



HAL
open science

Recherche Rapide du Mode de Prédiction Optimal basée Apprentissage via une Comparaison des Modes Intra/Inter en H.264/AVC

Maxime Bichon, Julien Le Tanou, Wassim Hamidouche

► **To cite this version:**

Maxime Bichon, Julien Le Tanou, Wassim Hamidouche. Recherche Rapide du Mode de Prédiction Optimal basée Apprentissage via une Comparaison des Modes Intra/Inter en H.264/AVC. 18ème Colloque COmpression et REprésentation des Signaux Audiovisuels (CORESA 2016), May 2016, Nancy, France. hal-01333662

HAL Id: hal-01333662

<https://hal.science/hal-01333662v1>

Submitted on 25 Apr 2017

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.

Recherche Rapide du Mode de Prédiction Optimal basée Apprentissage via une Comparaison des Modes Intra/Inter en H.264/AVC

M. Bichon^{†,‡}

J. Le Tanou[†]

W. Hamidouche[‡]

[†] Envivio France

[‡] INSA de Rennes, Institut d'Électronique et de Télécommunications de Rennes - UMR6164

maxime.bichon@ericsson.com, julien.le.tanou@ericsson.com, wassim.hamidouche@insa-rennes.fr

Résumé

Dans un contexte de codage vidéo hybride comme défini par la norme H.264/AVC, l'efficacité maximale de compression est atteinte par l'optimisation d'un ensemble de paramètres de codage. Parmi les plus cruciaux figure le choix du mode de prédiction. Ce dernier est communément sélectionné suivant un critère débit-distorsion RDO¹ au cours d'une recherche exhaustive, d'une complexité de calcul importante notamment due à l'étape d'estimation de mouvement (ME). La solution proposée ici adapte un algorithme de décision rapide basé analyse d'image, de manière à limiter la complexité introduite par la ME. En effet, la mise en concurrence des corrélations spatiales et temporelles relatives au macrobloc permet une décision rapide et fiable du mode de prédiction, évitant les MEs inutiles et coûteuses en calcul. Basé sur un apprentissage hors-ligne, le modèle proposé dans cet article permet d'accélérer systématiquement l'encodeur temps-réel x264, de 11,29% en moyenne et jusqu'à 20% selon les séquences, pour une perte en efficacité de codage négligeable (<1%).

Mots clefs

Choix de Mode Rapide, Apprentissage, Corrélation Spatio-Temporelle, H.264/AVC.

1 Introduction

La norme H.264/AVC [1, 2] est le standard de compression vidéo majoritairement déployé sur les services vidéos. Communément, la sélection du mode optimal du macrobloc (MB) se fait par une recherche exhaustive complexe basée sur l'évaluation d'un critère débit-distorsion (RDO[3]). L'objectif est de minimiser une distorsion D sous une contrainte de débit $R \leq R_t$. Cette minimisation sous contrainte se traduit par le calcul du coût débit-distorsion (coût R-D), noté J , que l'on cherche à minimiser:

$$J = D + \lambda R \quad (1)$$

avec λ le multiplicateur de Lagrange [4] permettant de réécrire le problème sans contrainte. Le coût R-D réel de chaque mode s'obtient par un processus de reconstruction, étape de l'encodage qui simule le décodage afin d'estimer précisément le coût R-D. Certaines méthodes accélèrent cette étape grâce à des métriques de distorsion sous-optimales telles que la SAD² ou SATD³, mais entraînent un biais due à l'approximation de la distorsion réelle. Une solution plus souple consiste à introduire une hiérarchie entre les modes. L'analyse RDO d'un sous-ensemble de modes, construit selon un critère robuste, est théoriquement plus rapide au détriment de pertes R-D négligeables. Trois catégories d'algorithmes de décision rapide utilisant cette stratégie se distinguent:

1. L'accélération de l'analyse des modes Intra a été largement traitée dans la littérature [5, 6]. Les algorithmes se basent généralement sur les caractéristiques spatiales du MB telles que le gradient [5] ou la variance de ses pixels [6]. Les gains, généralement faibles, s'expliquent par une analyse des modes Inter beaucoup plus complexe à cause de l'estimation de mouvement (ME). Cette étape, occupe à elle seule jusqu'à 30% du temps d'encodage total.
2. L'accélération de l'analyse Inter est populaire parmi les algorithmes de décision rapide en raison de la complexité et de l'efficacité de codage de la ME. [7, 8] proposent deux accélérations possibles de l'estimation elle-même. En limitant les tailles de partitions à analyser et donc l'usage de la ME, la complexité peut aussi être réduite de manière non négligeable, au détriment de l'efficacité R-D. Deux exemples définissent une hiérarchie des partitions Inter, préconisant une analyse plus fine, en fonction des Coded Block Pattern CBP [9] ou d'une pseudo-ME [10].
3. L'accélération des analyses Intra/Inter, consiste en un algorithme de décision rapide sur l'ensemble de la recherche exhaustive. [11] propose une solution où les modes sont hiérarchisés en fonction de l'activité temporelle. Une approche innovante [12] considère les modes Intra et Inter comme deux classes distinctes. En mesurant les corrélations spatiales et temporelles relatives au MB courant, les

2. Sum of Absolute Differences

3. Sum of Absolute Transformed Differences

1. Rate Distortion Optimization

auteurs utilisent la notion de risque Bayésien pour éliminer avec fiabilité l'ensemble des modes Intra ou Inter des choix possibles. Cependant, le modèle repose sur le choix de seuils fixe, méthode jugée trop rigide.

Dans cet article, nous proposons une évolution de [12] en un modèle simple et flexible. L'objectif est de palier aux cas où la ME est un "poids mort" pour l'encodage, en particulier dans le cas de séquences à mouvements complexes, p.ex. *Touchdown Pass* et *Pedestrian Area*. La suite de l'article est structurée de la manière suivante. Une rapide revue de l'algorithme initialement proposé est donnée dans la Section 2. Les nouveaux critères de décisions seront présentés dans la Section 3. La Section 4 fournit les explications justifiant la flexibilité du nouveau modèle. Les simulations et résultats seront discutés dans la Section 5, avant de conclure sur cette étude dans la Section 6.

2 Choix de mode Intra versus Inter basé analyse d'image

Le principe de [12] est d'estimer en amont de la *RDO* l'efficacité des modes Intra et Inter. Le calcul d'un descripteur de bloc, composé de trois critères robustes f_{intra} , f_{inter} et $|MV|$, permet d'estimer si l'ensemble des modes Intra ou Inter peuvent être éliminés des candidats potentiels, selon une notion de risque Bayésien. Le domaine de représentation du descripteur définit trois zones:

- La zone "sans risque": Les modes Intra ou Inter peuvent être ignorés de la *RDO* avec un impact faible sur l'efficacité de codage
- La zone "à risque tolérable": Les pertes R-D attendues du mode qui minimise le risque Bayésien sont acceptables, donc les modes Intra ou Inter sont ignorés de la *RDO*
- La zone "à risque intolérable": Les pertes attendues sont trop élevées, il faut appliquer la *RDO* sur l'ensemble de ces modes

La délimitation de ces zones se fait par un apprentissage supervisé hors-ligne, utilisant une base de MBs encodées en *RDO*, afin d'anticiper les pertes R-D en cas de mauvaise décision. De plus, l'estimation du risque Bayésien nécessite également un apprentissage au préalable. f_{intra} est une mesure de l'amplitude de la corrélation spatiale, calculée comme suit:

$$f_{intra} = \sum_{k=1}^{16} \min_m \{SATD(m_k)\} \quad (2)$$

avec k l'indice du bloc 4x4 du MB courant, m le mode de prédiction à déterminer parmi cinq des neuf disponibles en H.264/AVC et $SATD(\cdot)$ la mesure de distorsion, définie comme une erreur de prédiction dans le domaine transformé.

f_{inter} mesure l'amplitude de la corrélation temporelle. Afin d'être comparable à f_{intra} , elle se mesure par une SATD résultant d'une ME rapide "*MVFAST*" [13].

$$f_{inter} = \min_{MV} \{SATD(m_{MV})\} \quad (3)$$

avec MV le vecteur de mouvement cherché selon [13] et m_{MV} la compensation en mouvement dépendant du vecteur MV . La comparaison des deux variables détermine si le MB est plus fortement impacté par la corrélation spatiale ou temporelle et, a fortiori, quels modes sont à prioriser entre l'Intra et l'Inter.

