donc réduit le choix des projections à $q > 0$. Le nombre total de bin est simplement la somme des bins contenus dans toutes les projections. Ce nombre total de bin sera d'autant plus faible à q constant et pour des dimensions de supports rectangulaires donnés que les valeurs de p_i seront faible (Éq.

3). Autrement dit, la contrainte horizontale est fortement relâchée pour minimiser la redondance. Aussi, on a :

$$\prod_{i=1}^P p_i \ll P \quad (6)$$

On choisit pour ce faire des valeurs de p consécutives sur Z et centrées en 0.

Aussi, l'incrémement de la largeur du support P entraîne une augmentation de q bins dans les N projections alors que la capacité augmente de Mq ixels. La redondance exprimée à l'équation 5 devient :

$$Red = \frac{\prod_{i=1}^P (|p_i| (Q - 1) + q(P - 1) + 1)}{Q \times P} - 1$$

$$= \frac{N(q(P - 1) + 1) + (Mq - 1) \prod_{i=1}^P |p_i|}{MqP} - 1 \quad (7)$$

Sur la figure 1 est représenté le comportement de *Red* pour $q = \{1, 5, 11, 27\}$ en fonction de la capacité d'un buffer géométrique rectangulaire (en nombre d'ixels) limité à 5Ki (5120 ixels). Le buffer géométrique est reconstituable ici par 4 projections parmi 7 prises dans l'ensemble $S = \{(-4, q), (-2, q), (-1, q), (0, q), (1, q), (2, q), (4, q)\}$ soit $M = 4$ et $N = 7$ (Moj(7,4)). On peut remarquer sur la figure

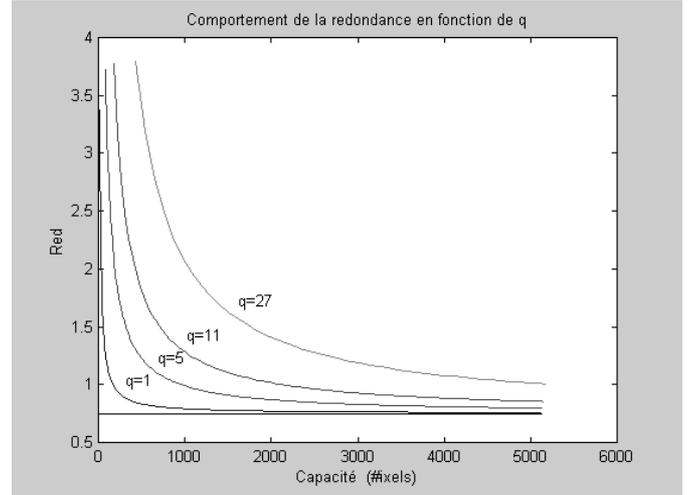


Fig. 1: Comportement de la redondance en fonction de la capacité du buffer et pour $q = \{1, 5, 11, 27\}$.

1 que toutes les courbes ont une même asymptote horizontale. Ce comportement se justifie par la limite de *Red* lorsque la capacité est infinie *i.e.* P infini. En effet, on a :

$$\lim_{P \rightarrow \infty} Red = \frac{N}{M} - 1$$

Dans l'exemple de la figure 1, on a bien l'asymptote en $\frac{7}{4} - 1$ soit à 75 % de redondance qui supporte la perte d'au plus 3 projections parmi 7 soit 43 % de perte de projections.

Cette asymptote est atteinte plus rapidement lorsque $q = 1$. La redondance est donc minimale à $q = 1$, pour un niveau de protection donné et pour un volume d'information

à protéger de même taille. Du point de vue géométrie discrète, prendre $q > 1$ revient à sommer des ixels placés sur des droites non connexes. Chaque sommation correspond en fait à des parcours sur des supports distincts que l'on appelle épais. À $q = 1$, Un bin contient un ixel de chaque ligne du buffer. Cependant, on n'interdira pas l'utilisation d'autres valeurs de q dans un comportement asymptotique *i.e.* pour des buffers de très grande capacité.

