

الجمهورية الجزائرية الديمقراطية الشعبية
وزارة التعليم العالي والبحث العلمي

UNIVERSITE DE BADJI MOKHTAR - ANNABA



جامعة باجي مختار - عنابة

Faculté : Sciences de l'ingénierat
Département : Electronique

Année : 2018

THÈSE

Présentée en vue de l'obtention du diplôme de Doctorat en Sciences

Intitulé

**Etude et évaluation de la qualité de service
d'un streaming vidéo dans un réseau ad hoc
(égale à égale)**

Option

Télécommunications

Par

BOUMEHREZ Farouk

Rapporteur : M. DOGHMANE Noureddine Prof Université d'Annaba
Co-rapporteur : M. MANSOURI Khaled Prof Université d'Annaba

Devant le jury

Président : M. MESSADEG Djemil Prof Université d'Annaba
Examineur : M. BAARIR Zine-eddine Prof Université de Biskra
Examineur : M. BOUDEN Toufik Prof Université de Jijel
Examineur : M. BEKHOUCHE Amara MCA Université de Souk Ahras
Invitée : Mme. HARIZE Saliha MCA Université d'Annaba

Remerciements

D'abord et avant tout, je remercie mon DIEU « Allah » pour m'avoir accordé la santé, la patience, la connaissance et tout ce qui est nécessaire pour terminer cette thèse.

Je tiens particulièrement à remercier mon directeur de thèse le Professeur DOGHMANE Noureddine, qui m'a proposé le sujet et qui m'a encadré pendant ces années dans mes activités de recherche. Merci donc de m'avoir écouté et guidé sans oublier ses qualités humaines.

Je tiens également à remercier Monsieur MANSOURI Khaled, Professeur à l'université de Badji Mokhtar d'Annaba pour m'avoir fait confiance pendant ces années et avoir me guider avec gentillesse pendant ces années.

Toute ma gratitude va à Monsieur MESSADEG Djemil, Professeur à l'université de Badji Mokhtar d'Annaba, qui m'a fait l'honneur de présider le jury.

Mes remerciements vont également à Monsieur BAARIR Zine-eddine Professeur à l'université Mohamed Khider de Biskra, Monsieur BOUDEN Toufik Professeur à l'université Mohammed Seddik Ben Yahia de Jijel et Monsieur BEKHOUCH Amara Maître de Conférences à l'université Mohamed Cherif Messaadia de Souk Ahras et Mme HARIZE Saliha Maître de Conférences à l'université de Badji Mokhtar d'Annaba pour avoir accepté d'examiner ce travail et de l'enrichir par leurs grands savoirs et leurs grandes expériences.

Enfin, à mes proches les plus chers, un grand merci à ma femme évidemment qui m'a soutenu jour après jour dans cette aventure et qui m'a permis de garder mon optimisme et ma volonté dans les moments de stress et de difficulté. A ma famille pour le support, pour m'avoir appris les valeurs de la vie. Je remercie aussi ma petite fille Jouri mon source de joie, mes amis Fouad, Chamseddine, pour leur soutien constant.

Résumé

Le streaming vidéo est une application exigeante, où de grandes quantités de données doivent être traitées et transmises. Dans cette thèse, nous avons évalué la nouvelle norme de codage vidéo (HEVC [Codage vidéo à haute efficacité] / H.265) avec certaines variables, en particulier le contenu vidéo et le paramètre de quantification (QP). Dans le but d'améliorer les performances de la transmission vidéo à travers des réseaux IP jugés hostiles nous avons utilisé le codage de description multiple (MDC). A cette fin, des expériences et des simulations ont été menées en utilisant la plateforme HM16.11.

Une fois les principes fondamentaux de HEVC/H.265 seront maîtrisés, il devient important de comprendre l'impact de perte des paquets de la transmission et du contenu vidéo sur la qualité de l'expérience (QoE) dans une transmission à travers des réseaux sans fil ADHOC. Nous proposons ensuite un modèle QoE basé sur le schéma d'adaptation de codeur HEVC avec le MDC pour le streaming vidéo. Les principales contributions sont (1) une évaluation des performances de la nouvelle norme de codage vidéo HEVC / H.265, qui est basée sur la variation des valeurs des paramètres de quantification (QP) en fonction du contenu vidéo; (2) Les applications multimédias avec le support QoE dans les réseaux sans fil adhoc sont étudiées, ainsi nous étudions l'impact de perte des paquets sur la QoE des séquences vidéo transmises (3) Le modèle d'adaptation de codeur HEVC basé sur MDC est modélisé avec les paramètres de l'encodeur et QoE. Une étude comparative a révélé que l'approche MDC proposée est efficace pour améliorer la transmission. Les résultats montrent qu'un bon choix de valeurs de QP peut compenser les effets de canal de transmission et améliorer la qualité vidéo reçue et montrent aussi l'efficacité de notre méthode proposée.

Mots clés: HEVC/H.265, Codage par Description Multiple (MDC), Qualité d'expérience (QoE), paramètre de Quantization(QP), Taux de perte des paquets (PLR), réseaux Ad-Hoc sans fil.

Abstract

Video streaming is a demanding application where large amounts of data must be processed and transmitted. In this thesis, we evaluated the new video coding standard “High Efficiency Video Coding” (HEVC / H.265) using some settings or features, in particular video content and quantization parameter (QP) based on Multiple Descriptive Coding (MDC). The experiments and simulations were conducted using the HM16.11 platform.

Once the HEVC fundamental principles are defined, it becomes important to understand the impact of transmission packet losses and video content on the quality of experience (QoE) in ADHOC wireless networks. . In addition, we propose a QoE model based on the HEVC encoder adaptation scheme with MDC for video streaming.

The main contributions are (1) a performance evaluation of the new and emerging video coding standard HEVC/H.265 based on the variation of Quantization Parameter (QP) values depending on different video content to deduce their influence on the sequence to be transmitted; (2) QoE support multimedia applications in wireless networks are investigated, so we inspect the packet loss impact on the QoE of transmitted video sequences; (3) HEVC encoder parameter adaptation scheme based on MDC is modeled with the encoder parameter and objective QoE model. A comparative study revealed that the proposed MDC approach is effective for improving the transmission. Results obtained show that a good choice of QP values can compensate for transmission channel effects and improve received video quality. These results show also the efficiency of our proposed method.

Key words: HEVC/H.265, Multiple Description Coding (MDC), Quality of Experience (QoE), Quantization Parameter (QP), Packet Loss ratio (PLR), Wireless Ad-Hoc Networks.

الملخص

إن تدفق الفيديو في عصرنا الحالي هو تطبيق يتطور باستمرار ولا يخف على أحد أنه لديه متطلبات صعبة حيث يجب معالجة كميات كبيرة من البيانات ونقلها بالاعتماد على اختيار معيار ضغط الفيديو المناسب لذلك، في البداية يقدم هذا العمل تحليلاً للمعيار الجديد لتشفير الفيديو HEVC / H.265 من خلال تجربة بعض المتغيرات بالأخص نوعية محتوى الفيديو و معملية التكميم (QP). من أجل تحسين أداء نقل الفيديو عبر شبكات IP، استخدمنا ترميز الوصف المتعدد (MDC) لتدفق الفيديو وتحقيقاً لهذه الغاية، أجريت التجارب والمحاكاة باستخدام منصة HM16.11.

في هذه الأطروحة بمجرد إتقاننا أساسيات HEVC / H.265 ، قمنا بدراسة تأثير محتوى الفيديو والتكميم (Quantification) وكذلك فقدان الحزم (Perte des paquets) خلال إرسالها عن طريق الشبكات اللاسلكية (شبكات أدهوك اللاسلكية) ونهتم بالأخص بجودة التجربة (QoE) وذلك للطبيعة الخاصة لهذه التقنية وحساسيتها الشديدة لفقدان الحزم والوقت الحقيقي للإرسال . نقترح بعد ذلك نموذج جودة التجربة (QoE) الذي يعتمد على ترميز الفيديو عالي الكفاءة (HEVC) ، وذلك بالاعتماد على تقنية MDC.

المساهمات الرئيسية هي : (1) تقييم أداء معيار ترميز الفيديو الجديد والناشئ HEVC / H.265 ، الذي يستند على تباين قيم معملية التكميم (QP) اعتماداً على محتويات الفيديو المختلفة لاستنتاج تأثيرها على التسلسل المراد؛ (2) التحقيق في تطبيقات الوسائط المتعددة لدعم QoE في الشبكات اللاسلكية ، ومن ثم فحص تأثير فقدان الحزم على QoE من التتابعات الفيديوية المرسله، (3) تصميم مخطط تكيف HEVC المبني على MDC مع معملية الترميز ونموذج QoE الموضوعي.

وكشفت الدراسة أن الآلية المقترحة (MDC) فعالة في تحسين الإرسال. كذلك تظهر النتائج أن الاختيار الجيد لقيم QP يمكن أن تعوض تأثيرات قناة الإرسال ويحسن جودة الفيديو المستقبلية. النتائج التي تم الحصول عليها تظهر كفاءة طريقتنا المقترحة.

الكلمات المفتاحية: ضغط الفيديو HEVC/H.265، ترميز الوصف المتعدد (MDC) ، جودة التجربة (QOE)، معملية التكميم (QP)، نسبة فقدان الحزم (PLR) ، شبكات أدهوك اللاسلكية.

LISTE DES ACRONYMES

A

AI: All Intra
ALF: Adaptive Loop Filtering
AMVP: Advanced Motion Vector Prediction
AP: Access Point
ARQ: Automatic Repeat Request
ATM: Asynchronous Transfer Mode
AVC: Advanced Video Coding

B

BA: Block Based Adaptation
BO: Band Offset

C

CABAC: Context Adaptive Binary Arithmetic Coding
CAVLC: Context Adaptive Variable Length Coding
CB: Coding Block
CCD: Charge Coupled Devices
CIE: Commission Internationale de l'Eclairage
CIF: Common Intermediate Format
CPU: Central Processing Unit
CRA : Clean Random Access
CTB: Coding Tree Block
CTU: Coding Tree Unit
CU: Coding Unit

D:

DARPA: Defense Advanced Research Projects Agency
DCT: Discrete Cosine Transform
DST: Discrete Sine Transform
DV: Digital Video

E :

EO: Edge Offset
ERC: Error Resilient Coding

F

FCIF: Full Common Intermediate Format
FEC: Forward Error Correction
FMO: Flexible Macroblock Ordering
FPS: Frames per Second

G

GOP: Group of Picture
GPL: General Public License
GSM: Global System for Mobile

H

HD: High Definition
HEB: High Efficiency Binarization
HEVC: High Efficiency Video Coding
HM: HEVC Test Model
HTB: High Throughput Binarization

I

IDR: Instantaneous Decoder Refresh
IEEE: Institute of Electrical and Electronics Engineers
IETF: Internet Engineering Task Force
IP: Internet Protocol
IPv6: Internet Protocol version 6
IPTV: Internet protocol Television
ISO: International Standardization Organization
ITU: International Telecommunication Union

J

JCT-VC: Joint Collaborative Team-Video Coding
JPEG: Joint Photographic Experts Group
JSCC: Joint Source Channel Coding
JVT: Joint Video Team

L

LAN: Local Area Network
LC: Layered Coding
LD: Low Delay
LTE: Long Term Evolution

M

MAC: Medium Access Control
MANET: Mobile Ad hoc Networks
MB: Macro Blocks
MDC: Multiple Description Coding
MDSQ: Multiple Description Scalar Quantizer
MOS: Mean Opinion Score
MPEG: Moving Picture Experts Group

N

NTSC: National Television System Committee

O

OSI: Open System Interconnection

P

PAL: Phase Alternation Line
PB: Prediction Block
PDA: Personal Digital Assistant
PLR: Packet Loss Ratio
PPS: Picture Parameter Set
PSNR: Peak Signal to Noise Ratio
PSS: Polyphase Spatial Subsampling
PU: Prediction Unit

Q

QCIF: Quarter Common Intermediate Format
QoE: Quality of Experience
QoS: Quality of Service
QP: Quantization Parameter

R

RA: Random Access
RA: Region Based Adaptation
RDOQ: Rate-Distortion Optimized Quantization
RGB: Red Green Blue
RNIS: Réseau Numérique à Intégration de Services
RTP: Real Time Protocol
RVB: Rouge Vert Bleu

S

SAO: Sample Adaptive Offset
SD: Standard Definition
SDC: Single Description Coding
SECAM: Séquentiel Couleur avec Mémoire
SI: Spatial Perceptual Information
SIF: Source Intermediate Format
SPS: Sequence Parameter Set
SQCIF: Sub Quarter Common Intermediate Format
SSIM: Structural Similarity
SURAN: Sur Vivable Radio Networks
SVC: Scalable Video Coding

T

TB: Transform Block
TCP/IP: Transfer Control Protocol/ Internet protocol
TI: Temporal Perceptual Information
TU: Transform Unit

U

UDP : User Datagram Protocol
UHD: Ultra High Definition
URQ: Uniform Reconstruction Quantization

V

VCEG: Video Coding Experts Group
VLC: Variable Length Coding
VoD: Video on Demand
VOIP: Voice over Internet Protocol
VPS: Video Parameter Set

W

WECA: Wireless Ethernet Compatibility Alliance
WLAN: Wireless Local Area Network
WMV: Windows Media Video

LISTE DES FIGURES

Chapitre 01

Figure I.1 Décomposition d'une image couleur.....	11
Figure I.2 Le signal vidéo monochrome.....	12
Figure I.3 Le signal vidéo couleur.....	12
Figure I.4 Signal vidéo composite.....	12
Figure I.5 Image fixe d'une scène vidéo.....	13
Figure I.6 Échantillonnage spatial et temporel d'une séquence vidéo.....	14
Figure I.7 Séquence vidéo entrelacée.....	15
Figure I.8 (a) Lignes paires, (b) Lignes impaire, (c) Une image pleine.....	16
Figure I.9 Représentation proportionnelle de toutes les définitions.....	18
Figure I.10 Encodeur / Décodeur.....	18
Figure I.11 Numérisation d'un signal vidéo.....	21
Figure I.12 Echantillonnage couleur.....	22
Figure I.13 Bloc diagramme d'un encodeur vidéo.....	23
Figure I.14 Structure d'un GOP dans un codec vidéo.....	25

Chapitre 02

Figure II.1 Comparaison de la structure de codage entre H.264 / MPEG-4 AVC et HEVC.....	31
Figure II.2 Schéma de principe du codeur HEVC.....	33
Figure II.3 Picture, Slice, Unité d'arbre de codage (CTU), Unité de codage (CU).....	34
Figure II.4 Unité d'arbre de codage (CTU) et Bloc d'arbre de codage (CTB).....	35
Figure II.5 Structure de CTB et CU.....	36
Figure II.6 Unité de prédiction (PU) et Bloc de prédiction (PB).....	37
Figure II.7 Bloc de transformation (TB).....	37
Figure II.8 Slice et CTU d'une trame vidéo.....	39
Figure II.9 Un exemple montrant une image partitionnée en neuf tuiles.....	39
Figure II.10 Les trames d'une séquence vidéo.....	40
Figure II.11 Mécanisme HEVC.....	41
Figure II.12 Les images de base de la DCT 8x8.....	43
Figure II.13 Codage entropique dans HEVC.....	44
Figure II.14 An exemple d'intensité et groupes de bands dans 80 modes pour 8 bits.....	46
Figure II.15 Mode EO.....	46
Figure II.16 Une forme de ALF.....	47
Figure II.17 Directions de prediction H.265/HEVC Vs H.264/AVC.....	48
Figure II.18 Structure d'un bitstream HEVC.....	49
Figure II.19 Structure d'un bitstream HEVC (source Parabola research).....	50

Chapitre 03

Figure III.1 La structure principale de la technique MDC d'une séquence vidéo.....	55
Figure III.2 Sous-échantillonnage MDC spatial polyphasé.....	57
Figure III.3 Simple MDC basé sur MDSQ.....	57
Figure III.4 MDC-FEC avec nombres égaux de descriptions.....	58
Figure III.5 (a) Codage à description unique (b) MDC dans le domaine temporel.....	59

Chapitre 04

Figure IV.1 Schéma du support d'interconnexion matériel.....	61
Figure IV.2 Les quatre couches de TCP/IP.....	62
Figure IV.3 Le format du Paquet IP.....	65
Figure IV.4 L'adressage IP.....	67
Figure IV.5 La décomposition des réseaux sans fil.....	69
Figure IV.6 Le modèle des réseaux sans fil avec infrastructure.....	69
Figure IV.7 Le modèle des réseaux sans fil sans infrastructure.....	70

Chapitre 05

Figure V.1 Un système de communication vidéo typique.....	77
Figure V.2 Un système de communication vidéo simplifié.....	78
Figure V.3 Diagramme d'évaluation typique de la vidéo.....	79
Figure V.4 Réseau point à point	82
Figure V.5 VLC media player.....	83
Figure V.6 Echantillons de vidéos de test.....	85
Figure V.7 Plans TI et SI pour les vidéos brutes.....	85
Figure V.8 Quelques frames de la séquence vidéo "SOCCER".....	86
Figure V.9 Quelques frames de la séquence vidéo "HARBOUR".....	86
Figure V.10 Quelques frames de la séquence vidéo " PARKJOY".....	87
Figure V.11 Quelques frames de la séquence vidéo "DUCKSTAKEOFF".....	87
Figure V.12 Quelques frames de la séquence vidéo "ICE".....	88
Figure V.13 Quelques frames de la séquence vidéo " CROWDRUN".....	88
Figure V.14 Schéma synoptique de codeur HM.....	90
Figure V.15 Mode All Intra.....	92
Figure V.16 Mode Random Access.....	93
Figure V.17 Mode Low Delay.....	94
Figure V.18 Schéma d'adaptation pour l'évaluation objective/subjective des séquences vidéo.....	94
Figure V.19 Schéma d'adaptation pour l'évaluation objective/subjective des séquences vidéo basée sur la méthode MDC.....	95
Figure V.20(a) PSNR de en fonction de QP.....	97
Figure V.20(b) PSNR en fonction de QP (cas de MDC)	97
Figure V.20(c) SSIM de en fonction de QP.....	97
Figure V.20(d) SSIM de en fonction de QP (cas de MDC)	97
Figure V.20(e) MOS Estimé en fonction de QP.....	97
Figure V.20(d) MOS Estimé de en fonction de QP (cas de MDC)	97
Figure V.21(a) Débit en fonction de QP.....	101
Figure V.21(b) Débit de MDC vs SDC.....	101
Figure V.22 Système proposé.....	103

Figure V.23 Echantillons des trames des séquences montrant les artefacts produits par PLR..	104
Figure V.24 Schéma d'adaptation pour l'évaluation objective de la séquence vidéo.....	105
Figure V.25 Impact du PLR sur la qualité perçue sans schéma MDC.....	109
Figure V.26 Impact du PLR sur la qualité perçue avec le schéma MDC.....	111
Figure V.27 Impact de QP et PLR sur les moyens PSNR pour toutes les séquences vidéo.....	114

LISTE DES TABLEAUX

Tableau I.1 Les différents standards de la vidéo analogique.....	10
Tableau I.2 Les formats vidéo.....	17
Tableau I.3 Exemple de débit mesuré pour une vidéo HD.....	19
Tableau II.1 Comparaison entre HEVC/H.265 / H.264/MPEG-4 AVC.....	49
Tableau IV.1 Modèle TCP/IP et OSI.....	63
Tableau V.1 Conversion PSNR en MOS.....	80
Tableau V.2 Analyse comparative des approches objectives.....	81
Tableau V.3 Échelle de qualité.....	81
Tableau V.4 Échelle de dégradation.....	81
Tableau V.5 Base de données EPFL-PoliMI d'évaluation de la qualité vidéo.....	84
Tableau V.6 Séquences vidéo originales avec résolution, fréquence image (fps) et longueur.....	84
Tableau V.7 Conditions d'essai de HM16.11.....	96
Tableau V.8 Variation du PSNR dans le cas de codeur HEVC/H.265.....	99
Tableau V.9 Variation du PSNR dans le cas de codeur HEVC/H.265 basée sur MDC.....	99

Table des matières

Remerciement	II
Résumé	III
Liste des Acronymes	VI
Liste des figures	X
Liste des tableaux	XIII
Introduction générale	1
Chapitre I: La compression de la Vidéo	7
I.1 Introduction.....	7
I.2 Image numérique.....	7
I.3 Définition de la vidéo.....	9
I.3.1 vidéo analogique.....	9
I.3.1.1 Standards de vidéo analogiques.....	9
I.3.1.2 Le standard YUV.....	10
I.4 Vidéo numérique.....	13
I.4.1 Scènes vidéo naturelles.....	13
I.4.2 Capture.....	14
I.4.3 Échantillonnage spatial.....	14
I.4.4 Échantillonnage temporel.....	15
I.5 Vidéo entrelacée.....	15
I.6 Vidéo progressif.....	16
I.7 Les formats vidéo.....	16
I.8 La compression vidéo.....	18
I.9 La Quantification.....	22
I.10 Les codecs vidéo.....	23
I.11 Groupe d'images (GOP).....	24
I.12 Les normes de compression vidéo.....	25
I.12.1 JPEG.....	25
I.12.2 MPEG-1.....	26
I.12.3 MPEG-2.....	26
I.12.4 H.261.....	27
I.12.5 H.263.....	27
I.12.6 H.264 : MPEG-4 AVC.....	28
I.13 Conclusion.....	29
Chapitre II: High Efficiency Video Coding	30
II.1 Introduction.....	30
II.2 HEVC / H.265 (High Efficiency Video Coding).....	30
II.3 Les Versions de la norme HEVC / H.265.....	32
II.4 Fonctionnement de l'encodeur HEVC (H.265).....	33
II.5 Partition des images (Picture partitioning).....	33
II.6 Structure des SLICES et des TILES.....	38
II.7 Prédiction.....	40
II.8 Transformation et Quantification.....	40
II.8.1 La transformée en cosinus discrète (DCT).....	41

II.8.2	Quantification.....	43
II.9	Codage entropique.....	43
II.10	Filtrage de boucle.....	45
II.11	Filtre de déblocage.....	45
II.12	Décalage adaptatif d'échantillon (Sample Adaptive Offset).....	45
II.13	Filtrage de boucle adaptatif (Adaptive Loop Filtering).....	46
II.14	Performance par rapport au H.264 / AVC.....	47
II.15	Les unités NAL (NAL unit).....	49
II.16	Conclusion.....	50
Chapitre III: Multiple Description Coding.....		51
III.1	Introduction.....	51
III.2	Effets des erreurs et les erreurs de propagation.....	51
III.2.1	Effets d'erreurs binaires.....	51
III.2.2	Effet des pertes de paquets.....	52
III.3	Techniques de codage source/ canal pour la protection des bitstreams vidéo.....	52
III.3.1	Requête automatique de répétition (Automatic Repeat reQuest ARQ).....	52
III.3.2	Correction d'erreur directe (Forward Error Correction FEC).....	53
III.3.3	Erreur de codage résilient (Error Resilient Coding ERC).....	53
III.4	MDC (Codage à description multiple).....	54
III.5	Méthode.....	54
III.6	Schémas de codage vidéo à description multiple.....	56
III.6.1	MDC dans le domaine spatial.....	57
III.6.2	MDC dans le domaine fréquentiel.....	57
III.6.3	MDC dans le domaine compressé.....	58
III.6.4	MDC dans le domaine temporel.....	58
III.6	Conclusion.....	60
Chapitre IV: Réseaux IP (Vidéo IP).....		61
IV.1	Introduction.....	61
IV.2	Architecture OSI.....	61
IV.3	Architecture TCP/IP.....	62
IV.4	Les principaux protocoles utilisés en transmission vidéo.....	64
IV.4.1	Protocole TCP.....	64
IV.4.2	Protocole UDP.....	64
IV.4.3	Protocole IP (Le paquet IP).....	65
IV.4.4	Adressage IP.....	66
IV.5	Modes de transmission.....	68
IV.5.1	Mode Unicast.....	68
IV.5.2	Mode Broadcast.....	68
IV.5.3	Mode Multicast.....	68
IV.6	Les environnements mobiles et sans fils.....	69
IV.6.1	Les réseaux avec infrastructure.....	69
IV.6.2	Les réseaux sans infrastructure.....	70
IV.7	Les réseaux mobiles ad hoc.....	70
IV.7.1	Bref historique.....	70
IV.7.2	Les caractéristiques des réseaux ad hoc.....	71
IV.7.2.1	Une topologie dynamique.....	71
IV.7.2.2	Une bande passante limitée.....	71
IV.7.2.3	Les contraintes d'énergie.....	72
IV.7.2.4	L'hétérogénéité des nœuds.....	72

IV.8	Réseaux locaux sans fil (WLAN) IEEE 802.11.....	72
IV.9	Qualité de Service (QOS).....	72
IV.10	Qualité d'Expérience (QoE) de la télévision / IPTV.....	75
IV.11	Utilisation de la qualité d'expérience.....	75
IV.12	Qualité de service et réseau ad hoc.....	76
IV.13	Conclusion	76
Chapitre V: Analyse des performances et résultats.....		77
V.1	Introduction.....	77
V.2	Fonctionnement général d'une transmission.....	77
V.2.1	La phase d'émission.....	78
V.2.2	La phase de réception.....	78
V.3	Méthodes d'évaluation	79
V.4	Visioconférence entre deux participants.....	82
V.4.1	Les logiciels de vidéo Player.....	82
V.4.2	VLC media player.....	82
V.5	Bases de données vidéo.....	83
V.6	L'implémentation de codeur et décodeur sous HEVC.....	89
V.6.1	Codeur x265.....	89
V.6.2	Codeur HM	89
V.6.3	Groupe d'images (GOP) dans HM	90
V.6.4	Modes de configuration de l'encodeur.....	91
V.7	Paramètres de configuration	94
V.7.1	Modèle de paramètre d'encodeur basé sur MDC.....	95
V.8	Simulations, résultats et discussions.....	96
V.8.1	Analyse des performances des H.265/HEVC.....	96
V.8.2	Comparaison de paramètre d'encodeur basé sur MDC.....	98
V.8.3	Variation du débit en fonction du pas de quantification.....	100
V.8.4	Impact de perte des paquets	102
V.8.4.1	FFmpeg.....	102
V.8.5	Configuration expérimentale.....	102
V.8.6	Système proposé.....	103
V.8.7	L'impact de perte de paquets sur le streaming vidéo HEVC.....	107
V.8.8	L'impact de perte de paquets sur le streaming vidéo HEVC Basé sur MDC.....	107
V.8.9	Discussion.....	112
V.9	Conclusion.....	115
Conclusion et perspectives.....		116
Bibliographie.....		119

Introduction Générale & Problématique

La vidéo est l'un des médias la plus importante pour les communications, le divertissement et de nombreuses autres applications. A l'origine, la vidéo était analogique depuis sa génération, en utilisant des capteurs, sa transmission par radio ou par câble jusqu'à sa réception. L'évolution des nouvelles techniques et technologies numériques a conduit à la numérisation de ce type de contenu et nous a permis d'avoir des vidéos de haute qualité et haute définition [1] [2].

En effet, depuis la fin des années 1970, la résolution de la vidéo n'a cessé de croître pour passer de quelques centaines de kilopixels par image à plusieurs dizaines de mégapixels par image, comme le UHD 8K (Ultra High Definition ou 4320p) avec une résolution de 7680×4320 . Ceci était cependant accompagné d'une énorme quantité de données numériques pour traiter, transmettre et stocker sur des ordinateurs de bureau, des smartphones, des tablettes et d'autres appareils. En effet, selon le Cisco Visual Networking Index 2017 [1], la vidéo d'ici jusqu'à 2021 représentera plus des trois quarts des données échangées sur les réseaux. Les équipements mobiles fourniront 20% du trafic IP (Internet Protocol), qui atteindra alors plusieurs dizaines d'exaoctets (10^{18} octets) par mois [1]. La solution retenue pour ces contraintes est bien sûr la compression qui s'est imposée comme domaine de recherche et était vivement anticipé depuis la fin des années 1980 [3].

Depuis le milieu des années 1990, la communauté internet a permis la transmission de vidéo numérique sous IP et tous types de réseaux [1]. Cependant, la diffusion sous IP a conduit à plusieurs problèmes, dont la largeur de la bande passante utilisée par les réseaux IP très variable dans le temps et relativement étroite [2][4]. La transmission sous IP souffre d'autres problèmes tels que la perte de paquets et les retards...etc. La perte de paquets est plus prononcée sur les réseaux sans fil [4][5], en particulier sur les réseaux ad hoc.

La vidéo numérique non compressée contient des quantités significatives de données. Les codecs vidéo jouent donc un rôle majeur, car ils peuvent compresser des séquences vidéo avec de grandes tailles des trames et des fréquences d'images relativement faibles tout en maintenant une qualité perceptive acceptable [4][5]. Le terme «streaming vidéo» fait référence à la transmission de la vidéo (peut ne pas être préenregistrée) de manière presque en temps réel, permettant au public de consommer le contenu immédiatement à son arrivée [6][7][8]. Actuellement, le streaming vidéo est sans doute l'une des applications multimédia les plus courantes [3][9][10] et il se déroule principalement sur des réseaux sans fil.

Les communications vidéo sur les réseaux sans fil ont toujours été un domaine difficile [2]. Il n'est pas aisé de réaliser un flux vidéo avec un niveau de qualité requis à partir des réseaux qui induisent des erreurs et avec une bande passante limitée [2] [11]. Ceci est encore plus difficile lorsque la transmission est en temps réel avec des exigences strictes [5].

La qualité de la vidéo transmise est influencée par deux types de distorsion qui sont plus prononcés dans les réseaux sans fil:[10]

1. La distorsion de la source résultant de la compression avec perte fournie par le codeur.
2. La distorsion causée par un canal sans fil avec perte.

Traditionnellement, la qualité vidéo a été mesurée en termes de qualité de service (QoS), un concept qui a commencé à être utilisé dans les années 1990[6] pour désigner un ensemble de techniques de réseau garantissant le routage du trafic sensible comme la vidéo et la voix. L'abréviation QoS a été utilisée partout pour souligner les améliorations des performances apportées par les nouveaux produits matériels et logiciels.

L'International Telecommunication Union (ITU) a défini la qualité de service comme «un ensemble des caractéristiques du service de télécommunication axés sur la satisfaction des utilisateurs» [6][11] tandis que l'Internet Engineering Task Force (IETF) caractérise la QoS comme «L'ensemble des exigences qui doivent être satisfaites par le flux de données de transfert de service particulier» [7]. La bande passante, le délai, la gigue et le taux de perte de paquets (PLR) sont des paramètres les plus couramment utilisés pour mesurer la qualité de service.

L'abréviation QoE (Quality of Experience) a été présentée plus récemment comme une approche globale de la qualité (mesure de bout en bout). La QoE consiste à mesurer les performances directement avec l'utilisateur [4][12]. Elle dépend de paramètres subjectifs qui déterminent l'interaction entre la vidéo présentée et la perception de l'utilisateur (de la couleur, de l'intensité de la lumière, etc.) et ensuite l'exprimer en mots, comme «excellent», «bon», «juste», «médiocre» ou «mauvais». Qualinet (European Network on Quality of Experience in Multimedia Systems and Services) a défini QoE comme « le degré de satisfaction ou de gêne de l'utilisateur pour l'application ou le service. Il résulte de réaliser ses attentes concernant le bénéfice et/ou le plaisir de l'application ou du service en fonction de la personnalité de l'utilisateur et de la situation actuelle.»[13]. La QoE est basée sur une mesure subjective du service fourni. La QoS repose sur la mesure objective des composants constituant le service. Au cours des dernières années, le terme QoS a été jugé fragile en termes d'aptitude à déterminer la qualité et le terme QoE est devenu l'alternative dominante [11].

Objectif de la thèse:

De ce contexte, cette thèse vise quatre objectifs:

- Une étude de performances de la nouvelle et émergente norme du codage vidéo (le codage vidéo à haute efficacité HEVC).
- Analyser l'impact du contenu vidéo (caractérisé par les activités de mouvement et la complexité des séquences vidéo) et les paramètres de codage (notamment Quantization Parameter (QP), résolution et fréquence d'images) et de présenter un modèle pour les paramètres du codeur HEVC/H.265.
- Le modèle proposé traite les problèmes d'une transmission fiable de la vidéo HEVC/H.265 sur des environnements de communication sans fil. Le système de transmission est une méthode conçue pour la transmission vidéo numérique en temps réel sur un canal hostile avec la concentration sur la protection au niveau de la source tel que Multiple Description Coding (MDC) en tant qu'approche qui lutte contre la perte de paquets vidéo. Dans ce système, les vidéos ne sont pas soumises au codage de canal.
- la qualité vidéo est évaluée objectivement selon un modèle de prédiction de QoE dans des scénarios de réseau IP sans fil et les métriques les plus utilisées telles que Peak Signal to Noise Ratio (PSNR), structural similarity (SSIM) et Mean Opinion Score estimé (MOS_Estimated) sont étudiées.

Etat de l'art:

Divers articles ont exploré l'impact des différentes déficiences sur l'expérience, les approches et les méthodologies utilisées pour évaluer la qualité de la vidéo dans les services multimédias.

Dans ce contexte, plusieurs travaux ont étudié la qualité vidéo perceptuelle pour différents codecs. Hadar et al.[14] ont utilisé des paramètres de compression, d'activités temporelles et spatiales, taille de trame, fréquence d'image et PLR pour un schéma d'encodage afin d'obtenir une qualité vidéo perceptuelle élevée. Cependant, l'effet du contenu vidéo n'a pas été étudié dans ce travail. Ils ont fait varier la perte de paquets dans la même séquence vidéo et ont trouvé que le PLR avait le plus grand effet. En outre, Hands et al.[15] suivent l'évaluation de la qualité vidéo en utilisant le codec MPEG-1 (Moving Picture Experts Group-1) et en modifiant la perte de paquets et la taille de la rafale de paquets. Ils ont démontré que la configuration des caractéristiques de perte de paquets peut améliorer la QoS dans des conditions de congestion du réseau.

D'autre part, Paudyal et al. [4][5] ont expliqué l'impact du contenu vidéo et des dégradations de transmission sur la QoE du flux vidéo MPEG-2 (Moving Picture Experts Group-2) transmis via un réseau filaire. La norme vidéo MPEG-2 est impliquée et a été constaté que le paramètre perte de paquets a une influence significative sur l'expérience de l'utilisateur et que l'impact du codec et des paramètres réseau sur la QoE dépend du contenu vidéo.

Des travaux antérieurs ont également porté sur des études utilisant le codec H.264/AVC (Advanced Video Coding), en utilisant des méthodes d'évaluation objectives et subjectives. Atzori et al.[16] ont travaillé sur la consommation de la qualité vidéo sur les tablettes. Dans cette étude, les séquences vidéo codées en utilisant le H.264/AVC sont transmises via un canal sans fil corrompu typique avec interruption de transmission. Les résultats de la qualité vidéo subjective ont été enregistrés et analysés. Tommasi et al.[17] ont présenté quelques études expérimentales proposées pour évaluer l'impact des pertes de paquets sur la qualité vidéo dans les systèmes de streaming. Ils ont manipulé l'effet des pertes de paquets isolés sur certaines métriques de qualité, grâce à l'analyse hors ligne des métriques de qualité. Quelques remarques utiles peuvent être déduites pour le réglage de la source vidéo afin de garantir la meilleure QoE possible aux derniers utilisateurs. En outre, ils ont étudié comment minimiser l'impact de dégradation sur une métrique de qualité vidéo lorsque les pertes de paquets augmentent en utilisant un débit binaire variable.

De plus, Maia et al.[18] ont discuté les évolutions et l'utilisation des métriques objectives et subjectives pour la mesure de la qualité vidéo. Ainsi, Pozueco et al. [19] ont expliqué le rôle de la QoE dans les réseaux multimédias. Ils évaluent nombreux facteurs qui affectent la fonction QoE dans les services multimédias en temps réel. Tavakoli et al.[12] ont analysé l'impact de QoE et quelques paramètres techniques importants pour les services de streaming adaptatifs.

Certaines publications ont abordé la qualité vidéo pour la norme HEVC/H.265. La majorité de ces travaux sont concentrés sur la conception des outils d'encodage. Kokkonis et al. [20] ont présenté les avantages de HEVC sur son prédécesseur H.264, pour la transmission de flux vidéo. Ils ont analysé une architecture proposée pour un système de surveillance avec des fonctionnalités 3D-HEVC pour améliorer la QoE en temps réel pour une communication sans fil. Ceci, en utilisant l'évaluation objective comme la métrique PSNR pour mesurer la qualité de la vidéo. Milicevic et al.[21] ont cherché à analyser les performances du modèle de test HM (HEVC Test Model) pour différentes séquences de test et structures d'image. Ils ont mené des tests pour une valeur fixe de QP et pour des configurations en mode faible retard (Low Delay), Les résultats obtenus ont montré les différences de débit et de temps de codage,

en particulier avec l'impact du contenu vidéo (résolution, mouvement, etc.) et des structures d'image. L'article écrit par Anegekuh et al. [22] a examiné l'impact des types de contenu vidéo et les paramètres de codage basés sur QP et le débit binaire dans la vidéo HEVC, avec des tests objectifs et subjectifs. Par conséquent, les résultats obtenus ont montré que les paramètres de codage et une variété de types de contenu vidéo affectent la qualité vidéo. Dans ce contexte, Nightingale et al. [2] ont analysé l'impact des dégradations de réseau sur la QoE en utilisant le streaming vidéo H.265/HEVC. En outre, l'impact des différentes déficiences sur l'expérience de l'utilisateur est étudié, mais en prenant en considération certaines déficiences spécifiques. Ainsi, les évaluations de plusieurs types de dégradations sur différents types de contenus vidéo ne sont pas discutées. L'influence de ces dégradations sur la QoE doit être analysée pour prendre des décisions dans la conception des systèmes multimédias adaptatifs.

Kazemi et al.[23] ont présenté un nouveau système, dont l'idée principale est l'utilisation du H.264/AVC basé sur le MDC pour la transmission vidéo sans fil. Pereira et al.[24] ont démontré que le schéma MDC temporel est une alternative simple pour le streaming vidéo où les méthodes qui utilisent des schémas de contrôle d'erreur, telles que Forward Error Correction (FEC) ou Automatic Repeat Request (ARQ), ne conviennent pas aux dégradations de transmission. Ainsi, ils ont évalué la performance du codeur image/vidéo MDC proposé pour deux descriptions. En outre, Chen et al.[3] ont discuté l'effet de la perte de paquets due au transfert de canal sur la distorsion de la qualité vidéo via un schéma de protection contre les erreurs adaptatives utilisant MDC sur des réseaux locaux sans fil. Zang et al. [25] ont proposé l'utilisation des méthodes MDC et FEC pour le streaming vidéo MPEG-4 (Moving Picture Experts Group-4) à travers les réseaux mobiles sans fil ad hoc (Mobile Ad hoc Networks MANET) avec perte de paquets. Ils ont démontré que les méthodes MDC atteignent une meilleure qualité de streaming vidéo, en termes de fréquence d'images, résolution et performances en temps réel. Dans [26] Kazemi et al. ont publié plusieurs études sur le MDC.

Plan de la thèse:

Cette thèse est organisée en cinq chapitres présentés de la manière suivante:

Dans le premier chapitre, les notions de base relatives à l'image et la vidéo sont présentées. Le principe général de fonctionnement des codeurs vidéo est traité et il est suivi par les principales normes de compression.

Le deuxième chapitre présente les fondements de base de la norme H.265/HEVC où l'ensemble de ses aspects sont détaillés. Il inclut les différents profils et niveaux, les modes de codage (I, P et B), la transformée en cosinus discrète (DCT), la transformée en sinus discrète (DST), l'algorithme de codage CABAC (Context Adaptive Binary Arithmetic Coding), ainsi, la différence entre HEVC/H.265 et H.264/MPEG-04 AVC.

Le principe général de fonctionnement des méthodes de protection de l'information et la lutte contre les pertes sont traités au troisième chapitre et une étude détaillée sur le MDC est présentée.

Dans le quatrième chapitre les fondamentaux de base sur les réseaux IP, video IP, réseaux sans fil et QoS/QoE où l'ensemble de ses aspects sont détaillés.

Le dernier chapitre est consacré aux simulations menées sur la plateforme HM 16.11 afin d'analyser les performances du standard H.265/HEVC. Différents paramètres tels que le contenu de la vidéo, les dimensions des frames, le nombre de frames de la séquence à coder, le QP, ... etc., sont étudiés et leurs effets sur la qualité et le débit sont mesurés et testés. Les résultats obtenus sont analysés, discutés et comparés. Aussi, nous avons également énuméré quelques mesures offrant la QoS/QoE de la vidéo (standard H.265/HEVC) sur IP en particulier la perte de paquets. Les émulations ont été menées sur le système d'exploitation "Linux".

Enfin, les résultats et les performances obtenus sont examinés et discutés et les perspectives susceptibles de compléter le présent travail.

Chapitre I

La compression de la Vidéo

I.1. Introduction

Le monde de la vidéo connaît actuellement une évolution vers le mode numérique, et ce à tous les niveaux. Lorsque l'œil humain perçoit une séquence d'images séquentielles, un phénomène étonnant se produit, si les images sont affichées assez rapidement, l'œil ne les distingue pas séparément, mais perçoit une légère animation. C'est sur cette base que les films et les vidéos sont développés. Le rythme de l'animation est appelé le nombre d'images par seconde. Pour qu'une légère animation soit perceptible à l'œil, une vitesse d'environ 10 images par seconde est nécessaire. Cependant, la qualité des vidéos ne dépend pas seulement du nombre d'images par seconde. La quantité d'informations contenues dans chaque image est également déterminante. Elle est désignée sous le terme de résolution d'image. La résolution correspond en règle générale au nombre d'éléments individuels constituant l'image (pixels) affichés à l'écran. Une résolution plus élevée permet d'obtenir une vidéo de meilleure qualité.

Ainsi, le nombre d'images par seconde et la résolution sont des paramètres très importants en matière de vidéo numérique, car ils déterminent la quantité de données à transmettre et à enregistrer pour la diffusion. Souvent des compromis entre la qualité optimale requise pour la vidéo et les limitations qui y sont imposés.

I.2 Image numérique

Une époque marquante pour le monde de l'image numérique a été le début des années 1990 qui a vu apparaître les premiers standards internationaux concernant la compression des données audiovisuelles [27] [28]. Ces standards, JPEG (Joint Photographic Experts Group) [29] et MPEG (Moving Picture Experts Group) [30][31], ont rapidement réussi à s'imposer dans les domaines informatiques et télévisuels [30], ce qui a permis l'échange et la diffusion des images.

Une image numérique est une image dont le support est stocké sous forme binaire dans un fichier informatique. Celle-ci peut être obtenue soit à partir de capteurs optiques (Appareil photo, Scanner, Caméra, ...) ou créée à partir de logiciels (Photoshop, Libre office...).

Chaque image numérique est constituée d'un nombre donné de lignes. Chaque ligne comporte un nombre de point déterminé. L'ensemble constitue une matrice. Ces points sont appelés pixels (de l'anglais picture element et noté : px). Chaque "case" de cette matrice contient des nombres caractéristiques avec la couleur assignée au pixel.

Habituellement, il existe trois principaux types de couleurs pour une image numérique :

- **Le noir et blanc** : Le noir et blanc est le plus simple. Le contenu de chaque case de la matrice est soit un 0 (noir) soit 1 (blanc), appelé souvent image binaire. Le nombre de couleurs n'est que de 2 et le rendu de l'image le moins performant mais parfois suffisant [32].
- **Les niveaux de gris** : Le codage dit en niveaux de gris permet d'obtenir plus de nuances que le simple noir et blanc. Il offre des possibilités supplémentaires pour coder le niveau de l'intensité lumineuse. La couleur est codée souvent sur un octet soit 8 bits ce qui offre la possibilité d'obtenir 256 niveau de gris (0 pour le noir et 255 pour le blanc). On peut aussi le faire avec 16 niveaux de gris (4bits) [32].
- **La couleur**: [33][34]
 1. **Le RGB** (*Red, green, blue*, pour *Rouge Vert Bleu*, en français *RVB*), mis au point en 1931 [33] par la *Commission Internationale de l'Eclairage* (CIE) consiste à représenter l'espace des couleurs à partir de trois rayonnements monochromatiques de couleurs:
 - ✓ rouge (de longueur d'onde égale à 700,0 nm),
 - ✓ vert (de longueur d'onde égale à 546,1 nm),
 - ✓ bleu (de longueur d'onde égale à 435,8 nm).

Cet espace de couleur correspond à la manière dont les couleurs sont habituellement codées informatiquement, ou plus précisément à la façon dont les tubes cathodiques expriment les couleurs.