La variable $|MV|$ concerne l'amplitude du mouvement estimé par la ME. Elle est utilisée comme critère de confiance de la ME et de f_{inter} . Son utilisation comme troisième dimension permet un découpage plus fiable du domaine de représentation du descripteur.

Les modifications que nous proposons sont présentées dans la suite de cet article. Les changements effectués pour la prise de décision, c'est-à-dire les nouvelles mesures de corrélations et de précision, sont exposées dans la Section 3. Dans la Section 4, l'espace de représentation est redéfini, selon deux frontières ($Bounder_{intra}$ et $Bounder_{inter}$) dépendantes de plusieurs paramètres de flexibilité.

3 Nouveaux Critères

Les critères utilisés pour mesurer l'amplitude des corrélations sont discutables, notamment due au biais induit par certaines approximations. D'une part, l'utilisation de la *SATD* comme mesure de la corrélation ne tient pas compte de la distorsion réelle et du débit. D'autre part, les variables utilisées sont obtenues grâce à des analyses rapides (et potentiellement imprécises) des modes Intra et Inter. Nous proposons ici plusieurs modifications des critères précédemment utilisés.

3.1 Corrélations Spatiale et Temporelle

Le calcul de la variable f_{intra} est réalisé par une analyse *RDO* du mode Intra 16x16, l'équation (2) devient alors:

$$f_{intra} = \min_m J_{SSD}(m) = D_{SSD}(m) + \lambda R_{SSD}(m) \quad (4)$$

avec J_{SSD} , D_{SSD} et R_{SSD} correspondant au coût R-D réel, à la distorsion et au débit associé au mode m . En remplaçant la *SATD* initiale par une distorsion mesurée sur le MB reconstruit et en se basant sur le coût R-D plutôt que sur la distorsion, le critère f_{intra} est une mesure plus fiable de la corrélation spatiale.

La variable f_{inter} est définie comme le coût R-D d'une analyse *RDO* du mode Inter 16x16.

$$f_{inter} = \min_{MV} J_{SSD}(MV) = D_{SSD}(MV) + \lambda R_{SSD}(MV). \quad (5)$$

Le principal changement apporté par ces deux nouvelles variables est la prise en compte de la distorsion réelle et du débit réel. f_{intra} et f_{inter} sont respectivement des mesures plus fiables de l'efficacité de codage des modes Intra et Inter, suivant un critère R-D.

3.2 Paramètres de Précision

Afin de contrôler la précision du modèle, l’algorithme initial utilise l’amplitude du mouvement $|MV|$. Néanmoins, dans le cas d’un mouvement important mais uniforme, la ME peut être précise malgré une forte amplitude. En effet, l’algorithme de Block-Matching utilisé pour la ME recherche le bloc le plus similaire dans une fenêtre de recherche autour d’une “graine”. Classiquement, cette graine correspond simplement à la position du bloc courant dans l’image de référence, mais dans l’encodeur x264 elle correspond à un vecteur de mouvement prédit calculé comme suit:

$$pMV(D) = Median[MV(A), MV(B), MV(C)] \quad (6)$$

avec $pMV(.)$ la fonction estimant la graine pour la ME, $Median[.]$ le vecteur médian parmi ceux proposés et $MV(.)$ le vecteur mouvement utilisé par le MB concerné. Le voisinage (A , B et C) utilisé pour ce calcul correspond respectivement aux MBs gauche, supérieur, et supérieur-droite du MB courant D . Sur la Figure 1, on peut voir la fenêtre obtenue avec ou sans graine.