Le buffer présenté délivre des paquets de tailles variables. Aussi, on préférera des buffers géométriques de forme hexagonale (figure 2(b)) qui ont la propriété de fournir des descriptions de tailles identiques. Le comportement de ces buffers, très similaire à l'exemple rectangulaire n'est pas détaillé dans ce papier.

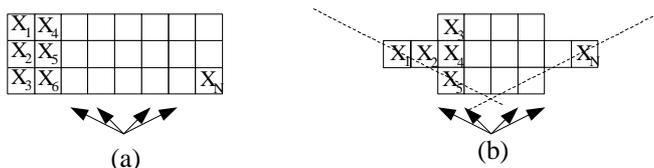


Fig. 2: 2 exemples de buffers géométriques reconstruits avec 3 projections parmi 5 prises dans l'ensemble $\{(-2,1)(-1,1)(0,1)(1,1)(2,1)\}$. Sans contrainte sur la taille des paquets, le support (a) peut être utilisé. Le support (b) fournit des projections de tailles constantes.

2.3 Description multiple Mojette

Les buffers décrits ci-dessus permettent la protection pour un seul niveau de priorité. Cependant la représentation d'une source sous forme scalable prédomine aujourd'hui dans la transmission sur les réseaux de paquets. Du JPEG2000 pour les images fixes au "MP3" pour l'audio, la plupart des standards offre une représentation hiérarchique de la source. Si des spécifications sont données pour réaliser un codage robuste aux effacements dans quelques standards récents [3], des systèmes de transmission à protection inégales sont d'une plus grande fiabilité (*Priority Encoding Transmission* PET). D'un point vu purement codage canal, Albanese et al [4] préconisent l'utilisation des codes systématiques *Maximum Distance Separable* (MDS) pour la protection des sous-flux d'une source scalable. Les deux techniques employées par le codage à description multiple et le système PET poursuivent le même objectif puisque chaque description ou paquet généré (qu'il soit paquet message ou paquet redondant) sont équivalents. En effet la propriété des codes MDS est de générer n paquets de codes à partir de k paquets de message tel que n'importe quels k paquets parmi n suffisent pour retrouver l'information initiale. Ces codes sont par définition optimaux en termes de redondance mais au prix d'une complexité de calcul non négligeable. Afin de réduire la complexité de ces codes, on préfère l'utilisation des codes $(1+\epsilon)$ MDS ($(1+\epsilon)k$ paquets sont nécessaires pour le décodage) plus simples à décoder. Les codes Tornado présenté dans [5] utilisent les graphes irréguliers et bi-partite pour faire chuter les temps de codage/décodage par rapport aux codes MDS construits à partir de matrices de Cauchy [6].

La différence est énorme pour des "gros" volumes de données : un fichier de 16 Mo a un temps d'encodage de 30802 sec. contre 3.93 sec. (voir [7] pour plus de comparaisons). La performance des deux systèmes est utilisée pour protéger des données audio avec les codes Reed-Solomon [8] ou de la vidéo dans des serveurs professionnels. Cependant, la complexité de ces algorithmes est proportionnelle avec le nombre de paquets manquants. En protégeant les flux scalables d'une manière hiérarchique, la transformation Mojette peut être de complexité constante quelque soit le *scenario* de réception et ce sans transmission de métadonnées. Afin d'assurer la protection de la représentation scalable d'une source, on concatène plusieurs buffers de niveaux de priorité distincts et on projette entièrement le nouveau buffer obtenu. La représentation du support utilisé est montrée à la figure 3. Les propriétés de ce support particulier permettent d'affecter trois niveaux de protection. L'information plus importante *i.e.* la basse fréquence (resp. la haute fréquence) peut être allouée dans la région de gauche (resp. la région de droite) qui est reconstruite par trois projections (resp. cinq) parmi cinq prises dans la série $S = \{(-2,1)(-1,1)(0,1)(1,1)(2,1)\}$. L'opérateur