Enfin, la version RGB représente chaque composant de couleur par un octet, ce qui est conforme à 256 intensités de rouge (28), 256 intensités de vert et 256 intensités de bleu. Ceci correspond donc à 16777216 possibilités théoriques des couleurs, ce qui signifie plus que ne peuvent percevoir l'œil humain (environ 2 millions).

2. Composantes : La recombinaison linéaire des signaux primaires RVB convertissent les signaux en composantes. Le triplet (rouge, vert, bleu) en un triplet (Y: représentant l'image en niveaux de gris et deux signaux représentent la différence des couleurs R-Y et B-Y notés Cr et Cb). Les pertes en qualité sont assez minimes et la bande passante requise est beaucoup moins importante qu'en RVB car les signaux Cr et Cb, portant uniquement la

couleur, sont peu encombrants. Dans les années 80, le format Betacam de Sony a établi cette représentation du signal vidéo en composantes. Depuis lors, elle est devenue la base au standard numérique 4:2:2, ainsi qu'à la grande majorité des formats d'enregistrement vidéo [35].

3. Y/C (ou composite séparé): c'est un type de représentation qui code la luminance et la chrominance chacune comme en composite, mais restent séparées l'une de l'autre au niveau de la connectique et des équipements. La liaison Y/C est utilisée par les formats grand public S-VHS (Super-Video Home System) et Hi8 [32], ainsi que sur certaines cartes de numérisation pour micro-ordinateurs.

I.3 Définition de la vidéo

Une vidéo est une succession d'images à un certain rythme, véhiculant une grosse quantité d'informations. L'œil humain peut être capable de distinguer environ 20 images par seconde. Donc, en affichant plus de 20 images par seconde, il est possible de tromper l'œil et de lui donner l'impression qu'il s'agit d'images animées avec mouvement. On décrit la fluidité d'une vidéo par le nombre d'images par secondes (de l'anglais frame rate), exprimée en fps (Frames per second, en français trames par seconde). On distingue la vidéo analogique et la vidéo numérique [33].

I.3.1 vidéo analogique

La vidéo analogique représentant l'information transmise sous forme d'un signal analogique, par l'intermédiaire d'ondes ou câble (la télévision support d'affichage vidéo le plus communément répandu). Une image vidéo analogique en couleurs a comme caractéristique d'être reconstituée par le mélange, à quantités variables, des trois couleurs initiales rouge, vert, bleu. Mais en pratique, le signal vidéo ne se présente pas toujours sous la forme de trois signaux RVB. Il peut en fait revêtir d'autres formes différentes.

I.3.1.1 Standards de vidéo analogiques

Il existe 3 standards de base de codage de la couleur de part le monde [32], le NTSC (*National Television System Committee*), le PAL (*Phase Alternating Line*) et le SECAM (*Séquentiel couleur à mémoire*). Cette hétérogénéité est liée à l'histoire de la télévision [32] et à la volonté de certains pays à prouver leur supériorité technologique en adoptant chacun un système différent.

Format de télédiffusion	Pays	Lignes horizontales	Lignes affichées	Fréquence
NTSC	Etats-Unis, Canada, Japon, Corée, Mexique	525 lignes	480 affichées	29,97 images/s
PAL	l'Algérie, Australie, Chine, majeur partie de l'Europe, Amérique du Sud.	625 lignes	575 affichées	25 images/s
SECAM	France, Moyen Orient, majeure partie de l'Afrique	625 lignes	575 affichées	25 images/s

Tableau I.1 Les différents standards de la vidéo analogique [37]

1 - Le standard NTSC

Comité du système de télévision nationale (NTSC) est un standard de codage analogique de la vidéo en couleurs lancée aux États-Unis en 1953 [38]. Il procure un format vidéo de 525 lignes et 30 images par seconde en niveaux de gris (29.97 en couleur). Il peut être exploité pour les DVD-vidéo avec une résolution de 720×480 lignes [38].

2- Le PAL

Alternance de phase suivant les lignes (PAL) est l'une des trois principales normes universelles de codage de couleurs utilisées lors de la diffusion d'un signal vidéo analogique. Historiquement, les concurrents du PAL sont NTSC, d'origine américaine, et SÉCAM, d'origine française [32].

3- Le standard SECAM

Séquentiel Couleur À Mémoire, désigne un mode standardisé de codage vidéo analogique en couleurs, commercialisé à partir de 1967. Adapté aux formats vidéo 625 lignes et 25 images par seconde. Le SECAM a été développé par l'ingénieur Français Henri de France. Il fournit une image en couleur d'aussi bonne qualité que le PAL, il a été adopté pour éviter la pénétration des téléviseurs étrangers (PAL, NTSC) sur le marché français [38]

I.3.1.2 Le standard YUV [33][36]

Le standard YUV (*le format CCIR 601*), auparavant nommé YCrCb (Y Cr Cb), est un modèle de représentation de la couleur consacré à la vidéo analogique.

Il est basé sur un mode de transmission vidéo à composantes séparées utilisant trois câbles différents pour transmettre les informations de luminance (luminosité) et les deux composantes de chrominance (couleur). Il s'agit du format utilisé dans les standards PAL et SECAM.

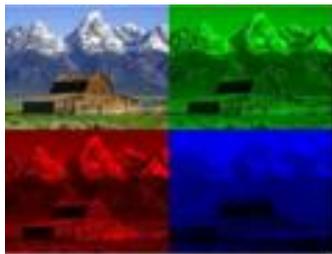
Le paramètre Y permet de représenter la luminance (l'information en niveaux de gris), tandis que U et V représentent la chrominance, c'est-à-dire l'information sur la couleur.

Ce modèle a été mis au point pour permettre d'envoyer des informations colorées aux moniteurs de télévision couleurs, tout en faisant en sorte que les moniteurs de télévision monochromes existant continuent de diffuser une image en tons de gris.

Voici les relations liant Y à R, G et B, U à R,G,B et à la luminance, et enfin V à R,G,B et à la luminance :

- $Y = 0.299R + 0.587 G + 0.114 B$
- $U = -0.147R - 0.289 G + 0.436B = 0.492(B - Y)$
- $V = 0.615R - 0.515G - 0.100B = 0.877(R - Y)$

Ainsi U est parfois noté Cr et V noté Cb, d'où la notation YCrCb.



R V B (rouge vert bleu)



YCbCr (appelé YUV)
 Y luminance = $0.3R+0.6V+0.1B$
 Cb Cr chrominance
 Cb ~ Y-B
 Cr ~ Y-R
 Compatible TV noir et blanc
 Oeil + sensible à Y que CbCr

Figure I.1 Décomposition d'une image couleur

Une connexion YUV permet d'offrir une qualité de vidéo optimale en envoyant simultanément les 575 lignes de l'image, sans entrelacement (en une seule fois).

Les deux composantes de chrominance modulent deux sous-porteuses couleur placées dans le haut du spectre de la luminance.

L'ajout de la couleur se fait donc au détriment du « piqué » de l'image en niveaux de gris.

Image N & B: 8 bandes verticales de plus en plus sombres



Signal vidéo correspondant (1 ligne)

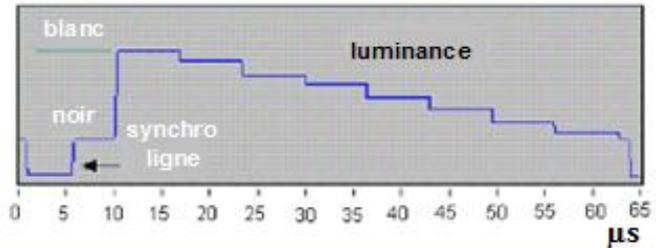


Figure I.2 Le signal vidéo monochrome

Allure d'une ligne du signal vidéo correspondant à 8 bandes verticales colorées :

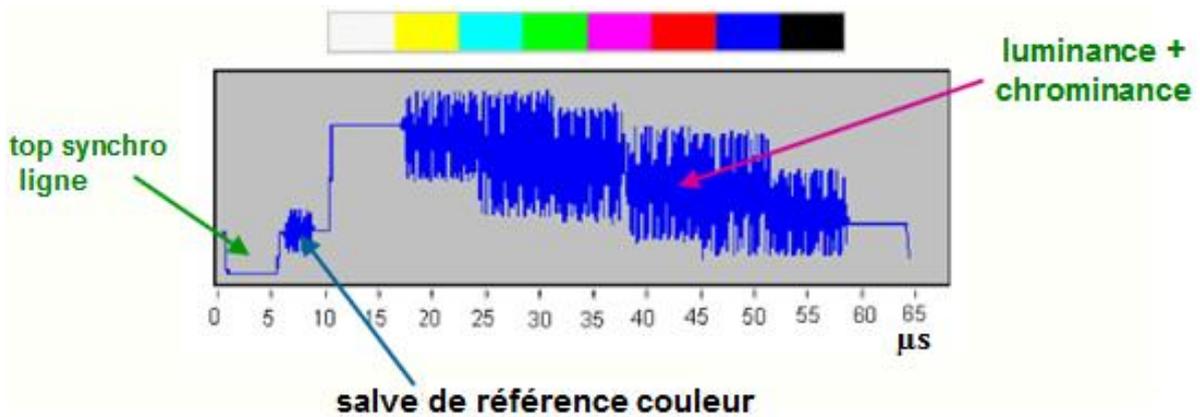


Figure I.3 Le signal vidéo couleur

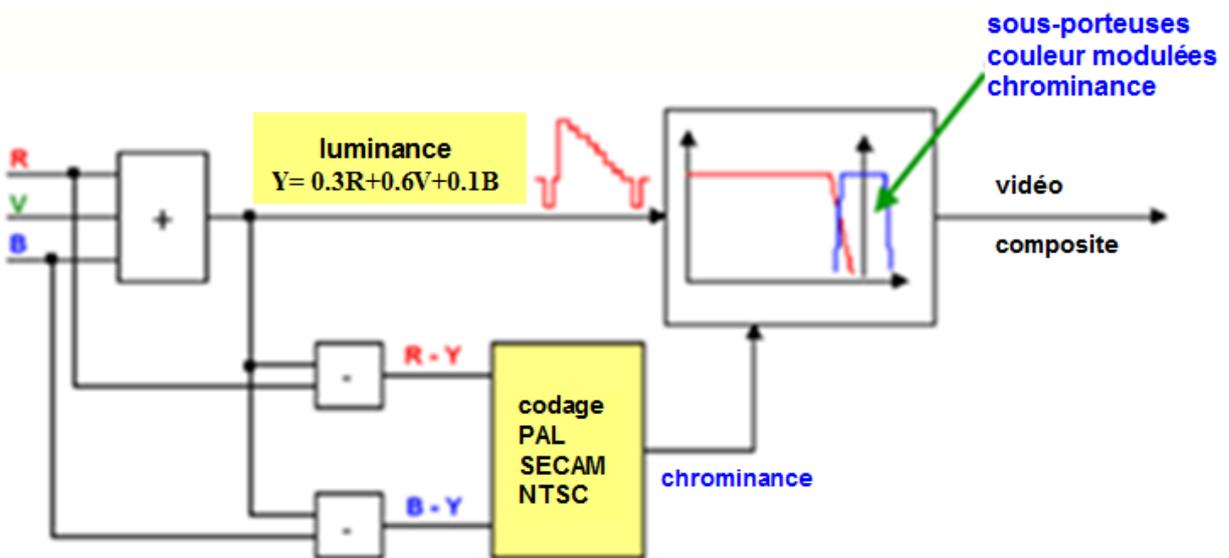


Figure I.4 Signal vidéo composite

La salve de référence couleur donne la fréquence de la sous-porteuse et permet de connaître:

- le standard : PAL, SECAM ou NTSC ?
- la couleur transmise en SECAM : R-Y ou B-Y ?

I.4 Vidéo numérique [31][37]

La vidéo numérique est une représentation d'une scène visuelle, naturelle ou réelle, échantillonnée spatialement et temporellement. Une scène est typiquement échantillonnée à un moment donné pour produire une image, qui représente la scène visuelle complète à ce moment précis, ou un champ, qui consiste typiquement en des lignes d'échantillons spatiaux impaires ou pairs.

L'échantillonnage est répété à intervalles réguliers (intervalles de 1/25 ou 1/30 seconde, par exemple) pour produire un signal vidéo en mouvement. Trois composants ou ensembles d'échantillons sont généralement requis pour représenter une scène en couleur. Les formats les plus répandus pour représenter la vidéo sous forme numérique comprennent la norme ITU-R 601 [15].

I.4.1 Scènes vidéo naturelles

Une scène vidéo est généralement composée de plusieurs objets ayant chacun leur propre forme, profondeur, texture et illumination. La couleur et la luminosité d'une scène vidéo naturelle changent avec des degrés variables de douceur tout au long de la scène, c'est-à-dire qu'elle a un ton continu. Les caractéristiques d'une scène vidéo typique (Figure I.5) qui sont pertinentes pour le traitement et la compression vidéo incluent des caractéristiques spatiales telles que la variation de texture dans la scène, le nombre et la forme des objets, la couleur et autres caractéristiques temporelles et le mouvement de la caméra ou du point de vue [31] [[33].



Figure I.5 Image fixe d'une scène vidéo.

I.4.2 Capture

Une scène visuelle est continue dans l'espace et le temps. La représentation d'une scène visuelle sous forme numérique consiste à échantillonner la scène réelle spatialement, habituellement sur une grille rectangulaire dans le plan image/vidéo, et temporellement, sous la forme d'une série d'images fixes ou composantes d'images échantillonnées à intervalles réguliers (Figure I.6). La vidéo numérique est la représentation d'une scène vidéo échantillonnée sous forme numérique. Chaque échantillon spatio-temporel, un élément d'image ou un pixel, est représenté par un ou plusieurs nombres qui décrivent la luminosité ou la luminance et la couleur de l'échantillon.

Pour obtenir une image échantillonnée en 2D, une caméra focalise une projection en 2-D de la scène vidéo sur un capteur, tel qu'une matrice de dispositifs à couplage de charge (CCD). Dans le cas de la capture d'images couleur, chaque composant de couleur est filtré séparément et projeté sur une matrice CCD [30].

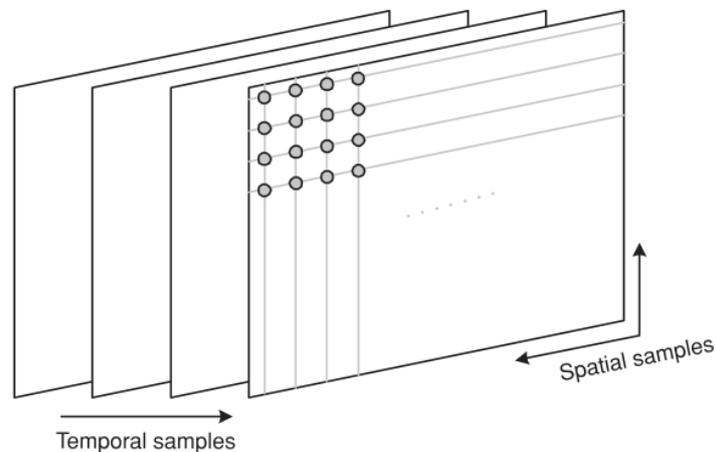


Figure I.6 Échantillonnage spatial et temporel d'une séquence vidéo.

I.4.3 Échantillonnage spatial

La sortie d'une matrice CCD est un signal vidéo analogique, un signal électrique variable qui représente une image vidéo. L'échantillonnage du signal à un moment donné produit une image échantillonnée ou une image qui a des valeurs définies à un ensemble de points d'échantillonnage. Le format le plus courant pour une image échantillonnée est un rectangle avec les points d'échantillonnage positionnés sur une grille carrée ou rectangulaire. L'échantillonnage a lieu à chacun des points d'intersection sur la grille et l'image échantillonnée peut être reconstruit en représentant chaque échantillon comme un élément d'image carré ou un pixel. Le nombre de points d'échantillonnage influence la qualité visuelle de l'image.

I.4.4 Échantillonnage temporel

Une image vidéo en mouvement est formée en prenant un "instantané" rectangulaire du signal à des intervalles de temps périodiques. La lecture de la série d'instantanés ou d'images produit l'apparence d'un mouvement. Un taux d'échantillonnage temporel supérieur ou un taux de trame donne un mouvement apparemment plus fluide dans la scène vidéo, mais nécessite plus d'échantillons à capturer et à stocker. Des fréquences de la vidéo inférieures à 10 images par seconde peuvent être utilisées pour des communications vidéo à très faible débit binaire, car la quantité de données est relativement faible, mais le mouvement est clairement saccadé et non naturel à ce débit. Entre 10 et 20 images par seconde est plus typique pour les communications vidéo à faible débit binaire; l'image est plus lisse mais un mouvement saccadé peut être visible dans les parties de la séquence qui se déplacent rapidement. L'échantillonnage temporel à 25 ou 30 images complètes par seconde est la norme pour les images de télévision en définitions standard (Standard Definition SD), avec entrelacement pour améliorer l'apparence du mouvement, En outre; 50 ou 60 images par seconde produit un mouvement apparent très lisse au détriment d'un débit de données très élevé.

Un signal vidéo peut être échantillonné sous la forme d'une série de trames complètes, d'un échantillonnage progressif ou d'une séquence de trames, d'un échantillonnage entrelacé.

I.5 Vidéo entrelacée

Lorsque vous regardez votre télévision, votre oeil regarde en réalité des demi-images que l'on appelle des trames. L'écran affiche par exemple 50 trames par seconde dont l'une est composée de lignes paires et la suivante de lignes impaires. Lorsque ces trames s'entrelacent elles créent une image pleine. Ce système appelé "entrelacement", chacun représentant la moitié des informations dans une trame vidéo complète (destinée à doubler le nombre d'images perçues par seconde sans augmenter le débit). Par exemple l'abréviation 50i désigne un balayage entrelacé de 50 trames par seconde (i pour interlaced en anglais) [37].

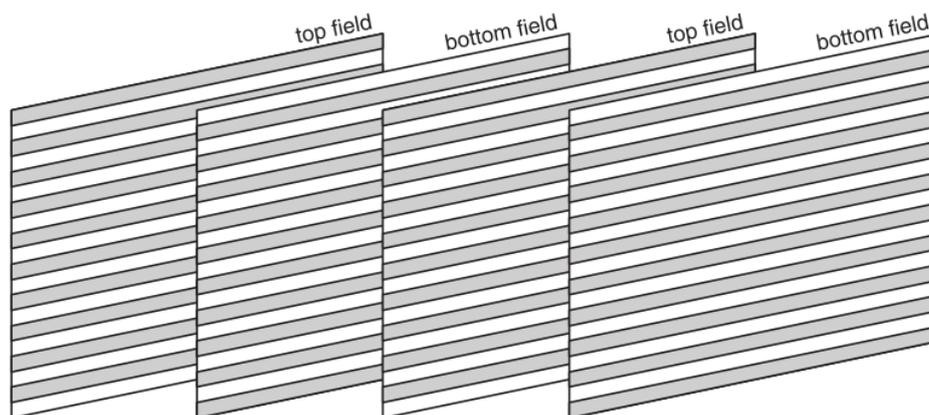


Figure I.7 Séquence vidéo entrelacée.

I.6 Vidéo progressif

Le mode progressif ajoute la précision de la vidéo puisque chaque image s'affiche d'un coup avec toutes ses lignes. En progressif, l'image est alors moins fluide. Pour mentionner le mode progressif, les spécialistes ajoutent la lettre "p" après le nombre de lignes ou d'images par seconde

L'avantage de la méthode d'échantillonnage entrelacée est qu'il est possible d'envoyer deux fois plus de champs par seconde que le nombre de trames dans une séquence vidéo progressif équivalente avec le même débit. Ceci donne l'impression d'un mouvement plus fluide. De plus en plus, le contenu vidéo peut être capturé et / ou affiché dans un format progressif. Lorsque la vidéo est capturée dans un format (par exemple entrelacé) et affichée dans une autre (par exemple progressive), il est nécessaire de convertir entre les formats [33].

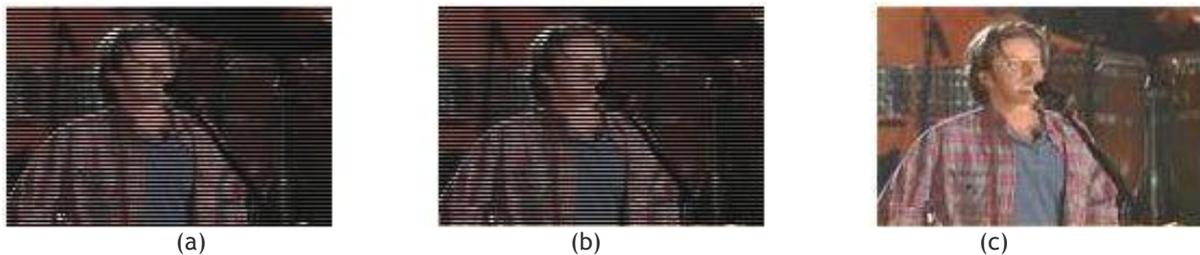


Figure I.8 (a) Lignes paires, (b) Lignes impaire, (c) Une image pleine

I.7 Les formats vidéo [39] [40]

Il y a plusieurs types de résolutions vidéo:

CIF (Common Intermediate Format) : également connu sous le nom FCIF (Full Common Intermediate Format), est un format utilisé pour normaliser les résolutions horizontales et verticales en pixels de séquences Y Cr Cb dans les signaux vidéo, couramment utilisés dans les systèmes de téléconférence vidéo et les services audiovisuels. Il est le format de base du codage des normes H.261 et H.263. CIF définit une séquence vidéo avec une résolution de 352×288 , un frame rate de $30000/1001$ (environ 29,97) comme NTSC, avec une couleur encodée en utilisant Y Cr Cb 4: 2: 0.

SIF (Source Intermediate Format) : SIF est pratiquement identique à CIF, mais pris de MPEG-1 plutôt que des normes ITU. SIF sur les systèmes à base de 525 lignes (NTSC) est de 352×240 et sur les systèmes à 625 lignes (PAL), il est identique à CIF (352×288).

QCIF (Quarter CIF) : Pour avoir un quart de la superficie, comme l'indique le «quart», la hauteur et la largeur de la trame sont divisées par deux.

Les termes également utilisés sont **SQCIF** (Sub Quarter CIF, parfois subQCIF), **4CIF** ($4 \times \text{CIF}$) et **16CIF** ($16 \times \text{CIF}$). Les résolutions de la luminance et la chrominance pour tous ces formats sont résumées dans le tableau ci-dessous.

Format de l'image	Résolution de la luminance (horiz x vert)	Résolution de la chrominance (horiz x vert)
Sub-QCIF	128 × 96	64 × 48
QCIF	176 × 144	88 × 72
CIF	352 × 288	176 × 144
4CIF	704 × 576	352 × 288
16CIF	1408 × 1152	704 × 576

Tableau I.2 Les formats vidéo

L'image en télévision est formée par un certain nombre de pixels, de forme carrée. La définition correspond au nombre de pixels horizontaux et verticaux qui forment l'image :

- ❖ **SD**: la définition standard de 720×576 pixels, c'est celle que proposent les DVD
- ❖ **720p** : un format intermédiaire faisant partie intégrante de la HD (Haute Définition), avec ses 1280×720 pixels. Il reste toujours employé par les prestations de vidéo à la demande (Video on Demand VoD) du fait de ses besoins en bandes passantes moins importants.
- ❖ **Full HD** : avec ses 1920×1080 pixels, c'est l'autre définition de la HD. La vraie, la pleine la totale, celle que propose tout les téléviseurs actuels. C'est également la définition employée par le Blu-ray, successeur du DVD qui remplacé l'autre format HD : le HD-DVD [40].
- ❖ **Quad HD - Ultra HD** : 3840×2160 pixels, souvent faussement appelé 4K parce que celui-ci est un format cinéma de 1920×1080 pixels. Le terme Quad HD est le plus juste puisque cette définition fournit effectivement quatre fois plus de pixels que le Full HD. Mais actuellement il faut employer le terme Ultra HD.
- ❖ **8K** : cette définition offre en tout cas 16 fois plus de points que Full HD, et quatre fois plus que l'Ultra HD. Le 8K, deux fois plus de lignes verticales et horizontales que l'Ultra HD, et donc quatre fois plus de lignes que Full HD. C'est quatre fois plus de pixels que l'Ultra HD, et 16 fois plus que Full HD [40].

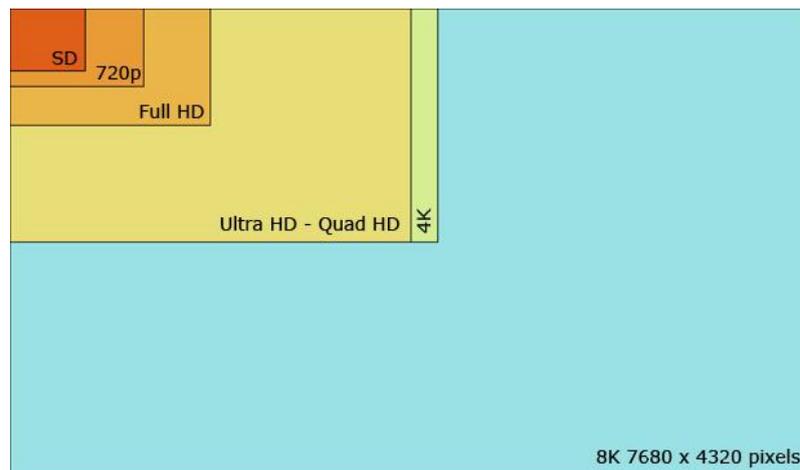


Figure I.9 Représentation proportionnelle de toutes les définitions

I.8 La compression vidéo [31] [34] [37]

Le codage vidéo (codage source) est le processus de compression et de décompression d'un signal vidéo numérique. Le codeur (un compresseur) convertit les données source en une forme comprimée occupant un nombre réduit de bits, avant la transmission ou le stockage, et le décodeur (un décompresseur) convertit la forme comprimée en une représentation des données vidéo originales. Le pair encodeur / décodeur est souvent décrite comme un CODEC (enCOder / DECOder)

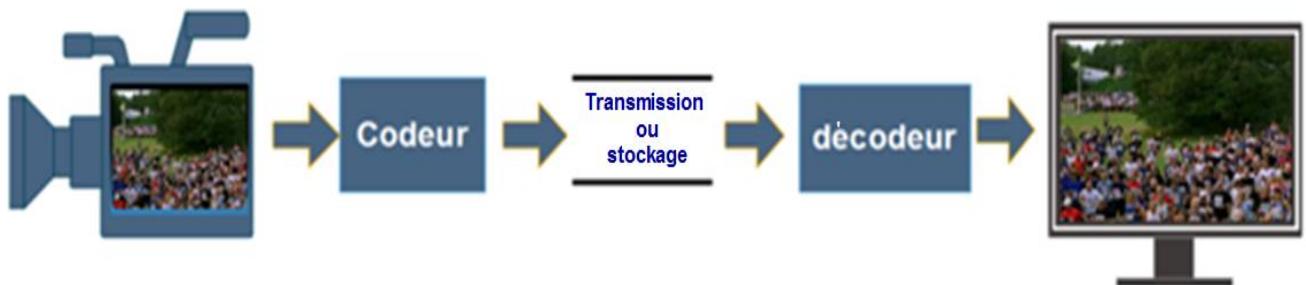
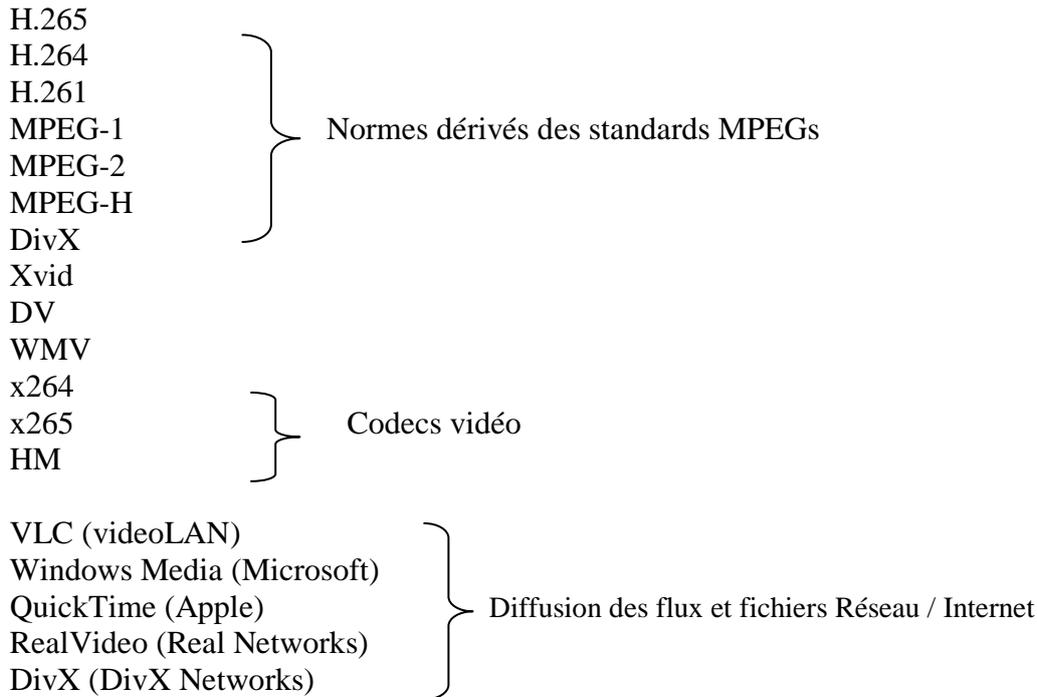


Figure I.10 Encodeur / Décodeur

Standards de compression vidéo : fournis par :

- ISO International Organization for Standardization.
- ITU International Telecommunication Union.
- JVT Joint Video Team (H.264/MPEG-4-AVC en 2003).
- JCT-VC Joint Collaborative Team on Video Coding : pour développement de HEVC codec.

Il existe plusieurs normes et codecs de compression vidéo:



Il est utile de noter que norme et codec sont deux choses différentes. La norme délimite un format, alors que le codec, de son côté, est un équipement ou une application (logiciel) capable de coder et de décoder (compresser et décompresser) un flux de données. Ainsi, il est par exemple inexact de parler de codec H.265, car plusieurs codecs existent, à l'image du HM [41], x265[42], ou libde265 [43]. Il arrive, néanmoins, que certains codecs, en l'absence d'alternative, empruntent directement le nom de la norme correspondante.

Les fichiers vidéo sont volumineux et exigent des débits importants comparant aux fichiers textes.

Soit une vidéo au format :

HD 1080p = 1920 × 1080 à 50 Hz (images/sec)

	Y luminance	Cb chrominance	Cr chrominance
Nbre pixels	1920×1080= 2.1 M	960×540= 0.5 M	960×540= 0.5 M
Nbre bits par pixel	8 bit	8 bit	8 bit
Débit	1.2 Gbps		

Tableau I.3 Exemple de débit mesuré pour une vidéo HD

La compression de données est obtenue en supprimant la redondance, c'est-à-dire des composants qui ne sont pas nécessaires pour une reproduction fidèle des données. De nombreux types de données contiennent une redondance statistique et peuvent être efficacement compressés en utilisant une compression sans perte. Ainsi, les données reconstruites à la sortie du décodeur constituent une copie parfaite des données d'origine.

Malheureusement, la compression sans perte des informations d'image et de vidéo ne donne qu'une compression modérée.

Une compression avec perte est nécessaire pour obtenir une compression plus élevée. Dans un système de compression avec perte, les données décompressées ne sont pas identiques aux données source et des taux de compression beaucoup plus élevés peuvent être obtenus au détriment d'une perte de qualité visuelle. Les systèmes de compression vidéo avec perte sont basés sur le principe de suppression de la redondance subjective, des éléments de l'image ou de la séquence vidéo qui peuvent être supprimés sans affecter de manière significative la perception de la qualité visuelle par le spectateur.

- les redondances spatiales (intra-images) : une seule image contient beaucoup de redondances, a fortiori si elle présente des zones uniformes plus ou moins grandes, c'est à dire où il existe de fortes corrélations entre des pixels voisins, inutile dans ces cas de coder intégralement chacun de ces points.

- les redondances temporelles (inter-images) : dans une séquence vidéo, les images successives sont pratiquement identiques, si ce n'est les points qui sont en mouvement. On ne codera que ces déplacements.

- les redondances subjectives (vision des détails fins) : compresser, c'est pondérer ce qui est plus ou moins visible par l'œil et ne pas coder ce qui n'est pas visible.

- Redondance statistique : Cette redondance s'applique aux codes eux-mêmes. Certains sont plus employés que d'autres, on leur réservera alors les mots les plus courts. Cette opération est dite codage entropique et n'entraîne aucune perte.

Toutes ces techniques peuvent s'ajouter les unes aux autres pour parvenir au résultat attendu, réduire le débit, et conserver autant que ce peut la qualité des images originales. Plus les taux de compression sont importants, plus les dégradations sont perceptibles.

Les signaux vidéo à numériser pour supprimer les différents systèmes de couleurs liés aux codages composites et favoriser l'échange international des programmes, il a été décidé de numériser des données communes à tous les pays. C'est donc logiquement la numérisation des composantes Y, Cr, Cb qui a été retenue selon la convention adoptée [44].

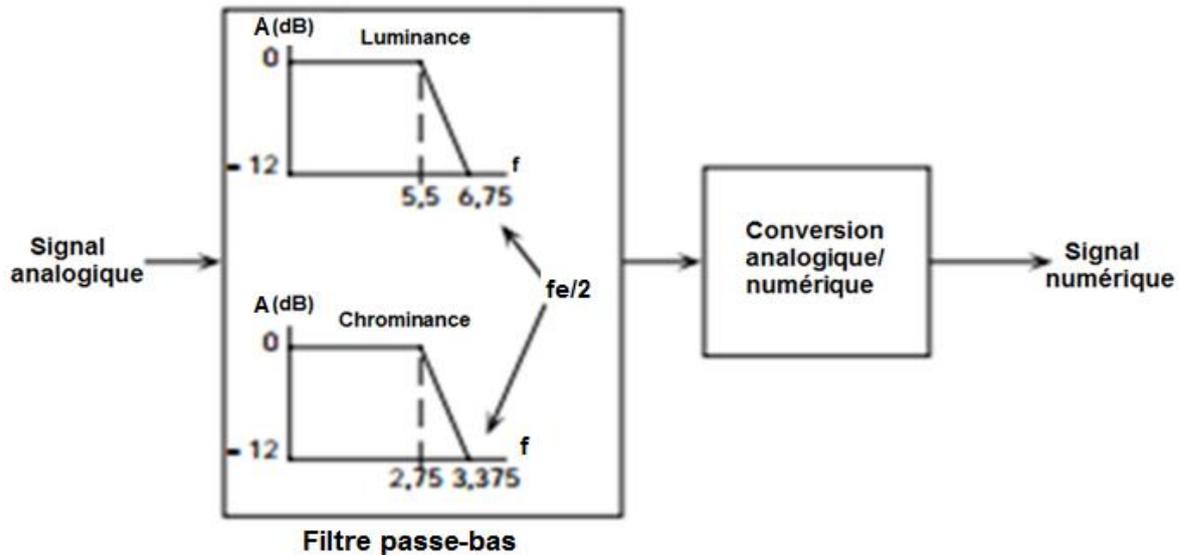


Figure I.11 Numérisation du signal vidéo

• **La luminance** : Les fréquences d'échantillonnage des composantes du signal vidéo ont été choisies communes aux systèmes à 625 et 525 lignes, de manière à s'affranchir définitivement des problèmes d'incompatibilité entre les différentes zones géographiques du globe. L'information de luminance dispose d'une bande passante de l'ordre de 6 MHz. Celle-ci est nominalement plate jusqu'à 5,5 MHz, avec un affaiblissement d'au moins 12 dB à 6,75 MHz. La fréquence d'échantillonnage du signal vidéo doit donc, pour respecter le critère de Shannon-Nyquist, être au moins égale à 12 MHz. Une fréquence d'échantillonnage du signal de luminance égale à : $f_e(Y) = 13,5$ MHz, cette valeur est égale à 864 fois la fréquence ligne des systèmes à 625 lignes et à 858 fois celle des systèmes à 525 lignes. Dans les deux cas, une ligne active numérique contient 720 échantillons de luminance.

• **La chrominance**: Les signaux de différence de couleurs ont une bande passante nominalement plate jusqu'à 2,75 MHz, avec un affaiblissement d'au moins 12 dB à 3,375 MHz. Ils sont échantillonnés à une fréquence deux fois plus faible que le signal de luminance : $f_e(Cr) = f_e(Cb) = 6,75$ MHz. Cette valeur est égale à 432 fois la fréquence ligne des systèmes à 625 lignes et à 429 fois celle des systèmes à 525 lignes.

- 4:2:2 la luminance est échantillonnée à 13,5 MHz, les deux composantes de chrominance à 6,75 MHz.
- 4:1:1 La luminance est échantillonnée à 13,5 MHz, mais les composantes de chrominance sont échantillonnées à 3,375 MHz (soit $\frac{1}{4}$ de la luminance).
- 4:2:0 Luminance échantillonnée à 13,5 MHz, chrominance à 6,75 Mhz mais une ligne sur deux.

- 4:4:4 Toutes les composantes sont échantillonnées à 13,5 MHz. Ce type de codage est employé par les stations graphiques.

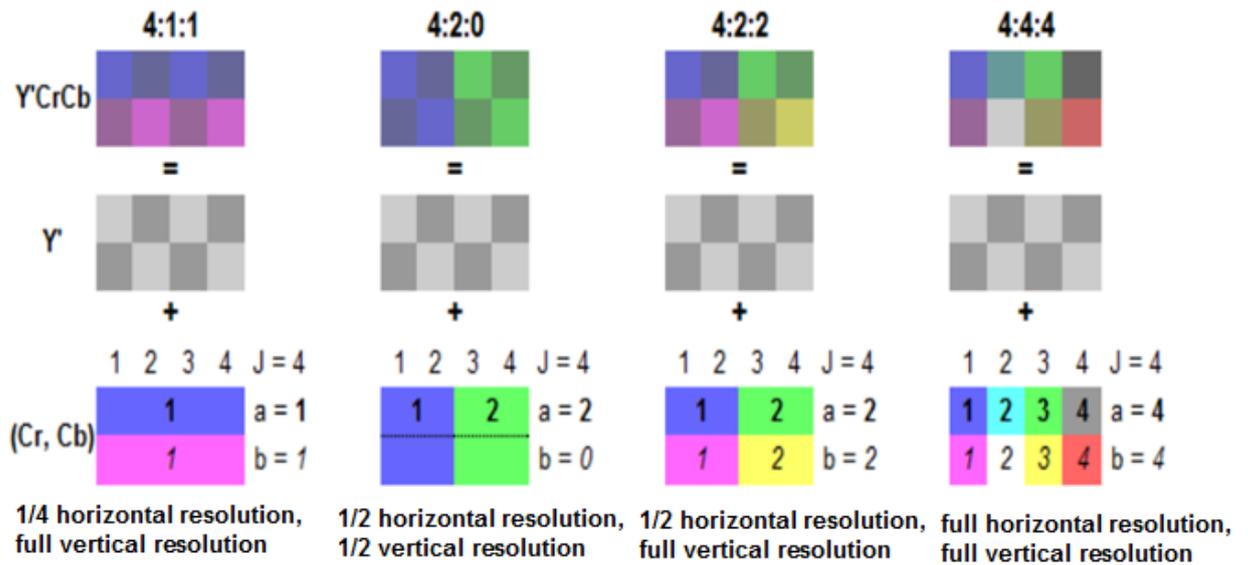


Figure I.12 Echantillonnage couleur.

Ces nombres représentent le rapport entre la fréquence d'échantillonnage de la luminance et des composantes de chrominance. A l'origine, les normes 4 : * : * spécifiait une quantification sur 8 bits des composantes du signal vidéo, ce qui permet de disposer de 256 niveaux possibles. Par la suite, pour les productions haut de gamme, la norme a évolué vers le 10 bits, ce qui permet 1024 niveaux.

Soit une vidéo au format 720 × 576 à 25 Hz (images/sec)

- En 4 : 2 : 0, en 8 bits on a un taux de 3 à 10 Mbits/s (diffusion)
- En 4 : 2 : 0 ou 4 : 1 : 1 en 8 bits un taux de 25 Mbits/s (DV)
- En 4 : 2 : 2, en 8 ou 10 bits un taux de 50 Mbits/s (postproduction)

I.9 La Quantification

La quantification est le deuxième procédé du système de numérisation. Elle a pour objectif de faire correspondre chaque amplitude échantillonnée à un nombre entier indiqué en base 2, n chiffres permettent de coder $N = 2^n$, dans cette état, comme n égale à 3 bits, le nombre de niveaux possibles est de 7 (8 en comptabilisant le 0) [31].

Dans cette étape le signal analogique qui peut prendre une infinité des valeurs, est converti en un signal constitué d'un nombre fini de valeurs numériques. Des erreurs d'arrondis par défaut ou par débordement sont alors inévitables parce que plusieurs valeurs proches mais différentes, correspondra une seule valeur binaire. La précision du signal converti sera donc directement liée au nombre des valeurs disponibles.

I.10 Les codecs vidéo [30] [31] [33][34]

Le numérique s'impose dans tous les domaines, qu'il s'agisse de l'image, du son ou de la transmission de données, et la vidéo ne peut qu'emprunter le même chemin. Pour passer de la vidéo analogique traditionnelle à la vidéo numérique, il est indispensable de créer des formats spécifiques, entièrement compatibles avec le monde informatique.

Un codec vidéo permet de coder une image ou une séquence vidéo source sous une forme compressée et la décode pour achever une copie ou une approximation de la séquence source ; Si la séquence vidéo décodée est identique à l'original, le processus de codage est sans perte; Si la séquence décodée diffère de l'original, le processus est à perte.

Le CODEC représente la séquence vidéo originale par un modèle, le modèle devrait représenter la séquence en utilisant le moins de bits possible et avec la plus grande fidélité possible. Ces deux objectifs, l'efficacité de la compression et la haute qualité, sont généralement contradictoires, c'est-à-dire qu'un débit binaire compressé inférieur produit typiquement une qualité d'image réduite au niveau du décodeur. Un encodeur vidéo est constitué de trois unités fonctionnelles principales: un modèle de prédiction, un modèle spatial et un codeur entropique (figure 1.13) [31].

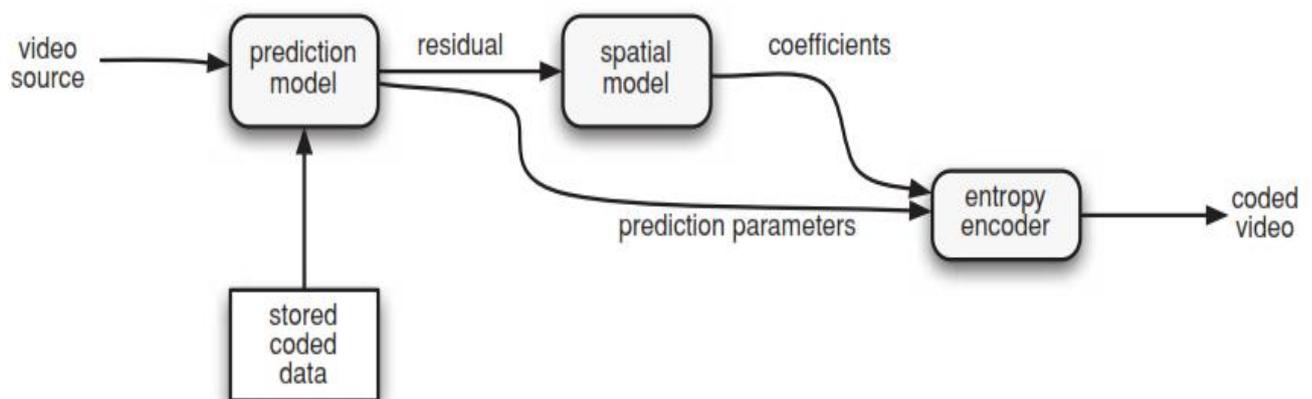


Figure I.13 Bloc diagramme d'un codeur vidéo [37]

L'entrée du modèle de prédiction est une séquence vidéo «brute» non compressée. Le modèle de prédiction tente de réduire la redondance en exploitant les similitudes entre les trames vidéo voisines et / ou les échantillons d'image voisins, typiquement en construisant une prédiction de la trame vidéo courante ou du bloc de données vidéo. En H.264 / AVC, la prédiction est formée à partir de données dans la trame actuelle ou dans une ou plusieurs trames précédentes et / ou futures. Il est créé par extrapolation spatiale à partir d'échantillons d'image voisins, intra-prédiction, ou en compensant les différences entre les trames, inter ou

prédiction compensée par mouvement. La sortie du modèle de prédiction est une trame résiduelle, créée en soustrayant la prédiction de la trame actuelle réelle, et un ensemble de paramètres de modèle indiquant le type d'intra-prédiction ou décrivant comment le mouvement a été compensé. La trame résiduelle constitue l'entrée du modèle spatial qui utilise des similitudes entre les échantillons locaux dans la trame résiduelle pour réduire la redondance spatiale. En H.264 / AVC, ceci est effectué en appliquant une transformation aux échantillons résiduels et en quantifiant les résultats. La transformée convertit les échantillons en un autre domaine dans lequel ils sont représentés par des coefficients de transformation. Les coefficients sont quantifiés pour éliminer les valeurs insignifiantes, laissant un petit nombre de coefficients significatifs qui fournissent une représentation plus compacte de la trame résiduelle. La sortie du modèle spatial est un ensemble de coefficients de transformation quantifiés.