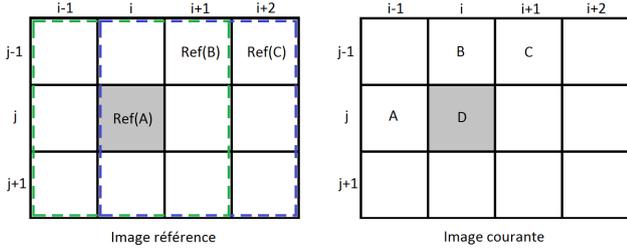


Figure 1 – Fenêtres de recherche de la ME avec vecteur prédit (zone bleue) et sans (zone verte)

L’intérêt de ce vecteur prédit est une convergence plus rapide du résultat optimal de la ME. De plus, cette prédiction est aussi utilisée dans la norme H.264/AVC afin d’obtenir un gain, en termes de débit, en ne transmettant que l’écart MVD entre le vecteur médian et le vecteur final. Nous proposons d’utiliser $|MVD|$ au lieu de $|MV|$ comme mesure de confiance de la ME.

4 Frontières Flexibles

Les différentes zones proposées dans l’article initial recouvrent de manière non-uniforme l’espace de représentation du descripteur partitionné en plusieurs cellules. Le découpage de cet espace se fait en fonction des données d’apprentissage qui servent également à évaluer les pertes R-D envisagées en cas d’erreur pour chaque cellule. Les cellules dont les pertes R-D estimées sont inférieures à 0.5% appartiennent à la zone sans risque. Ce choix de seuil arbitraire implique une certaine rigidité dans le modèle. Nous proposons à la place de calculer deux frontières flexibles, $Bounder_{intra}$ et $Bounder_{inter}$, définies ci-après

par les équations (8) et (9). Quatre observations sont faites et intégrées au modèle pour le rendre adaptatif:

1. On observe que le coût R-D est une valeur qui varie en fonction des caractéristiques de la source. Un MB sans texture et avec peu de mouvement aura un coût R-D relativement faible. Cela est due au fait qu’une forte corrélation spatiale et temporelle entraîne une prédiction de très bonne qualité. Inversement, si un MB est fortement texturé ou affecté par un mouvement complexe, la prédiction sera de mauvaise qualité et le coût R-D élevé. Une variable intermédiaire est donc calculée:

$$df_{ratio} = \frac{|f_{intra} - f_{inter}|}{MAX(f_{intra}, f_{inter})} \quad (7)$$

où df_{ratio} est une valeur comprise dans l’intervalle $[0, 1[$ qui estime l’importance de l’écart entre les deux corrélations.

2. Remarqué par [12] lors des expérimentations, commettre une erreur en choisissant de privilégier l’Intra plutôt que l’Inter est plus critique que l’inverse. Afin de prendre en compte cet effet, deux frontières $Bounder_{intra}$ et $Bounder_{inter}$ sont définies. La variable df_{ratio} présentée ci-dessus sera donc comparée, pour chaque MB, à la borne correspondant à la classe de plus faible coût.

3. Expliquée dans la section 3.2, la précision de la variable f_{inter} est dépendante des résultats de la ME. La variable $|MVD|$ est introduite dans les calculs de frontière afin de tenir compte de cette précision. Basés sur une étude expérimentale, nous proposons les modèles suivants:

$$Bounder_{intra} = a * \ln(|MVD|) + b. \quad (8)$$

$$Bounder_{inter} = c * \ln(|MVD|) + d. \quad (9)$$

Les constantes a , b , c et d sont estimées par un apprentissage hors-ligne et sont différentes selon le type d’image auxquels le modèle est appliqué: P, B ou B référence.