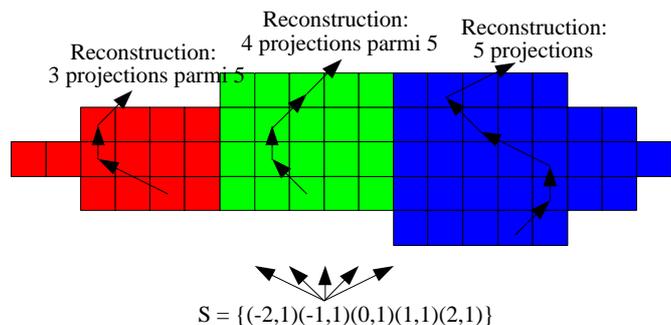


Fig. 3: Buffer géométrique avec 3 niveaux de protection.

Mojette sur le support entier permet une bonne distribution des ixels dans chaque projection. La solution est presque optimale. En effet, dans la transmission d'un JPEG progressif, ϵ est d'environ 1 % en comparaison avec le cas MDS alors que les codes Tornado nécessite 5 % d'information en plus.

Cette représentation de l'information scalable peut être étendue à autant de niveaux de protection que nécessite la granularité de la source. Notre codage à description multiple suppose en amont une opération de structuration de cette granularité en décomposant le flux source en sous-flux (*e.g.* décomposition en sous-bandes pour une image, images I,P,B pour un flux MPEG).

3 Allocation optimale de redondance

3.1 Formalisme d'optimisation

On suppose un codeur en couches en amont du système de protection Mojette qui délivre une série de L flux scalables. Pour tout $s = 1, 2, \dots, L$, et pour tout $k = 1, 2, \dots, K_s$, $f_{s,k}$ représente le $k^{\text{ième}}$ élément du flux s . La structure

en couches implique que l'ensemble f_s ne peut être obtenu que si f_{l-1} est décodé. Soit $\rho(l)$, la fonction de priorité à calculer qui indique pour un flux l , le nombre de projections nécessaires à sa reconstruction. Pour la source considérée, ρ est une fonction monotone croissante, *i.e.* :

$$\rho(1) \leq \rho(2) \leq \dots \leq \rho(L). \quad (8)$$

Soit Q_s , la qualité estimée au décodage du flux s . Par la scalabilité de la source, on a nécessairement $Q_s \geq Q_{s-1}$. On peut définir un incrément de qualité q_s positif ou nul par

$$q_s = Q_s - Q_{s-1}. \quad (9)$$

Pour déterminer la fonction de priorité ρ , on utilise une estimation du profil des pertes du canal. Cette estimation est donnée par la densité de probabilité p_i ; $i = 0, 2, \dots, N$, telle que p_i est la probabilité que exactement i paquets soit perdus. On estime que la probabilité de perdre au plus k paquets est donnée par la fonction de distribution cumulée $c(k) = \sum_{i=0}^k p_i$; $k = 0, 1, \dots, N$. La probabilité $c(\rho(s))$ est donc la probabilité de recevoir suffisamment de projections pour décoder le flux d'indice s .

L'espérance mathématique de la qualité à la réception est alors donnée par

$$\overline{Q}(\rho) = \sum_{s=1}^{\times} c(\rho(s)) q_s. \quad (10)$$

Afin de déterminer la fonction de priorité ρ , on maximise $\overline{Q}(\rho)$ fonction de la densité de probabilité p_i . L'incrément de qualité peut être calculé simplement à partir du PSNR, de l'erreur quadratique moyenne ou encore de critères perceptuels de qualité (note objective à partir d'évaluations subjectives de la qualité [9]).