Les paramètres du modèle de prédiction, c'est-à-dire le mode d'intra-prédiction et les vecteurs de mouvement et le modèle spatial, les coefficients sont compressés par le codeur entropique. Ceci supprime la redondance statistique dans les données, par exemple en représentant des vecteurs et des coefficients courants par de courts codes binaires. Le codeur entropique produit un train de bits compressé ou un fichier qui peut être transmis et / ou stocké. Une séquence compressée est constituée de paramètres de prédiction codés, de coefficients résiduels codés et d'informations d'en-tête [37].

Le décodeur vidéo reconstruit une trame vidéo à partir du train de bits compressé. Les coefficients et les paramètres de prédiction sont décodés par un décodeur entropique après quoi le modèle spatial est décodé pour reconstruire une version de la trame résiduelle. Le décodeur utilise les paramètres de prédiction, conjointement avec des pixels d'image précédemment décodés, pour créer une prédiction de la trame courant et la trame elle-même est reconstruite en ajoutant la trame résiduelle à cette prédiction.

I.11 Groupe d'images (GOP)

Dans toutes les normes de codage vidéo, les GOP sont utilisés pour définir des relations de codage dans une séquence de trames vidéo. De manière générale, tous les modes GOP [31], [33] commencent par un I / Key Intra-image qui est codé de manière indépendante sans faire référence à d'autres images. L'accès aléatoire dépend de l'utilisation de la première image clé dans le GOP. En général, nous prévoyons que l'utilisation de plus grandes tailles de GOP conduira à plus d'efficacité de compression [36].

Un GOP peut contenir les types d'images suivants :

- Image I ou I-frame (image à codage interne): Image de référence. C'est une image fixe et indépendante des autres types d'image. Chaque GOP commence avec une image de ce type.
- Image P ou P-frame (image à codage prédictif): Cette image contient des informations de différence (par prédiction compensée de mouvement) avec l'image I (ou image P) passée. C'est également une image de référence.
- Image B (image à codage prédictif bidirectionnel): Cette image comprend des informations de différence avec les images I (ou image P) passées et futures à l'intérieur d'un GOP. Afin d'éviter une trop grande propagation d'erreur de prédiction, les images B ne sont généralement pas exploitées en tant qu'image de référence.

Un GOP comprend une seule image I qu'est en conformité à la première image à encoder. Elle est suivie d'une succession d'images P et B dont le motif se répète jusqu'à la fin du GOP. Les images P se situent à intervalle régulier et les images B complètent ces intervalles. On définit ce motif de sous-GOP. Quelques codecs vidéo permettent plus d'une image I dans un GOP [8]. Le schéma ci-dessous décrit la structure d'un GOP dont le sous-GOP est composé d'une image P et de deux images B:

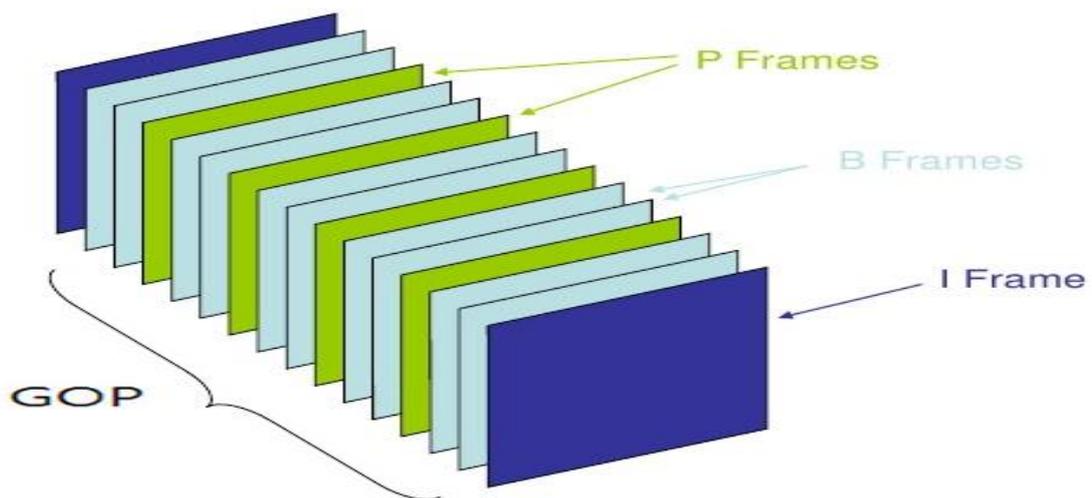


Figure I.14 Structure d'un GOP dans un codec vidéo

I.12 Les normes de compression vidéo

I.12.1 JPEG [29]

Apparu en 1989, JPEG est une norme de compression des images fixes. Cette norme a été développée à la base en ce qui concerne l'informatique. Le JPEG exploite les redondances spatiales. La compression se fait avec perte d'information

Dans ce standard une séquence vidéo est une succession d'images fixes. Mais chaque fabricant a développé son propre format dérivant du JPEG. Le principe adopté pour la compression d'images fixes en JPEG qui est appliqué aux séquences animées. Il a donné naissance au Motion JPEG, appelé aussi MJPEG. Il est employé par la plupart des cartes d'acquisition vidéo et dans la technologie DV (Digital Vidéo). La compression temporelle utilise une toute autre méthode, en constatant qu'il y a peu de différences entre des images consécutives. C'est le principe utilisé par le MPEG qui s'avère environ 4 fois moins gourmand en espace disque que le MJPEG, à qualité visuelle égale. Le codage du JPEG et du M-JPEG est intra-image.

I.12.2 MPEG-1

MPEG-1 est une norme pour la compression avec perte de la vidéo et de l'audio destinée aux applications multimédia. Il est conçu pour compresser des images animées et des données audio associées pour les supports de stockage numérique jusqu'à environ 1,5 Mbit/s[33], rendant les CD vidéo, télévision numérique par câble et satellite et diffusion audio numérique (DAB) possible[15] [33]. Le MPEG-1 reprend l'élimination des redondances spatiales (comme le JPEG, le codage est dit intra-image, et exploite aussi les redondances temporelles entre les images (le codage est dit inter-image)

Aujourd'hui, MPEG-1 est utilisé dans un grand nombre de produits et de technologies. La partie la plus connue de la norme MPEG-1 est peut-être le format audio MP3. La norme MPEG-1 est publiée en tant que norme ISO / CEI 11172 [45]

La norme se compose des cinq parties suivantes: [33]

- Systèmes (stockage et synchronisation de la vidéo, de l'audio et d'autres données ensemble)
- Vidéo (contenu vidéo compressé)
- Audio (contenu audio compressé)
- Test de conformité (test de l'exactitude des implémentations de la norme)
- Logiciel de référence (exemple de logiciel montrant comment encoder et décoder selon la norme).

I.12.3 MPEG-2 [5] [33] [45]

Créé en 1994, il reprend les techniques de bases du MPEG-1. D'un autre côté, la qualité d'image est nettement supérieure puisqu'elle s'étend de la vidéo standard (3 à 10 Mbps à la haute définition 300 Mbps). La norme MPEG-2 est utilisée par tous les opérateurs de télévision numérique. Comme la norme MPEG-1, MPEG-2 a un codage inter-image, les systèmes MPEG (1 et 2) effectuent leur traitement sur des GOP.

La norme MPEG-2 est conçue pour recouvrir toutes les applications de diffusion d'images et de sons par satellite, par câble et par voie terrestre, ainsi que les supports tels que les DVD. Cette norme a ensuite été mise au point pour des applications de diffusion en production et post-production, édition, ... etc.

Par rapport à la norme MPEG-1, MPEG-2 a pour différences :

- Support de formats d'image de qualité supérieur (jusqu'à la HD)
- Traitement de balayage entrelacé (50 trames/s) et du balayage progressif (50 images complètes/s)
- Compatibilité avec le MPEG-1 c'est à dire un décodeur MPEG-2 peut décoder un signal MPEG-1.

Cependant la norme MPEG-2 se décrit seulement les outils de compression des données ainsi que la syntaxe du signal. La technique employée dans le codeur est créée par le constructeur et peut constamment évoluer. La qualité de l'image dépend donc aussi de la qualité du codeur et du décodeur.

I.12.4 H.261

La norme H.261 est une ITU-T norme de compression vidéo réalisée en 1990, d'abord il s'agit du premier membre de la famille de normes de codage vidéo H.26x dans le domaine du VCEG (Video Coding Experts Group) de l'ITU-T. Il s'agissait de la première norme de codage vidéo utile en pratique.

Le modèle H.261 a été conçu à l'origine pour la transmission sur des lignes RNIS (Réseau Numérique à Intégration de Services) sur lesquelles les débits de données sont des multiples de 64 kbit / s. L'algorithme de codage a été conçu pour pouvoir fonctionner à des débits vidéo compris entre 40 kbit / s et 2 Mbit / s. La norme prend en charge deux tailles d'images vidéo: CIF et QCIF (en utilisant un échantillonnage 4: 2: 0. Il a également une astuce rétro compatible pour l'envoi d'images fixes avec une résolution de 704×576 luminance et une résolution de 352×288 chrominance (qui a été ajoutée dans une révision ultérieure en 1993). Les images peuvent être codées I ou P mais pas B [31].

I.12.5 H.263

H.263 est une norme de compression vidéo conçue à l'origine comme un format compressé à faible débit binaire pour la vidéoconférence. Il a été développé par VCEG dans le cadre d'un projet achevé en 1995/1996 en tant que membre de la famille de normes de codage vidéo H.26x dans le domaine de l'ITU-T. L'ajout de diverses fonctions améliorées

supplémentaires en 1998 et 2000. Des ajouts plus modestes ont également été faits en 1997 et en 2001, et une édition unifiée a été produite en 2005.

Les formats supportés par cette norme sont le Sub-QCIF et le QCIF et optionnellement le CIF le 4CIF et le 16CIF. L'utilisation d'image B est possible. En plus de cette configuration de base, six options de codage ont été ajoutées afin d'augmenter les performances du codage et ses fonctionnalités. La version 2 de la recommandation H.263 (1998), souvent appelée H.263+ [46] met en œuvre douze options supplémentaires et permet désormais de définir des formats et des fréquences de la vidéo personnalisées. Les options ajoutées améliorent la qualité et la résistance aux erreurs.

La dernière version de H.263 (2000) appelé H.263++ [36] ajoute trois options. Outre amélioration en termes de qualité et de taux de compression, elle prend mieux en compte la transmission vidéo en temps réel sur un réseau à qualité de service non garantie (perte de paquets...).

I.12.6 H.264 : MPEG-4 AVC

H.264 est un format ou norme de codage vidéo dont une première version. La première version a été approuvée en mai 2003 et la plus récente date d'avril 2012 et on retrouve cette norme très simplement autour de nous : les chaînes de TNT qui diffusent en 1080i, par exemple, font appel à cette forme de MPEG-4[10] [31].

MPEG-4 est une méthode de la compression des données numériques audio et vidéo. Il a été introduit à la fin de 1998 et a désigné une norme pour un groupe de formats de codage audio et vidéo et de technologie connexe convenus par le Groupe d'experts des images animées (MPEG) ISO / CEI (ISO / IEC JTC1 / SC29 / WG11). ISO / CEI 14496 [31]. Les utilisations du format MPEG-4 comprennent la compression de données audiovisuelles pour la diffusion sur le Web (diffusion en continu) et la distribution de CD, la voix (téléphone, vidéophone) et la télédiffusion. La norme permet le codage d'une grande variété de format vidéo (taille, résolution, fréquence vidéo) mais aussi le codage d'objets vidéo de forme arbitraire, d'images fixes ainsi que d'objets synthétiques 3D. les usages de MPEG-4 englobent toutes les nouvelles applications multimédias comme le téléchargement et le streaming sur Internet, le multimédia sur téléphone mobile, la radio numérique, les jeux vidéo, la télévision HD .

MPEG-4 AVC Part 10 a été développé en collaboration avec ITU-T en tant que recommandations H.264. Le codec H.264/AVC est donc adapté à une très grande variété de réseaux et de systèmes (par exemple, pour la diffusion de la télévision, le stockage HD DVD et Blu-ray, le streaming RTP/IP, et des systèmes de téléphonie propre à l'UIT-T). La norme UIT-T H.264 et la norme ISO/CEI MPEG-4 Part 10 (ISO/CEI 14496-10) sont

techniquement identiques, et la technologie employée est aussi connue sous le nom AVC, pour Advanced Video Coding.

Le JVT a ensuite travaillé sur le concept d'extensibilité en élaborant une extension à la norme H.264 (annexe G) : les spécifications Scalable Video Coding (SVC), puis sur la norme HEVC.

I.13 Conclusion

Ce chapitre examine la structure et les caractéristiques des images numériques et des signaux vidéo et introduit des concepts fondamentaux tels que les formats d'échantillonnage et les méthodes de compression vidéo ainsi que les normes et les codecs vidéo.

Nous bénéficions aujourd'hui de recherche et développement liés à l'optimisation de standards de compression pour la vidéo numérique. Ces standards peuvent dorénavant se déployer sur différentes plateformes technologiques universelles et interopérables, répondant ainsi à une multitude d'applications et de besoins en termes de production et de diffusion de contenus vidéo numérique.

Chapitre II

High Efficiency Video Coding

II.1 Introduction

La demande croissante de la vidéo Haute Définition et Ultra Haute Définition, associée à un désir croissant de vidéo streaming, a entraîné une augmentation exponentielle de la demande en bande passante et en stockage. Ces défis peuvent être relevés par la nouvelle norme HEVC, également appelée H.265. Le projet d'encodeur HEVC a été lancé par MulticoreWare [42] en 2013, dans le but de fournir l'encodeur vidéo HEVC le plus performant.

Fort apprécié des diffuseurs de tous types, le H.264 s'est imposé en l'espace de 10 ans comme le standard le plus largement accepté pour la compression de contenus vidéo de définition Full HD et inférieure. Il se heurte, néanmoins, à une limite : alors qu'on nous parle aujourd'hui d'Ultra HD et même de 8K ainsi que de fréquences de 100, 200 ou même 600 Hz, sa compatibilité se borne à une image de 4 096 x 2 304 px pour moins de 60 image/seconde. Un format "maximum" qui réclamerait, qui plus est, un débit binaire de l'ordre de 50 Mb/s.

La couche de codage vidéo de HEVC [2], [47] utilise la même approche "hybride" (prédictions inter-image et intra-image) et codage de transformation (2D) utilisée dans toutes les normes de compression vidéo depuis la naissance de MPEG qui peut créer un train de bits (bitstream) conforme à la norme HEVC [48].

II.2 HEVC / H.265 (High Efficiency Video Coding)

HEVC est une nouvelle norme de codage vidéo qui a été développée pour améliorer l'efficacité du codage avec une meilleure transmission réseau. Ce codec permet une réduction de 50% du stockage par rapport à son prédécesseur, le H264 / AVC, tout en conservant le même niveau de qualité d'image élevé [47] [48] [49]. Néanmoins, il entraîne une surcharge de calcul supplémentaire de 2 à 10 fois plus importante. La norme HEVC/ H.265 a été développée par le Groupe (MPEG) ISO / IEC et le groupe VCEG de l'ITU-T. HEVC est également connu comme ISO / IEC 23008-2 MPEG-H Partie 2 et ITU-T H.265. HEVC peut prendre en charge la vidéo 8K UHD, avec une taille d'image allant jusqu'à 8192x4320 pixels.

Le nouveau standard promet une économie de bande passante (débit binaire) de l'ordre de 20 à 60 % par rapport aux flux confiés aux soins de son prédécesseur. Cette différence d'efficacité varie certes en fonction du contenu (une image fixe en 1280×720 px ne permet pas les mêmes gains qu'un film en 1080p), mais se montre dans tous les cas sensible. Bien que HEVC/H.265 permet d'envisager la distribution de contenus UHD/4K/8K avec bien moins de contraintes. En effet, le HEVC/ H.265 est prévu pour accompagner nos contenus jusqu'à une "véritable" définition 8K de 8192×4320 px, et supporte techniquement jusqu'à 300 images par seconde.

Ce codec est approprié pour les images HD avec des tailles de blocs pouvant atteindre 64×64 et des fréquences d'images élevées avec une prédiction de vecteur de mouvement avancée.

Figure II.1 montre une comparaison de la structure de codage entre H.264 / MPEG-4 AVC et HEVC [49].

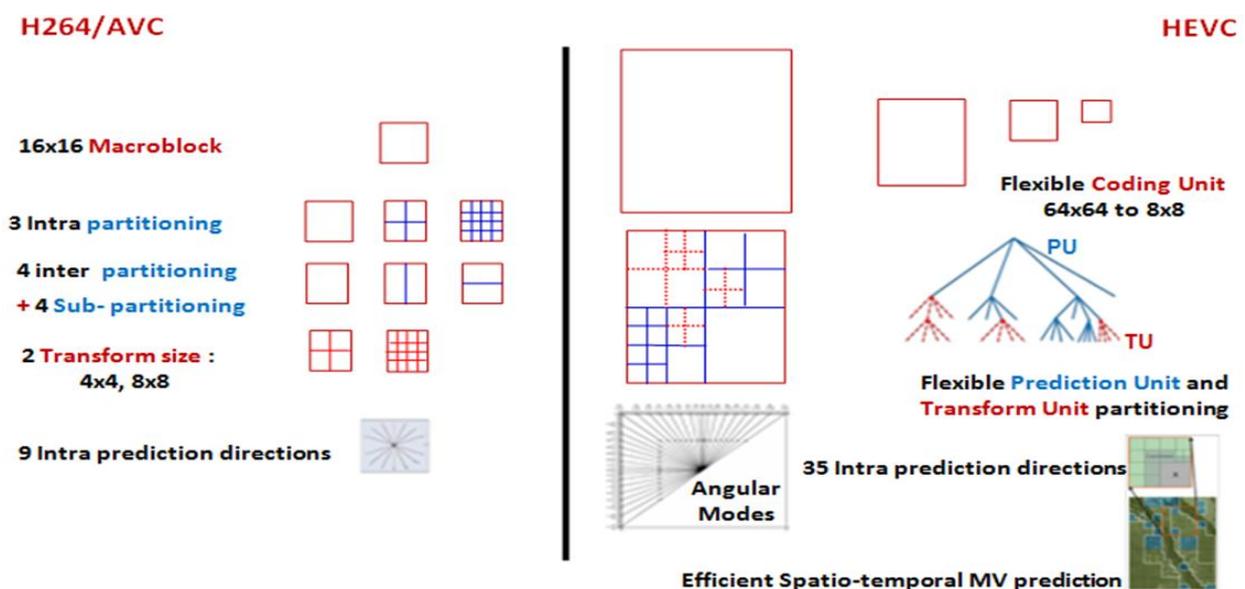


Figure II.1 Comparaison de la structure de codage entre H.264 / MPEG-4 AVC et HEVC/H.265

Dans le codeur HEVC, chaque trame d'une séquence vidéo est fractionnée en blocs de tailles différentes (16, 32 ou 64). Dans un GOP, l'intra prédiction est utilisée pour la première trame qui exploite la redondance spatiale de l'image, puis le reste des trames qui exploitent la redondance temporelle sont traitées en utilisant l'inter-prédiction [49], [50].

II.3 Les Versions de la norme HEVC / H.265

Les versions de la norme HEVC / H.265 selon l'UIT-T: [51]

Version 1: (13 avril 2013) Première version approuvée de la norme HEVC / H.265 contenant les profils Main, Main 10 et Main Still Picture. [51] [52]

- le profil principal « Main profile » : Le profil principal permet une profondeur de couleur de 8 bits par échantillon avec sous-échantillonnage de la chrominance 4:2:0, qui est le type le plus commun de la vidéo avec des appareils grand public.
- le profil « Main 10 » : Le Main 10 permet une profondeur de couleur de 8 bits à 10 bits par échantillon avec sous-échantillonnage chroma 4:2:0, garantissant une profondeur de couleur de 10 bits.
- le profil « Main Still Picture »: Le profil Main Still Picture permet le codage d'une seule image fixe en utilisant les mêmes contraintes que le profil Main profile. En tant que sous-ensemble du profil Main profile, le profil Main Still Picture offre une profondeur de couleur sur 8 bits avec une colorimétrie échantillonnée en 4:2:0.

Version 2: (29 octobre 2014) Deuxième version approuvée de la norme HEVC/H.265 qui ajoute 21 profils d'extensions de plage, deux profils d'extensions évolutives (SVC) et un profil d'extensions multi-vues. [52]

Monochrome, Monochrome 12, Monochrome 16, Main 12, Main 4:2:2 10, Main 4:2:2 12, Main 4:4:4, Main 4:4:4 10, Main 4:4:4 12, Monochrome 12 Intra, Monochrome 16 Intra, Main 12 Intra, Main 4:2:2 10 Intra, Main 4:2:2 12 Intra, Main 4:4:4 Intra, Main 4:4:4 10 Intra, Main 4:4:4 12 Intra, Main 4:4:4 16 Intra, Main 4:4:4 Still Picture, Main 4:4:4 16 Still Picture, High Throughput 4:4:4 16 Intra, Scalable Main, Scalable Main 10, and Multiview Main [50]. Tous les profils d'extensions inter-frames ont un profil Intra [52].

Version 3: (29 avril 2015) Troisième version approuvée de la norme HEVC / H.265 qui ajoute le profil 3D Main. [41] [42] [52]

Version 4: (22 décembre 2016) Quatrième version approuvée de la norme HEVC / H.265 qui ajoute sept profils d'extensions de codage de contenu d'écran, trois profils d'extension à haut débit et quatre profils d'extensions évolutives. [52]

II.4 Fonctionnement de l'encodeur HEVC (H.265)

Comme le codage H.264 et MPEG-2, le codage HEVC utilise trois types d'images I-B et P dans un ensemble d'images, combinant des éléments de compression entre une trame et une autre.

La couche de codage vidéo HEVC comporte un certain nombre de blocs de traitement, comme décrit dans la figure II.2 ci-dessous. La trame vidéo brute est divisée en blocs et chaque région de bloc est passée à travers divers éléments du codeur. La première image est codée en utilisant la prédiction intra-image, et est appelée I-Frame. Les images restantes dans le même groupe peuvent être codées en utilisant la prédiction inter-image, où la prédiction peut être faite (P-Frames) ou bidirectionnelle (B-Frames). La prédiction inter-image fait appel à des images de référence et des vecteurs de mouvement (MV) sélectionnés qui indiquent le mouvement des divers blocs [8], [41].

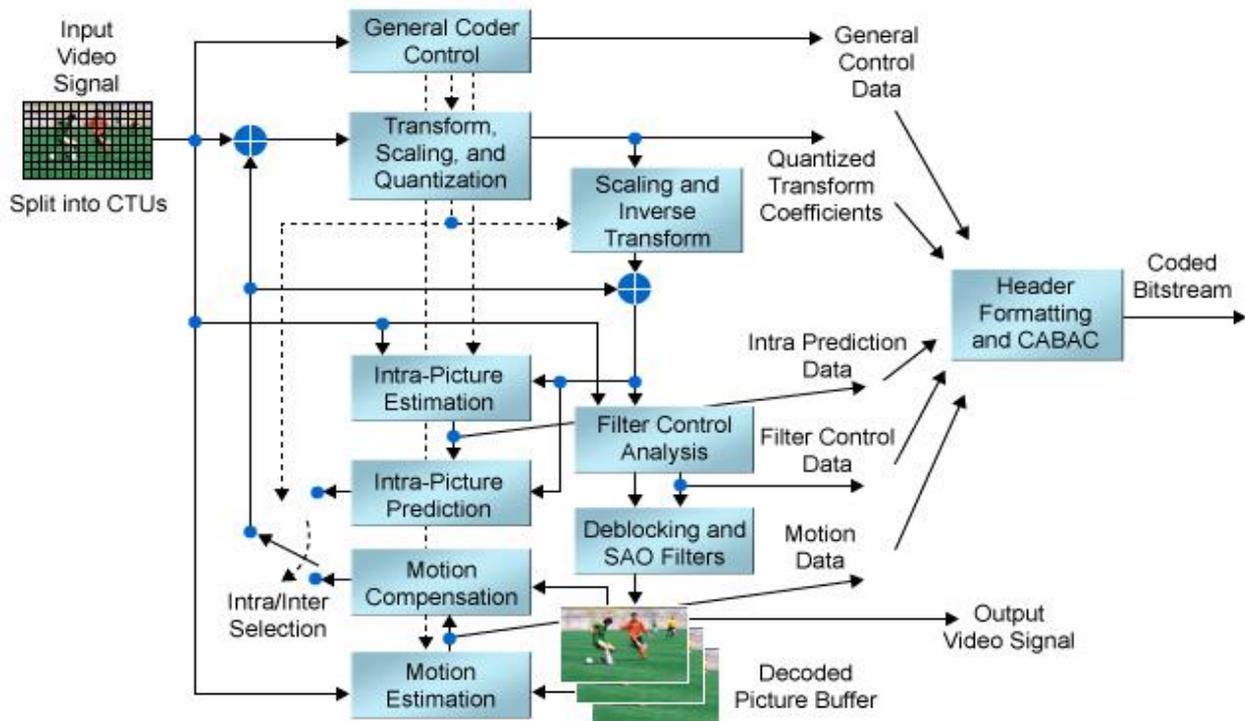


Figure II.2 Schéma de principe du codeur HEVC (avec des éléments de modélisation du décodeur ombrés en gris clair)

II.5 Partition des images (Picture partitioning) [47][49][53][54]

L'une des plus hautes performances du HEVC est l'introduction des structures des blocs avec des mécanismes de sous-partitions flexibles. Le bloc de base dans HEVC est connu comme la plus grande unité de codage (CTU) et peut être divisé récursivement en plus petites

unités de codage (CUs), qui à leurs tour peuvent être divisées en petites unités de prédiction (PUs) et unités de transformation (TUs).

Ce qui suit est une terminologie importante de HEVC.

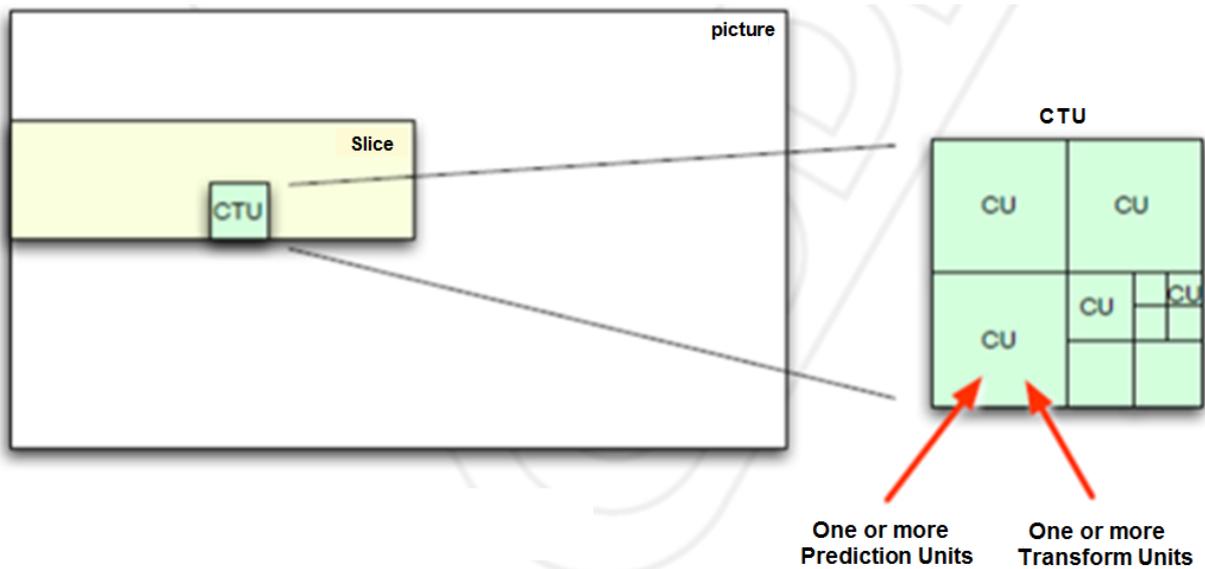
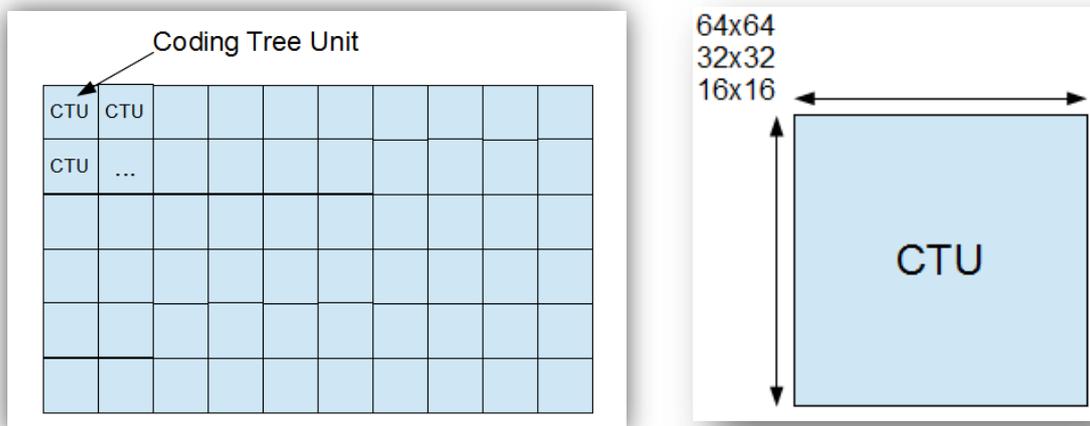


Figure II.3: Picture, Slice, Unité d'arbre de codage (CTU), Unité de codage (CU)

- **Coding Tree Unit (CTU) et Coding Tree Block (CTB):**

La CTU est le bloc de pixels de base, similaire au macro-bloc dans H264. L'image (frame) est divisé en CTU de taille 64×64 , 32×32 ou 16×16 .



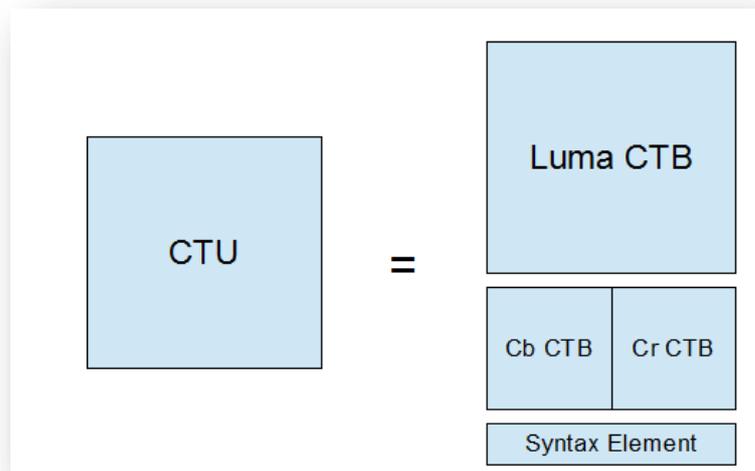


Figure II.4 Unité d'arbre de codage (CTU) et Bloc d'arbre de codage (CTB)

La CTU est une unité logique. Elle se compose généralement de trois blocs, luma (Y) et deux échantillons de chrominance (Cb et Cr), et des éléments de syntaxe associés. Chaque bloc est appelé CTB (Coding Tree Block).

- **Coding Unit (CU) et Coding Block (CB):** CTU est divisé en un bloc de CB de luminance et deux CB de chrominance, chacun formé comme un arbre quadruple. Ensemble, ces CBs appelés unité de codage (CU).

CB est assez bon pour la décision de type prédiction, mais il pourrait être encore trop grand pour stocker des vecteurs de mouvement en mode inter prédiction ou en mode intra prédiction. . Ainsi, Prediction Block (PB) a été introduit. Chaque CB peut être divisé en PB différemment en fonction de la prévisibilité temporelle et / ou spatiale.

Chaque CTB a toujours la même taille que le CTU 64×64 , 32×32 ou 16×16 . Selon une partie de la trame vidéo, CTB peut être trop grand pour décider si nous devons effectuer une prédiction inter-image ou une prédiction intra-image. Ainsi, chaque CTB peut être divisé différemment en plusieurs CB et chaque CB devient le point de décision du type de prédiction. Par exemple, certains CTB sont divisés en 16×16 CB tandis que d'autres sont divisés en 8×8 CB. HEVC prend en charge la même taille de CB que celle du CTB (aussi petit que 8×8).

La figure suivante montre comment 64×64 CTB peuvent être divisés en CB.

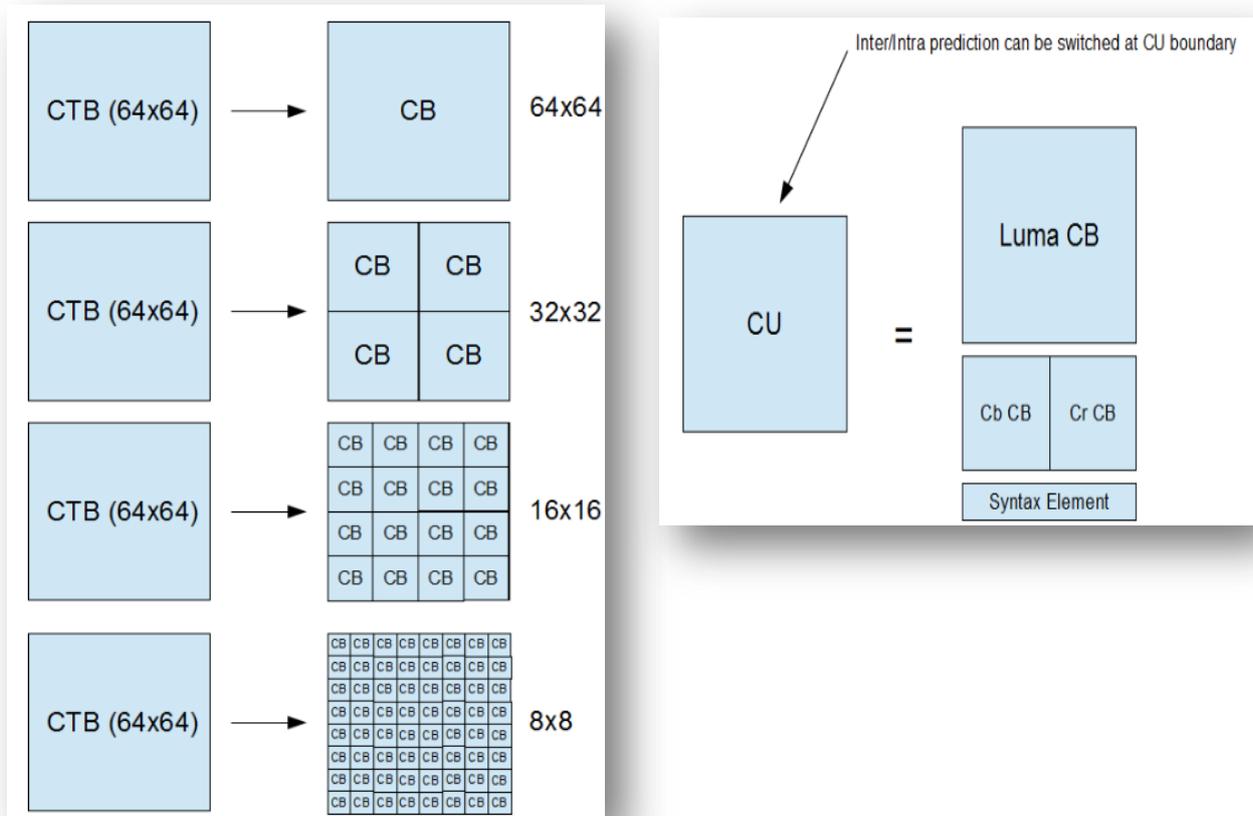


Figure II.5 Structure de CTB et CU

Une fois la prédiction faite, nous devons coder le résidu (différence entre l'image prédite et l'image réelle) avec une transformation de type DCT (Discrete Cosine Transform). Encore une fois, CB pourrait être trop grand; pour cela un CB peut contenir à la fois une partie détaillée (haute fréquence) et une partie plate (basse fréquence). Par conséquent, chaque CB peut être divisé différemment en Transform Block (TB), TB n'a pas besoin d'être aligné avec PB. Il est possible et souvent logique d'effectuer une transformation unique à travers les résidus de plusieurs PB, et inversement. Pour la transformée 4×4 des résidus de prédiction intra-image luma, une transformée entière dérivée d'une forme de transformée sinus discrète (DST) est spécifiée de manière alternative.

- **Prediction Unit (PU) et Prediction Block (PB):** la décision d'une prédiction inter-image ou une prédiction intra-image sera utilisée et décidée au niveau du CU. Le CU est partitionné en unités de prédiction (PUs). Les CBs sont des blocs de prédiction divisés en (PBs) pour les composantes de luminance et de chrominance. La taille de PB varie de 64×64 à 4×4 .

CB est le point de décision que ce soit pour effectuer une prédiction inter-image ou intra-image. Plus précisément, le type de prédiction est codé en CU.

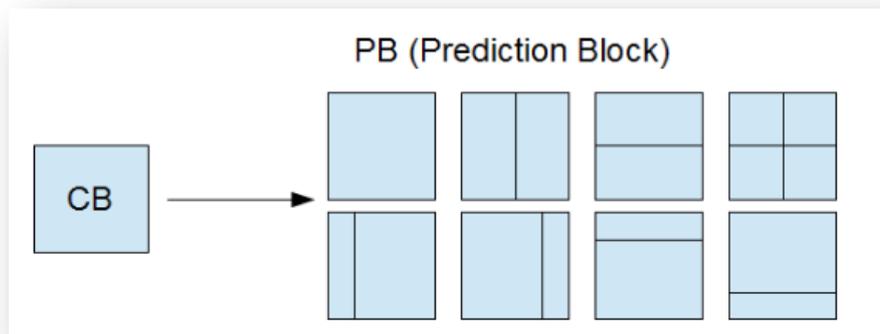


Figure II.6 Unité de prédiction (PU) et Bloc de prédiction (PB)

- **Transform Unit** (Unité de transformation) (TU) et **Transform Block** (TB) (bloc de transformation): Une arborescence de TU a une racine au niveau de l'CU et effectue une transformation spatiale. La partie de la luminance et la chrominance de l'TU es appelée TB. La taille de TB peut être 4×4 , 8×8 , 16×16 ou 32×32 .

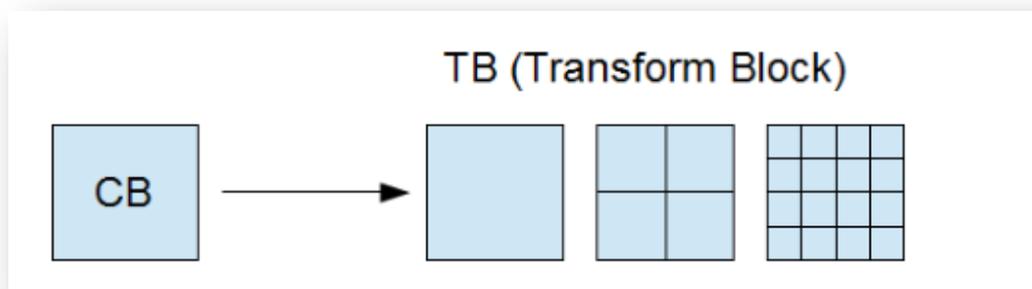


Figure II.7 Bloc de transformation (TB)

- **Motion vector signaling:** Un nouvel algorithme appelé Advanced Motion Vector Prediction (AMVP) est utilisé. De plus, par rapport à l' H.264 / MPEG-4 AVC, l'inférence par sauts et à mouvement direct améliorée est également spécifiée.
- **Compensation de mouvement (Motion compensation):** Précision de quart d'échantillon pour les MV et des filtres à 7 ou 8 dérivations sont utilisés pour l'interpolation des positions de l'échantillon fractionnaire. Similaire au H.264 / MPEG-4 AVC, plusieurs images de référence sont utilisées. Pour chaque PB, un ou deux vecteurs de mouvement peuvent être transmis, résultant soit en un codage unipredictive ou un codage bipredictive respectivement. Comme dans H.264 / MPEG-4 AVC, une opération de mise à l'échelle et de décalage peut être appliquée au signal (s) de prédiction d'une manière connue sous le nom de prédiction pondérée.

- **Intrapicture Prediction:** Les échantillons décodés des limites des blocs adjacents sont utilisés comme données de référence pour la prédiction spatiale dans les régions où la prédiction d'image n'est pas effectuée. La prédiction intra image prend en charge 33 modes directionnels (comparés à huit de ces modes dans H.264 / MPEG-4 AVC). En plus des modes de prédiction planaire et DC (flat), les modes de prédiction intra images sélectionnés sont codés en dérivant les modes les plus probables (par exemple, les directions de prédiction) sur la base de ceux des PBs voisins précédemment décodés.
- **Contrôle de quantification:** Comme dans H.264 / MPEG-4 AVC, la quantification de reconstruction uniforme (URQ) est utilisée dans HEVC, avec des matrices de mise à l'échelle de quantification supportées pour les différentes tailles des blocs de transformation.
- **Codage entropique (Entropy coding):** l'algorithme de codage utilisé dans HEVC est le CABAC (Context Adaptive Binary Arithmetic Coding), similaire au H.264. mais a subi plusieurs améliorations pour améliorer sa vitesse de traitement (en particulier pour les architectures de traitement en parallèle), ses performances de compression, et pour réduire ses besoins en mémoire.
- **Filtre de déblocage en boucle (In-loop deblocking filter):** le filtre de déblocage en boucle similaire à celui utilisé dans H.264 / MPEG-4 AVC est utilisé dans la boucle de prédiction inter-image. Cependant, la conception est simplifiée en ce qui concerne ses processus de prise de décision et de filtrage, et est rendue plus conviviale au traitement parallèle.

II.6 Structure des SLICES et des TILES

Une slice (tranche) est une structure de données qui peut être décodée indépendamment des autres slices de la même image, en termes de codage entropique, de prédiction de signal et de reconstruction de signal résiduel. Une slice peut être une image entière ou une région d'une image. L'un des principaux objectifs des slices est la resynchronisation en cas de perte de données. Dans le cas d'une transmission par paquets, le nombre maximal de bits de charge utile dans une slice est généralement limité et le nombre d'unités CTU dans la slice est souvent modifié pour minimiser le sur-débit (overhead) de mise en paquets tout en conservant la taille de chaque paquet dans cette limite [31] [55].

La norme de codage vidéo H.264 / AVC utilise des slices pour soutenir le codage / décodage parallèle et offre également une résilience aux erreurs. Chaque slice peut être décodée de manière indépendante, c'est-à-dire sans utiliser d'informations provenant d'une autre slice. H.264 / AVC introduit l'ordonnancement de macro-blocs flexible (FMO) comme un outil pour organiser les macro-blocs en slices d'une manière très flexible. Alors que le FMO

pouvait améliorer la résilience aux erreurs dans les situations de perte élevée [53], il n'était pas largement utilisé dans la pratique en raison de la complexité et d'une efficacité de codage plus faible provoquée par la prédiction invalidante entre les limites des slices.

La figure II.8 montre une trame vidéo partitionnée en slices, avec une slice surlignée en bleu. La slice en surbrillance contient six CTU 64×64 .



Figure II.8 Slice et CTU d'une trame vidéo (source Parabola research).

HEVC a introduit les tiles comme moyen de prise en charge du traitement parallèle, avec plus de flexibilité que les tranches normales en H.264 / AVC mais d'une complexité considérablement inférieure à celle de FMO. Les tiles sont spécifiées par des limites verticales et horizontales avec des intersections qui partitionnent une image en régions rectangulaires [47]. La figure II.9 montre un exemple de partitions de tiles qui contiennent des slices. L'espacement des limites de ligne et de colonne des tiles n'a pas besoin d'être uniforme. Cela offre une plus grande flexibilité et peut être utile pour les applications de résilience aux erreurs. Dans chaque tile, les CTU sont traités dans un ordre de balayage raster. De même, les tiles elles-mêmes sont traitées dans un ordre de balayage raster dans une image.