4. Un paramètre de quantification (QP) élevé entraîne une distorsion plus élevée mais un débit plus faible et peut donc modifier les coûts R-D d’une manière non négligeable. La valeur de QP est donc introduite dans le modèle. Au travers d’observations, les relations suivantes ont pu être établies entre les différentes constantes des modèles précédemment introduites (a , b , c et d) et QP :

$$a = \alpha_0 QP^2 + \beta_0 QP + \gamma_0. \quad (10)$$

$$b = \alpha_1 QP^2 + \beta_1 QP + \gamma_1. \quad (11)$$

$$c = \alpha_2 QP + \beta_2. \quad (12)$$

$$d = \alpha_3 QP + \beta_3. \quad (13)$$

Quand la ME est suffisamment précise, c.à.d. $|MVD| = 0$, les équations (8) et (9) sont remplacées par deux fonctions affines de QP . On visualise sur la Figure 2 que si df_{ratio} est supérieure (zone verte) à la frontière calculée, alors la précision du modèle est suffisante pour éliminer un ensemble de mode. Ces deux frontières sont vues comme

“flottantes” dans l’intervalle de df_{ratio} , selon le couple $(|MVD|,QP)$ et les équations (8~13).

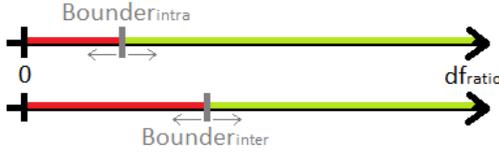


Figure 2 – Frontières en fonction de df_{ratio}

5 Expériences et Résultats

5.1 Environnement Expérimental

L’encodeur de référence JM [14], associé à la norme H.264/AVC, permet la comparaison d’outils sur une base commune. L’importante complexité de calcul du JM en fait une plateforme non-réaliste pour le développement et l’évaluation d’algorithmes d’accélération. L’efficacité et la stabilité reconnues de l’encodeur x264 [15] en font, a contrario, une plateforme cohérente pour ces travaux.

Nous utilisons la configuration suivante: une I par seconde; 3 B hiérarchique; la ME a une précision au 1/4 de pixel, recherche hexagonale; aucune partition Inter inférieure au 8x8. Le x264 est configuré en *RDO*.

Un ensemble de séquences aux formats 4K, 1080p et 720p, sont utilisées pour évaluer l’algorithme. Trois séquences 720p non-standards sont ajoutées pour enrichir l’ensemble de test: *Danse*, *PubSony* et *Home*. Des images de ces séquences sont données Figure 3.

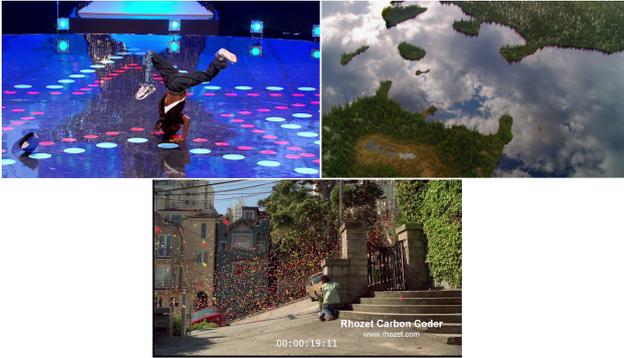


Figure 3 – En haut *Danse*(720p) et *Home*(720p); en bas *PubSony*(720p)

Les trois séquences 4K (4096x2160), *SquareTimelapse*, *TunnelFlag* et *ToddlerFountain*, sont fournies par Netflix et publiquement disponible [16]. Les séquences *Cactus*, *MobileCalendar* et *Home* sont définies comme ayant des mouvements simples à prédire. A contrario, les séquences restantes ont des mouvements complexes.

L’évaluation de l’algorithme se fait selon deux critères. Le premier est la réduction relative de complexité donnée par:

$$T\% = \frac{Time(x264_{rapide}) - Time(x264_{original})}{Time(x264_{original})} \quad (14)$$

avec $T\%$ le temps relatif, $Time(\cdot)$ le temps d’encodage et $x264_{original}$ et $x264_{rapide}$, respectivement le x264 avec et sans l’algorithme que nous proposons.

Le second critère représente les pertes en efficacité de codage par la mesure de Bjontegaard [17], qui consiste à mesurer la différence d’aire entre deux courbes R-D. Elle est ici basée sur une métrique PSNR pour cinq valeurs de QP (25,28,31,34,37). Les scores négatifs représentent une économie de débit et donc une amélioration de l’efficacité de codage.