3.2 Création des unités de transport

La fonction de priorité détermine pour chaque flux le nombre de paquets nécessaires à leur reconstruction. La description multiple s'effectue sur une source de taille réduite (sortie d'un schéma de compression). Aussi une recherche exhaustive des nuplets qui maximisent l'espérance mathématique de la qualité est envisagée. Comme il est supposé dans la partie précédente, on contraint $\rho(i) \leq \rho(i+1)$. De plus, on suppose que la perte d'un élément d'information issu d'un flux binaire progressif entraîne la perte des éléments à venir. La recherche se termine dès lors qu'un seuil de qualité est atteint ou que la capacité du canal est dépassée. Il est à noter enfin que le calcul du nombre de projections nécessaires se fait en supposant que les paquets ont une charge utile "MDS". A savoir si lp mesure la longueur d'un paquet et m_l le nombre de paquets nécessaires à sa reconstruction, on a :

$$lp = \sum_{l=1}^{\times} \frac{f_l}{m_l}. \quad (11)$$

La répartition des L flux dans N paquets est schématisée sur la figure 4.

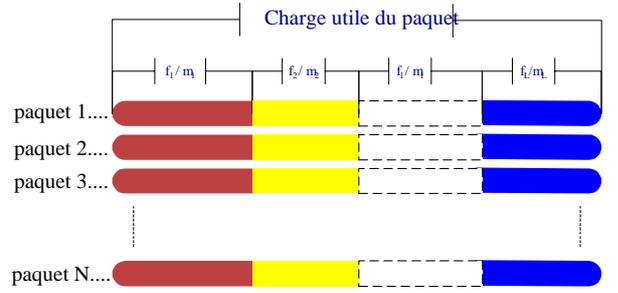


Fig. 4: Répartition des sous-flux dans N paquets.

4 Conclusion

La protection scalable fait de la transformation Mojette un codage à description multiple particulier. Ce schéma de codage fournit à sa sortie des descriptions équivalentes transmissibles sur un canal ou aucune distinction n'est faite entre les unités de transports. Il est donc applicable pour la transmission sur des réseaux de paquets tel que l'Internet. En tenant compte des caractéristiques de la source et du canal, le schéma présenté est également un codage conjoint source-canal.

Références

- [1] N. Normand, B. Parrein, "Description multiple et codage source-canal sur un canal à effacement avec la transformation Mojette", GRETSI 99, pp. 869-872, vol. 3, sept. 1999, Vannes(France).
- [2] M. Katz, Questions of Uniqueness and Resolution in Reconstruction from projections, Lecture Notes in Biomathematics, Springer-Verlag, Vol. 26, 1979.
- [3] I. Moccaganta, S. Soudagar, J. Liang, H. Chen, "Error-Resilient Coding in JPEG-2000 and MPEG-4", IEEE Journal on Sel. Areas in Comm., vol.18, no.6, pp.899-914, June 2000.
- [4] A. Albanese, J. Blömer, J. Edmonds, M. Luby, M. Sudan, Priority Encoding Transmission, IEEE Trans. on Information Theory, vol. 42, no. 6, pp. 1737-1744, Nov.1996.
- [5] M.Luby, M. Mitzenmacher, A. Shokrollahi, D. Spielman, V. Stemann, "Practical Loss-Resilient Codes", 29th ACM Symposium on Theory of Computing, 1997.
- [6] J.Blömer, M. Kalfane, R. Karp, M. Karpinski, M. Luby, D. Zuckerman, "An XOR-Based Erasure-Resilient Coding Scheme", ICSI Tech. Rep. TR-95-048, August 1995.
- [7] J.W. Byers, M. Luby, M. Mitzenmacher, A.Rege, "A Digital Fountain Approach to Reliable Distribution of Bulk Data", ICSI Tech.Report, TR-98-013, 1998.
- [8] J.Rosenberg, H. Schulzrinne, "RTP Payload format for generic forward error correction", RFC 2733, Dec. 1999.
- [9] P. Le Callet, D. Barba, "Critère de qualité visuelle d'images couleurs: application à l'évaluation réaliste de codeurs d'images", GRETSI 2001.