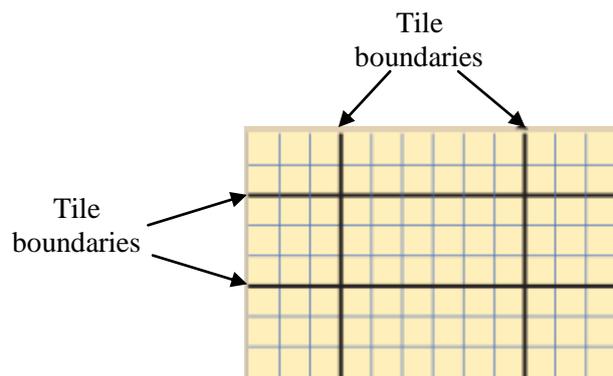


Figure II.9 Un exemple montrant une image partitionnée en neuf tiles

HEVC prend également en charge les Slices, similaires aux slices trouvées dans H.264 / AVC, mais sans FMO. Les slices et les tiles peuvent être utilisés ensemble dans la même image. Pour prendre en charge le traitement parallèle, chaque slice du HEVC peut être subdivisée en slices plus petites appelées slices d'entropie. Chaque slice d'entropie peut être indépendamment décodée par entropie sans référence aux autres slices d'entropie. Par conséquent, chaque cœur d'une CPU peut gérer un décodage entropique traité en parallèle [57].

II.7 Prédiction

Les trames de la vidéo sont codées en utilisant la prédiction Intra ou Inter. La figure II.10 montre une séquence vidéo codée. La première image (0) est codée en utilisant uniquement la prédiction intra, à travers la prédiction spatiale d'autres régions de la même image. Les images suivantes sont prédites à partir d'une, deux ou plusieurs images de référence, en utilisant la prédiction Inter et / ou Intra pour chaque unité de prédiction (PU) [56]. Les sources de prédiction pour chaque image sont indiquées par des flèches.

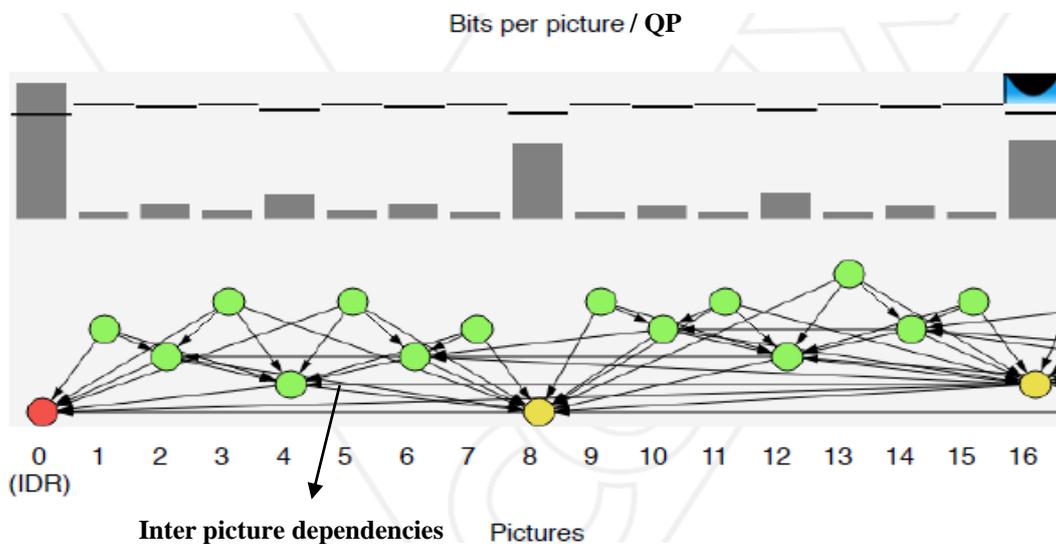


Figure II.10 Les trames d'une séquence vidéo codée par HEVC

(Exemple Sintel.hevc - source Parabola research)

II.8 Transformation et Quantification

Similaire au H.264 / AVC, HEVC applique une transformation entière de type DCT sur le résidu de prédiction. HEVC inclut des transformations qui peuvent être appliquées à des blocs de tailles allant de 4×4 à 32×32 pixels. Les vecteurs de base des transformations 4×4 , 8×8 , 16×16 et 32×32 sont disponibles dans [58].

HEVC prend également en charge les transformations sur des blocs rectangulaires (non carrés), où les transformations de lignes et de colonnes ont des tailles différentes. Les

transformations entières utilisées dans HEVC sont de meilleures approximations de la DCT que les transformations utilisées dans H.264 / AVC. Les vecteurs de base des transformées HEVC ont une énergie égale, il n'est donc pas nécessaire de compenser les différentes normes, comme dans H.264 / AVC. HEVC incorpore également une transformation sinusoïdale discrète (DST) 4×4 , qui est utilisée pour les blocs codés avec un certain sens - modes d'intra-prédiction. Lorsqu'on utilise l'intra-prédiction, les pixels proches de ceux utilisés pour la prédiction (c'est-à-dire près des limites supérieures ou gauches) seront généralement prédits plus précisément que les pixels plus éloignés. Par conséquent, les résidus ont tendance à être plus grands pour les pixels éloignés des limites. La DST sera généralement meilleure pour coder ces types de résidus, parce que les fonctions de base DST commencent bas et augmentent, par rapport aux fonctions de base DCT qui commencent haut et diminuent [59].

II.8.1 La transformée en cosinus discrète (DCT):

Discrete Cosine Transform (DCT) joue un rôle essentiel dans la compression vidéo en raison de son efficacité presque optimale et de son compactage à haute énergie. HEVC soutient DCT de différentes tailles telles que 4, 8, 16 et 32, c'est une technique de compression avec perte. DCT intègre à la fois le schéma de perte et le système de codage sans perte.

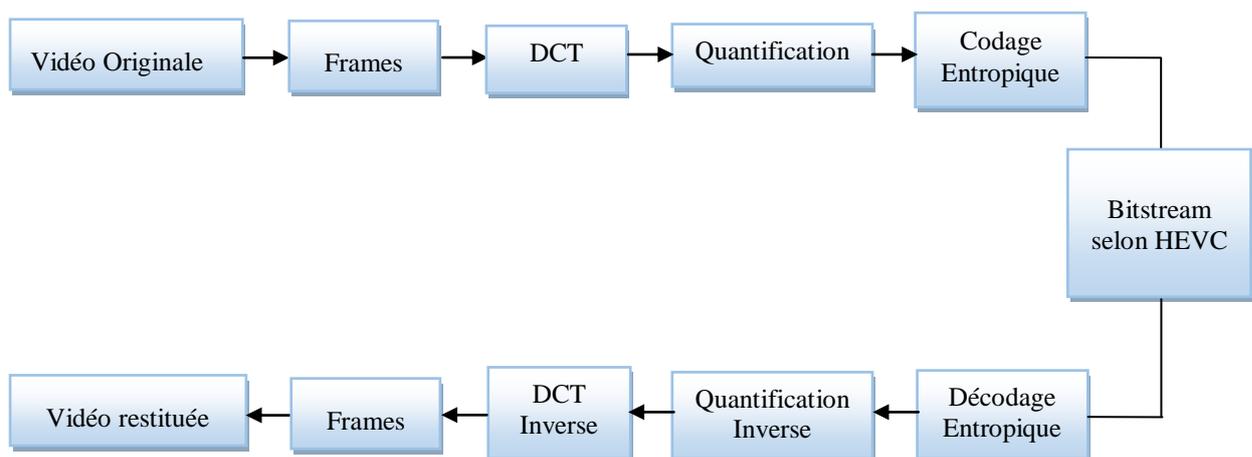


Figure II.11 Mécanisme HEVC

La DCT est une transformée travaillant sur un signal discret. Elle prend un ensemble de points d'un domaine spatial et les transforme en une représentation équivalente dans le domaine fréquentiel. Nous allons opérer la DCT sur un signal en trois dimensions. En effet, le signal est une image graphique, les axes X et Y étant les deux dimensions de l'écran, l'axe des Z reprenant l'amplitude du signal, et la valeur du pixel en un point particulier de l'écran. La DCT transforme un signal d'amplitude (chaque valeur du signal représente "l'amplitude" d'un

phénomène, ici la couleur) discret bidimensionnel en une information bidimensionnelle de "fréquences" [60].

La transformée DCT est une transformation numérique qui est appliquée à chaque bloc et pour chaque « couleur ». Cette transformée est une variante de la transformée de Fourier. Cette méthode permet de décrire chaque bloc en une carte de fréquences et en amplitudes plutôt qu'en pixels et couleurs. La valeur d'une fréquence reflète l'importance et la rapidité d'un changement, tandis que la valeur d'une amplitude correspond à l'écart associé à chaque changement de couleur.

Principe

Projeter un bloc d'une image dans une base de cosinus : conversion d'une image en élément fréquentiel.

Les basses fréquences représentent les éléments les plus importants

Les hautes fréquences correspondent aux détails de l'image.

À chaque bloc de $N \cdot N$ pixels sont ainsi associées $N \cdot N$ fréquences

La transformée DCT s'exprime mathématiquement par :

$$DCT(i, j) = \frac{2}{N} C(i) C(j) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} \text{pixel}(x, y) \cos \left[\frac{(2x+1)i\pi}{2N} \right] \cos \left[\frac{(2y+1)j\pi}{2N} \right], \quad (1)$$

Et la transformée DCT inverse s'exprime par :

$$\text{pixel}(x, y) = \frac{2}{N} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} C(i) C(j) DCT(i, j) \cos \left[\frac{(2x+1)i\pi}{2N} \right] \cos \left[\frac{(2y+1)j\pi}{2N} \right], \quad (2)$$

Dans les deux cas, la constante C vaut :

$$C(x) = \begin{cases} \frac{1}{\sqrt{2}} & \text{pour } x = 0 \\ 1 & \text{pour } x > 0 \end{cases}$$

Pour clarifier la transformation, un exemple provenant de Digital Image Compression Techniques [48] a été utilisé.

Matrice (bloc de pixels) de base :

$$f = \begin{bmatrix} 139 & 144 & 149 & 153 & 155 & 155 & 155 & 155 \\ 144 & 151 & 153 & 156 & 159 & 156 & 156 & 156 \\ 150 & 155 & 160 & 163 & 158 & 156 & 156 & 156 \\ 159 & 161 & 162 & 160 & 160 & 159 & 159 & 159 \\ 159 & 160 & 161 & 162 & 162 & 155 & 155 & 155 \\ 161 & 161 & 161 & 161 & 160 & 157 & 157 & 157 \\ 162 & 162 & 161 & 163 & 162 & 157 & 157 & 157 \\ 162 & 162 & 161 & 161 & 163 & 158 & 158 & 158 \end{bmatrix}$$

Matrice d'origine

$$F = \begin{bmatrix} 1260 & -1 & -12 & -5 & 2 & -2 & -3 & 1 \\ -23 & -17 & -6 & -3 & -3 & 0 & 0 & -1 \\ -11 & -9 & -2 & 2 & 0 & -1 & -1 & 0 \\ -7 & -2 & 0 & 1 & 1 & 0 & 0 & 0 \\ -1 & -1 & 1 & 2 & 0 & -1 & 1 & 1 \\ 2 & 0 & 2 & 0 & -1 & 1 & 1 & -1 \\ -1 & 0 & 0 & -1 & 0 & 2 & 1 & -1 \\ -3 & 2 & -4 & -2 & 2 & 1 & -1 & 0 \end{bmatrix}$$

Matrice transformée DCT

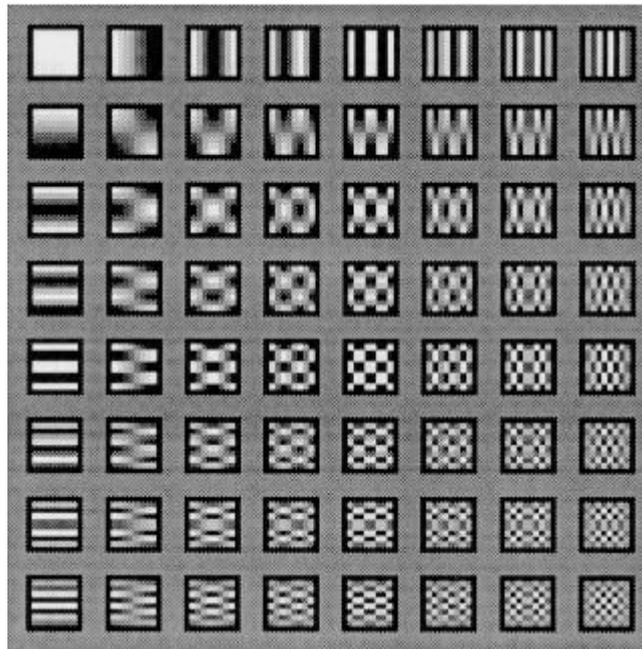


Figure II.12 Les images de base de la DCT 8x8

II.8.2 Quantification

Les blocs de la DCT obtenues seront quantifiés, la quantification est l'étape essentielle responsable de la dégradation de l'image et la vidéo introduit des pertes dans le schéma de codage, grâce à cette étape la vidéo va être réellement compressée et l'utilisateur choisit au départ la qualité de la compression.

Comme dans H.264 / AVC, HEVC utilise principalement la même quantification surveillée par un paramètre de quantification (QP), l'intervalle de valeurs de QP varie entre 0 et 51[41].

La quantification permet également de contrôler le débit de sortie du codeur. En effet, plus on prendra un grand QP, plus on baissera la quantité d'information à coder et donc le débit binaire, mais cela influe sur la qualité de la vidéo, il y a donc un compromis à trouver entre la qualité et le débit souhaité.

II.9 Codage entropique

Après transformation, le codage entropique est appliqué pour coder tous les éléments de syntaxe et les coefficients de transformation quantifiés. Dans H.264 / AVC, le codage CAVLC (Context Adaptive Variable Length Coding) est le codeur entropique de base, et le CABAC est facultativement utilisé dans les profils principal et haut. CABAC peut fournir une meilleure efficacité de codage que le CAVLC en raison de son moteur de codage arithmétique

et de sa modélisation contextuelle plus sophistiquée. Bien que CABAC améliore l'efficacité de codage de 10 à 15% par rapport au codage CAVLC, il augmente, par contre, la complexité de codage. Ceci est plus prononcé à des débits binaires plus élevés (petits paramètres de quantification (QP)), où les données de coefficient de transformation ont un rôle important dans les bitstreams codés [47][60].

Dans HEVC, pour améliorer le débit le plus défavorable, le codec utilise un mode alternatif à plus haut débit pour coder les données du coefficient de transformation. La figure II.13 illustre le bloc diagramme du codage entropique HEVC. Comme on peut le voir, il existe deux modes de codage entropique HEVC: la binarisation à haut rendement (HEB) et la binarisation à haut débit (HTB) [61]. Le mode HEB est entièrement basé sur CABAC tandis que le mode HTB est partiellement basé sur le module de codage résiduel CAVLC bien connu. HTB est destiné à servir en tant que mode haut débit de HEVC, et son utilisation est signalée au niveau de la slice (identificateur d'un bit indiquant si HTB est utilisé). En mode HTB, tous les éléments de syntaxe à l'exception des coefficients résiduels sont codés en utilisant CABAC tandis que les coefficients résiduels sont codés en utilisant CAVLC. En utilisant cette conception harmonisée, le codage entropique HEVC utilise les meilleures caractéristiques des codages CABAC et CAVLC (c'est-à-dire, haute efficacité et faible complexité, respectivement).

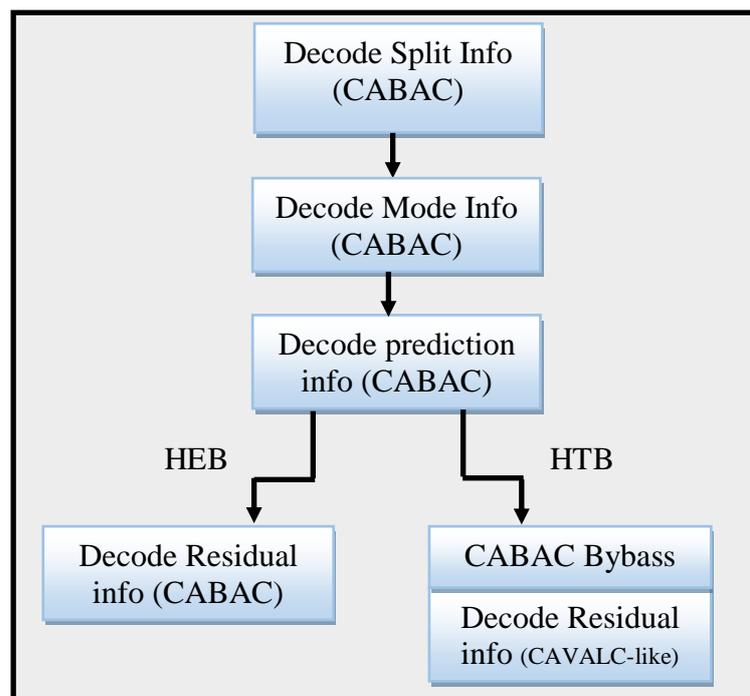


Figure II.13 Codage entropique dans HEVC

II.10 Filtrage de boucle

Le filtrage en boucle est appliqué après la quantification inverse et la transformation, mais avant que l'image reconstruite ne soit utilisée pour prédire d'autres images par compensation de mouvement. Le filtrage de boucle reflète le fait que le filtrage est effectué dans le cadre de la boucle de prédiction plutôt que du post-traitement. H.264 / AVC inclut un filtre de déblocage en boucle. HEVC utilise un filtre de déblocage similaire à celui utilisé dans H.264 / AVC mais développe également un traitement en boucle en introduisant deux nouveaux outils: SAO (Sample Adaptive Offset) et ALF (adaptive loop filtering). Ces techniques sont destinées à annuler la distorsion introduite dans les principales étapes du processus de codage (prédiction, transformation et quantification). En incluant le filtrage dans la boucle de prédiction, les images serviront de meilleures références pour la prédiction compensée par le mouvement car elles ont moins de distorsion de codage [62].

II.11 Filtre de déblocage

Le blocage est connu comme l'un des artefacts les plus visibles et répréhensibles des méthodes de compression par blocs. Pour cette raison, en H.264 / AVC, les filtres passe-bas sont appliqués de manière adaptative pour bloquer les limites en fonction de la force des limites. Cela améliore la qualité subjective et objective de la vidéo, HEVC utilise un filtre de déblocage en boucle similaire à celui utilisé dans H.264 / AVC. Dans HEVC, il existe plusieurs types des limites de bloc, telles que les CU, PU et TU. L'ensemble des limites qui peuvent être filtrées dans HEVC est l'union de toutes ces limites (à l'exception des 4x4 blocs, qui ne sont pas filtrés pour réduire la complexité). Pour chaque limite, il est décidé d'activer ou de désactiver le déblocage et d'appliquer ou non un filtrage fort ou faible. Cette décision est basée sur les gradients de pixels à travers la frontière et les dérivés des seuils de la QP dans les blocs. Pour plus de détails sur le filtre de déblocage, reportez-vous à [47].

II.12 Décalage adaptatif d'échantillon (Sample Adaptive Offset)

SAO est un nouvel outil de codage introduit dans HEVC, qui consiste à classer les pixels dans différentes catégories et à ajouter une valeur de décalage simple à chaque pixel en fonction de sa catégorie [63].

SAO classe les pixels reconstruits en différentes catégories en fonction de l'intensité ou des propriétés des bords. Il ajoute ensuite un décalage, soit le décalage de bande (BO) ou le décalage de bord (EO) aux pixels de chaque catégorie dans une région pour réduire la distorsion. BO classe tous les pixels d'une région en plusieurs bandes, chaque bande contenant des pixels dans le même intervalle d'intensité.

La gamme d'intensité est divisée en 32 intervalles égaux de zéro à l'intensité maximale. Par exemple, pour les données 8 bits, la valeur maximale est 255, donc les bandes auront une largeur de $256/32 = 8$ pixels. Les 32 bandes sont divisées en deux groupes. Un groupe se compose des 16 bandes centrales tandis que l'autre groupe comprend les 16 bandes restantes (voir la figure V.14). Le codeur décide à quel groupe de bandes il doit appliquer SAO, donc 16 décalages seront codés dans le flux de bits [62].

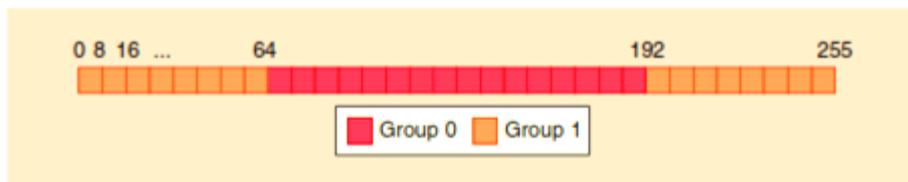


Figure II.14 Un exemple d'intensité et groupes de bands dans 80 modes pour 8 bits

EO utilise l'un des quatre motifs unidimensionnels de trois pixels pour classer les pixels en fonction de leur direction, comme illustré dans la Figure V.15. Chaque pixel peut être classé en tant que pic (s'il est supérieur à deux voisins), vallée (s'il est moins que les deux voisins), bord (s'il est égal à un voisin, catégories 2 et 3), ou aucun de ceux-ci. Quatre valeurs de décalage seront calculées pour ces quatre catégories.

Le codeur peut choisir d'appliquer BO ou EO à différentes régions d'une image. Il peut également signaler que ni BO ni EO ne sont utilisés pour une région.

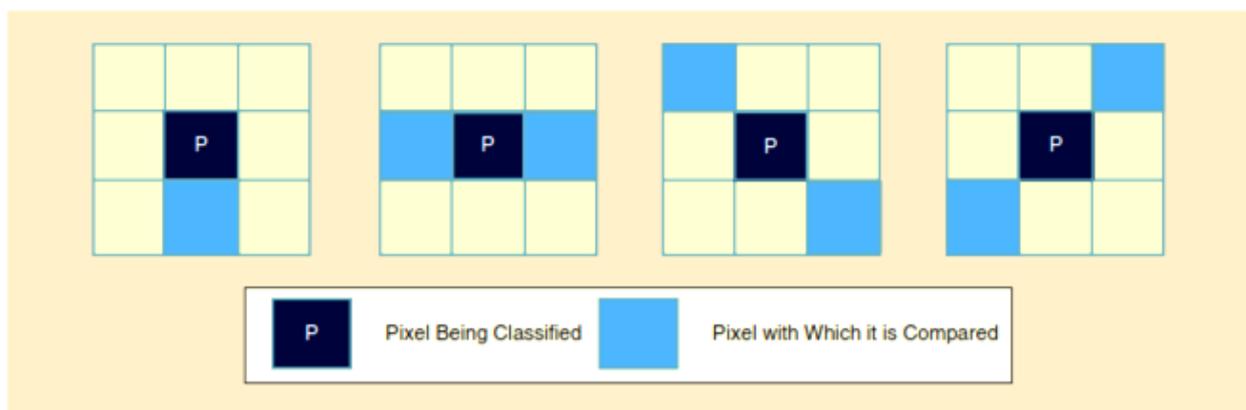


Figure II.15 Mode EO

II.13 Filtrage de boucle adaptatif (Adaptive Loop Filtering)

En HEVC, une ALF (Adaptive Loop Filtering) est appliquée au signal reconstruit après le filtre de déblocage et SAO. Le filtre est adaptatif dans le sens que les coefficients sont signalés dans les bitstreams et peuvent donc être conçus sur la base du contenu de l'image et de la distorsion de l'image reconstruite. Le filtre est utilisé pour restaurer l'image reconstruite de telle sorte que l'erreur quadratique moyenne entre l'image source et l'image reconstruite soit minimisée.

La version 1 de HEVC [47] utilise une seule forme de filtre, une croix superposée sur un carré de 3×3 avec neuf coefficients à coder dans le bitstream (figure II.16) [47][62]. Noter que le nombre de prises dans le filtre est supérieur à neuf en raison de la symétrie.

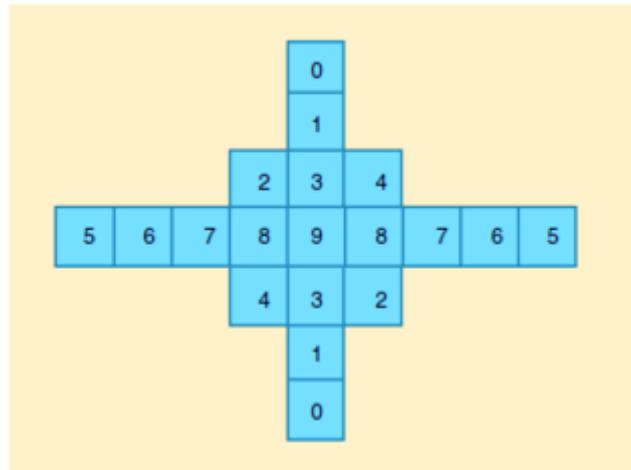


Figure II.16 Une forme de ALF

Deux modes peuvent être utilisés pour appliquer différents filtres à différents pixels dans chaque image: l'adaptation basée sur la région (RA) et l'adaptation basée sur le bloc (BA).

En mode RA, l'image est divisée en 16 régions de même taille. Ces régions peuvent être fusionnées et chaque région restant après la fusion aura son propre filtre (avec un ensemble unique de coefficients). En mode BA, les blocs 4×4 sont classés dans l'une des 16 catégories basées sur l'activité et la direction du bord.

Ces catégories peuvent être fusionnées, et à la fin, un filtre sera conçu pour chacune des catégories restantes après la fusion.

Les coefficients de filtre pour chaque région peuvent être calculés sur la base de l'auto-corrélation et de la corrélation croisée des pixels d'origine et des pixels de reconstruction dans la région (en utilisant les équations de Wiener-Hopf) [64].

L'ALF peut être activé ou désactivé pour différentes zones d'image en fonction du partitionnement des CTU en CUs (dans une structure de segmentation à quatre arbres).

II.14 Performance par rapport au H.264 / AVC

Le codage HEVC comprend de nombreux développements, notamment:

- Codage de bloc de structure: Bien que le codage H.264 utilise de petits clusters avec une taille maximale de 16×16 , le codage HEVC utilise un bloc de châssis ou un codage CTB avec une taille maximale de 64×64 pixels. Les tailles de blocs plus grandes sont plus efficaces lors du codage de grandes images telles que 4K. Comme le montre la figure V.17.

- le codage H.264 utilise jusqu'à 9 tendances de prédiction interne. Alors que le codage HEVC peut utiliser plus de 35 directions, ajoutant plus de blocs de pixels de référence potentiels qui donnent une plus grande efficacité à la compression de trame interne. Cela se fait au prix d'un temps d'encodage supplémentaire nécessaire pour rechercher des tendances supplémentaires.
- **Luma : 35 prediction directions (33 + Planar + DC)**

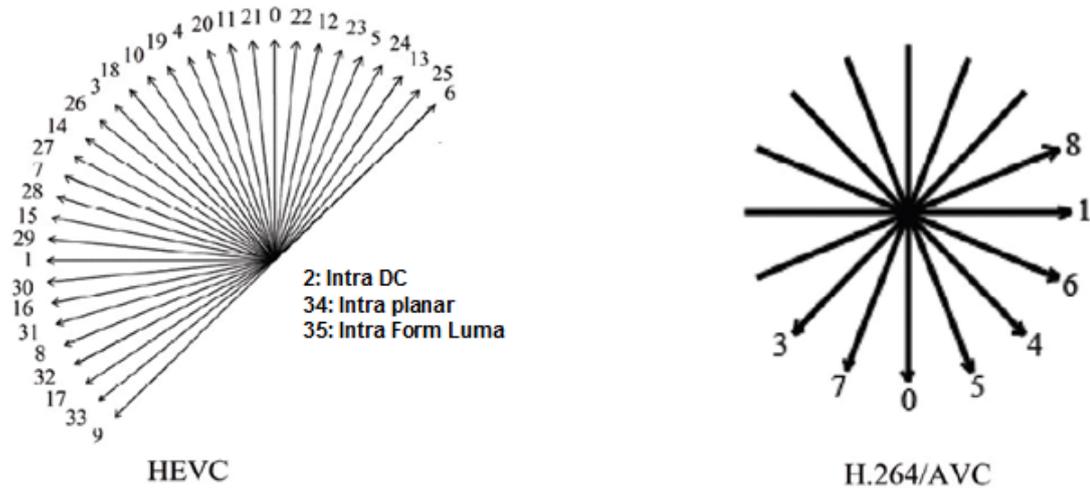


Figure II.17 Directions de prédiction H.265/HEVC Vs H.264/AVC

D'autres développements incluent:

- Prédiction de la force adaptative orientée vers le mouvement, qui permet au codec de trouver plus de répétitions entre les trames.
- Outils de parallélisme supérieurs, qui incluent le traitement parallèle du front d'onde pour fournir un codage plus efficace dans un environnement multi-noyaux.
- Le codage entropique est un calcul binaire compatible avec le contexte (CABAC) seulement, contrairement au codage de longueur variable qui est compatible avec le contexte (CAVLC).
- Améliorations du filtre de suppression de bloc et création d'un second filtre appelé l'équation d'échantillonnage d'adaptation, qui détermine ensuite la sortie le long des bords du bloc.

En général, H.265 offre plusieurs avantages par rapport à H.264, notamment une meilleure compression, une image délicate et une économie de bande passante. Le tableau II.1 présente la comparaison entre H.265 / H.264.

	H.265 / HEVC	H. H.264 / AVC
Noms	MPEG-H, HEVC, Part 2	MPEG 4 Part 10, AVC
Date confirmée	2013	2003
Progression	Successeur du H.264/AVC	Successeur du MPEG-2 Part
Amélioration	Réduction de 40 à 50% du débit par rapport au H.264 avec la même qualité visuelle * Il est susceptible de mettre en œuvre Ultra HD, 2K, 4K pour la diffusion et en ligne	Réduction de 40 à 50% du débit par rapport à la partie MPEG-2 * Disponible pour fournir des sources HD pour la diffusion et en ligne
Soutenir jusqu'à 8K	Oui	Non. Support jusqu' à 4K
Soutenir jusqu'à 300 fps	Oui	No. support up to 59.94 fps only.

Tableau II.1 Comparaison entre HEVC/H.265 / H.264/MPEG-4 AVC.

II.15 Les unités NAL (NAL unit) [41][47]

Les unités NAL dans le flux binaire HEVC sont séparées par une certaine séquence de bits réservé, spécifiquement 0x000001 (3 octets) ou 0x00000001 (4 octets).

Le flux binaire contient divers paramètres, organisés en ensembles de paramètres, et des valeurs de pixel codées organisées en slices. La structure du bitstream est illustrée à la figure II.18.



Figure II.18 Structure d'un bitstream HEVC

Les jeux de paramètres sont les suivants:

- **Video Parameter Set (VPS)** : contient les informations des différentes couches (Slices), par exemple dans les extensions de codage vidéo évolutif ou codage vidéo 3D.
- **Sequence Parameter Set (SPS)** contient les informations des sous-couches (Slices) temporelles, profondeur de couleur, taille du bloc de codage, etc.
- **Picture Parameter Set (PPS)** contient les informations des paramètres de quantification initiale (QP), résolution, nombre des tiles (mosaïques), etc.
- **Slice header** contient l'index de trame (nombre de commandes d'image), l'adresse de la première CTU, et ainsi de suite.

Ces paramètres permettent au décodeur HEVC d'interpréter et de décoder la vidéo bitstream.

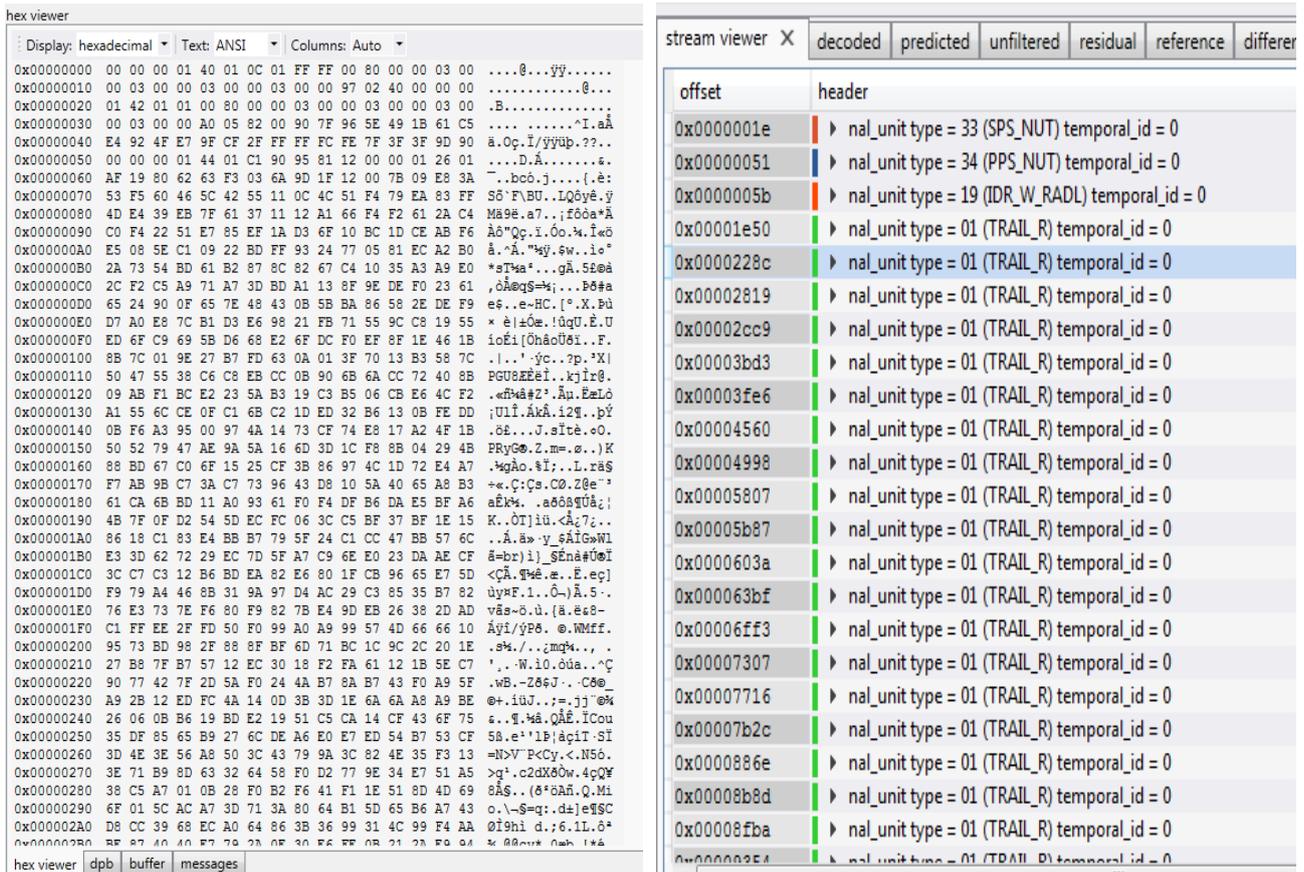


Figure II.19 Structure d'un bitstream HEVC (source Parabola research)

II.16 Conclusion

Le nouveau standard de compression HEVC permet d'augmenter le niveau de compression de 50% en moyenne tout en conservant le même niveau de qualité objective et subjective que le précédent, H.264.

Le HEVC / H.265 a non seulement donné une meilleure qualité visuelle avec un taux de stockage et une bande passante réduits, mais également un algorithme de codage habile en codant des vecteurs de mouvement avec une précision beaucoup plus grande et des erreurs résiduelles minimales. Ce nouveau codec présente également un filtre de déblocage amélioré et un décalage adaptatif de l'échantillon pour réduire encore plus d'artéfacts. Outre l'amélioration en termes de qualité et de taux de compression, elle prend en compte la transmission vidéo en temps réel sur un réseau à qualité de service non garantie (perte de paquets...).

Chapitre III

Multiple Description Coding

I.1. Introduction

Ces dernières années, la transmission de la vidéo en temps réel via des canaux perturbés a été un défi. La transmission de la vidéo brute est impossible en raison de l'énorme bande passante requise, c'est pourquoi la compression vidéo est indispensable.

Par ailleurs, la vidéo compressée est très sensible à la perte de paquets dans les réseaux de "meilleur effort" (Best effort) comme Internet.

Le codage à description multiple (MDC) résout le problème de codage d'une source en transmettant deux (ou plusieurs) bitstreams. Ainsi, une reconstruction de haute qualité est décodable à partir des deux (ou plusieurs) bitstreams. Tandis qu'une reconstruction de qualité inférieure, mais acceptable, est décodable si l'un des deux bitstreams est perdu.

III.2 Effets des erreurs et les erreurs de propagation

La compression vidéo utilise généralement des Couches de codage vidéo (en anglais: Video Coding Layer - VCL) [2] afin d'obtenir des gains de compression élevés. Les VCL sont très sensibles aux erreurs de bits. Une erreur de bit unique peut se propager à de nombreux VCL, et la détection d'erreur peut ne pas être capable de localiser l'emplacement exact de l'erreur. De plus, le codage prédictif, employé dans toutes les normes de codage vidéo, se prête à la propagation d'erreurs. Dans cette section, nous discuterons les effets des erreurs sur les bits et les pertes de paquets, ainsi que la propagation d'erreurs spatiales et temporelles résultant de telles erreurs.

III.2.1 Effets d'erreurs binaires

Les erreurs sur les bits se produisent dans les réseaux orientés bits (par exemple les réseaux sans fil) et peuvent être détectées au niveau du transport ou au niveau du codage de canal. Au niveau du transport, la détection d'erreur est normalement effectuée sur un bloc de données. Si un décodeur vidéo décide de supprimer un segment vidéo contenant des erreurs, alors une erreur binaire est équivalente à la perte d'un paquet. Si la détection d'erreur au niveau du transport n'est pas disponible ou si le système décide de traiter des données connues comme étant incorrectes, des erreurs binaires seront présentes sur le décodeur vidéo. Les erreurs de bit unique dans le flux de bits vidéo compressé ne seront pas localisées dans une petite région

spatiale, mais conduiront à la perte de nombreux macro-blocs en raison de la perte de synchronisation entre le décodeur et de bits [26]. Comme le décodeur ne peut pas localiser l'emplacement exact des erreurs sur les bits pendant le décodage, les données entre deux mots de code de synchronisation (mots de synchronisation) sont généralement abandonnées.

Les erreurs de bits sont traitées de cette manière; elles sont des pertes de paquets équivalentes avec des limites de paquets aux emplacements des mots de synchronisation.

III.2.2 Effet des pertes de paquets

Les pertes de paquets se produisent dans les réseaux de paquets tels que l'ATM (Asynchronous Transfer Mode), réseaux sans fil ou Internet. De même, les erreurs sur les bits peuvent être vues comme des pertes de paquets au niveau du décodeur vidéo, en fonction de la manière dont ces erreurs sont traitées, comme vu ci-dessus. La taille des paquets, l'emplacement spatio-temporel des paquets dans la séquence vidéo et l'étendue du codage prédictif dicteront la localisation d'erreur des pertes de paquets.

Une perte de paquet peut affecter seulement un petit emplacement spatial d'une trame vidéo ou peut conduire à la perte de trame (s) vidéo, en fonction de la taille du paquet. Dans certains réseaux, les pertes de paquets peuvent se produire de manière explosive, ce qui peut entraîner la perte de nombreuses trames vidéo. Si la perte de paquets se produit dans un emplacement spatial limité dans une trame vidéo, une dissimulation d'erreur spatiale ou spatio-temporelle peut être appliquée aux données vidéo manquantes. Si des trames complètes sont perdues, le décodeur peut choisir de figer l'affichage vidéo avec la dernière trame reçue correctement, ou il peut tenter une interpolation de trame temporelle pour conserver une fréquence de la vidéo adéquate. Nous discuterons plus tard des techniques de résilience d'erreur qui améliorent la localisation et minimisent les effets des pertes de paquets [3][4][26].

III.3 Techniques de codage source/canal pour la protection des bitstreams vidéo

Il existe trois catégories d'approches utilisées pour compenser l'effet de perte de données (paquets) dans la vidéo pour la transmission vidéo sur des réseaux bruités [26]:

1. Un protocole de couche de transport basé sur ARQ (Automatic Repeat reQuest) fiable, tel que TCP.
2. FEC (Forward error correction)
3. ERC (error resilient coding).

III.3.1 Requête automatique de répétition (Automatic Repeat reQuest ARQ)

La demande de répétition automatique est un protocole de contrôle d'erreur dans la transmission de données. Lorsque le récepteur détecte une erreur dans un paquet, il demande automatiquement à l'émetteur de renvoyer le paquet. Ce processus est répété jusqu'à ce que le paquet soit exempt d'erreur ou que l'erreur continue au-delà d'un nombre prédéterminé de

transmissions. ARQ est parfois utilisé avec la communication GSM (Global System for Mobile) pour garantir l'intégrité des données.

En ARQ, le récepteur demande par l'intermédiaire d'un canal retour, la retransmission d'un paquet perdu. Les paquets perdus sont les paquets non reçus ou reçus avec une erreur de bit. En vérifiant les symboles de parité, les paquets erronés sont détectés et demandés pour la retransmission.

III.3.2 Correction d'erreur directe (Forward Error Correction FEC)

Dans cette méthode (la correction d'erreur directe), certaines données de correction sont ajoutées au message principal de manière redondante. Si les données perdues sont dans la capacité de correction des codes FEC, le message entier peut être récupéré. La capacité de correction dépend du nombre de symboles de parité ajoutés. Par exemple, FEC dans Reed Solomon, en ajoutant $2t$ parités, t sont les symboles erronés pouvant être corrigés [26].

III.3.3 Erreur de codage résilient (Error Resilient Coding ERC)

ERC permettent au décodeur de masquer et de régénérer les données perdues à partir des données reçues, en exploitant la corrélation qui existe parmi les blocs des images.

ARQ et FEC sont des protections au niveau du canal, tandis que ERC peut être utilisée comme une protection au niveau de la source, telle que :

- MDC est l'une des solutions les plus prometteuses pour la vidéo en temps réel sur les réseaux avec perte [2] [24].
- Codage Conjoint Source-Canal connue sous le nom de codage JSCC (Joint Source Channel Coding) tel que LC (Layered Coding).

Layered Coding: également appelé codage vidéo évolutif (SVC), la source est codée en couche de base et en une ou plusieurs couches d'amélioration. Au niveau du récepteur, les couches sont superposées hiérarchiquement l'une sur l'autre. La qualité de la vidéo est améliorée par le nombre de couches d'amélioration reçues. La couche de base est généralement protégée à l'aide de codes FEC et, par conséquent, LC est catégorisée comme une méthode JSCC. L'avantage de MDC devant Layered Coding est qu'il ne nécessite pas de dispositions spéciales dans le réseau pour fournir un sous-canal fiable.

- ERC comme options de H.264 / AVC: Ces options de codage telles que MB (macro blocks) Intra-refreshment, codage de la trame B, Slice et Flexible Macroblock Ordering (FMO) permettent au décodeur de masquer et de régénérer les données perdues à partir des données reçues, en exploitant la corrélation qui existe. Ces options ne fonctionnent que si le taux de perte est très faible ou au plus bas. Dans les environnements à taux de perte modéré ou élevé, MDC est plus bénéfique [26].

III.4 MDC (Codage à description multiple)

Le codage à description multiple était à l'origine utilisé pour la communication de la parole sur un réseau à commutation de circuits dans les années 1970. Traditionnellement, pour éviter les interruptions de communication, une liaison de transmission supplémentaire était en attente et serait activée en cas de panne du lien principal. Cependant, cette approche n'était pas rentable et d'où l'idée de scinder l'information sur deux ou plusieurs canaux.

MDC, a été proposé lors de l'atelier sur la théorie de l'information de l'IEEE en 1979. Le problème de MD a été posé par Gersho, Witsenhausen, Wolf, Wyner, Ziv et Ozarow [24]. Supposons qu'une source est décrite par deux descriptions codées respectivement aux débits binaires R_1 et R_2 . Chaque description peut-elle être décodée individuellement avec une distorsion D_1 et D_2 , respectivement, alors que le décodage des deux descriptions entraîne une distorsion D_0 , le problème de MD est de caractériser les quintuples réalisables $\{D_0, D_1, D_2, R_1, R_2\}$

Le MDC est un moyen efficace développé pour gérer la transmission de données sur des réseaux perturbés. MDC est un codage alternatif à un codage évolutif (Scalable coding). MDC produit deux ou plusieurs bitstreams ayant une égale importance.