5.2 Résultats

Le Tableau 1 montre la réduction de complexité induite par l’algorithme. Les résultats sont beaucoup plus probants sur les séquences dont la complexité du mouvement est plus forte. Les séquences *Cactus*, *MobileCalendar* et *Home* avec des mouvements simples présentent les gains les plus faibles (<6%). En effet pour ce type de séquence, la ME converge très rapidement vers le résultat optimal, la réduction potentielle de complexité étant ainsi faible.

Séquences de test		QP			
		25	31	37	Moy
720p	Danse	-15.57%	-19.09%	-21.54%	-18.73%
	PubSony	-6.93%	-9.50%	-10.8%	-9.08%
	MobileCalendar	-1.72%	-3.03%	-4.81%	-3.19%
	Home	-4.51%	-5.62%	-6.16%	-5.43%
	Moyenne	-7.18%	-9.31%	-10.83%	-9.11%
1080p	PedestrianArea	-12.37%	-14.22%	-15.24%	-13.94%
	TouchdownPass	-10.33%	-10.58%	-10.9%	-10.60%
	RedKayak	-18.24%	-21.53%	-21.32%	-20.36%
	Cactus	-4.16%	-5.61%	-5.67%	-5.15%
	Moyenne	-11.28%	-12.99%	-13.28%	-12.51%
4K	SquareTimelapse	-10.77%	-14.89%	-16.55%	-14.07%
	TunnelFlag	-7.68%	-16.94%	-20.66%	-15.09%
	ToddlerFountain	-4.28%	-9.5%	-11.73%	-8.5%
	Moyenne	-7.58%	-13.78%	-16.31%	-12.56%
Tous	Moyenne	-8.78%	-11.86%	-13.22%	-11.29%
	Maximum	-18.24%	-21.53%	-21.54%	-20.36%
	Minimum	-1.72%	-3.03%	-4.81%	-3.19%

Tableau 1 – Complexité relative au x264

Avec une réduction de complexité maximale de 21.54% (pour *RedKayak*) et une moyenne de 11.29%, l’algorithme réponds aux objectifs. Les gains maximums sont atteints pour les séquences jugées difficiles pour la ME. Les gains minimums sont observés sur *MobileCalendar* avec une réduction moyenne de complexité de 3.19%. Autre point remarquable, les gains tendent à augmenter pour un QP élevé, s’expliquant par l’augmentation de la complexité des décisions relativement à la complexité de codage de l’information résiduelle.

Les résultats en efficacité de codage sont présentés dans le Tableau 2. Les pertes R-D sont minimales (0.91% dans le pire cas) pour l’ensemble des séquences testées.

Séquences de test		BD-Rate
720p	Danse	0.83%
	PubSony	0.36%
	MobileCalendar	0.18%
	Home	0.26%
	Moyenne	0.41%
1080p	PedestrianArea	0.91%
	TouchdownPass	0.71%
	RedKayak	0.35%
	Cactus	0.39%
	Moyenne	0.61%
4K	SquareTimelapse	-0.3%
	TunnelFlag	-0.66%
	ToddlerFountain	-0.01%
	Moyenne	-0.32%

Tableau 2 – BD-Rate en comparaison au x264

6 Conclusion

La solution que nous proposons permet de définir des seuils de décision de manière adaptative ou *flexible*, plutôt qu'arbitrairement comme proposé par [12]. Les résultats sont satisfaisants au vu de l'objectif initial, qui était d'accélérer les décisions des modes de codage dans un contexte RDO, en particulier en compensant la complexité d'une ME mise en échec. Les gains systématiques observés, atteignent jusqu'à 20% sur certaines séquences, pour des pertes R-D négligeables (<1%). Ces gains dans un encodeur quasi temps-réel tel que le x264 démontrent l'efficacité de l'algorithme proposé. Néanmoins, certaines séquences peuvent présenter des gains limités, dans le cas où l'estimation de mouvement converge rapidement. L'approche proposée peut-être efficacement couplée avec d'autres algorithmes de décisions rapides des modes Intra ou Inter [5, 9]. Enfin, les hypothèses avancées lors de la conception de l'algorithme sont valables en HEVC [18], la norme qui succède à H.264/AVC. Son adaptation au standard suivant est donc la prochaine étape à considérer.