MDC a plusieurs propriétés importantes:

- 1. Des descriptions indépendantes d'une source vidéo, décodables et redéfinies mutuellement, sont générées.*
- 2. Chaque description peut être décodée indépendamment pour donner une reproduction utilisable du signal original.*
- 3. les descriptions sont transmises séparément, éventuellement via des chemins de réseaux différents.*
- 4. L'ensemble des descriptions permettent au décodeur de décoder avec succès la vidéo.*
- 5. les descriptions multiples contiennent des informations complémentaires de sorte que la qualité du signal décodé s'améliore avec le nombre de descriptions reçues.*

III.5 Méthode

MDC code un signal dans plusieurs flux binaires. Chaque flux binaire est une description et chaque description est décodable indépendamment, comme le montre la figure II.1. Le signal reçu peut être reconstruit avec plusieurs descriptions, ce qui donne une amélioration de la qualité.

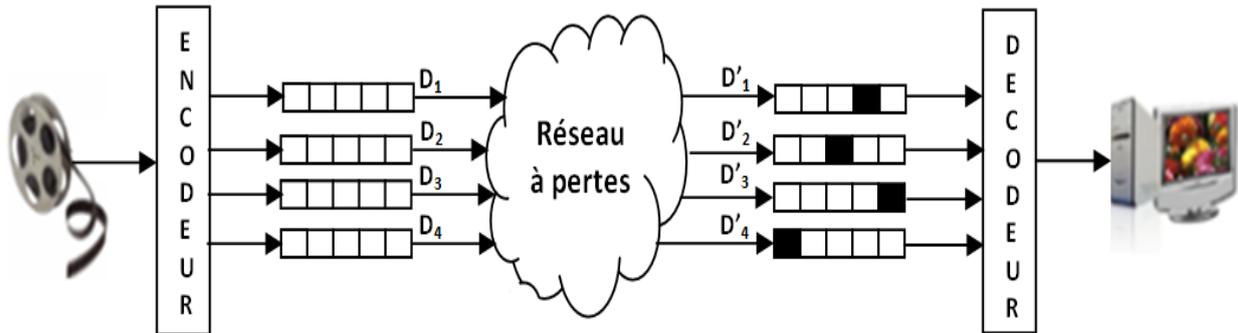


Figure III.1 La structure principale de la technique MDC d'une séquence vidéo.
Les cases noires indiquent les informations perdues[26].

Le MDC [26] prend le pari de répartir l'information différemment. En effet le codeur fournit non pas une seule mais plusieurs descriptions. Chacune d'entre elles constitue une unité indépendamment transmissible. De plus, différents canaux peuvent être utilisés dans ce but. Chaque description ne contient qu'une partie du signal vidéo, mais de manière "équivalente" à toutes les autres descriptions. Autrement dit, la reconstruction peut débuter quelles que soient la ou les descriptions reçues en premier.

Alors que dans le cas d'une description unique, la perte d'un paquet bloque la reconstruction et nécessite sa retransmission. Dans le cas de descriptions multiples la perte de paquet est moins grave. La réception d'un nombre suffisant de paquets (quels qu'ils soient) permet de reconstituer une partie du message initial. Il s'agit donc d'une forme de codage conjoint. D'une part, la possibilité de reconstruire partiellement et progressivement l'information implique que le couple codeur/décodeur a une connaissance de la sémantique de l'information (caractéristique du codage source), les descriptions sont indépendamment manipulables (transmissibles) et peuvent naturellement comporter la redondance (caractéristique du codage canal). Le codage MDC a été introduit dans un contexte d'envoi ininterrompu de données à travers un réseau.

La fiabilité des liens n'étant jamais parfaitement garantie, la méthode alors utilisée consistait à rediriger l'information dès la détection de la perte d'un lien via une autre route prévue à cet effet. Outre des temps de détection de perte et de redirection indésirable, ce système nécessitait donc l'emploi d'un ou plusieurs canaux parallèles inutiles en temps normal, mais pouvant potentiellement être mis à contributions à tout moment.

Actuellement le codage multiple présente un intérêt dans un contexte où s'applique au moins l'une des conditions suivantes:

- la perte de données en cours de transmission est fréquente ;
- l'ordre de réception des données transmises peut différer de l'ordre de l'émission ;

– différents niveaux de qualité de l'information sont acceptables.

En effet, MDC réduit les effets du PLR en transmettant les descriptions à travers des chemins différents. De plus, diverses techniques de dissimulation d'erreurs peuvent être développées pour récupérer l'information perdue. [4] [24] [26]

Les avantages du MDC sont obtenus au détriment de la redondance supplémentaire dans les descriptions. Par conséquent, l'un des principaux objectifs de la conception des systèmes MDC est de minimiser la redondance. La complexité de calcul est une caractéristique importante du MDC [26] en plus, HEVC / H.265 est plus compliqué que H.264 / AVC [50]

Il est vrai que la puissance de calcul des processeurs augmente sans cesse et la consommation augmente. Aussi, en raison de la croissance rapide de nombreux dispositifs alimentés par batterie, tels que les tablettes, les smartphones, les nœuds de capteurs sans fil et les ordinateurs portables, la réduction de la consommation d'énergie présente un grand intérêt.

En outre, chaque description afin d'être décodable individuellement doit avoir quelques informations de base de la source. Au décodeur central, une seule copie de cette information de base est utilisée et les autres sont redondantes. Le débit binaire de MDC est supérieur à celui du codage à description unique (SDC) et il offre la même qualité sur un canal sans perte. Ce taux excédentaire est appelé redondance. En d'autres termes, les descriptions, puisqu'elles présentent la même source (mais avec une distorsion), sont corrélées et codées séparément, ce qui produit une redondance. Et puisque les descriptions sont plus proches de la source et donc plus corrélées, la redondance augmente.

Le coût de la redondance est toléré en raison de la résilience aux erreurs atteinte par MDC. Pour les canaux ayant un PLR élevé, la probabilité de décodage latéral est plus élevée et par conséquent, une qualité supérieure est intéressante et la redondance est nécessaire. Cependant, pour les cas où toutes les descriptions sont reçues, cette redondance doit être minimisée. Ainsi, l'accordabilité de la redondance est l'une des principales caractéristiques de MDC, et l'optimisation de la quantité de redondance est le principal défi pour les concepteurs de MDC.

III.6 Schémas de codage vidéo à description multiple [3][24][26]

Les schémas MDC sont regroupés et analysés en fonction de leur domaine d'application. Nous présentons ici les détails de ces groupes MDC et leurs méthodes MDC spécifiques. Nous nous concentrons sur la vidéo, mais les algorithmes sont également proposés et appliqués aux images.

Le processus MDC, générant et divisant l'information dans les descriptions, peut être effectué dans plusieurs domaines. Les domaines sont spatiales, temporels, fréquentiels

et compressés, dans lesquels les descriptions sont générées en partitionnant respectivement les pixels, les trames, les données transformées et les données compressées.

III.6.1 MDC dans le domaine spatial

Dans cette catégorie, le processus MDC est effectué dans le domaine des pixels. L'approche la plus simple consiste à diviser l'image / trame en plusieurs sous-images et à les coder chacune indépendamment. La figure III.2 montre un sous-échantillonnage spatial polyphasé (PSS : Polyphase spatial subsampling) d'une trame générant quatre sous-images pour obtenir un codage à quatre descriptions. Au niveau du décodeur, si toutes les descriptions sont reçues, les sous-images sont fusionnées et l'image en pleine résolution est reconstruite. Sinon, toute description manquée doit être récupérée en utilisant une interpolation ou des techniques similaires. Cependant, dans cette approche de base, il n'y a aucun moyen d'ajouter une redondance et donc pas de disposition pour améliorer la qualité des descriptions latérales.

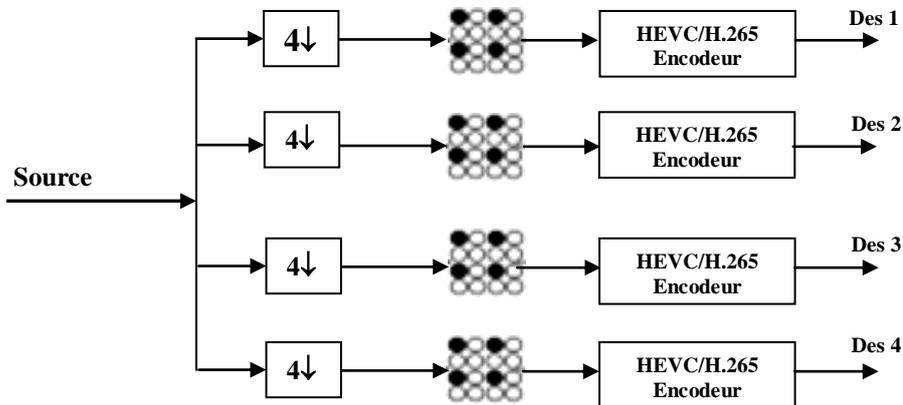


Figure III.2 sous-échantillonnage MDC spatial polyphasé

III.6.2 MDC dans le domaine fréquentiel

Plusieurs approches existent, comme MDSQ (Multiple Description Scalar Quantizer). Le concept de MDSQ consiste à utiliser différentes méthodes de quantification de sorte qu'elles s'affinent au niveau du décodeur central. Le moyen le plus simple de MDSQ est de décaler la moitié des intervalles de quantification de chaque codeur latéral; comme montré sur la figure III.3.

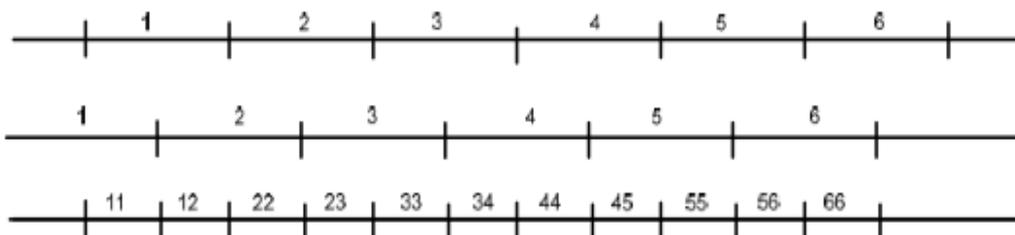


Figure III.3 Simple MDC basé sur MDSQ

III.6.3 MDC dans le domaine compressé

Ces schémas MDC sont appliqués après que la vidéo ait déjà été codée. Cela peut également être appelé MDC de domaine de paquets. L'idée principale ici est de partitionner une couche en K segments qui sont ensuite étendus à $N > K$ segments en utilisant des codes FEC. En raison de l'utilisation de codes de canal, ce schéma peut être supposé être une méthode JSCC. Ensuite, ces N segments sont à nouveau partitionnés en plusieurs segments qui sont alors envoyés indépendamment. On sait qu'avec K de ces N segments, les K segments de sources sont récupérables. L'idée a été inspirée de [65], puis introduite et personnalisée pour le codage à descriptions multiples. Comme cela est représenté sur la figure III.4, la vidéo ou l'image est codée dans un flux de bits évolutif où la couche i est partitionnée en i segments. Pour le codage de la description k , les codes de Reed-Solomon, $RS(N,i)$, sont générés pour ces segments. Les descriptions sont composées en collectant un et un seul segment (source ou code FEC) de chaque couche. Cette configuration et le débit binaire optimal pour chaque couche sont traités dans [66].

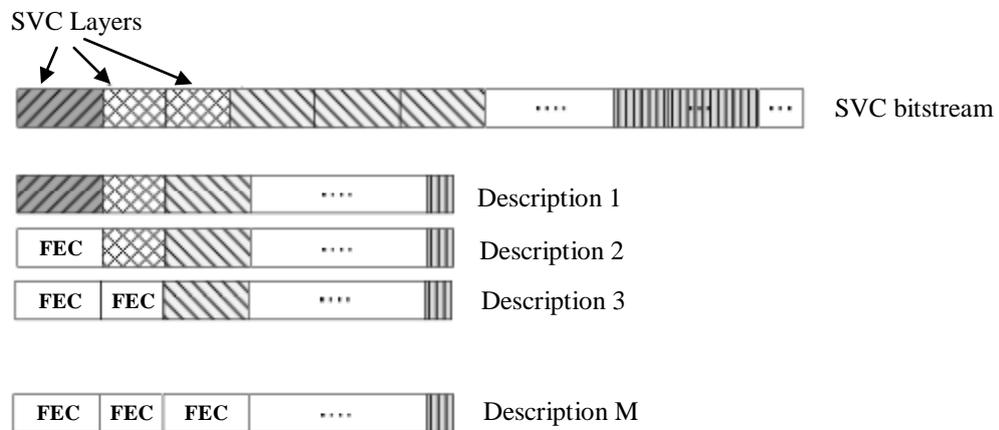


Figure III.4 MDC-FEC avec nombres égaux de descriptions

III.6.4 MDC dans le domaine temporel

La trame est la clé de la granularité dans la catégorie MDC dans le domaine temporel. Principalement les trames sont réparties entre les descriptions (figure III.5b). La figure III.5 compare les méthodes MDC et SDC.

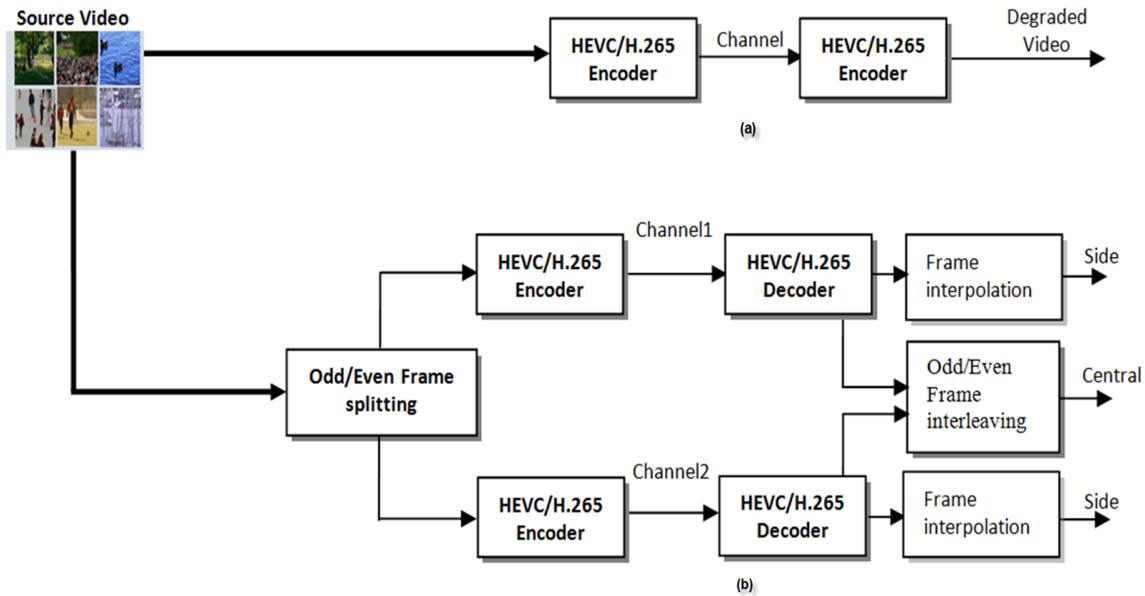


Figure III.5 (a) Codage à description unique (b) MDC dans le domaine temporel

Dans SDC, la séquence vidéo brute est codée avec HEVC / H.265 en une seule description et transmise sur un seul chemin (canal). Le décodeur décode le flux binaire reçu comme représenté sur la figure III.5(a).

En revanche, la figure III.5(b) présente la méthode MDC, où la séquence vidéo brute est divisée en deux descriptions ayant une importance et des débits égaux, les trames impaires dans une description et les trames paires dans l'autre description. La séquence vidéo obtenue paire (ou impaire) est codée en utilisant HEVC / H.265 et mise en paquets de manière indépendante, puis transmise sur deux chemins indépendants (canaux). Du côté du décodeur, chaque bitstreams est décodé par le standard HEVC / H.265. Si toutes les descriptions sont reçues avec succès, le décodeur décode et entrelace pour obtenir la séquence vidéo reconstruite avec le décodeur central avec une résolution maximale et une plus petite distorsion centrale. Si un seul flux (bitstream) est reçu, le décodeur HEVC / H.265 peut décoder la description reçue en interpolant ou en répétant les trames reçues en utilisant le décodeur latéral et reconstruit les données avec une qualité acceptable. La distorsion résultante des données est appelée distorsion latérale. [23][24][25][26].

De plus, le décodage des trames perdues est remplacé par des trames interpolées ou estimées par des méthodes de dissimulation. L'estimation / compensation de mouvement est effectuée par intra-description, ce qui signifie que les trames paires (impaires) sont prédites à partir de trames paires (impaires).

Cependant, l'efficacité de cette approche dépend de l'inter-corrélation temporelle. En d'autres termes, si le mouvement inter-trame est rapide, la version interpolée de la trame perdue

du côté du décodeur est entièrement différente de l'originale. Ce problème a quelques solutions comme l'insertion de trames supplémentaires dans les descriptions. Ces nouvelles images peuvent être identiques dans toutes les descriptions [26].

III.7 Conclusion

MDC est l'une des solutions prometteuses pour la diffusion vidéo en temps réel sur des réseaux avec perte. Le processus MDC générant et divisant l'information dans les descriptions, peut être effectué dans plusieurs domaines. Les domaines sont spatiales, temporels, fréquentiels et compressés, dans lesquels les descriptions sont générées en partitionnant respectivement les pixels, les trames, les données transformées et les données compressées. Sous chaque domaine, nous avons certains nombre de catégories de MDC.

Ainsi, nous cherchons à améliorer et interpréter une vidéo compressée HEVC basée sur MDC à travers un réseau IP en termes de qualité de service non garantie (perte de paquets...) et de taux de compression, elle prend mieux en compte la diffusion vidéo en temps réel.

Chapitre IV

Réseaux IP (Vidéo IP)

IV.1 Introduction

Dans ce chapitre nous allons rappeler les protocoles de communication TCP/IP qui coopèrent avec l'objectif d'offrir une interface à leurs usagers. Nous allons aussi présenter l'architecture générale du réseau (modèle OSI, protocole TCP/IP, adressage IP).

Les environnements mobiles et les principaux concepts liés à ces environnements sont étudiés, nous commençons par la définition de ces environnements. Nous abordons le concept des réseaux ad hoc, leurs caractéristiques et leurs domaines d'applications. Enfin nous exposons les métriques de la qualité de service.

IV.2 Architecture OSI

OSI signifie Open System Interconnection, ce qui se traduit par Interconnexion de systèmes ouverts. Ce modèle a été proposé par l'ISO (International Organization for Standardization) afin de mettre en place une référence de communications entre les machines d'un réseau. Autrement dit, les réglementations qui gèrent les communications entre les machines (Portable, Ordinateurs...etc). Il est donc utilisé pour modéliser toute communication entre les nœuds de réseaux et tout apprentissage sur les réseaux passent par la connaissance du modèle OSI. Ainsi, les différents constructeurs soit en mesure d'élaborer des produits (logiciels ou matériels) compatibles [67].

Le modèle OSI est un modèle qui comporte 7 couches (figure IV.1):

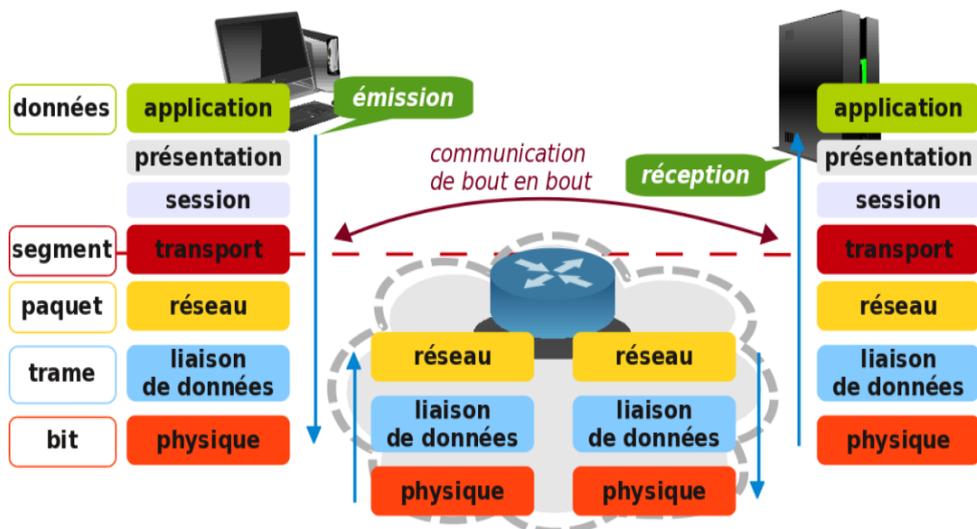


Figure IV.1 Schéma du support d'interconnexion matériel [68]

- La couche physique: définit la façon de laquelle les données sont converties en signaux numériques.
- La couche liaison donnée: Gère les communications entre 2 hôtes reliés directement par un support physique ou connectés à un équipement qui émule une connexion directe.
- La couche réseau : permet de gérer les adresses et le routage des données (adresse IP).
- La couche transport: Gère les communications de bout en bout entre processus (TCP et UDP).
- La couche session: définit l'ouverture des sessions sur les machines du réseau (adresse MAC Medium Access Control).
- La couche présentation: chargée de la conversion du format en un format standard des données applicatives transmises
- La couche application: Fournit un point d'accès aux services réseaux. Le modèle ne donne aucune spécification à cette couche.

IV.3 Architecture TCP/IP [67][68]

Les protocoles TCP/IP (signifie Transmission Control Protocol/Internet Protocol) s'inspire du modèle OSI auquel il reprend l'approche modulaire mais réduit le nombre à quatre construites au-dessus d'une cinquième (matérielle). Le schéma suivant montre les couches conceptuelles ainsi que la structure des données échangées par chaque couche.

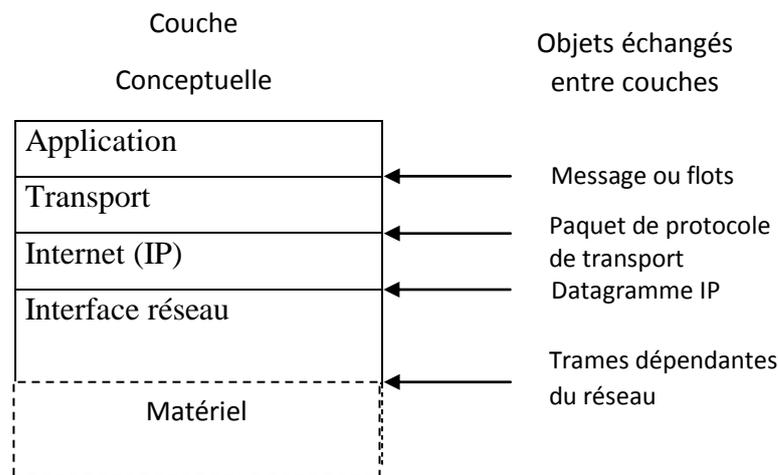


Figure IV.2 Les quatre couches de TCP/IP

L'adressage IP est une opération de la couche IP du protocole TCP/IP c'est-à-dire le fait de fournir à chaque machine du réseau une adresse IP afin de pouvoir acheminer des paquets de données. TCP/IP a été créé à l'origine dans un but militaire. Les couches du modèle TCP/IP

sont plus générales que celles du modèle OSI. Elles doivent répondre à un certain nombre de critères parmi lesquels:

- ❖ Fractionnement des messages en paquets.
- ❖ Utilisation d'un système d'adresse.
- ❖ Acheminement des données sur le réseau (routage).
- ❖ Contrôle des erreurs de transmission de données.

La connaissance du système de protocole TCP/IP n'est pas nécessaire pour un simple usager, au même titre qu'un téléspectateur ne nécessite pas d'apprendre la manière dont marche son téléviseur. Mais sa connaissance reste essentielle pour les personnes souhaitant administrer ou maintenir un réseau fonctionnant au sein d'un système de protocoles TCP/IP.

Protocoles utilisés	Modèle TCP/IP	correspondance en OSI
	Couche application	Application
		Présentation
		Session
TCP / UDP, gestion des erreurs	Couche Transport	Transport
IP / ARP et RARP /ICMP / IGMP	Couche Internet	Réseau
FTS, FDDI, PPP, Ethernet, Anneau à jeton	Couche Accès réseau	Liaison de donnée
		Physique

Tableau IV.1 Modèle TCP/IP et OSI [67]

Il est clair que les couches du modèle TCP/IP ont des tâches beaucoup plus variées que les couches du modèle OSI. De fait que, certaines couches du modèle TCP/IP correspondent à plusieurs couches du modèle OSI.

Le rôle de chaque couche est le suivant :

La couche application:

Au plus haut niveau, les utilisateurs appellent des programmes d'applications qui accèdent aux services disponibles sur TCP/IP. Une application interagit avec les protocoles de couche transport pour émettre et recevoir des données.

La couche transport:

Le premier devoir d'une couche transport est de permettre des communications de programme à programme. La couche transport peut réguler le flux des informations. Elle peut assurer un transport fiable (le mode connecté avec le protocole TCP). Comme elle peut

aussi transporter des paquets avec un service non fiable et sans récupération des paquets perdus (le mode non connecté avec le protocole UDP). Les protocoles de transport découpent le flux de données en paquets (terminologie ISO) et transmettent chaque paquet avec son adresse de destination à la couche suivante.

La couche Internet:

La couche Internet (IP) est chargée de fournir le paquet des données. Elle définit les datagrammes et gère la décomposition / recombinaison des segments. Elle accepte la demande de transmission de paquet provenant de la couche transport accompagnée de l'identification de l'ordinateur destinataire, encapsule le paquet dans un datagramme IP.

La couche Accès réseau:

Il s'agit de la couche de plus bas niveau des protocoles TCP/IP. C'est la couche d'interface avec le réseau, elle est chargée d'accepter des datagrammes IP et de les spécifier la forme sous laquelle les données doivent être acheminées lorsque le réseau est local ou d'un sous-système complexe qui utilise son propre protocole de liaison de données.

IV.4 Les principaux protocoles utilisés en transmission vidéo

IV.4.1 Protocole TCP

TCP «Transmission Control Protocol » [69] est un protocole de la couche transport du modèle OSI. TCP est orienté connexion et il est duplex intégral (full-duplex). Il permet le transfert des données avec un contrôle et une reprise sur erreur. La taille minimum de son en-tête est de 20 octets qui contiennent en plus un nombre d'options.

TCP est incompatible avec le multimédia temps réel et les jeux en réseau. Plusieurs champs dans l'en-tête TCP sont indispensables (dont numéro de port et numéro de séquence), et devront se retrouver dans les protocoles de substitution UDP et RTP (Real Time Protocol).

IV.4.2 Protocole UDP

UDP « User Datagram Protocol » [67][69] est un protocole de transport du modèle OSI utilisé souvent en streaming vidéo. Contrairement à TCP, UDP a très peu d'options à gérer et l'en-tête des paquets UDP est beaucoup plus petit, ce qui contribue à rendre son traitement plus rapide. Le transport est en mode non connecté est donc non fiable et sans garantie de séquençement des datagrammes. Il n'y a pas de contrôle de reprise sur erreur, mais UDP indique qu'il y a eu erreur. Il apporte la possibilité de distinguer plusieurs destinations sur un même ordinateur.

Un programme d'application qui utilise UDP doit gérer les problèmes de fiabilité, pertes de messages, duplications, retards, déséquencelement et perte de connectivité. Chaque message UDP est appelé datagramme (datagramme UDP). Du point de vue conceptuel, le datagramme comporte deux parties ; un en-tête fixe (8 octets) et une zone de données.

IV.4.3 Protocole IP (Le paquet IP) [67] [68] [69]

Le protocole IP a été conçu pour réaliser l'interconnexion de réseaux et permettre ainsi les communications entre systèmes. Il fait partie de la couche Internet de la suite des protocoles TCP/IP. En réalité, le segment une fois encapsulé dans la couche Internet prend le nom de datagramme. Le protocole IP traite les datagrammes IP indépendamment les uns des autres.

Chaque paquet de chaque réseau transporte donc en son sein un paquet commun, le paquet IP (Internet Protocol). Cette structure est illustrée à la figure IV.3, En fait, comme la taille du paquet IP est fortement variable, elle peut ne pas être adaptée, en étant, par exemple, trop grande. Dans ces circonstances, on commence par découper le paquet IP en fragments, et l'on encapsule chaque fragment dans un paquet spécifique.

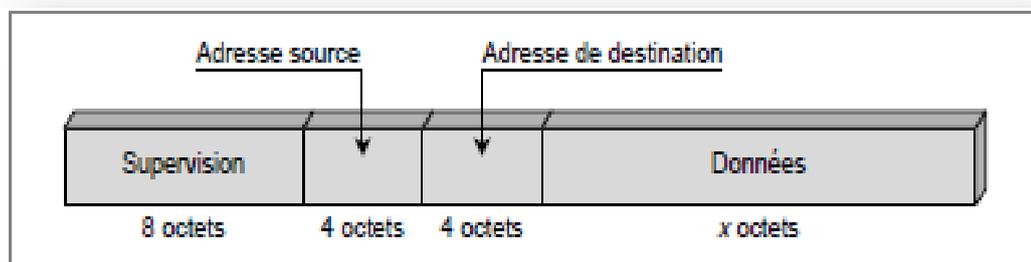


Figure IV.3 Le format du Paquet IP

Les nœuds intermédiaires où transitent les paquets s'appellent des *routeurs*. Ce système s'adapte parfaitement aux applications informatiques, qui acceptent en général des contraintes temporelles moins sévères. En revanche, la qualité nécessaire pour l'acheminement de la vidéo ou de la parole téléphonique est le plus souvent impossible à obtenir dans l'état actuel de la technologie. L'une des raisons à cela tient au traitement « premier arrivé-premier servi » des paquets IP dans les nœuds de transfert : un petit paquet urgent qui se trouve derrière un gros paquet non urgent est obligé d'attendre. De plus, Internet, en tant que réseau de réseaux, ne possède pas d'administrateur ayant une vision globale ni d'opérateur unique ayant une connaissance complète du réseau et étant capable de le gérer ou d'adapter ses infrastructures en fonction du nombre d'utilisateurs.

L'intelligence permettant de contrôler les paquets qui transitent dans le réseau se trouve dans les équipements terminaux. Ces éléments situés aux extrémités du réseau doivent adapter leurs applications à l'état du réseau par une compression plus ou moins importante [67].

Les données circulent sur Internet sous forme de datagrammes (nous parlons aussi de paquets). Les datagrammes sont des données encapsulées, c'est-à-dire des données auxquelles ils ont ajouté des en-têtes correspondant à des informations sur leur transport (telles que l'adresse IP de destination,...) [69].

Les données contenues dans les datagrammes sont analysées (et éventuellement modifiées) par les routeurs permettant leur transit.

IV.4.4 Adressage IP [67][69]

Chaque équipement impliqué dans une communication doit posséder une adresse IP unique, codée sur 4 octets. Cette adresse est représentée par quatre chiffres décimaux séparés par un point.

L'adresse IP : 10101100.00010000.11111110.00000001 s'écrira 172.16.254.1

Types d'adresses IP

Adresse "unicast": permet d'identifier un équipement IP de façon unique.

Adresse "multicast": adresse de diffusion vers un groupe d'équipements IP.

Adresse "broadcast": adresse de diffusion vers toutes les adresses IP d'un même sous réseau.

Les classes d'adresses

À l'origine, plusieurs groupes d'adresses ont été définis avec l'objectif d'optimiser le cheminement (ou le routage) des paquets entre les différents réseaux. Ces groupes ont été baptisés classes d'adresses IP. Ces classes correspondent à des regroupements en réseaux de même taille. Les réseaux de la même classe ont le même nombre d'hôtes maximum.

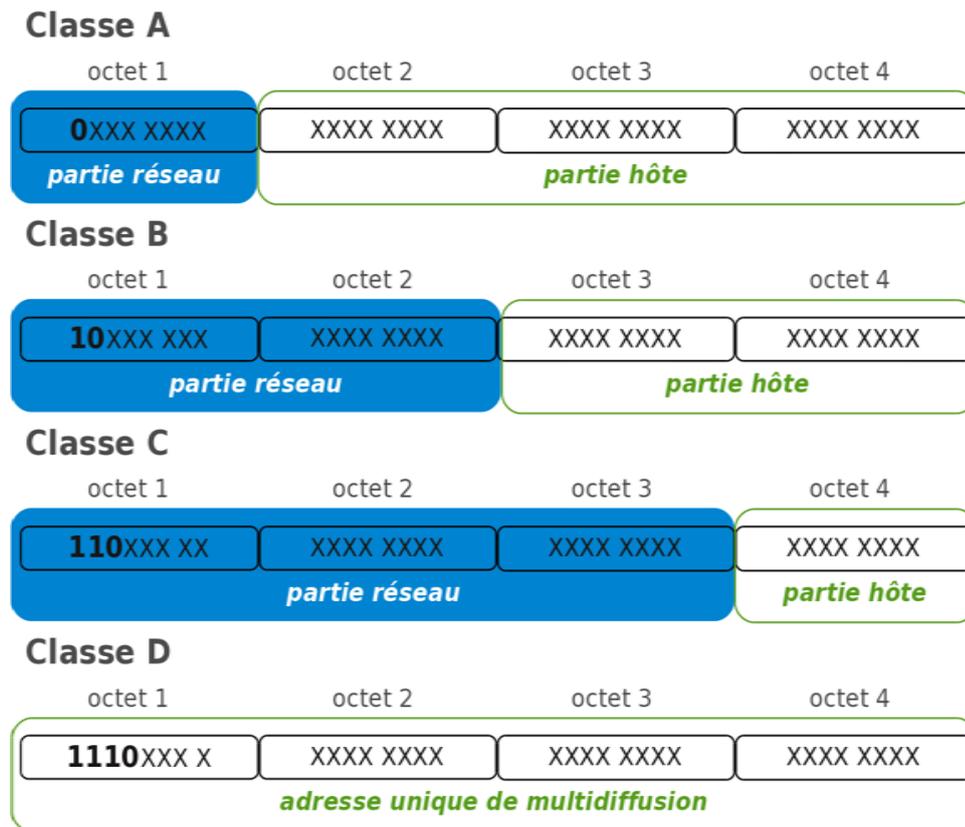


Figure IV.4 L'adressage IP

Classe A

Le premier octet a une valeur comprise entre 1 et 126 ; soit un bit de poids fort égal à 0. Ce premier octet désigne le numéro de réseau et les 3 autres correspondent à l'adresse de l'hôte. L'adresse réseau 127.0.0.0 est réservée pour les communications en boucle locale.

Les adresses de classe A sont attribuées aux réseaux comportant un nombre élevé d'hôtes.

Classe B

Le premier octet a une valeur comprise entre 128 et 191 ; soit 2 bits de poids fort égaux à 10. Les 2 premiers octets désignent le numéro de réseau et les 2 autres correspondent à l'adresse de l'hôte.

Les adresses de classe B sont attribuées à des réseaux de taille moyenne à grande.

Classe C

Le premier octet a une valeur comprise entre 192 et 223 ; soit 3 bits de poids fort égaux à 110. Les 3 premiers octets désignent le numéro de réseau et le dernier correspond à l'adresse de l'hôte.

Les adresses de classe C sont généralement employées pour de petits réseaux locaux.

Classe D

Le premier octet a une valeur comprise entre 224 et 239 ; soit 3 bits de poids fort égaux à 111. Il s'agit d'une zone d'adresses dédiées aux services de multidiffusion vers des groupes d'hôtes (*host groups*).

Classe E

Le premier octet a une valeur comprise entre 240 et 255. Il s'agit d'une zone d'adresses réservées aux expérimentations. Ces adresses ne doivent pas être utilisées pour adresser des hôtes ou des groupes d'hôtes.

IV.5 Modes de transmission

IV.5.1 Mode Unicast

C'est le principe le plus utilisé et le plus simple. Les ordinateurs possédant chacun une adresse IP, nous pouvons envoyer les trames en spécifiant l'adresse IP de l'ordinateur à qui nous voulons envoyer les informations. Les éléments actifs et passifs du réseau (commutateurs, répéteurs, routeurs,...) dirigent l'information dans la bonne direction pour que les trames arrivent au bon endroit. Seule la machine ayant l'adresse contenue dans la trame regarde et traite l'information [67].

IV.5.2 Mode Broadcast

Le principe du broadcast est d'envoyer une information à tous les ordinateurs du réseau où nous l'avons. Au lieu d'envoyer en unicast vers l'adresse IP de chaque machine, nous envoyons la trame à tous les ordinateurs du sous-réseau en utilisant l'adresse de broadcast. Cette adresse est réservée à cet usage. L'ensemble des ordinateurs du sous-réseau traite la trame comme si elle leur était personnellement envoyée. Les trames de broadcast ont une particularité précise: c'est de ne pas pouvoir passer les routeurs car il s'adresse uniquement à la totalité des ordinateurs d'un même sous-réseau.

IV.5.3 Mode Multicast

Plutôt que d'envoyer les fichiers du serveur vers chacune des machines clientes (unicast) nous pouvons n'envoyer l'information qu'une seule fois et chaque ordinateur client la récupère. En effet, au sein d'un réseau Ethernet par exemple, toutes les trames qui circulent traversent tous les ordinateurs. C'est le principe du multicast : nous envoyons l'information à une adresse et tous les clients écoutent cette adresse.

Chaque client multicast s'enregistre avec une adresse IP multicast de classe D (entre 224.0.0.0 et 239.255.255.255 sauf 224.0.0.0 non utilisée et 224.0.0.1 qui correspond au

"broadcast du multicast"). C'est sur cette adresse que les informations vont être envoyées. Les clients écoutent ce qui arrive sur cette adresse et suivent la procédure décrite par le protocole multicast implémenté.

IV.6 Les environnements mobiles et sans fils

Un système, composé des sites mobiles et qui fournit à ses utilisateurs la possibilité de l'accès à des informations sans aucunes exigences de leurs positions géographiques, présente l'environnement mobile. Les réseaux sans fils proposent deux architectures différentes, réseau avec infrastructure fixe et réseau sans infrastructure fixe dues à leurs façons de communication entre les sites mobiles.

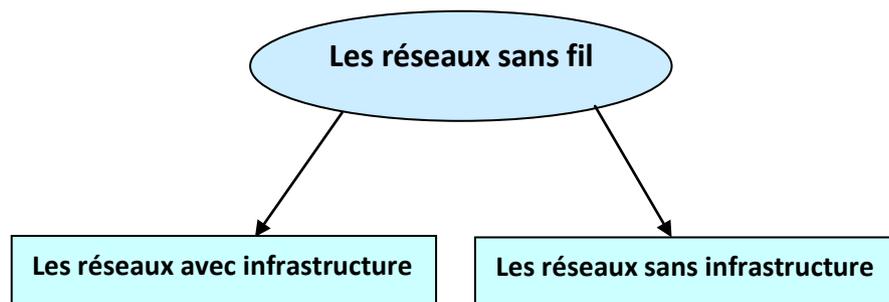


Figure IV.5 La décomposition des réseaux sans fil

IV.6.1 Les réseaux avec infrastructure

En mode avec infrastructure, toute transmission de données doit passer par un point d'accès (AP: Access Point). Cette entité particulière joue le rôle de routeur à l'intérieur du réseau sans fil et souvent de passerelle vers un réseau filaire.

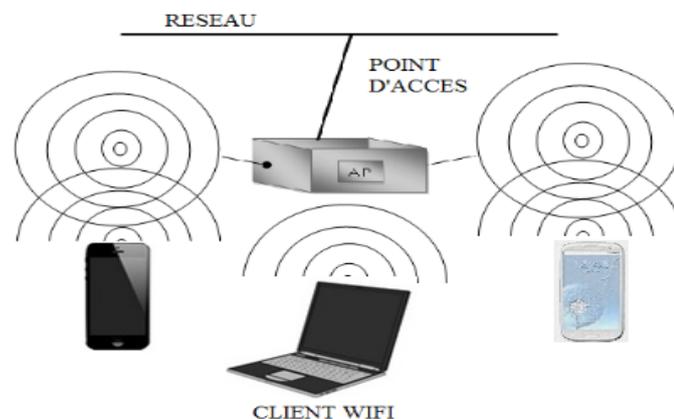


Figure IV.6 Le modèle des réseaux sans fil avec infrastructure [70]

Chaque point d'accès administre alors une zone géographique et garantit la liaison avec différents zones, avec un réseau filaire ou avec internet. Les réseaux cellulaires mobiles

peuvent être pris en compte comme étant une forme spécifique de réseau avec point d'accès. Les liaisons sans fils se caractérisent par une bande passante limitée qui réduit brutalement le volume des informations échangées.

IV.6.2. Les réseaux sans infrastructure

Dans ce mode de communication, chaque nœud (appelé aussi station) du réseau a la possibilité de communiquer directement avec tous ses voisins, c'est à dire tous les nœuds capables de recevoir le signal envoyé et de le comprendre. Chaque nœud peut se connecter, se déplacer ou se déconnecter du réseau à tout moment. Le réseau mobile sans infrastructure ne comporte pas la notion « site fixe », tous les sites sont mobiles et se communiquent d'une manière directe en utilisant leurs interfaces de communication sans fil. Il n'y a aucune infrastructure. Ce manque d'infrastructure filaire ou de réseau câblé oblige les unités mobiles à se comporter comme des routeurs impliqués dans la découverte et la maintenance des chemins pour les autres hôtes du réseau. On parle alors des réseaux ad hoc ou MANET.

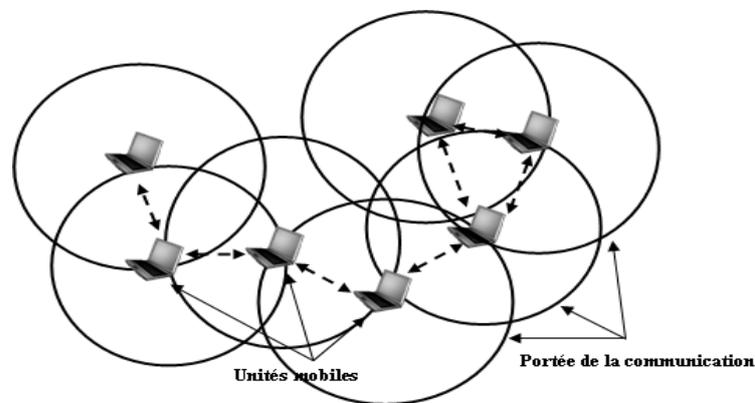


Figure IV.7 Le modèle des réseaux sans fil sans infrastructure [70]

IV.7 Les réseaux mobiles ad hoc

IV.7.1. Bref historique

Une définition des réseaux ad hoc est donnée formellement dans la RFC informatif 2501 (RFC : requests for comments) [71][72]: « Un réseau ad hoc comprend des plates-formes mobiles (par exemple, un routeur interconnectant différents hôtes et équipements sans fil) appelées nœuds qui sont libres de se déplacer sans contrainte. Un réseau ad hoc est donc un système autonome de nœuds mobiles. Ce système peut fonctionner d'une manière isolée ou s'interfacer à des réseaux fixes au travers de passerelles». Le projet militaire Américain DARPA (*Defense Advanced Research Projects Agency*)[70], qui a eu lieu au début des années

1970 a évoqué la naissance des premiers réseaux utilisant le médium radio. Ces réseaux sont définis par deux composantes :

- la disposition d'une architecture distribuée,
- le partage d'un canal de diffusion en répétant des paquets pour élargir la zone de couverture globale.

Dans le même axe des applications militaires, dans les années 1983, les Survivable Radio Networks (SURAN) furent développés par le DARPA. Leur objectif était de dépasser les limitations. Autrement dit, permettant le passage à des réseaux comportant énormément de nœuds, gérant le domaine de la sécurité et l'énergie.

L'arrivée du protocole 802.11 de l'IEEE (Institute of Electrical and Electronics Engineers) était le point de départ des réseaux sans fils autour des bases fixes, et qui a permis l'apparition des problématiques liées à ces réseaux par la recherche civile dans les années 90.

IV.7.2 Les caractéristiques des réseaux ad hoc

Le réseau ad hoc présente beaucoup de simplicité et plusieurs avantages face aux autres réseaux (filaire et cellulaires) en raison de sa simplicité de déploiement en cas d'urgence ou de travaux temporaires alors que d'autres réseaux génèrent des coûts importants. Cependant de nouveaux problèmes apparaissent. Dans la plupart des cas, le nœud destination ne se trouve pas obligatoirement dans la portée du nœud source ce qui implique que l'échange des données entre les deux nœuds, doit être effectué par des stations intermédiaires.

Parmi les principales caractéristiques des réseaux ad hoc, nous exposons :

IV.7.2.1 Une topologie dynamique

Comme nous l'avons mentionné, les unités mobiles du réseau, se déplacent et se communiquent d'une façon libre et arbitraire. Par conséquent la topologie du réseau est dynamique. Autrement dit, elle peut changer, à des instants imprévisibles, d'une manière rapide et aléatoire aussi. Il est utile de mentionner que les liens de la topologie des réseaux ad hoc peuvent être unis ou bidirectionnels.