References

- [1] T. Wiegand, G. J. Sullivan, G. Bjontegaard, et A. Luthra. Overview of the H.264/AVC video coding standard. *IEEE Transactions on Circuits and Systems for Video Technology*, 13(7):560–576, Juillet 2003.
- [2] I. E. Richardson. *The H.264 Advanced Video Compression Standard*. Wiley publications, 2010.
- [3] G. J. Sullivan et T. Wiegand. Rate-distortion optimization for video compression. *Signal Processing Magazine, IEEE*, 15(6):74–90, Novembre 1998.
- [4] T. Wiegand et B. Girod. Lagrange multiplier selection in hybrid video coder control. Dans *Image Processing, 2001. Proceedings. 2001 International Conference on, IEEE*, pages 542–545, Thessaloniki, Octobre 2001.
- [5] A. Elyousfi. An improved fast mode decision method for H.264/AVC intracoding. *Advances in Multimedia*, 2014(7):1–8, Janvier 2014.
- [6] Y. H. Huang, T. S. Ou, et H. H. Chen. Fast decision of block size, prediction mode and intra block for H.264 intra prediction. *IEEE Transactions on Circuits and Systems for Video Technology*, 20(8):1122–1132, Juillet 2010.
- [7] A. M. Tourapis, O. C. Au, et M. L. Liou. Predictive motion vector field adaptive search technique (pmvfast): enhancing block-based motion estimation. Dans *Proc. Visual Communications and Image Processing 2001*, pages 883–892, 2001.
- [8] I. Ahmad, W. Zheng, J. Luo, et M. Liou. A fast adaptive motion estimation algorithm. *IEEE Transactions on Circuits and Systems for Video Technology*, 16(3):420–438, Mars 2006.
- [9] Z. Shi, W. A. C. Fernando, et A. M. Kondoz. A hybrid coded block patterns based fast mode decision in H.264/AVC. Dans *Multimedia and Expo Workshops, 2012 IEEE International Conference on*, pages 13–18, Melbourne, VIC, Juillet 2012.
- [10] W. Zhang, Y. Huang, et J. Peng. Fast mode decision for H.264/AVC based on local spatio-temporal coherency. Dans *Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2013 Asia-Pacific, IEEE*, pages 1–6, Kaohsiung, Octobre 2013.
- [11] H. Zeng, C. Cai, et K.K. Ma. Fast mode decision for H.264/AVC based on macroblock motion activity. *IEEE Transactions on Circuits and Systems for Video Technology*, 19(4):491–499, Avril 2009.
- [12] C. Kim et C.-C. J. Kuo. Feature-based intra-/intercoding mode selection for H.264/AVC. *IEEE Transactions on Circuits and Systems for Video Technology*, 17(4):441–453, Avril 2007.
- [13] P.I. Hosur et K.K. Ma. Motion vector field adaptive fast motion estimation. Dans *Second International Conference on Information, Communications and Signal Processing (ICICS)*, pages 7–10, Singapore, Décembre 1999.
- [14] ITU-T et ISO/IEC JTC. Reference software for advanced video coding. <https://www.itu.int/rec/T-REC-H.264.2>, Juin 2015.
- [15] VideoLAN. x264 software revision 2495. <http://www.videolan.org/developers/x264.html>, Avril 2014.
- [16] Video Test Media [derf's collection]. <https://media.xiph.org/video/derf/>, Février 2016.
- [17] G. Bjontegaard. Calculation of average PSNR differences between RD-curves. Dans *ITU-T VCEG, Texas, USA, Proposal VCEG-M33*, Austin, Texas, Avril 2001.
- [18] High efficiency video coding. Dans *Rec. ITU-T H.265 and ISO/IEC 23008-2*, Sapporo, JP, Janvier 2013.