IV.7.2.2 Une bande passante limitée

Les réseaux basés sur la communication sans fil sont caractérisés par l'utilisation d'un médium de communication partagé. Ce partage fait que la bande passante réservée à un hôte soit modeste.

IV.7.2.3 Les contraintes d'énergie

Les équipements mobiles ont des batteries limitées, et dans certains cas très limitées, telles que les PDAs (Personal Digital Assistant), et par conséquent d'un temps de traitement réduite. Sachant qu'une partie de l'énergie est déjà exploitée par la fonctionnalité de routage, cela limite les services et les applications supportées par chaque nœud.

IV.7.2.4 L'hétérogénéité des nœuds

Un nœud mobile peut être équipé d'une ou plusieurs interfaces radio ayant des capacités de transmission variées et fonctionnant dans différentes plages de fréquences. Cette hétérogénéité de capacité peut conduire à des liens asymétriques dans le réseau. De plus, les nœuds peuvent avoir des différences en terme de capacité de traitement de logiciel et de mobilité (lent, rapide). Dans ce cas, une adaptation dynamique des protocoles s'avère nécessaire pour supporter de telles situations.

IV.8 Réseaux locaux sans fil (WLAN) IEEE 802.11

Le réseau local sans fil (WLAN Wireless Local Area Network) est un réseau couvrant l'équivalent d'un réseau local (LAN) d'une société, avec une portée d'approximativement des dizaines de mètres. Il offre la possibilité de relier les terminaux présents dans la zone de couverture.

WIFI : (ou IEEE 802.11)

WIFI (ou IEEE 802.11) est un standard international exposant les caractéristiques d'un réseau local sans fil WLAN, elle est confortée par l'alliance WECA (Wireless Ethernet Compatibility Alliance).

Avec le Wi-Fi, il est possible de créer des réseaux locaux sans fil haut débit tant que l'ordinateur à connecter n'est pas trop éloigné du point d'accès. En pratique, le WiFi connecte des ordinateurs portables, ordinateurs de bureau, PDA ou tout type d'appareil avec une liaison haut débit (11Mbps ou plus) sur un rayon de plusieurs dizaines de mètres à l'intérieur (généralement entre 20 et 50 mètres) à quelques centaines de mètres dans un environnement ouvert.[70]

IV.9 Qualité de Service (QoS) [6][7][11]

Au début de l'apparition de l'Internet la préoccupation majeure était de transmettre les paquets à leurs destinations sans donner des garanties aux utilisateurs. En effet, l'Internet existant, fondé sur le principe du Best Effort, n'a pas été initialement conçu pour prendre en compte les informations de QoS. Le modèle Best Effort maximise l'utilisation de ressources

tout en simplifiant l'opération des équipements d'interconnexion. Au moment où la mémoire d'un routeur est surchargée, les paquets arrivants sont rejetés. Les mécanismes de retransmission de TCP couvrent ces pertes, assurant un transfert sans erreur de bout en bout.

L'application de vidéo sur IP est pour la plus part basée sur le protocole UDP. Ce protocole peut être utilisé pour obtenir un débit binaire presque constant, mais n'implique pas un mécanisme de gestion de la congestion. Une mise à disposition contrôlée de la bande passante permettrait aux fournisseurs de définir de nouveaux services qui pourraient être facturés en conséquence. S'il devient possible dans l'Internet d'assurer le débit ou le délai observés, les opérateurs seront en position d'implanter une politique tarifaire conforme aux ressources accordées.

L'évolution d'Internet pour prendre en compte les besoins des nouveaux flux et pour contrôler la distribution de ressources se fera en modifiant les mécanismes de file d'attente dans les routeurs plutôt qu'en modifiant TCP. En effet, les méthodes utilisées dans l'Internet jusqu'à aujourd'hui, consistant à fournir des réseaux surdimensionnés, ne peuvent plus s'appliquer indéfiniment. De plus la nature intrinsèque de l'Internet (mode sans connexion, niveau de service Best-effort) ne permet ni d'offrir une QoS constante, ni de donner des priorités à certains types de trafic. C'est pourquoi les architectes du réseau, les constructeurs et les fournisseurs de service concentrent aujourd'hui leurs efforts sur la définition et l'implémentation de nouvelles approches et architecture de QoS.

Généralement, cette qualité est axée sur le débit, le délai et la perte des paquets : la vidéo par IP a pour objectif de réussir à converser en temps réel (facteurs QoS), télécharger une application volumineuse ne nécessite pas plus que d'avoir une bande passante suffisamment importante pour récupérer le fichier le plus rapidement possible (facteur de flux); les deux applications sont exigeantes en termes de réception de l'ensemble des paquets (facteur de perte).

Critère de QoS [4][6]

La caractérisation de la QoE dans les réseaux est généralement exprimée par les critères suivants:

➤ Le délai (Delay):

C'est le temps écoulé entre l'envoi d'un paquet par l'émetteur et sa réception par le destinataire. Le délai prend en compte deux genres de délai :

- Le délai de propagation : le long du chemin parcouru par les paquets.
- Le délai de transmission, induit par la mise en file d'attente des paquets dans les systèmes intermédiaires (les routeurs). La plupart des applications et surtout les applications temps réel sont très sensibles aux valeurs élevées de délais et exigent un délai limité pour qu'elles puissent fonctionner correctement.

Le terme "délai" comprend en fait trois aspects temporels différents:

Le délai de propagation, induit par la distance physique qui sépare la source de la destination.

Le délai de transmission dépendant de la taille des flux. Ce paramètre est aussi étroitement lié à l'utilisation du réseau et au partage de la bande passante disponible.

Enfin, le délai d'attente et de traitement des paquets à l'intérieur des files d'attente, déterminé par la charge du réseau, et les stratégies de traitement de l'information dans les routeurs afin d'avoir une fluidité optimale de l'écoulement de l'information..

➤ **La Gigue (Jitter):**

La gigue est la variation de délai de bout en bout. Plusieurs applications (audio et vidéo) subissent une distorsion du signal avec des valeurs de gigue élevées. Des délais relativement importants éventuellement substitués par les traitements lents des routeurs nuisent automatiquement à la QoE par ce paramètre: des variations de délais apparaîtront et affecteront la qualité demandée. Des taux d'erreurs minimes, voire nuls caractérisent un certain rendement et paramètrent une bonne QoS en matière de disponibilité du réseau.

➤ **Le débit :**

C'est en fait le taux de transfert maximum pouvant être maintenu entre deux points terminaux (un émetteur et un récepteur). Ce facteur est influencé non seulement par les capacités physiques des liens, mais aussi par les autres flux partageant ces liens. Les applications d'aujourd'hui consomment de plus en plus de bande passante, ralentissant ou bloquant d'autres applications. De même, une utilisation massive du réseau conduit à ralentir le flux. La notion de bande passante d'un réseau intervient à ce niveau: une bande passante minimale est exigée afin de maintenir des garanties de QoS point à point, nécessaires à intervalles différents.

➤ **La fiabilité :**

C'est le taux moyen d'erreur d'une liaison, actuellement les supports de transmission deviennent de plus en plus fiables et ce taux devient de plus en plus négligeable par rapport aux pertes dues aux situations de congestion des réseaux.

➤ **Le taux de perte :**

Ce facteur est très important, les pertes sont généralement dues soit à la non fiabilité des liens physiques soit aux mécanismes de gestion des files d'attente dans les routeurs intermédiaires qui se trouvent toujours contraints de jeter les paquets pour faire face aux situations de congestion. En effet, certaines applications ne tolèrent pas des valeurs élevées de ce facteur.

IV.10 Qualité d'Expérience (QoE) de la télévision / IPTV

Dans le contexte de la télévision ou encore de la Télévision IP, la mesure de la qualité utilisateur entend s'affranchir des composantes techniques (marque, modèle de la TV, définition) et du type de contenu.

Les constructeurs de solutions de mesure de QoE proposent généralement une métrique «agrégée» relative à la perception de l'utilisateur appelée MOS (Mean Opinion Score) [73]. Parmi les facteurs influençant cette métrique pour la télévision on peut notamment trouver: Le "Block" (présence de carré à l'image), "Blur" (netteté de l'image), le "Jerkiness" (saccade et gel d'image) ainsi que le temps de zapping [74] [75]. Les facteurs "Block", "Blur" et "Jerkiness" sont aussi appelés artéfacts.

La mesure des paramètres de Qualité d'Expérience de l'IPTV a aussi fait l'objet d'une recommandation de l'ITU-T G.1080 [74].

En plus de la qualité d'image, les systèmes de mesure de la QoE [75] mettent en avant le confort d'utilisation d'un service. Par exemple la réactivité lors de la navigation dans les menus d'un portail de (VoD) Vidéo à la demande.

IV.11 Utilisation de la Qualité d'Expérience

Les outils QoE sont généralement utilisés en parallèle avec les outils QoS [76]. Ils permettent de mettre en avant une dégradation du service du point de vue de l'utilisateur et donc de déclencher des analyses dans le but de conduire à des actions correctives. De même, ils permettent aux ingénieurs (services TV, voix ...) ou aux responsables techniques ou métiers (business applications) de pré-orienter leurs recherches et ainsi de réduire le temps de résolution d'un problème.

La QoE propose une approche de bout en bout, soit en utilisant l'évaluation du service par une population d'utilisateurs, soit en simulant le comportement des utilisateurs en utilisant des outils spécifiquement élaborés pour les besoins des utilisateurs.

IV.12 Qualité de service et réseau ad Hoc

Les techniques mises en œuvre pour garantir cette QoS peuvent concerner tous les aspects de la transmission d'information, notamment l'accès au média, la réservation de route, la gestion des files d'attente,...etc. L'apparition de cet axe de recherche coïncide avec une nouvelle utilisation des réseaux informatiques. L'ancienne approche, dite best effort notamment employée par IP ne permet pas de garantir les contraintes attendues par les nouveaux services multimédias. Les systèmes à QoS participent généralement à lutter contre le caractère erratique du flot des paquets échangés. Une des causes est la congestion du réseau (saturation de la bande passante par une quantité d'information trop importante au même moment) qui aboutit au ralentissement ou à l'arrêt de certains flux en des nœuds intermédiaires. Ceci entraîne de fait une perte de qualité à la réception. Le principal mécanisme utilisé consiste à adapter le fonctionnement du réseau à la nature des flux qu'il doit traiter. Cela implique donc d'une part qu'il soit possible de caractériser les flux de données en transit, d'autre part que les informations relatives à cette classification soit accessibles par les nœuds routeurs. On procède par exemple, à la réservation de débit et de mémoire dans les files d'attente des routeurs pour privilégier certains flux ayant des contraintes temporelles importantes. Dans certaines références, des modèles de QoS ont été développés spécialement pour les réseaux ad hoc. [72][77].

IV.13 Conclusion

Les réseaux locaux sans fil, et le Wi-Fi en particulier, sont des technologies intéressantes et largement utilisées dans plusieurs domaines tels que l'industrie, la santé et le domaine militaire.

Mais la qualité dans ce domaine reste un thème très délicat, puisque depuis l'usage de ce genre de réseau plusieurs défauts ont été repérés. Lors du déploiement d'un réseau sans fil, le Wi-Fi (802.11) paraît être le remède qui apporte la meilleure solution pour répondre aux besoins des réseaux locaux sans fil en raison de l'avantage qu'il procure, à savoir son interopérabilité avec les réseaux sans fil.

En effet, Cette technologie est fréquemment employée dans les sociétés souhaitant héberger des utilisateurs mobiles ou désirant une alternative au réseau câblé tout en conservant des performances presque similaires.

Chapitre V

Analyse des performances et résultats

V.1 Introduction

L'amélioration de la transmission de la vidéo sur un réseau IP revient à trouver des solutions techniques aux problèmes qui affectent la QoS comme le délai, la gigue, la variation du retard, la perte de paquets. L'objectif final est de pouvoir offrir une qualité de service satisfaisante dans la plupart des conditions du réseau. D'une manière générale, au niveau de l'émetteur, la vidéo brute est recueillie et codée avant d'être transmise à travers les réseaux IP. Du côté du récepteur, le flux de données reçu est décodé et récupéré sous forme d'un signal vidéo reconstruit.

Par ailleurs, les transmissions de contenus multimédias (données audio et vidéo) sont de plus en plus demandées, surtout à travers les réseaux sans fil. Le développement des applications peut générer plusieurs types de codecs dont l'intérêt touchera la plupart des applications grand public, ce qui est un objectif très intéressant.

Dans ce chapitre, après une présentation général d'une transmission numérique, Nous démontrons quelques notions essentielles de la norme de compression HEVC /H.265 pour la vidéo numérique. Ensuite, nous proposons une nouvelle technique universelle utilisée pour le codage de source et leur application dans la transmission de la norme vidéo HEVC /H.265 par un réseau IP.

V.2 Fonctionnement général d'une transmission:

Plusieurs traitements pour la vidéo sont requis avant l'émission et après la réception. Nous allons survoler rapidement les principales étapes qui sont génériques indépendamment du protocole utilisé.

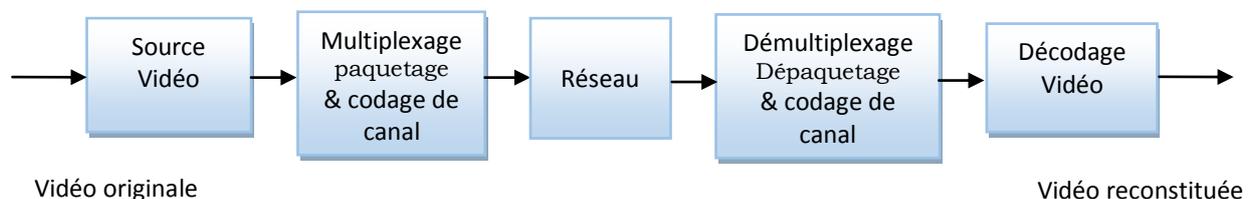


Figure V.1 Un système de communication vidéo typique



Figure V.2 Un système de communication vidéo simplifié

V.2.1 La phase d'émission

V.2.1.1 L'acquisition

Il s'agit en fait de capter l'information (signal analogique) de telle manière qu'elle puisse être transmise sur un réseau numérique comme Internet, réseau sans fil...etc.

V.2.1.2 Le codage et la compression

Il existe plusieurs types de codage dont le codage canal et le codage source. L'objectif du codage canal est de s'assurer de l'intégrité du signal reçu et éventuellement le corriger. Alors que le codage source comprime les données numériques selon un schéma de compression pour diminuer le débit émis, et favoriser ainsi le transfert de données en temps réel. Il y a un grand nombre de genres de compressions au niveau applicatif.

Ces différents codages de l'information influent principalement sur trois facteurs primordiaux dans une communication:

- Le débit en émission.
- La charge processeur pour coder/décoder l'information.
- Le niveau de qualités du média.

La plupart des algorithmes de compression favorise une caractéristique au détriment d'une autre.

V.2.1.3 La transmission des paquets

Le module de transmission assure la transformation de flot de bits à la sortie du codeur vidéo en un flot de paquets (ou datagrammes) [78]. Cette transformation peut s'accomplir aux travers de différents protocoles que nous avons étudiés dans le chapitre 04.

V.2.2 La phase de réception

V.2.2.1 La réception des paquets

Le module ou protocole de réception prend les paquets disponibles sur le réseau et place l'information contenue dans ces paquets à l'intérieur d'un buffer de réception. Les paquets peuvent provenir de différentes sources.

V.2.2.2 La phase décodage/décompression de l'information

Ce module utilise l'un des formats de décompression disponibles, pour restituer une information non codée, en fonction de la compression originale, la qualité obtenue en sortie se verra pour la plupart du temps plus ou moins diminuée.

V.3 Méthodes d'évaluation

Les méthodologies utilisées pour l'évaluation de la qualité vidéo sont généralement divisées en quatre étapes [19], comme le montre la figure V.3.

1. Sélection de vidéos de références.
2. Création de vidéos endommagées.
3. Évaluation des vidéos endommagées par les utilisateurs.
4. Évaluation statistique des vidéos endommagées.

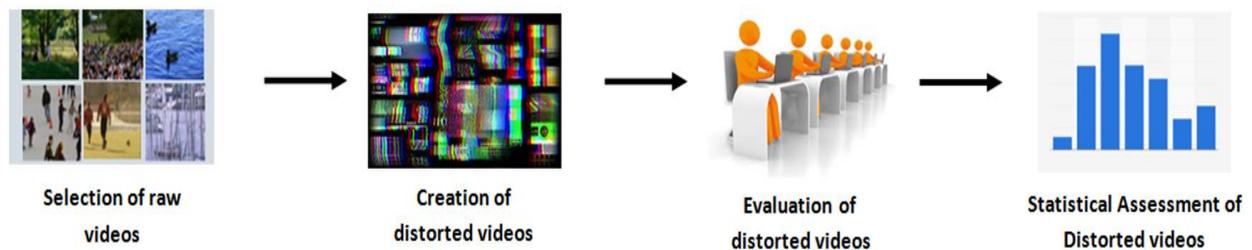


Figure V.3 Diagramme d'évaluation typique de la vidéo.

Les vidéos de références ne sont pas codées par aucun type de compression. Les résolutions spatiales couramment utilisées vont du CIF (352 x 288) à la HD (1280 x 720 et 1920 x 1080), récemment, à la Ultra HD (3840 x 2160) jusqu'à 8K.

Alors selon l'application utilisée, nous avons deux types de transmission vidéo :

- **Individuelle :** un face à face entre personnes à l'aide d'ordinateurs personnels. Elle est dédiée aux conférences personnelles.
- **Collective :** Une personne ou plusieurs personnes faisant face à un groupe de personnes via un système dédié à la visioconférence. Des téléviseurs peuvent être utilisés dans ce cas (téléconférence). Elle est dédiée aux conférences de groupes.

Les métriques d'évaluation de la qualité vidéo ont été classées en deux catégories : méthodes objectives et subjectives [79]. Il est à noter que les méthodes d'évaluation subjectives sont très coûteuses en temps et en effort par rapport aux autres méthodes. D'un autre côté, les méthodes objectives sont préférées car elles offrent des résultats directement utilisables pour une vérification et une comparaison immédiates [19] [80].

Une vidéo de référence est comparée image par image avec la vidéo déformée pour obtenir une mesure de la qualité vidéo qui prend en compte plusieurs aspects. Ces aspects comprennent le traitement de la couleur, le masquage spatio-temporel, le contraste perçu et l'adaptation à une luminance, une couleur spécifique et/ou à la sensibilité au contraste [5].

Dans ce travail de thèse, lors du traitement des vidéos, nous avons principalement utilisé deux types de métriques objectives, qui sont respectivement le PSNR et le SSIM. Ainsi, le PSNR est une évaluation objective de la qualité la plus utilisée ; il effectue une comparaison pixel par pixel entre la référence et le contenu déformé, ce dernier pouvant être presque mappé sur MOS en utilisant le tableau V.1 [18] ou l'équation (6)[81].

PSNR(dB)	MOS
> 37	Excellent
31-37	Good
25-31	Fair
20-25	Poor
< 20	Bad

Tableau V. 1 Conversion PSNR en MOS

PSNR est défini en utilisant la formule de l'erreur quadratique moyenne (MSE) comme suit:

$$MSE = \frac{1}{m \times n} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - k(i, j)]^2, \quad (3)$$

$$PSNR = 10 \log_{10} \left(\frac{MAX^2}{MSE} \right), \quad [dB] \quad (4)$$

Où MAX est la valeur maximale du pixel, ce qui peut prendre par exemple 255 pour l'image 8 bits. Ce critère fournit l'erreur entre la vidéo originale et la vidéo reconstruite; un PSNR plus élevé indique généralement une meilleure qualité.

En outre, SSIM représente la mesure de la perte d'information structurelle. Il peut fournir une bonne approximation de la distorsion de la vidéo perçue [18] et donne une meilleure indication de la qualité de l'image.

L'approche SSIM repose sur l'idée que les pixels spatialement proches ont de fortes interdépendances, reflétant la perception humaine [23][82]. La méthode SSIM considère la perception de l'image par l'œil humain; il mesure la luminance, le contraste et la similarité structurelle et les combine en une valeur qui détermine l'intervalle de qualité vidéo [0, 1],

où 0 est le pire et 1 est la meilleure qualité. L'expression du SSIM est une combinaison de trois paramètres, le signal d'origine x , et le signal codé y étant défini comme suit:

$$\text{SSIM}(x, y) = [I(x, y)]^\alpha [C(x, y)]^\beta [S(x, y)]^\gamma, \quad (5)$$

- $I(x, y)$ compare la luminance du signal.
- $C(x, y)$ compare le contraste du signal.
- $S(x, y)$ mesure la similarité structurelle.

$\alpha > 0, \beta > 0, \gamma > 0$ mesure le poids des éléments individuels.

Méthodes	Paramètres	Type de métriques	Ce qui est analysé
PSNR [83]	Full-référence (comparé image par image)	Donnée	Tous les pixels
SSIM [84]	Full-référence (comparé image par image)	Picture	Variations des modèles locaux d'intensités de pixels

Tableau V.2 Analyse comparative des approches objectives

L'approche subjective produit une évaluation de la qualité vidéo basée sur la perception d'observateurs humains. En fait, les expériences sont généralement effectuées dans des laboratoires où les séquences vidéo sont affichées à la télévision [85] [86] [87].

Le MOS est une métrique couramment utilisée et par laquelle un observateur évalue chaque vidéo en sélectionnant un score d'une échelle de qualité allant de 1 à 5 (Tableau V.3) [18], [88]. Une autre option est l'échelle de dégradation (tableau V.4) [18].

Quality scale		Degradation scale	
5	Excellent	5	Imperceptible
4	Good	4	Perceptible, but not annoying
3	Fair	3	Slightly annoying
2	Poor	2	Annoying
1	Bad	1	Very annoying

Tableau V.3 Échelle de qualité

Tableau V.4 Échelle de dégradation

Où l'on demande à l'observateur de rapporter le niveau de dégradation de la vidéo déformée en la comparant à la vidéo de référence précédemment montrée.

Une conversion du PSNR, qui est une mesure objective, en échelle MOS a été proposée dans la Réf. [68], en utilisant une fonction d'approximation appropriée:

$$\text{PSNR}_{\text{MOS}} = 0.0935 \text{ PSNR} + 0.152, \quad (6)$$

PSNR_{MOS} ou $\text{MOS}_{\text{Estimé}}$ est une estimation des scores MOS de PSNR mesurés.

L'outil de mesure de qualité vidéo MSU [89] est servi pour comparer la vidéo perçue avec l'originale, de sorte que l'impact de l'utilisation du réseau puisse être analysé. La mise en œuvre du modèle de calcul a été effectuée dans le langage de programmation C.

V.4 Visioconférence entre deux participants

Il suffit que leurs postes de visioconférence se connectent directement à travers le réseau :

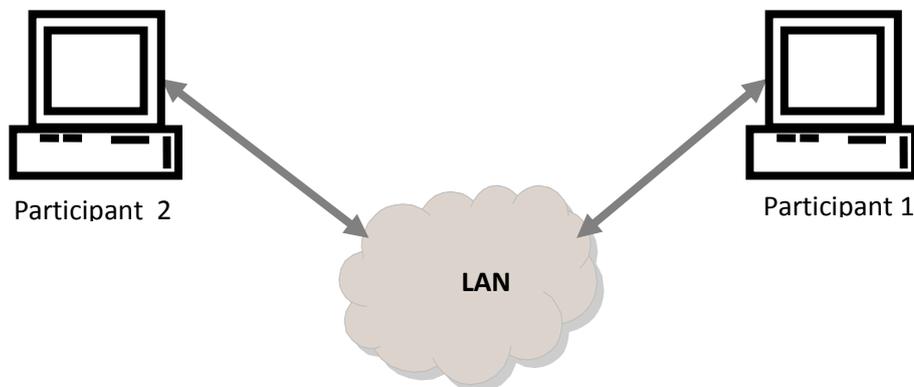


Figure V.4 Réseau point à point

L'émulation permet la communication en temps réel entre deux ordinateurs ou un groupe d'ordinateurs donc rapprochement de ressources via les réseaux de paquets. Pour réaliser une émulation en poste de travail il est nécessaire de disposer d'un logiciel de communication qui doit être installé sur chaque poste. Les logiciels proposés à l'heure actuelle sont peu performants.

V.4.1 Les logiciels de vidéo Player

Un certain nombre de logiciels de communication ont été conçus pour offrir les services de transmission vidéo. Ils offrent des services différents qui se caractérisent par le nombre de codec possibles, protocole utilisé, le débit de transmission, la qualité de l'image et la qualité du son. Certains logiciels sont offerts gratuitement (téléchargeable sur Internet) alors que d'autres qui sont plus performants sont souvent payants.

Pour notre application, nous avons utilisé le logiciel de communication suivant:

V.4.2 VLC media player

VLC à l'origine VideoLAN Client est un lecteur multimédia disponible tiré du programme VideoLAN4. Cette application est multiplateforme, étant donné qu'il marche sous windows, Linux..., GNU/Linux, BSD, OS X, iOS, BeOS, Solaris, Android5, QNX et Pocket Micro soit au total près de 20 plateformes. Il est distribué sous licence GNU GPL.

Un des grands avantages de VLC est qu'il inclut les codecs nécessaires à la lecture de la majorité des formats audio et vidéo. De plus, le lecteur est capable de lire une grosse quantité de flux réseaux. Il se montre par ailleurs très tolérant avec les flux légèrement endommagés, allant jusqu'à les réparer de son mieux.

VLC est l'un des principaux lecteurs multimédias, utilisé par des millions de personnes [4].

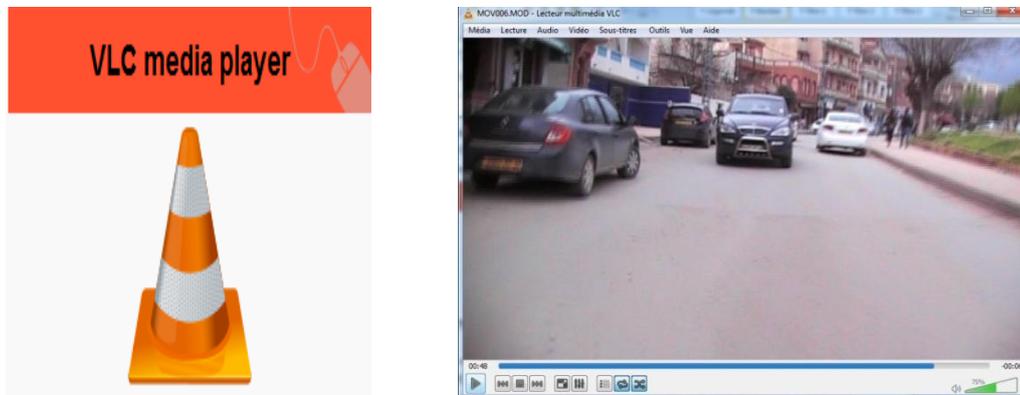


Figure V.5 VLC media player

Ce logiciel est basé sur l'adressage IP unicast, il offre une possibilité de communication point à point entre deux postes. On a utilisé ce logiciel pour une vidéoconférence entre deux postes dans le réseau de l'émulation. Ce logiciel possède un codec vidéo HEVC/H.265.

V.5 Bases de données vidéo

Pour mesurer la qualité vidéo plusieurs bases de données ont été proposées. La disponibilité de telles bases de données pour la comparaison des performances, le développement et le test de nouveaux modèles est d'une importance cruciale. Quelques bases de données sont actuellement disponibles et sont parmi les plus populaires dans le domaine de l'évaluation de la qualité. Les bases de données sur l'état l'art sont limitées au taux de perte de paquets, aux artefacts de codage et de compression seulement. Dans cette thèse, notre objectif est d'utiliser la base de données vidéo ReTRiEVED [4] [5], pour évaluer les performances des métriques de qualité vidéo en tenant compte des paramètres de transmission de QoS [6]. Différents groupes de recherche rendent les vidéos brutes disponibles sur leurs sites Web, comme le montre par exemple le tableau V.5.

Base de données	Institut	Résolution Spatial (Pixels)	Frame rate (fps)	Nombre totale des vidéos originales
EPFL-PoliMI video quality assessment database [90] [91][92]	Politecnico di Milano (PoliMI) et École polytechnique fédérale de Lausanne (EPFL)	352 × 288 704 × 576	25, 30	12

Tableau V.5 Base de données EPFL-PoliMI d'évaluation de la qualité vidéo.

La base de données ReTRiEVED contient 184 vidéos de test obtenues à partir de huit vidéos sources de différent contenu caractérisées par une grande étendue d'informations spatiales et temporelles, différents contextes, températures, mouvements, couleurs et les mouvements de la caméra... etc.

Parmi eux, six vidéos, Crowdrun, Duckstakeoff, Harbour, Ice, Parkjoy et Soccer, ont été prises de la base de données vidéo EPFL-PoliMI disponible sur [91], et ont été utilisées par la base de données ReTRiEVED avec une résolution spatiale 4CIF.

Les séquences vidéo non comprimées ont trois éléments (YUV), nous devons entreprendre l'échantillonnage de ces trois composantes en même temps, avec un échantillonnage de couleur YUV 4:2:0, avec 8 bits par échantillon. Le tableau V.6 et la figure V.6 présentent les caractéristiques des séquences vidéo de tests obtenus et les échantillons des trames extraites de chaque séquence vidéo.

Vidéos	Résolution (Pixels)	Frame Rate (fps)	Longueur (s)	Frames
Soccer	704 × 576	30	7	300
Harbour	704 × 576	30	9	300
Parkjoy	704 × 576	25	8	250
Duckstakeoff	704 × 576	25	9	250
Ice	704 × 576	30	7	240
Crowdrun	704 × 576	25	9	250

Tableau V.6 Séquences vidéo originales avec résolution, fréquence image (fps) et longueur (durée).



Figure V.6 Echantillons de vidéos de test

On peut remarquer que les indices de l'information perceptive spatiale (SI) et l'information perceptive temporelle (TI) recommandés par l'UIT varient de relativement petites à relativement grandes pour le contenu sélectionné [4] [5] [93]. Figure V.7 montre les plans SI et TI des différentes séquences vidéos [4] [5].

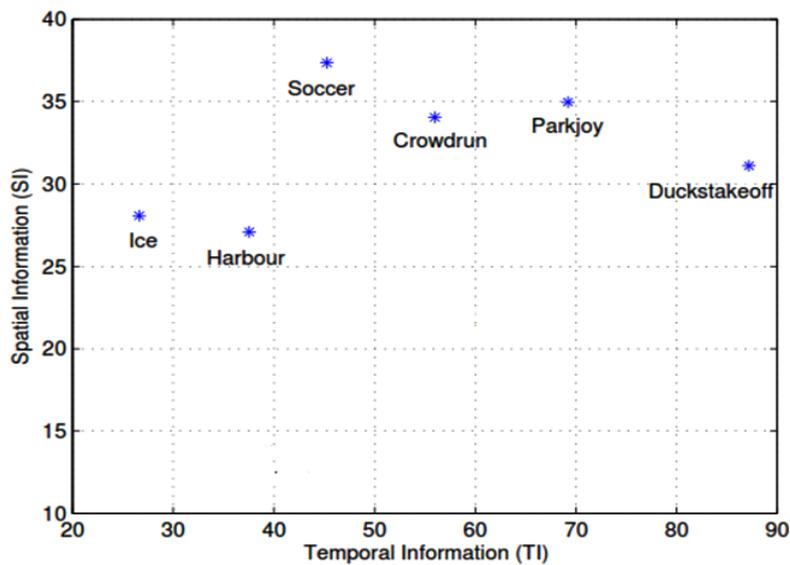


Figure V.7 Plans TI et SI pour les vidéos brutes

Séquence SOCCER:



Frame 1



Frame 11



Frame 90



Frame 200



Frame 280



Frame 300

Figure V.8 Quelques frames de la séquence vidéo "SOCCER".

Séquence HARBOUR:



Frame 1



Frame 50



Frame 130



Frame 210



Frame 260



Frame 300

Figure V.9 Quelques frames de la séquence vidéo "HARBOUR".

Séquence PARKJOY:



Frame 1



Frame 40



Frame 150



Frame 197



Frame 235



Frame 300

Figure V.10 Quelques frames de la séquence vidéo " PARKJOY".

Séquence DUCKSTAKEOFF:



Frame 1



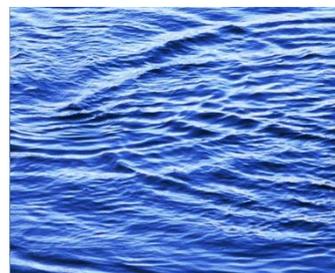
Frame 30



Frame 51



Frame 136



Frame 211



Frame 300

Figure V.11 Quelques frames de la séquence vidéo " DUCKSTAKEOFF".

Séquence ICE:

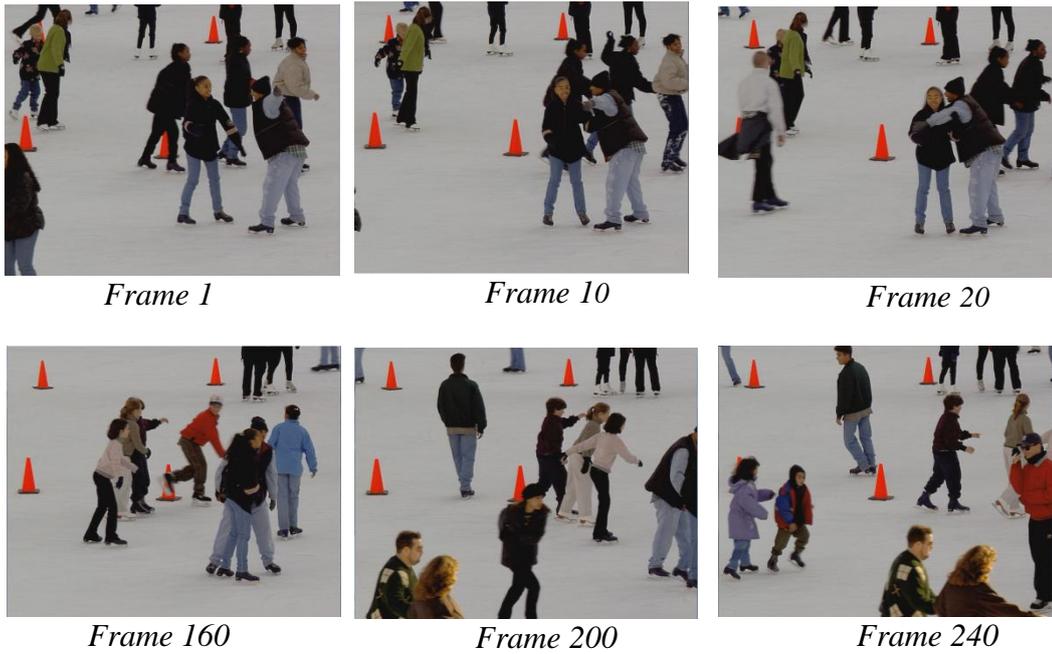


Figure V.12 Quelques frames de la séquence vidéo "ICE".

Séquence CROWDRUN:

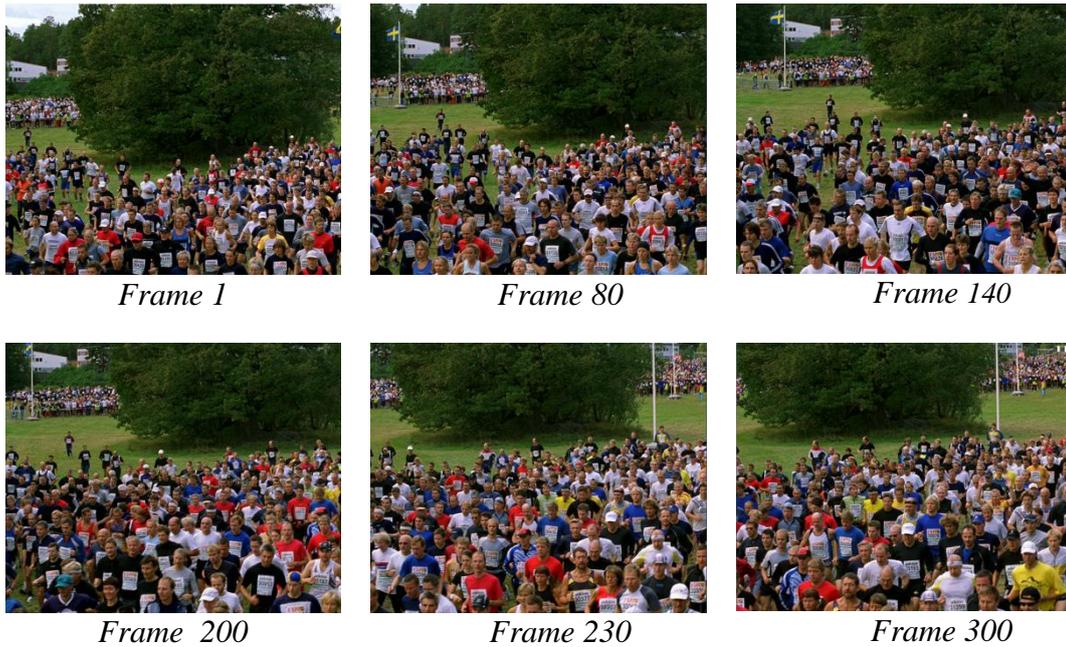


Figure V.13 Quelques frames de la séquence vidéo "CROWDRUN".

V.6 Implémentation de codeur et décodeur sous HEVC

Le but de HEVC-plateformes est de permettre à l'utilisateur d'explorer diverses options de codage/décodage de HEVC et de comparer les codeurs vidéo.

V.6.1 Codeur x265

x265 est développé par MulticoreWare [42], leader des solutions logicielles hautes performances, avec le soutien de fournisseurs de technologies vidéo de premier plan tels que Telestream et DoremiLabs et des contributions de développeurs open source. x265 exploite la plupart des fonctions d'encodage vidéo exceptionnelles et des optimisations du projet d'encodeur AVC x264.

Le logiciel x265 est disponible gratuitement sous la licence GNU GPL 2[42]. Alors que x265 est principalement conçu comme une bibliothèque logiciel d'encodage vidéo, un exécutable de ligne de commande est fourni pour faciliter les tests et le développement. De nos jours x265 est utilisé dans de nombreux produits et services vidéo et matériels de premier plan.

L'encodeur x265 permet à l'utilisateur de contrôler les options d'encodage suivantes:

- Résolution vidéo (largeur et hauteur)
- Débit binaire cible
- Profil de quantification pour les paramètres de quantification
- Nombre d'images à encoder

V.6.2 Codeur HM

Au cours du processus de normalisation HEVC, le JCT-VC a également développé un modèle de test (logiciel de référence) appelé HM (HEVC Test Model). Le but du logiciel de référence était de fournir une base permettant d'effectuer des expériences afin de déterminer les performances de codage. HM est utilisé pour compresser la vidéo YUV420. Une interface utilisateur était développée pour le codeur HM pour permettre à l'utilisateur de contrôler ce qui suit [47].

- Résolution vidéo (largeur et hauteur)
- Débit binaire ciblé
- Profil de quantification pour les paramètres de quantification
- Période intra-I

- Nombre d'images à encoder
- Taille maximale de l'unité de codage
- Profondeur maximale de l'unité de codage
- Taille de tranche (slice) déterminée par CTU, octets ou pixels.
- HM intègre diverses options pour l'estimation de mouvement, telles que la recherche complète.

Une illustration de l'interface utilisateur pour le codeur HM est donnée à la figure V.14.

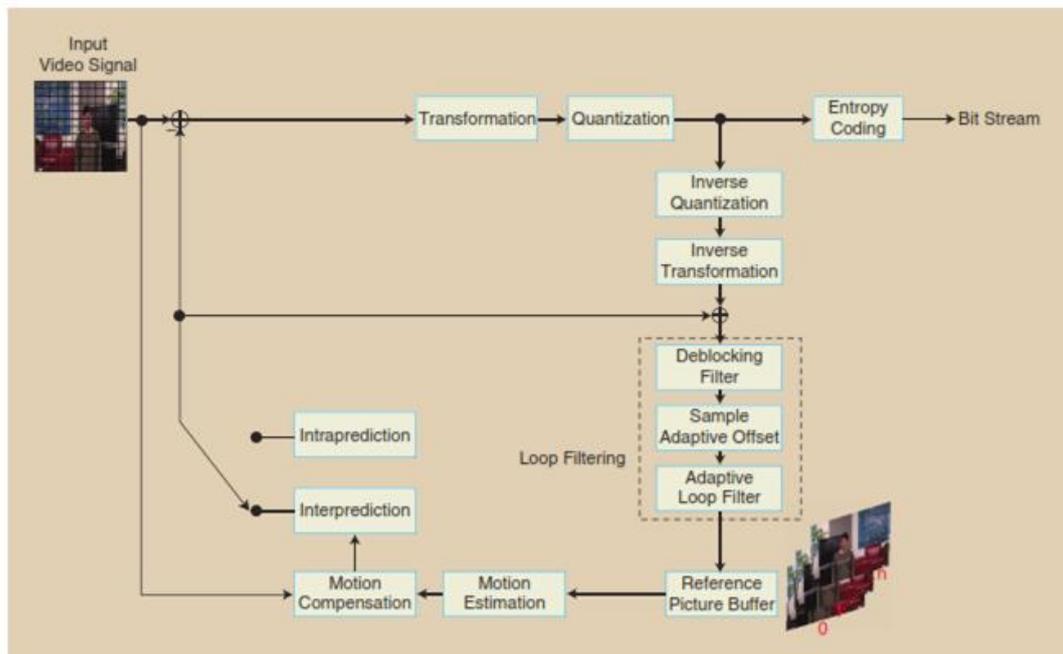


Figure V.14 Schéma synoptique de codeur HM [41]

De plus, des pertes de paquets ont été implémentées dans HM et x265 pour permettre d'étudier les effets de la perte de paquets sur la qualité de la vidéo décodée.

V.6.3 Groupe d'images (GOP) dans HM

Dans toutes les normes de codage vidéo, les GOP sont utilisés pour définir les relations de codage dans une séquence de trames vidéo. De manière générale, tous les modes GOP [41] [95] commencent par un I (Intra image) qui est codé de manière indépendante sans faire référence à d'autres images. L'accès aléatoire dépend de l'utilisation de la première image dans le GOP. En général, nous prévoyons que l'utilisation de plus grandes tailles de GOP conduirait à plus d'efficacité de compression. Pour les besoins de cette thèse, la taille du GOP sera modifiée à l'aide du logiciel de référence HM16.11 [41].

HEVC a deux façons principales d'organiser les GOP:

- Un GOP fermé qui utilise des types d'image IDR (Instantaneous Decoder Refresh) sans dépendance en dehors du GOP,
- Open GOP qui utilise des types d'image CRA (Clean Random Access)

HEVC [47] comme d'autres normes de codage vidéo utilise principalement trois types de trame: I (Intra), P (Prédicative), et B (Bidirectionnelle) [94] [95].

Plus précisément, nous avons:

- I-Pictures: Les images Intra (I), également appelées images de référence, contiennent toutes les informations vidéo nécessaires pour le décodage, sans faire référence à d'autres images.

En règle générale, les images I nécessitent plus de bits à coder que les autres types de trame. Chaque GOP a une image I qui est la première image de la séquence vidéo de codage. Les images I sont codées en utilisant une prédiction spatiale angulaire avec des modes allant de 0 à 34 et prennent moins de temps à coder par rapport aux images P ou B.

- P-Pictures: Les images prédites (P) sont codées en utilisant l'inter-prédiction d'un signal de compensation de mouvement par PU (c'est-à-dire unidirectionnel) basé sur la disponibilité d'une image I ou P précédente la plus proche. Les images P utilisent uniquement la liste d'images de référence 0 (L0) et les images P prennent la durée de codage modérée par rapport aux images I
- Images B: Les images bidirectionnelles (B) sont codées en utilisant l'inter-prédiction avec deux signaux compensés par mouvement par PU (c'est-à-dire bidirectionnel) en fonction de la disponibilité des images I et P qui les précèdent et les suivent.
Les images B utilisent la liste d'images de référence 0 (L0) et la liste 1 (L1).
Les images B prennent plus de temps à encoder par rapport aux images I et P.

V.6.4 Modes de configuration de l'encodeur

Le logiciel de référence HM [41] a trois configurations principales. Elles [41] [95] incluent: le mode tout intra (all Intra(AI)), le mode d'accès aléatoire RA (Random Access), et le mode faible retard LD (low delay).

All Intra Mode (AI): Dans ce mode, toutes les images sont codées uniquement à l'aide de tranches I (Slice I). Il peut s'agir d'une image IDR ou CRA selon le fichier de configuration. Il n'y a pas de prédiction temporelle faite dans ce mode car toutes les images sont intra-codées en utilisant la prédiction angulaire spatiale. Le paramètre de quantification pour chaque image I peut être modifié dans la séquence. La figure V.15 montre une représentation en mode tout intra.

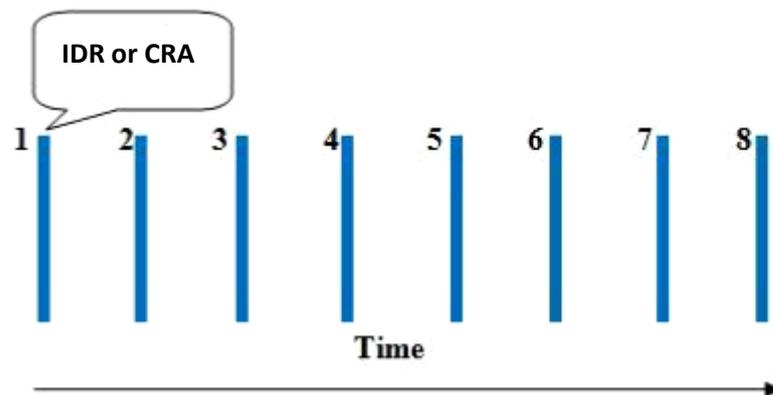


Figure V.15 Mode All Intra

Random Access (RA): Ce mode utilise une relation pyramidale ou dyadique parmi les tranches (slices) B. Les images B hiérarchiques sont utilisées pour le codage avec une image à accès aléatoire utilisée toutes les 1s. Ce mode peut être utilisé dans les applications de diffusion vidéo numérique. La Figure V.16 montre une configuration à accès aléatoire. Chaque image est représentée avec un numéro indiquant l'ordre d'encodage. La première image d'une séquence vidéo est toujours codée comme une image IDR et les autres images intra sont codées en tant qu'images intra non-IDR ("Open GOP"). Puisque cette configuration suit une structure semblable à une pyramide, il y a plusieurs couches temporelles dans les images et chacune des images intermédiaires est codée comme une image B. La période Intra, la taille et le type de GOP peuvent toujours être modifiés. Cette structure généralisée P et B (GPB) peut également être utilisée pour définir la couche temporelle la plus basse qui fait référence à l'image I ou GPB pour l'inter-prédiction.

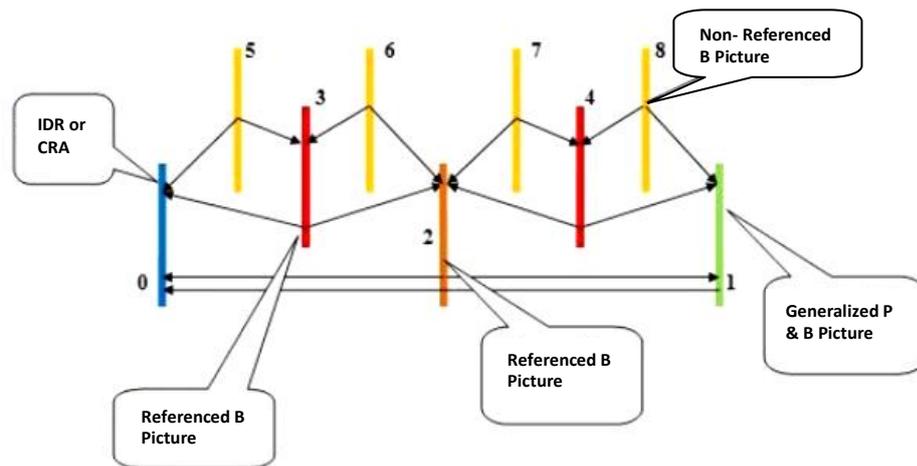


Figure V.16 Mode Random Access

Low Delay (LD): Cette configuration diffère légèrement du mode d'accès aléatoire car il n'y a pas de ré-ordonnancement d'image et seule la première trame est codée en utilisant des tranches (slices) I. Ce mode est adapté aux applications de streaming et de vidéoconférence en direct. Similaire à l'accès aléatoire, dans les conditions de codage à faible retard, la première image de la séquence vidéo est toujours codée comme une image IDR. Les autres images de la séquence sont codées sous la forme d'images P et B généralisées (GPB). Comme il n'y a pas de réorganisation d'image comme dans un accès aléatoire, le GPB utilise uniquement les images de référence, où l'ordre d'image ou l'ordre d'affichage est plus petit que l'image courante en cours d'encodage. En faible retard, il y a deux images de référence List0 et List1 dans les tampons décodés à utiliser lors de la reconstruction de l'image originale. Les contenus de List0 et List1 sont identiques et sont mis à jour avec un processus de gestion de fenêtre glissante.

Les images de référence présentées dans la figure V.17 démontrent qu'il n'y a pas de ré-ordonnancement de l'image ou, en d'autres termes, pas de relation cyclique (dyadique) entre les images de référence B. Chaque image est associée à un nombre représentant l'ordre d'encodage et un QP pour chaque image. Celle-ci peut être modifiée selon le décalage QP dérivé de l'image IDR ou CRA au début du GOP. Dans HM, il est prévu d'utiliser une configuration à faible retard avec un codage d'image P en utilisant uniquement l'image de référence List0 pour l'inter-prédiction.

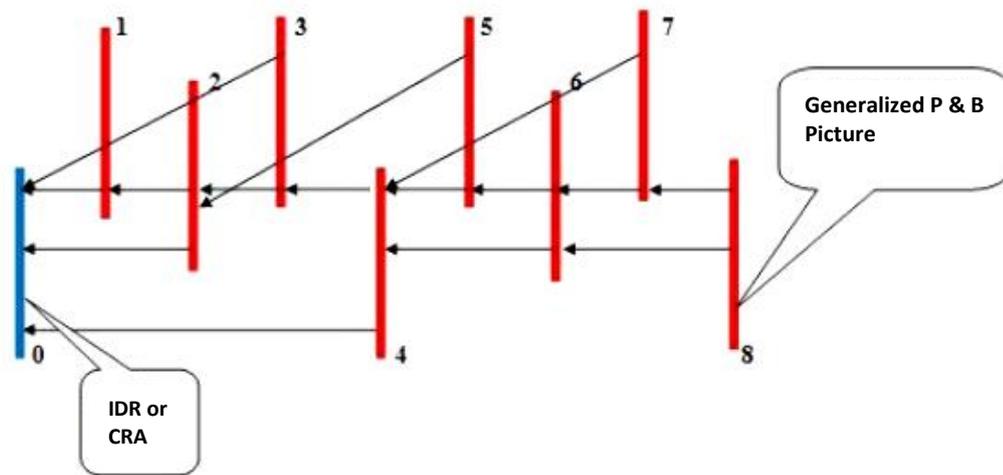


Figure V.17 Mode Low Delay

V.7 Paramètres de configuration

L'outil que nous utilisons pour coder les séquences vidéo brutes est l'encodeur HEVC/H.265 de référence, à savoir la plateforme HM 16.11. La figure V.18 présente le schéma d'adaptation du système qui a été conçu et utilisé pour fournir une étude réaliste de l'impact du processus d'encodage HEVC sur la qualité vidéo. Le type de contenu vidéo défini selon les différentes séquences vidéo est utilisé avec différentes valeurs de QP pour modéliser l'impact du processus de codage HEVC pour un scénario de test donné. Les séquences vidéo ont été codées à l'aide de la gamme de QP complète (de 17 à 42 avec un pas de 5) depuis la recommandation de JCT-VC [22] [96].

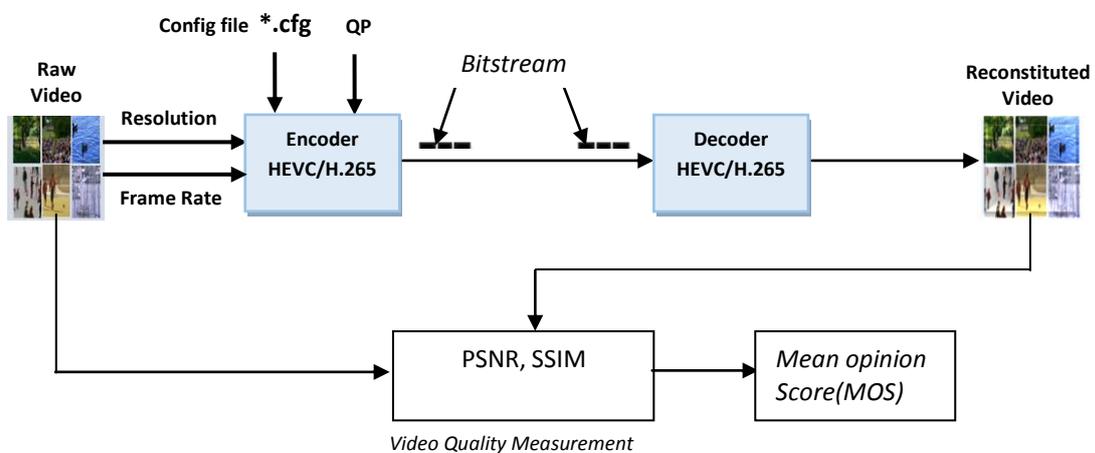


Figure V.18 Schéma d'adaptation pour l'évaluation objective/subjective des séquences vidéo

V.7.1. Modèle de paramètre d'encodeur basé sur MDC

À l'aide de l'outil de langage MATLAB, deux descriptions (bitstreams) de la même vidéo sont créées; ces deux dernières prennent respectivement les trames paires et les trames impaires, comme indiqué sur la figure III.5(b), par exemple, Ice_even.yuv et Ice_odd.yuv. Ensuite, dans les mêmes conditions que celles mentionnées au tableau V.7 ci-dessous, le codeur HEVC / H.265 a été appliqué à chaque description. Du côté décodeur, la description centrale est reconstruite à partir de bitstream(s) avec le décodeur HEVC / H.265, comme vu dans la figure III.5 (chapitre 3).

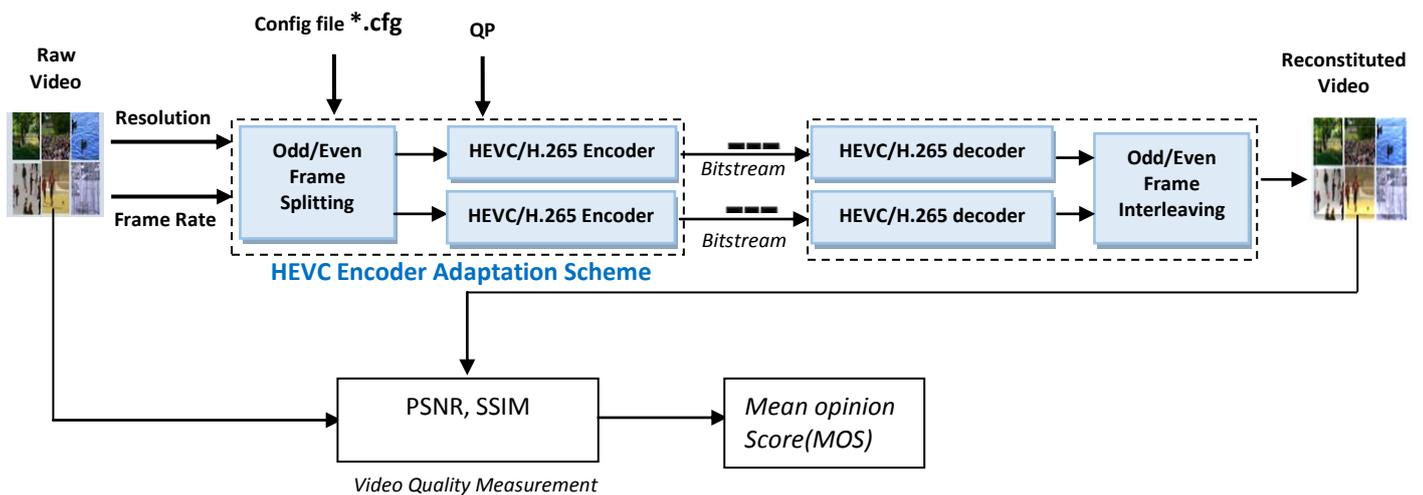


Figure V.19 Schéma d'adaptation pour l'évaluation objective/subjective des séquences vidéo basée sur la méthode MDC

Les conditions dans lesquelles HEVC doit être testé [47] [53] sont choisies pour obtenir différentes métriques objectives [97], telles que PSNR [calculé par Eq. (4)], SSIM [calculé par Eq. (5)], et MOS_Estimé [calculé par Eq. (6)].

Le JCT-VC définit de nombreuses configurations dans le logiciel de référence HM 16.11 [41] [67] [94] tels que le mode d'accès aléatoire (RA), mode tout intra-(AI), et le mode faible retard (LD). [86] [95]

Dans cette expérience, chaque trame des séquences vidéo est codée dans le mode faible retard avec des tranches (slices) B (LB), car elle convient aux applications de diffusion en direct (streaming) et de vidéo conférence. Le fichier de configuration: encoder_lowdelay_main.cfg est utilisé, comme indiqué dans le tableau V.7.

Paramètres	Valeur
MaxCUWidth	64
Height	64
MaxPartitionDepth	4
QuadtreeTULog2MaxSize	5
QuadtreeTULog2MinSize	2
QuadtreeTUMaxDepthInter	3
QuadtreeTUMaxDepthIntra	3
IntraPeriod	1
GOPSize	4
RDOQ	1
LoopFilterDisable	1(enabled)
InternalBitDepth	8
SAO	1
QP	17, 22, 27, 32, 37,42

Tableau V.7 Conditions d'essai de HM16.11

V.8 Simulations, résultats et discussions

V.8.1 Analyse des performances des H.265/HEVC

La figure V.20 montre les graphiques du PSNR, du SSIM et du MOS_Estimé selon la valeur de QP avec un codeur HEVC sans MDC et avec MDC respectivement, les mesures obtenues sont présentées sur les Figures V.20(a) –V.20(f).

Comme prévu, les courbes montrent que le PSNR, SSIM et MOS Estimé décroissent en fonction du QP. Tous les résultats ont démontré que le codage des différentes vidéos avec les mêmes paramètres de codec tels que QP, fréquence de la vidéo et résolution spatiale conduit à des valeurs PSNR différentes, comme le montre la figure V.20 (a). De plus, à mesure que le QP augmente, le débit binaire diminue, ce qui diminue la qualité vidéo, comme l'ont prouvé plusieurs études [10] [17] [22] [83].

En revanche, les valeurs QP les plus basses se traduisent par des débits binaires plus élevés, ce qui entraîne une amélioration de la qualité vidéo. Comme le PSNR varie avec différentes séquences vidéo, on peut affirmer que le type de contenu a un impact sur les valeurs de PSNR, ce qui indique, en plus des paramètres de codage, que d'autres facteurs peuvent également influencer sur les valeurs PSNR.

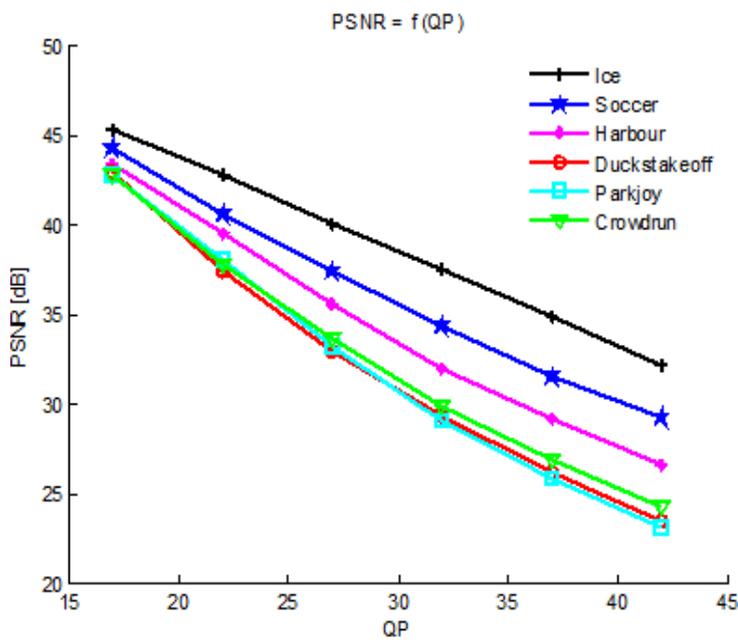
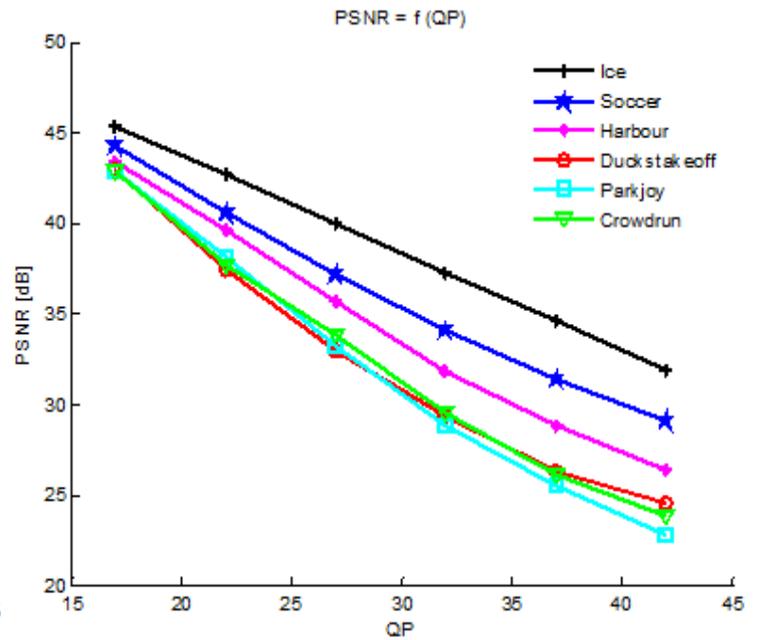


Figure V.20 (a) PSNR de en fonction de QP



V.20 (b) PSNR en fonction de QP (cas de MDC)

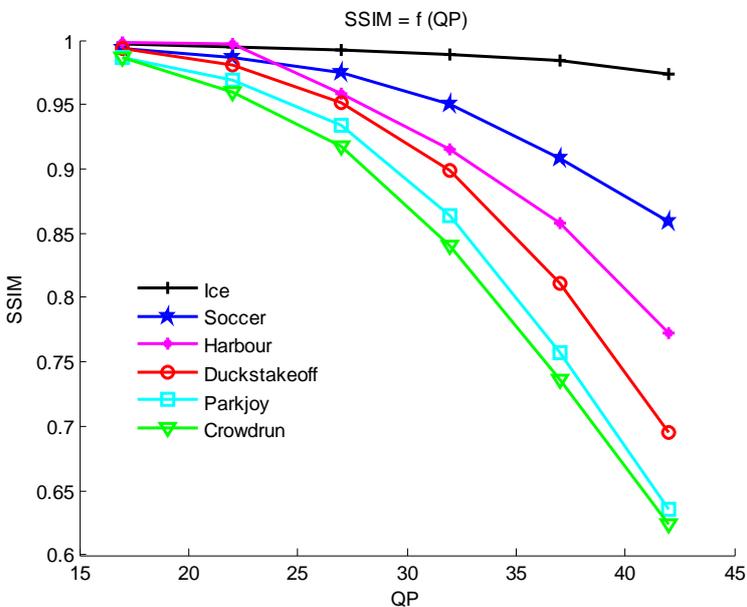


Figure V.20 (c) SSIM de en fonction de QP

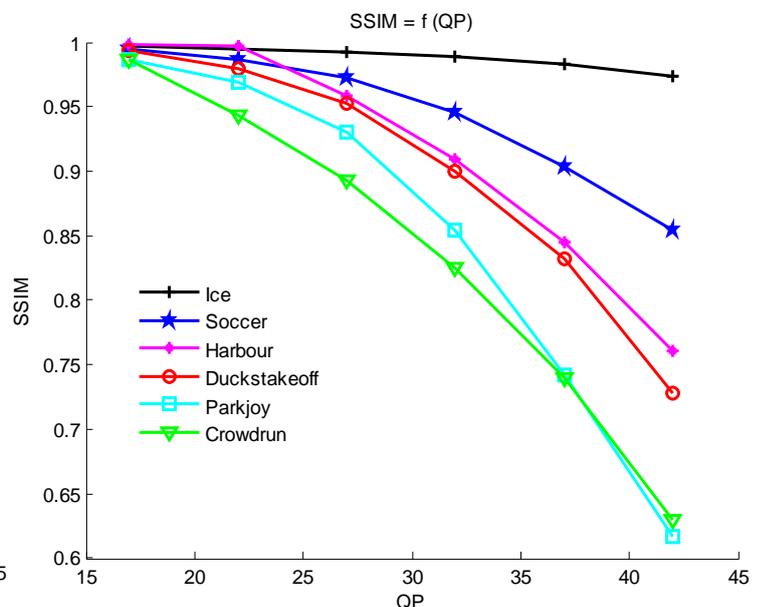


Figure V.20 (d) SSIM de en fonction de QP (cas de MDC)

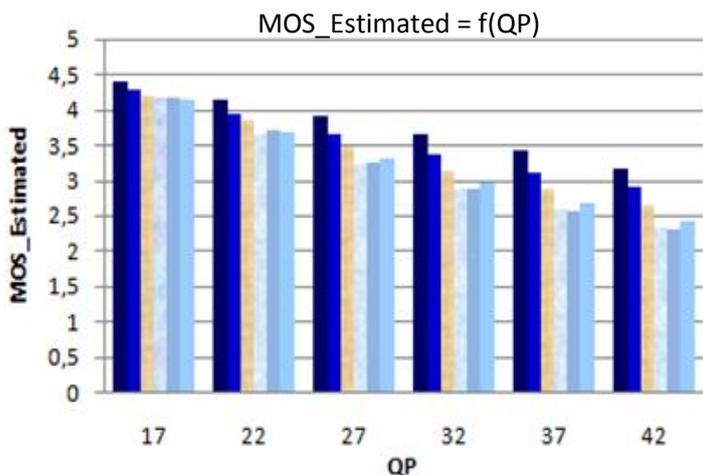


Figure V.20 (e) MOS Estimé en fonction de QP

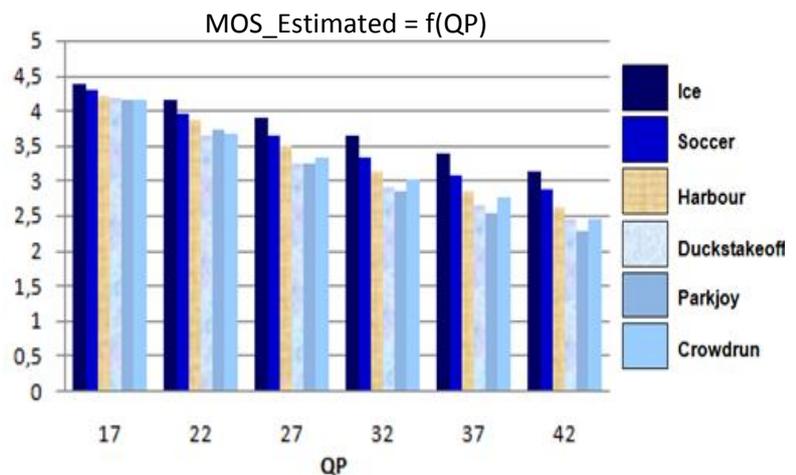


Figure V.20 (f) MOS Estimé de en fonction de QP (cas de MDC)

Dans la référence [8], le modèle de paramètre du codeur peut être approximativement dérivé comme suit: la caractérisation de la source vidéo et le modèle de paramètre du codeur peuvent être approximativement dérivés comme suit:

$$PSNR = \alpha e^{-\beta QP}, \quad (7)$$

Où α , β sont des coefficients dépendant des fichiers de configuration du codeur HEVC et des séquences vidéos comme indiqué dans l'équation suivante :

$$(\alpha, \beta) = f(fr, W, H, config\ file) \quad (8)$$

Où fr est la fréquence de la vidéo (frame rate), et $W \times H$ est la résolution de la séquence vidéo.

De même, $PSNR = \alpha$ si $QP \rightarrow 0$, qui fait référence au PSNR de la vidéo reconstruite en raison de la dégradation de la qualité du codeur HEVC à l'exception de la quantification. Les deux paramètres α et β sont principalement liés au fichier de configuration, Le paramètre α en mode low delay prend une valeur moyenne de l'ordre de 62.2. Quant au paramètre β c'est une constante dépendant de la configuration pour différentes séquences vidéo [8].

En d'autres termes, les résultats obtenus sont confirmés par le SSIM comme le montre la figure 20(c), qui est un meilleur indicateur de qualité vidéo. C'est ainsi qu'on peut démontrer la différence entre le codage des séquences vidéo lorsque la QP augmente. Cependant, on peut observer que les vidéos ayant une valeur élevée de QP sont nettement plus dégradées. De plus, la figure 20(e) indique que le MOS_Estimé décrit par l'équation (6) a une relation avec PSNR et SSIM. Ainsi, le PSNR et le SSIM sont positivement corrélés avec la qualité vidéo; c'est -à-dire que l'augmentation du PSNR ou du SSIM conduit à une augmentation de MOS_Estimé.

V.8.2 Comparaison des paramètres de l'encodeur basé sur MDC

Il n'y a pas de différence significative entre le modèle du codeur HEVC et le modèle du codeur HEVC basé sur MDC comme le montrent les Figures V.20.

A partir de la figure V.20(b), on peut observer que le PSNR diminue progressivement avec une augmentation des valeurs de QP telles qu'indiquées dans les résultats précédents comme le montre la figure V.20(a). En outre, le PSNR prend la même tendance décroissante dans les deux cas. On peut noter aussi que lorsque le QP augmente, il y a une légère différence, ce qui peut être considéré comme négligeable. Ce comportement a été rapporté par d'autres études [3] [23]. De même, SSIM et MOS_Estimé ont une relation avec PSNR, comme montré sur les figures V.20.

Les séquences vidéo avec le PSNR le plus bas présenté dans les deux cas sont respectivement, le Duckstakeoff, Parkjoy et Crowdrun. Le PSNR des séquences vidéo Ice, Soccer et Harbour est plutôt élevé. Il semble alors qu'il y ait une grande variation entre la séquence Crowdrun et la séquence Ice. Par exemple, dans un modèle HEVC basé sur MDC; la séquence Ice lorsque QP=17 PSNR=45.29 dB et PSNR de la séquence Crowdrun=42,82 dB, tandis que pour QP = 42, PSNR = 31,86 dB pour la séquence Ice, PSNR = 25,55 dB pour la séquence Crowdrun. Les tableaux V.8 et V.9 présentent les valeurs de PSNR des vidéos en fonction de QP dans les deux cas.

QP	Ice	Soccer	Harbour	Duckstakeoff	Crowdrun	parkjoy
17	45.31	44.22	43.41	42.97	42.74	42.79
22	42.72	40.54	39.55	37.45	37.76	38.06
27	40.07	37.38	35.58	32.92	33.59	33.22
32	37.50	34.29	31.98	29.30	29.90	29.08
37	34.88	31.52	29.20	26.17	26.88	25.81
42	32.11	29.29	26.65	23.48	24.24	23.13

Tableau V.8 Variation du PSNR dans le cas de codeur HEVC/H.265

QP	Ice	Soccer	Harbour	Duckstakeoff	Crowdrun	parkjoy
17	45.29	44.28	43.40	42.97	42.82	42.86
22	42.68	40.53	39.60	37.42	37.70	38.10
27	39.91	37.16	35.65	32.96	33.77	33.19
32	37.20	34.05	31.81	29.36	30.53	28.84
37	34.58	31.32	28.83	26.84	27.89	25.49
42	31.86	29.09	26.36	24.55	25.55	22.79

Tableau V.9 Variation du PSNR dans le cas de codeur HEVC/H.265 basée sur MDC

Cependant, toutes les vidéos codées contiennent différents types de contenu (activités de mouvement TI et complexité de la vidéo SI), il convient de noter que la qualité vidéo diminue avec l'augmentation de la valeur QP (compression). Cependant, cela apparaît clairement sur les figures ci-dessus que la baisse de qualité est plus importante pour les vidéos à mouvement plus rapide que pour les vidéos avec un mouvement lent. En outre, on doit garder à l'esprit le niveau de compression, car il détermine la qualité vidéo, c'est-à-dire qu'une compression trop élevée peut entraîner une dégradation importante de la qualité vidéo. D'autre part, les résultats présentés dans les figures précédentes montrent que l'encodage des vidéos avec un QP inférieures à 27 conduits à une meilleure qualité vidéo égale ou supérieure à 37 dB. De plus, toutes les séquences ont besoin d'un niveau de compression différent pour la même qualité, c'est pourquoi les séquences vidéo Ice, Soccer et Harbour nécessitent une compression plus faible et

une bande passante de transmission requise pour la même qualité par rapport aux séquences vidéo Duckstakeoff, Parkjoy et Crowdrun.

En outre, le SSIM donne une idée importante sur la structure et la complexité de la vidéo, comme indiqué sur les figures V.20(c) et V.20(d) qui démontrent tout ce qui a été dit à propos du PSNR. De plus, SSIM définit la différence structurelle entre les séquences vidéo Ice, Soccer et Harbour, qui diminue lentement pour la même qualité par rapport aux séquences vidéo Duckstakeoff, Parkjoy et Crowdrun qui déclinent rapidement dans la même direction. D'un autre côté, le processus d'encodage utilise la complexité de la vidéo pour obtenir une QP conduisant au débit binaire ciblé pour une qualité donnée. Cependant, différents types de contenu vidéo ont différents débits binaires en raison des caractéristiques de mouvement différentes. En effet, les vidéos avec des activités de mouvement lent ont des débits binaires inférieurs par rapport à ceux avec des activités de mouvement rapide. Ainsi, les exigences de bande passante de transmission de débit binaire sont plus élevées pour les activités de mouvement supérieur.

V.8.3 Variation du débit en fonction du pas de quantification

Cette simulation est conduite sur deux différentes vidéos et avec les paramètres de codage mentionnés ci-dessus (Section V.5). La figure V.21 présente les débits binaires de la séquence vidéo Parkjoy, qui a le plus haut SI et TI, et le débit binaire de la séquence vidéo Ice, qui a les plus bas TI et SI.

Le débit binaire dans le cas d'utilisation du MDC est plus élevé que le débit en SDC, comme nous l'avons vu dans les études théoriques. Cependant, pour certaines applications, la robustesse fournie par le MDC est plus importante que le débit binaire [26][98].

Le débit des deux composantes chrominances U et V est inférieur à celui de la composante luminance Y, parce que leurs caractéristiques sont différentes et doivent être modélisées séparément pour augmenter l'efficacité du codage [3].

On peut remarquer que lors du codage des séquences vidéo avec la même QP, les débits binaires résultants sont plus élevés pour des séquences avec des activités de mouvements rapides et une plus grande complexité des vidéos. Par exemple en HEVC basé sur SDC, lorsque QP = 32 le débit binaire de Ice = 337.49 kbps et le débit binaire de Parkjoy séquence = 2757,92 kbps. Tandis que pour l'encodeur HEVC basé sur MDC, le débit binaire = 900,12 kbps pour la séquence Ice et le débit binaire = 7106,68 kbps pour la séquence Parkjoy.

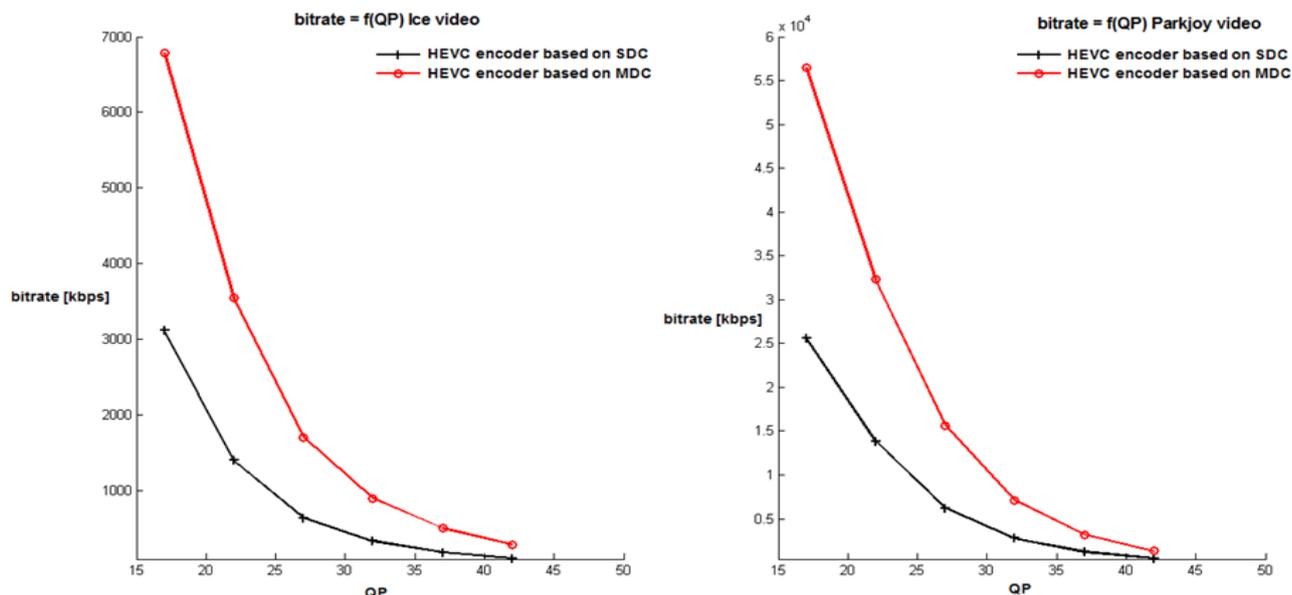


Figure V.21(a) Débit en fonction de QP

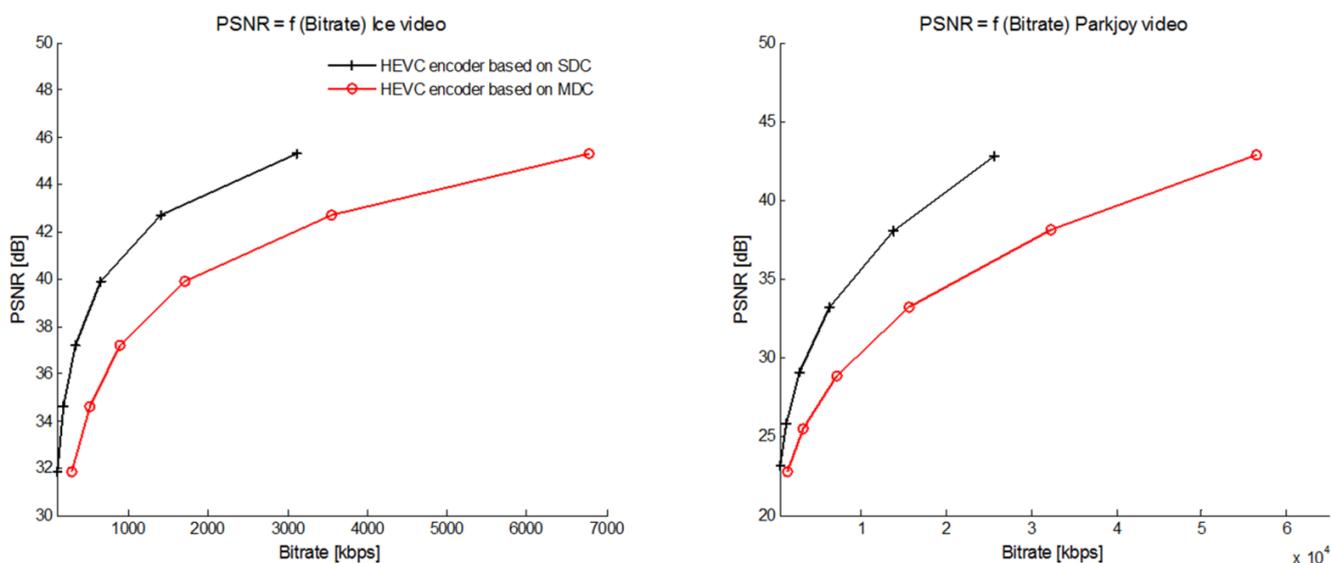


Figure V.21(b) Débit de MDC vs SDC

Par conséquent, ceci est important car il nous permet de voir l'effet de QP et comment il devrait être utilisé pour concevoir une qualité vidéo améliorée sans perturbations. En outre, la fréquence de la vidéo a un impact sur la vidéo codée. Ainsi, pour un même QP, la qualité vidéo augmente à mesure que la fréquence de la vidéo augmente. Par voie de conséquence, la qualité vidéo codée et conçue (V_Q) dans notre modèle de prédiction dépend de l'équation suivante:

$$V_Q = f(VC, VCT, QP), \quad (9)$$

1. Le codec vidéo HEVC / H.265 (VC).
2. Le type de contenu vidéo caractérisé par les activités de mouvement et la complexité des séquences vidéo (VCT), également indiqué par les auteurs dans les références [4] et [22].
3. Le QP en tant que paramètres de codage initial détermine la qualité vidéo initiale.

V.8.4 Impact des pertes de paquets

V.8.4.1 FFmpeg [99]

Le FFmpeg [99] est une multi plate-forme (collection de logiciels libres) pour traiter, enregistrer, convertir et diffuser des données audio et vidéo.

Cette multi plate-forme a été créée par Fabrice Bellard en 2000, et est actuellement supervisée par Michael Niedermayer. Elle est maintenant hébergée sur les serveurs du projet MPlayer. Elle a été développée sous GNU/Linux mais peut être compilée sur les autres systèmes d'exploitation [100], y compris Windows. Le choix de la licence GNU LGPL ou GNU GPL dépend du choix de configuration d'options.

Le projet est composé d'une multitude de composants, outils en ligne de commande où une variété des logiciels les plus importants:

- `ffmpeg`: outil pour convertir des formats vidéo. Il prend en charge l'encodage en temps réel à partir d'une carte vidéo ou carte TV.
- `ffserver`: est un serveur HTTP dédié à la diffusion de contenu vidéo en continu (streaming).
- `ffplay`: pour lire un fichier multimédia.
- `libavcodec`: comprend l'ensemble des encodeurs et décodeurs audio/vidéo de FFmpeg. La majorité des codecs ont été réélaborés à partir de zéro afin de garantir les principales performances et la réutilisabilité du code source, une partie d'entre eux ont été développés par rétro-ingénierie.
- `libavformat` contient un analyseur syntaxique (un parcourer en français, parser en anglais) et un générateur pour les formats audio/vidéo les plus communs.

V.8.5 Configuration expérimentale

Dans la configuration expérimentale, une architecture LAN sans fil a été utilisée pour diffuser la vidéo avec le fichier d'extension HEVC (.hevc) de l'émetteur au récepteur, c.-à-d, le processus d'encodage HEVC (coder/décoder) d'une séquence vidéo de l'émetteur au récepteur pour un scénario de test donné.

Le FFmpeg est utilisé du côté du décodeur, car dans certains cas, la plate-forme HM ne peut pas décoder les trames affectées par l'altération du réseau, ce qui rend l'évaluation de la qualité impossible. FFmpeg est très utile, en particulier pour copier la séquence vidéo HEVC et la diffusion sur UDP (en utilisant l'encodage HEVC @ 9000 Kbps et le protocole UDP).

Dans ce travail de thèse, la vidéo devrait être encodée. Ensuite, elle est envoyée avec FFmpeg en streaming vers la destination via un canal perturbé accumulé par l'émulateur de

réseau (NetEM) [101]. Les tests ne portent que sur des flux UDP: de ce fait, la problématique liée au routage des messages d'acquittement TCP n'est pas considérée.

V.8.6 Système proposé

Le domaine sans fil est configuré dans le mode "ad hoc" avec la fonctionnalité de point d'accès désactivé, la distance entre les deux stations de base (ordinateurs) est de 100 m. En travaillant avec un microordinateur dont les paramètres du système sont les suivants :

- Les ordinateurs sont des ordinateurs portables HP.
- Processeur Intel (R) Core (TM) i7-4790S, Processeur Speed @ 3. 20 GHZ.
- RAM 8 GO.
- Architecture de processeur 64 bits.
- Type système : système d'exploitation : Ubuntu 14.04LTS Linux.

La configuration d'émulation est illustrée par la Figure V.22.

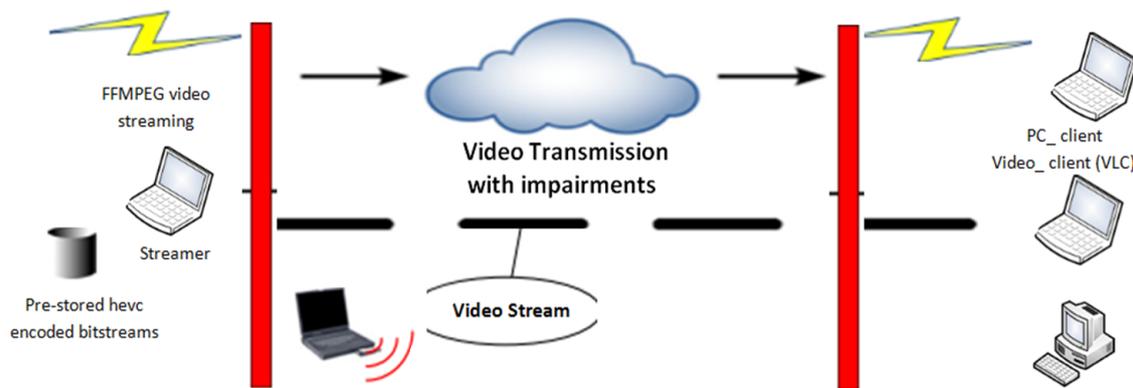


Figure V.22 Système proposé

Le streaming vidéo est un mode diffusion pratique et rapide pour écouter ou visionner ses contenus favoris sur le réseau IP ou internet. Pour cela, il faut que chaque personne dispose d'un outil qui gère les différents média (vidéo, ...etc). Ceci peut se faire grâce à une adjonction (logiciels et/ou matériels ..., etc.) à son poste de travail.

Les effets du PLR ont été introduits par la suppression aléatoire de paquets dans un nœud intermédiaire. Nous avons pris en considération sept valeurs différentes de PLR (0,1%, 0,4%, 1%, 3%, 5%, 8% et 10%). La sélection des valeurs de PLR est basée sur la recommandation de l'UIT [6] [7] [102]. Du côté du récepteur, un lecteur VLC (version 2.1.6) a été utilisé.

Par conséquent, pour analyser l'impact de la perte de paquets, du contenu vidéo et de la qualité perceptuelle, nous avons pris en compte l'effet de la variation du PLR à un moment donné. Les valeurs de délai et de gigue sont définies mis à zéro. La figure V.23 montre que le PLR peut provoquer des blocs brisés et des artefacts de lignes de répétition dans la vidéo décodée, dégradant ainsi la qualité perceptuelle.



Figure V.23 Echantillons des trames des séquences montrant les artefacts produits par PLR.

Des recherches antérieures ont démontré que le PLR représente la déficience du réseau la plus importante [5] [9] [10]. Ainsi, pour étudier l'évaluation de la qualité vidéo, la figure V.24 illustre le schéma d'adaptation du système qui a été conçu et utilisé pour fournir une étude réaliste de l'impact du PLR sur la qualité vidéo. Nous avons utilisé les paramètres du codeur HEVC qui sont discutés dans Section V.5 et les séquences vidéo qui sont présentées dans le tableau V.7. Ensuite, chaque vidéo est transmise sur un réseau IP sans fil émulé en tant que fichier *.hevc comme il est illustré dans la figure V.22.

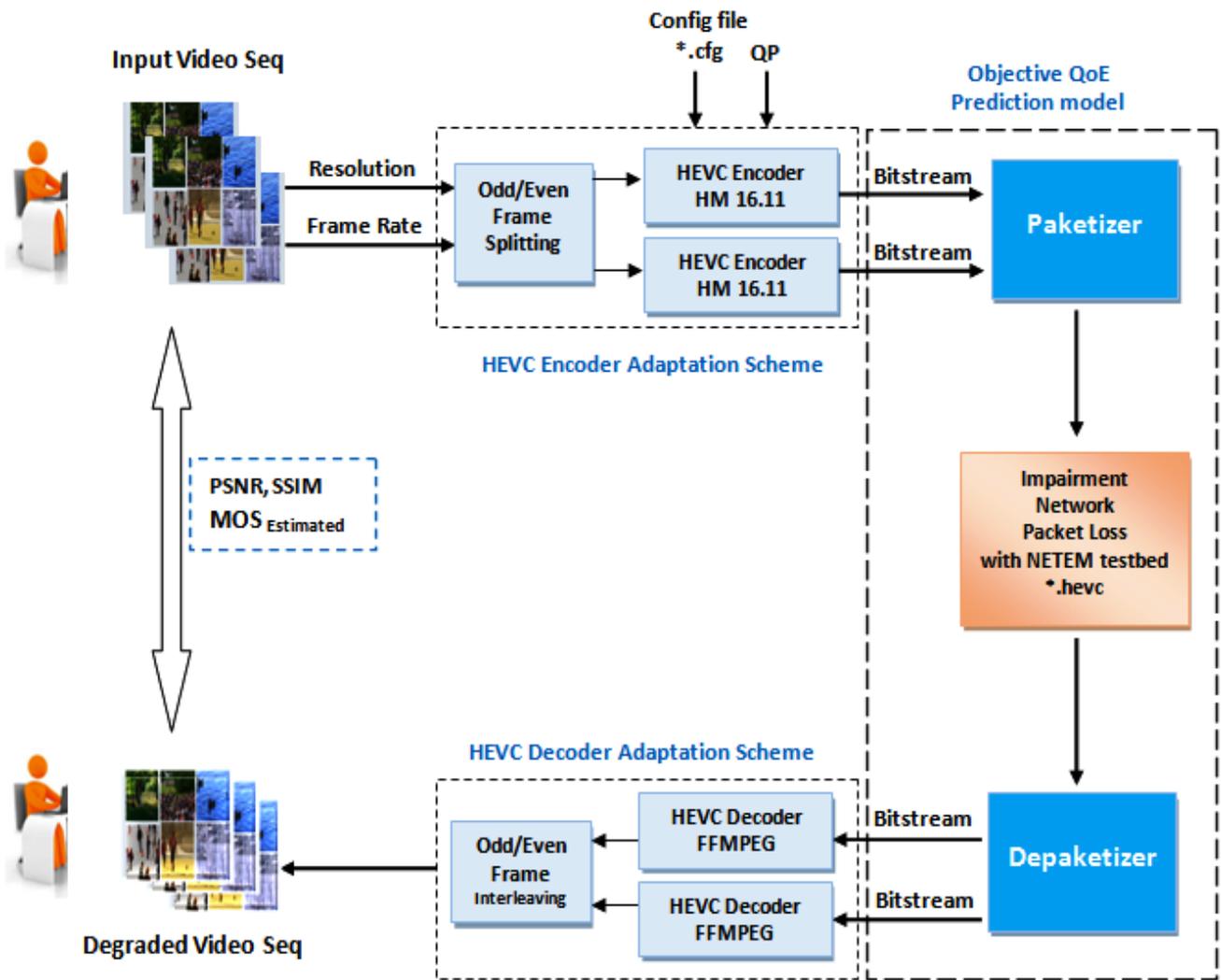
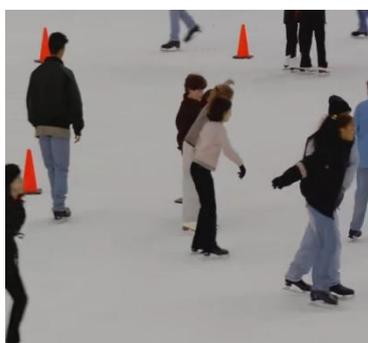


Figure V.24 Schéma d'adaptation pour l'évaluation objective de la séquence vidéo

Effets de perte des paquets sur la vidéo Ice



PLR 0%



PLR 3%



PLR 10%

Effets de perte des paquets sur la vidéo Crowdrun



PLR 0%



PLR 3%



PLR 10%

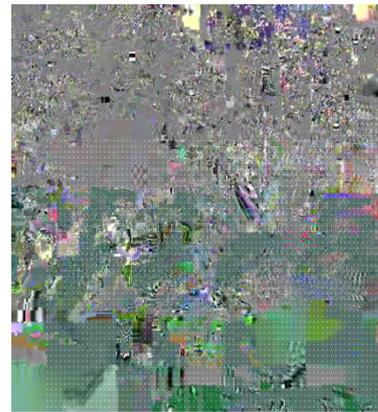
Effets de perte de paquets sur la vidéo Parkjoy



PLR 0%



PLR 3%



PLR 10%

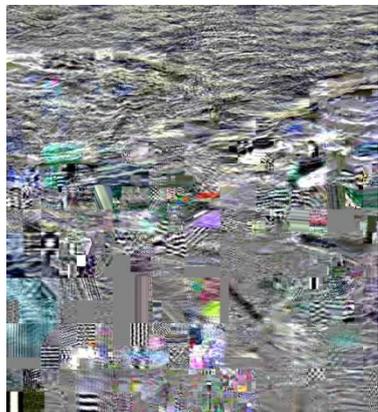
Effets de perte de paquets sur la vidéo Duckstakeoff :



PLR 0%



PLR 3%



PLR 10%

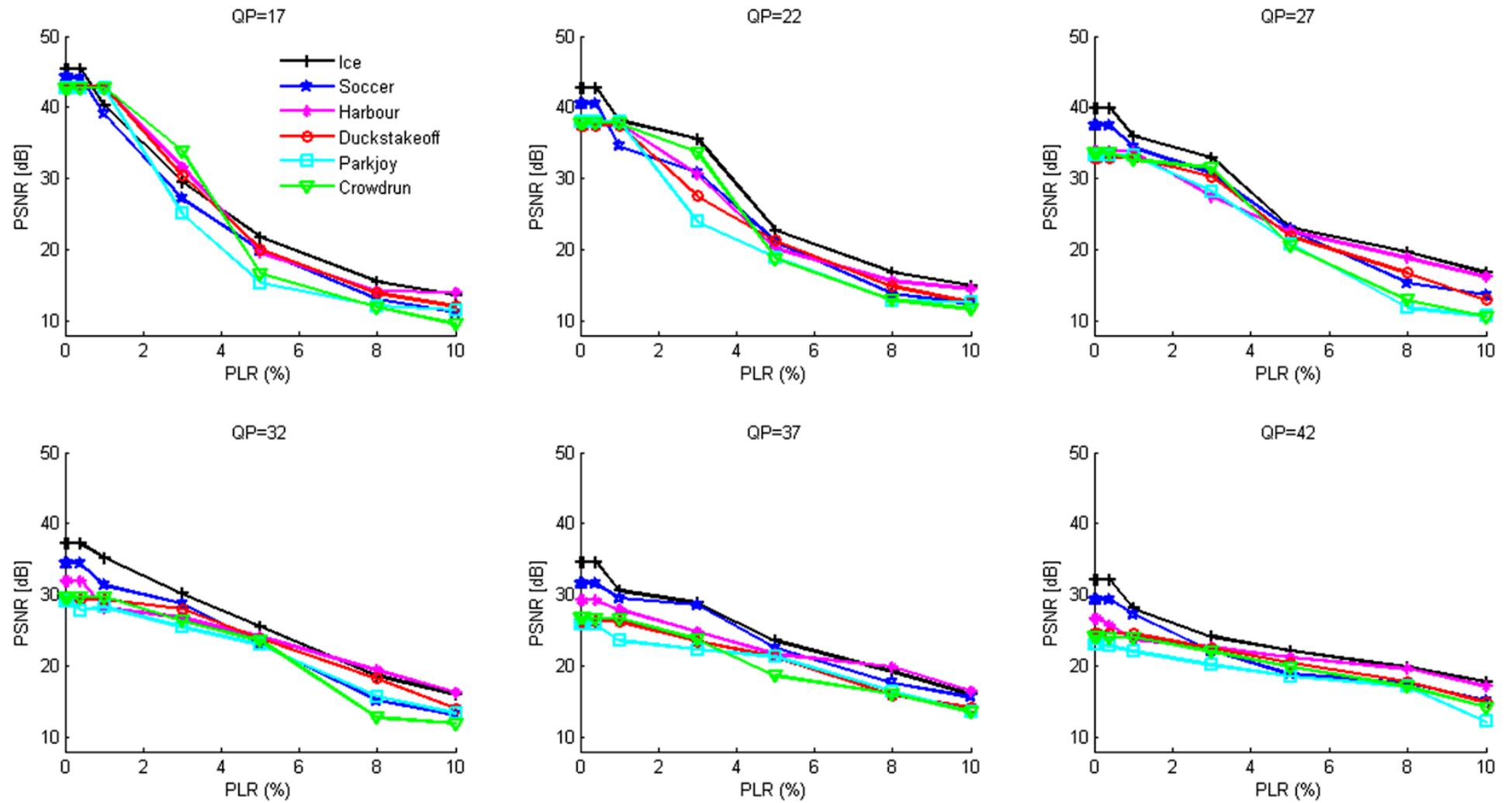
V.8.7 L'impact de perte des paquets sur le streaming vidéo HEVC

La figure V.25 montre que les plus petites valeurs de PLR donnent des valeurs plus élevées de PSNR et MOS_Estimé, indiquant une meilleure utilisation de la perception vidéo. Lorsque le PLR augmente, le PSNR diminue fortement, mais il reste acceptable jusqu'à ce que les valeurs de la perte soient supérieures à 3%. Au-delà de cette valeur, la vidéo sera dégradée. Par conséquent, les consommateurs qui utilisent des vidéos avec pourcentages de pertes élevées peuvent envisager de ne plus les visionner. Les vidéos sont plus sensibles à la perte de données.

Il est évident que lorsque les valeurs du PLR sont $> 3\%$, il existe des différences significatives dans les valeurs de PSNR pour toutes les séquences vidéo. En outre, nous avons noté que, si la vidéo est tellement affectée par PLR, la qualité perçue associée dépend du processus d'encodage HEVC. De plus, les résultats de l'impact de ceux-ci apparaissent clairement sur les graphiques. Par conséquent, le codec HEVC a un notable impact sur le taux de compression et la qualité visuelle des vidéos par rapport à H.264/AVC [16] [23] [49] et MPEG-2 [4] [5].

V.8.8 L'impact de perte des paquets sur le streaming vidéo HEVC Basé sur MDC

Comme mentionné précédemment, dans cette section, nous avons proposé un modèle de QoE utilisant le schéma d'adaptation du codeur HEVC basé sur le MDC, comme le montre la figure V.24. La figure V.26 présente les résultats en termes de PSNR, MOS_Estimé lors de l'application de la technique MDC (dans le domaine temporel) utilisant HM16.11 et transmettant les bitstreams des trames impaires / paires sur le réseau IP sans fil émulé en tant que fichier *.hevc. Ensuite, du côté du décodeur, les séquences vidéo sont reconstruites à partir des trames impaires / paires, comme cela a été indiqué sur la figure III.5(b). De plus, la figure V.26 montre que les variations de PSNR et de MOS_Estimé avec PLR ont des tendances à la baisse similaires dans la majorité des cas.



(a)

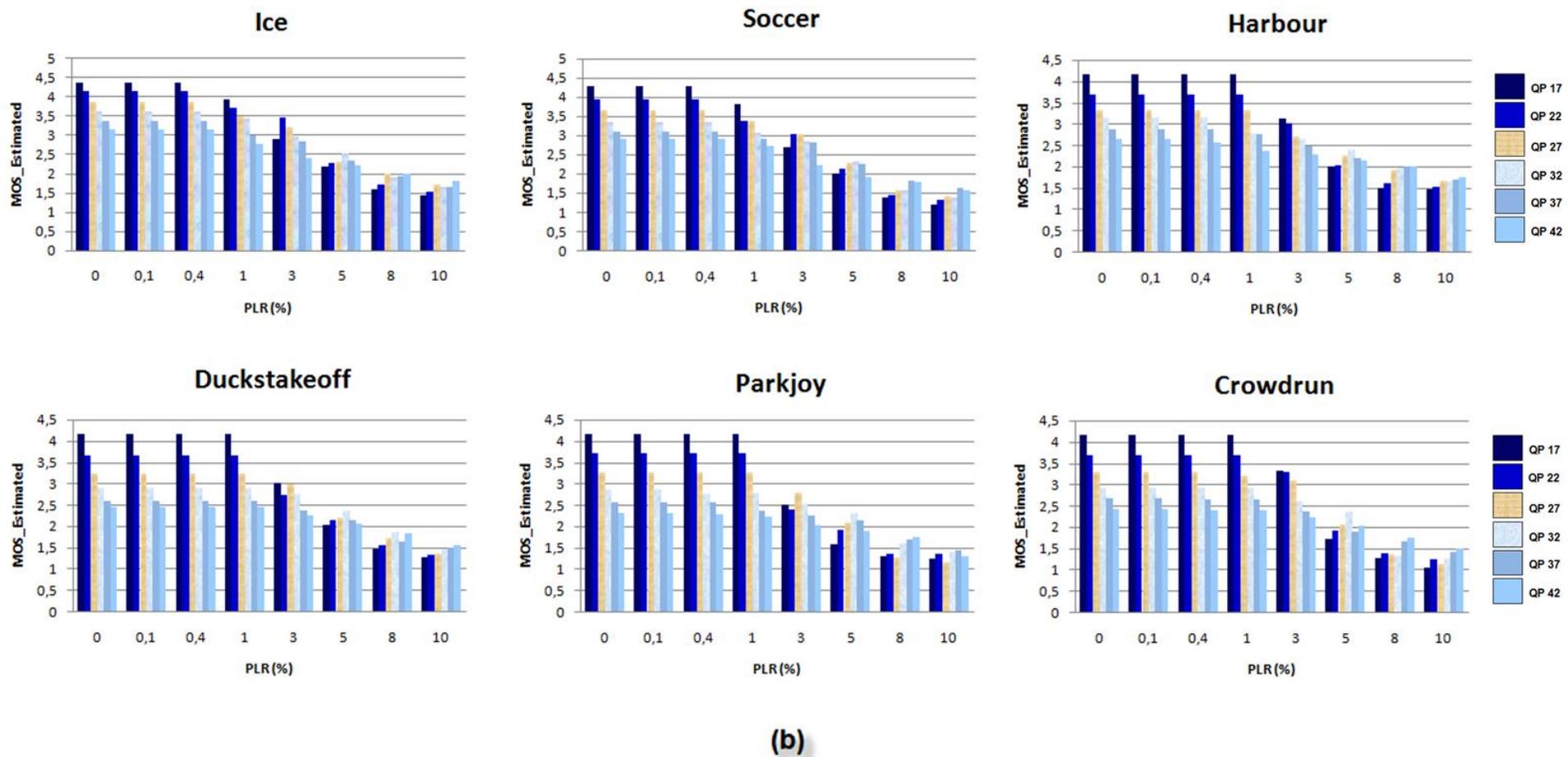
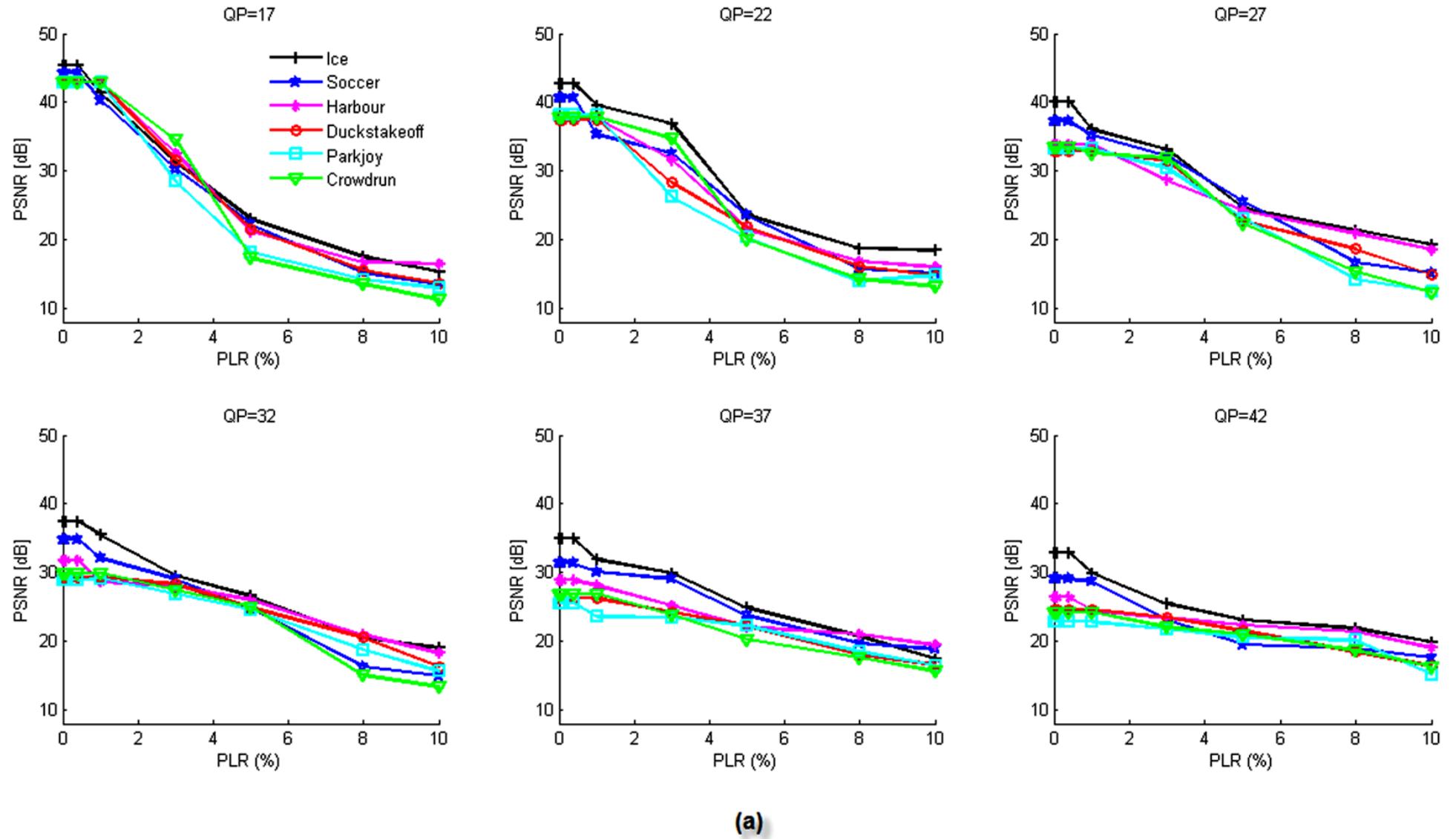


Figure V.25 Impact du PLR sur la qualité perçue sans schéma MDC
 (a) Relation entre PSNR et PLR. (b) Relation entre MOS_Estimé et PLR



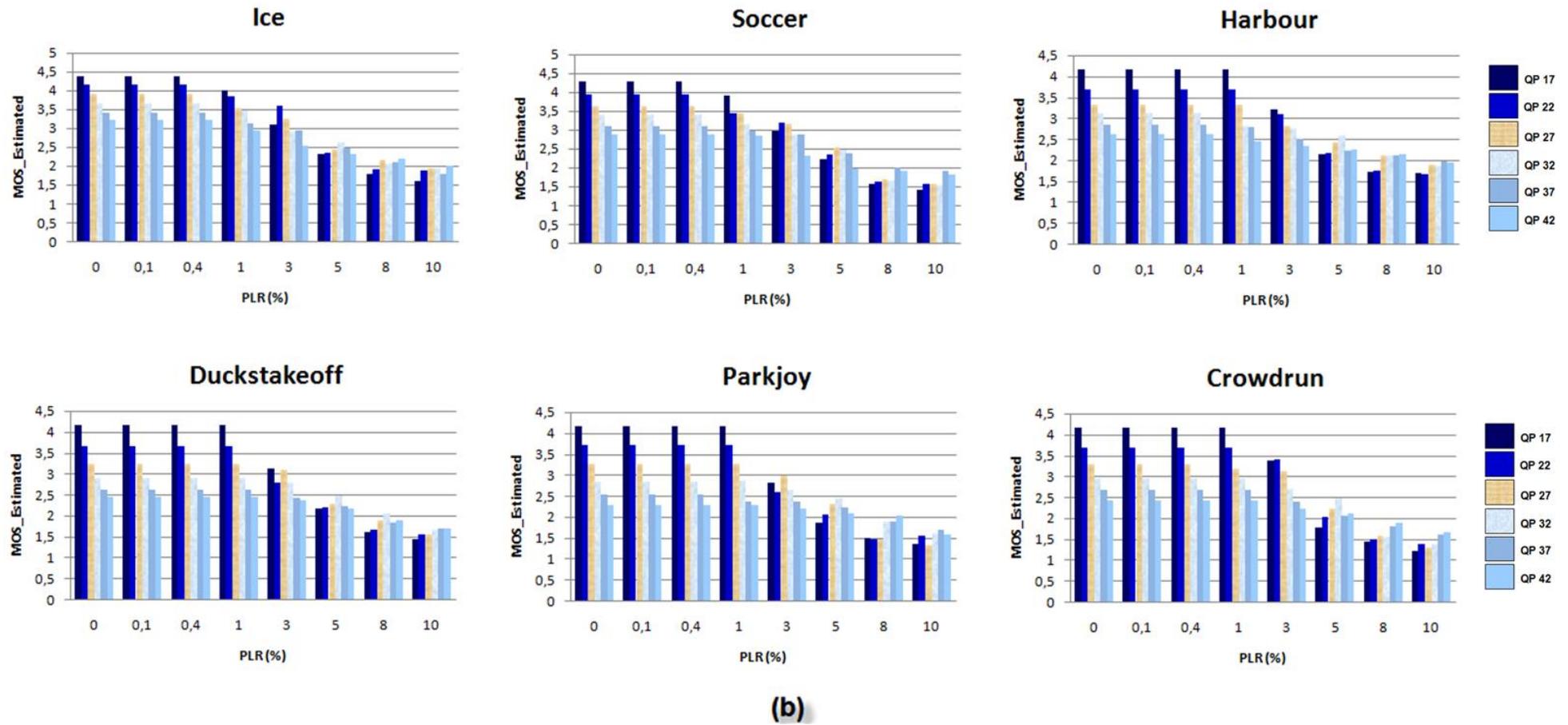


Figure V.26 Impact du PLR sur la qualité perçue avec le schéma MDC
 (a) Relation entre PSNR et PLR. (b) Relation entre MOS_Estimé et PLR.

Les Figures 25 et 26 montrent les effets des paramètres de codage HEVC, tels que QP, le débit binaire associé, le contenu vidéo, etc., avec les variations du PLR sur toutes les séquences. De plus, il y a une différence significative entre les différentes séquences vidéo; ces résultats indiquent que pour les grandes valeurs de PLR, le PSNR et le MOS_Estimé diminuent dans toutes les vidéos.

Généralement, pour un PLR supérieur à 5%, le PSNR baisse à moins de 20 dB, ce qui signifie que le MOS_Estimé est proche de 1 (mauvais) avec des valeurs MOS_Estimé légèrement plus élevées dans la séquence vidéo Ice. Par conséquent, à partir de la figure 26, on peut voir que la qualité perçue se dégrade de manière significative dans toutes les vidéos considérées pour des valeurs élevées du PLR associé à une amélioration significative par rapport au codeur HEVC sans schéma MDC, comme le montre la Figure V.25.

V.8.9 Discussion

Les résultats précédents ont montré que les types de contenu vidéo et QP affectent le PSNR et le MOS_Estimé. En outre, nous avons constaté que le PSNR diminue progressivement avec l'augmentation des valeurs de QP. Ceci, avec une variation significative selon les types de contenu vidéo, qui ont des caractéristiques différentes comme le niveau de mouvement, le contexte, la température, la couleur et les mouvements de caméra (voir figures V.25 et V.26), Ice, Soccer et Harbour sont plus résistants aux PLR que les séquences vidéo Duckstakeoff, Parkjoy et Crowdrun.

Pour illustrer la qualité vidéo avec différents débits selon le PLR, une comparaison a été effectuée et les résultats ont démontré que les flux vidéo à haut débit surpassent les flux à bas débit lorsque le PLR est faible dans les deux cas (SDC et MDC). Tandis qu'avec l'augmentation du PLR, un flux vidéo à haut débit binaire est bien pire que le flux vidéo à faible débit binaire, comme le montrent les Figures V.25 et V.26. Pour obtenir de meilleures performances, le débit doit être choisi en fonction du réseau et de la capacité disponible de la bande passante. Par conséquent, le QP a un impact majeur sur la qualité vidéo et le taux de compression, ce qui signifie que la régulation de la quantité de détails spatiaux est maintenue.

Pour tous les types de contenu vidéo, la qualité vidéo diminue lorsque les valeurs de QP et PLR augmentent. Par conséquent, il existe une relation entre QP, PLR et la dégradation de la qualité. Cela signifie que l'impact du PLR sur la qualité vidéo peut être considéré comme un facteur important d'évaluation de la qualité vidéo tout autant que le QP, le débit binaire, le type de contenu, etc.

Dans ce travail de thèse, nous avons démontré que lorsque le PLR augmente, il peut minimiser l'impact de dégradation sur une métrique de qualité vidéo par une augmentation de QP (réduisant le débit binaire de codage) (voir aussi les figures V.25 et V.26). De plus, le choix d'une valeur QP dans le codeur HEVC représente un facteur essentiel pour le contrôle du débit binaire et la dégradation de la qualité vidéo, comme dans les autres codecs vidéo, par exemple, H.264 / AVC [10] [17] [20] et MPEG-2 [4][5].

D'un autre côté, les vidéos codées avec le même QP et ayant subi le même PLR peuvent avoir des mesures de qualité différentes lorsqu'elles sont évaluées avec la même métrique de qualité, telle que PSNR. En outre, cette étude a montré que pour les plus petites valeurs de PLR (0,1%, 0,4%, 1% et 3%), le PSNR avec ou sans le MDC reste le même, comme le montrent les Figures 25(a) et 26(a). Ainsi que pour les valeurs MOS_Estimé (voir aussi les Figures 25(b) et 26(b)) comme il est démontré dans les résultats théoriques [26] [94]. On peut voir qu'à partir de 3% du PLR, la technique MDC conduit à de meilleurs résultats, indiquant qu'elle est une solution recommandée pour contrer les valeurs les plus élevées du PLR.

Pour prouver l'importance du MDC, nous avons comparé l'impact du PLR sur le streaming vidéo HEVC basé sur le MDC et sans MDC. A cette fin, les résultats peuvent nous aider à réaliser une comparaison entre les valeurs moyennes du PSNR obtenues pour chaque QP pour toutes les vidéos dans les deux cas. La figure V.27 ci-dessous montre que pour les valeurs PLR égales à 0,1%, 0,4%, 1% ou 3%, des valeurs plus élevées du PSNR peuvent être obtenues, alors que le PSNR diminue lorsque le PLR est supérieur à 3%. Le plus grands avantages de MDC est de lutter contre les pertes de paquets comme détaillé dans la Réf [26]. De plus, en présence du MDC, il existe des valeurs plus élevées de PSNR lorsque le PLR augmente. On peut noter que l'utilisation du MDC est un moyen efficace pour améliorer la transmission avec un gain du PSNR. Autrement, les figures V.25, V.26, V.27 démontrent que le SDC a un PSNR beaucoup plus élevé aux débits binaires supérieurs; la qualité vidéo décline rapidement avec l'augmentation du PLR. La technique MDC (les deux descriptions pair et impair sont reçues et fusionnées) peut améliorer la qualité vidéo par rapport au SDC. Pour les valeurs inférieures du PLR, seule une amélioration de 0,3 à 1 dB peut être détectée, mais lorsque le PLR augmente, la dégradation de la qualité vidéo peut être observée et l'approche MDC peut offrir une amélioration de 2 à 3 dB.

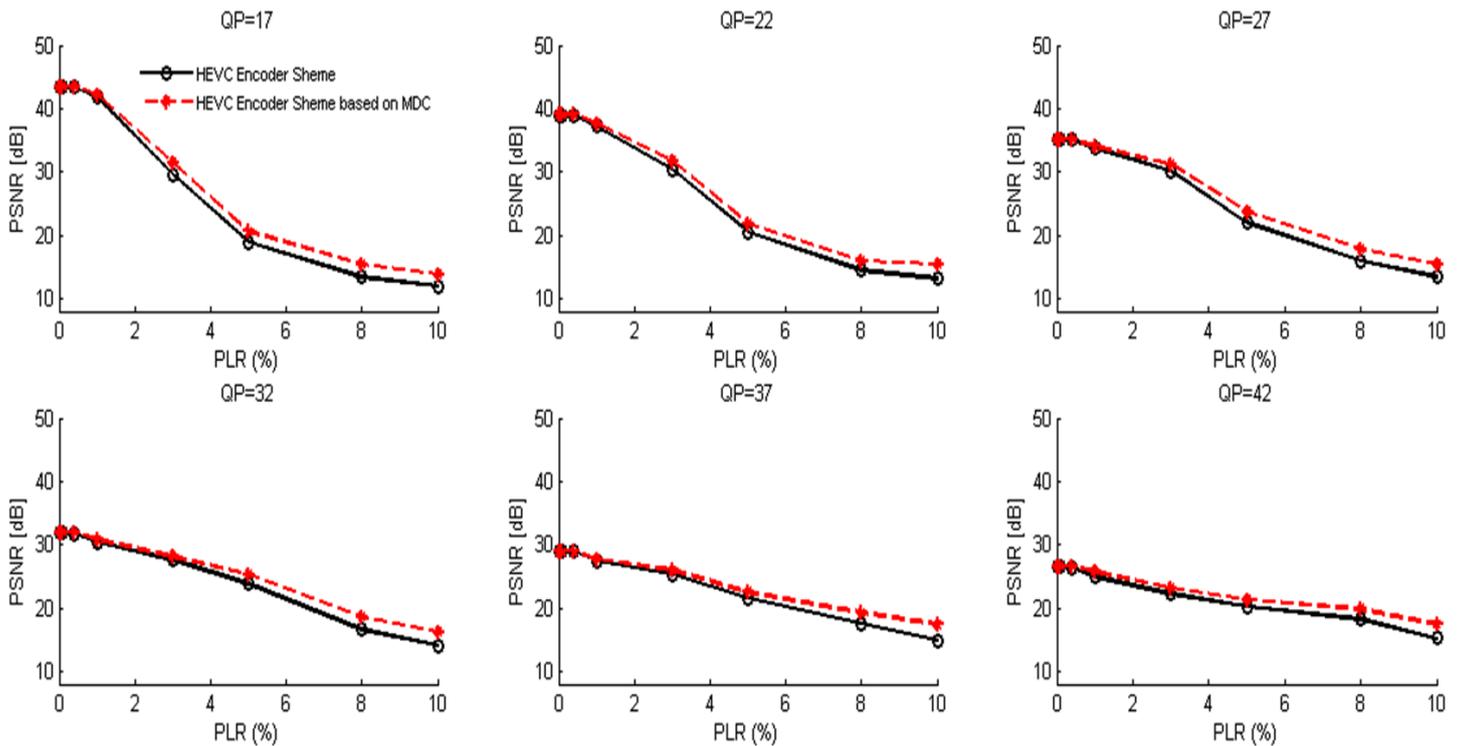


Figure V.27 Impact de QP et PLR sur les moyens PSNR pour toutes les séquences vidéo.

Selon les résultats précédents, MDC surpasse SDC pour des valeurs de PLR élevées en raison de la redondance (les débits binaires supplémentaires introduites par MDC). De plus, pour des PLR faibles, il y a toujours un débit binaire redondant en MDC. Ainsi, la flexibilité et le réglage de la redondance est la tâche clé des schémas MDC (débits binaires redondants). Pour les canaux à faible perte, la redondance n'est pas requise, mais elle doit être élevée dans les canaux à perte élevée. En fait, avec des canaux qui varient dynamiquement, les techniques MDC présentent plus d'intérêt avec un paramètre de redondance élevé. Ainsi, les techniques MDC résistent aux taux de perte variables et donnent de meilleures performances permettant une consommation moindre de la bande passante disponible

D'autre part, selon les figures V.25 et V.26, le contenu vidéo a différentes caractéristiques exprimé par SI et TI. La séquence vidéo Parkjoy qui a le SI et TI les plus élevés est plus sensible au PLR avec une dégradation de qualité supérieure comparativement à la séquence vidéo Soccer qui a le SI le plus élevé. La séquence vidéo Ice, qui a les plus bas TI et SI est plus résisté au PLR avec une qualité supérieure par rapport aux séquences vidéo Parkjoy et Soccer. On peut également observer que les vidéos Parkjoy et Soccer présentent les valeurs les plus faibles de PSNR et MOS_Estimé avec des valeurs légèrement plus élevées dans le cas de la séquence vidéo Ice. De plus, les résultats ont montré que l'effet de SI ne change pas

pour les faibles valeurs de PLR; mais pour les valeurs élevées de PLR, l'altération domine fortement l'effet de SI. Dans le cas de valeurs réduites de PLR, la qualité perceptive est significativement différente pour les vidéos avec divers TI. Cependant, lorsque le PLR augmente, la qualité perceptive se dégrade indépendamment de TI des séquences vidéo.

V.9 Conclusion

Pour comprendre l'impact de contenu vidéo, les paramètres de codage tel que QP, ainsi que la perte de paquets sur la qualité vidéo perçue, on utilise la vidéo compressée par la norme HEVC, en suite nous devons mettre à jour le système en utilisant le système d'exploitation LINUX, qui donne la possibilité d'introduire des effets de pertes dans notre lien d'émulation avec FFmpeg. Nous pouvons nous connecter à un réseau LAN (IP) en utilisant un protocole UDP.

Les résultats montrent qu'en ce qui concerne les petites valeurs de PLR, les valeurs de PSNR, SSIM, MOS_Estimé augmentent, et pour des valeurs élevées de PLR, ces métriques baissent pour les différentes séquences vidéo.

Les différents contenus vidéo n'affecte pas la qualité pour les faibles valeurs du PLR et inversement. Cependant, le SI et TI donne une idée sur la qualité de la vidéo. On peut conclure que les performances de streaming vidéo dépend du choix de l'application et du type de vidéo, un ajustement doit toujours être fait entre les variables: codec, contenu vidéo, débit et QoE.

Conclusion et perspectives

Nous nous trouvons à l'heure actuelle au carrefour de différentes nouvelles technologies. La possibilité de faire transiter la vidéo par les réseaux sans fil, d'une part, et la flexibilité offerte par ces canaux, d'autre part. Le terme QoS caractérise l'aptitude d'un réseau à assurer la transmission des vidéos.

De nos jours, la transmission vidéo sur toutes sortes de réseaux est un domaine qui ne cesse de croître. Les améliorations des propriétés du réseau ainsi qu'une plus grande efficacité des codecs vidéo et une augmentation de la capacité de traitement et de stockage de tous les types d'appareils (ordinateurs de bureau, smartphones, tablettes, ...) permettent de rendre le streaming vidéo réalisable. De nombreuses applications utilisent le streaming vidéo.

L'objectif de cette thèse est l'implémentation d'une vidéo codée par la nouvelle norme H.265/HEVC qui est l'une des normes de codage les plus récentes avec sa structure et ses éléments principaux afin d'avoir les caractéristiques employées dans tout type de système de compression vidéo sur un réseau sans fil Ad Hoc.

En outre, le streaming vidéo est devenu très utile dans la vie quotidienne et aussi dans le domaine professionnel. Dans cette thèse, une méthode simple comme MDC est proposée pour améliorer la performance de la qualité vidéo, qui aurait pu être dégradée pendant la transmission où la QoE est fortement recommandée. En effet, les réseaux sans fil tels que les réseaux ad hoc ont la particularité d'introduire des perturbations et des pertes de paquets.

Le système de codeur HEVC est d'abord appliqué pour coder et gérer la qualité vidéo de bout en bout. Ceci dépend essentiellement du type de contenu vidéo et de QP comme des paramètres d'encodage initiaux, cette méthode offre des meilleurs résultats en termes de PSNR, SSIM et MOS_Estimé. Par ailleurs, l'impact du codec vidéo, du type de contenu vidéo, des différentes valeurs de QP et des dégradations sur la qualité vidéo perçue sont analysés. La qualité perçue dépend davantage des déficiences que du contenu vidéo. Cependant, pour les faibles valeurs de PLR, la qualité vidéo perçue dépend également du contenu vidéo. Cette thèse a démontré que dans le cas où le PLR augmente, il peut minimiser la dégradation de la vidéo en termes de métriques de qualité vidéo en utilisant une QP accrue (réduisant le débit binaire d'encodage). Par conséquent, l'étude comparative montre que pour la transmission de vidéo sur des canaux

sans fil avec perte de paquets plus élevé, la technique MDC offre une solution améliorée pour lutter contre la perte de paquets, ce qui démontre qu'il s'agit d'une méthode robuste de distribution de flux.

Nous nous sommes intéressés également à la modélisation d'une transmission fiable de la vidéo codée en utilisant HEVC basée sur MDC pour les systèmes de communication sans fil. Ceci est très important car il est très difficile de mesurer toute la qualité du canal de transmission telle que la bande passante disponible. En revanche, il est assez aisé de mesurer en temps réel le taux de perte de paquets de la couche application. Ceci est essentiel car il permet la prédiction de la qualité vidéo en temps réel lorsque le type de contenu vidéo, le débit codé et le taux de perte de paquets du réseau sont connus.

Enfin, le nouveau codec HEVC peut être utilisé efficacement pour les applications de streaming vidéo, bien qu'il pose encore certains problèmes qui doivent être préalablement solutionnés. Nous pensons que cela n'est pas dans un futur lointain compte tenu de l'évolution de la technologie.

Certains points dans notre travail peuvent être complétés, dans des futurs travaux :

- Notre application se base particulièrement sur des réseaux LAN. Il est nécessaire de l'extrapoler vers d'autres réseaux mobiles par exemple (4G et 5G). En effet, dans ce type de réseaux mobiles, d'autres problèmes peuvent surgir liés entre autres au délai, à la gigue et aux collisions.
- Les applications « temps réel », de type télévision IP, VOIP pourraient trouver un réel intérêt dans l'adressage IPv6 mais aussi grâce à son potentiel de QoS qui priorise les flux multimédias vidéo, audio.

Pour les travaux futurs, l'effet des dégradations et du contenu vidéo sur la qualité perçue sera davantage examiné en réalisant de nouvelles expériences objectives et subjectives avec plusieurs classes de vidéos de référence 3D, HD et UHD, ... 8K avec un codage vidéo évolutif et un codage vidéo multi-vue sur HEVC/H.265. Nous proposons également en termes de perspectives, d'améliorer la robustesse de la transmission vidéo HEVC en temps réel dans les réseaux ad hoc véhiculaires. En effet, ces canaux sont les plus hostiles à la transmission vidéo caractérisée par une topologie variable, un nombre variable de nœuds, une mobilité variable et rapide, ...etc.

A cet effet, deux points doivent être abordés: streaming vidéo en temps réel et réduction du délai de streaming vidéo. Nous nous intéressons à la transmission vidéo à travers les VANETS pour plusieurs raisons, y compris son utilisation: pour plus de sécurité routière, le divertissement du conducteur et des passagers, la surveillance ... etc. Les réseaux véhiculaires utilisent des réseaux sans fil en particulier le 802.11p, mais il est également possible d'utiliser des réseaux mobiles tels que LTE (Long Term Evolution) voir même la 5G.

Bibliographie

- [1] CISCO, "CISCO visual networking index forecast and methodology, 2016-2021," in White Paper (2017), <http://www.cisco.com>
- [2] J. Nightingale et al., "The impact of network impairment on quality of experience (QoE) in H.265/HEVC video streaming," IEEE Transactions on Consumer Electronics. **60** (2), 242-250 (2014) [doi:10.1109/TCE.2014.6852000].
- [3] C. M. Chen et al., "Robust video streaming over wireless LANs using multiple description transcoding and prioritized retransmission," J. Vis. Commun. Image R. **18**, 191-206 (2007).
- [4] P. Paudyal, F. Battisti, and M. Carli, "Evaluation of the effects of transmission impairments on perceived video quality by exploiting ReTRiEVED dataset," J. Electron. Imaging(SPIE). **26**(2), 023003 (2017) [doi:10.1117/1.JEI.26.2.023003].
- [5] P. Paudyal, F. Battisti, and M. Carli, "Impact of video content and transmission impairments on quality of experience," Multimedia Tools Appl. **75**(23), 16461-16485 (2016) [doi:10.1007/s11042-015-3214-0]
- [6] ITU-T Recommendation E.800, "Definitions of Terms Related to Quality of Service" (2008).
- [7] IETF RFC 2386, "A Framework for QoS-based Routing in the Internet" (1998).
- [8] Z. Cheng et al., "Subjective QoE based HEVC encoder adaptation scheme for multi-user video streaming," in Proc. IEEE Int. Symp. Broadband Multimedia Syst. Broadcast. (BMSB), pp.1-6, Nara, Japan (2016) [doi:10.1109/BMSB.2016.7521987].
- [9] J. Frnda, M.V oznak, and L.Sevcik, "Impact of packet loss and delay variation on the quality of real-time video streaming", TelecommunSyst. **62**,265–275 (2016) [doi:10.1007/s11235-015-0037-2].
- [10] S. Paulikas et al., "The Impact of Packet Loss on Quality of H.264/AVC Video Streaming", ELEKTRONIKA IR ELEKTROTECHNIKA, ISSN 1392-1215.**22**(2), 81-85 (2016) [doi: <http://dx.doi.org/10.5755/j01.eie.22.2.14596>]
- [11] ITU-T Recommendation Y.1901, "Requirements for the Support of IPTV Services", 2009.
- [12] S. Tavakoli et al., "Quality of Experience of adaptive video streaming: Investigation in service parameters and subjective quality assessment methodology,"Signal Processing: Image Communication.**39**, 432-443 (2015) [doi: <http://dx.doi.org/10.1016/j.image.2015.05.001>].
- [13] "Qualinet white paper on definitions of quality of experience (2012)," in European Network on Quality of Experience in Multimedia Systems and Services (COST Action IC 1003), P. Le Callet, S. Moller, and A. Perkis, Eds (2013).
- [14] O. Hadar et al., "Effects of compression parameters on the perceived quality of video stream over a lossy internet protocol network," Optical Engineering.**45**(8), 087003(2006) [doi:<http://dx.doi.org/10.1117/1.2338591>]

- [15] D. Hands, and M. Wilkins, "A study of the impact of network loss and burst size on video streaming quality and acceptability," in *Interactive Distributed Multimedia Systems and Telecommunication Services*, M. Diaz, P. Owezarski, and P. Snac, Eds. Berlin, Heidelberg, Springer Berlin Heidelberg, **1718**, pp, 45-57 (1999) [[doi:10.1007/3-540-48109-5_5](https://doi.org/10.1007/3-540-48109-5_5)].
- [16] L. Atzori et al., "Quality perception when streaming video on tablet devices," *J. Vis. Commun. Image R.* **25**, 586–595 (2014).
- [17] F. Tommasi, V.D. Luca, and C. Melle, "Packet losses and objective video quality metrics in H.264 video streaming," *J. Vis. Commun. Image R.* **27**, 7-27 (2015).
- [18] B. Maia, H C. Yehia, and L. De Errico, "A concise review of the quality of experience assessment for video streaming," *Computer Communications.***57**, 1-12 (2015).
- [19] L. Pozueco et al., "Subjective video quality evaluation of different content types under different impairments," *New Review of Hypermedia and Multimedia*, 1-28 (2016) [[doi:10.1080/13614568.2016.1152310](https://doi.org/10.1080/13614568.2016.1152310)].
- [20] G. Kokkonis, K. E. Psannis, M. Roumeliotis and D. Schonfeld, "Real-time wireless multisensory smart surveillance with 3D-HEVC streams for internet-of-things (IoT)," *Joural of Supercomputing.* **73**(3), 1044-1062 (2017) [[doi:http://link.springer.com/article/10.1007%2Fs11227-016-1769-9](http://link.springer.com/article/10.1007%2Fs11227-016-1769-9)].
- [21] Z. Milicevic, and Z. Bojkovic, "High Efficient Video Coding (HEVC) performance analysis for different configurations in main profile," *International Journal of Communications.***10**, 109-112 (2016). Also, presented as Invited Paper at INASE International Conference Communications and Computers CC2016, Rome, Italy, 5-7 November (2016).
- [22] L. Anegekuh, L. Sun, and E. Ifeakor, "Encoding and video content based HEVC video quality Prediction," *Multimed Tools Appl.* **74**, 3715-3738 (2015) [[doi:10.1007/s11042-013-1795-z](https://doi.org/10.1007/s11042-013-1795-z)].
- [23] M. Kazemi, R.Iqbal, and S. Shirmohammadi, "Redundancy allocation based on the weighted mismatch-rate slope for multiple description video coding," *IEEE Trans on Multimedia.* **19**(1), 54-66 (2017) [[doi: 10.1109/TMM.2016.2607342](https://doi.org/10.1109/TMM.2016.2607342)].
- [24] M. Pereira, M. Antonini, and M. Barlaud, "Multiple description image and video coding for wireless channels," *Signal Processing: Image Communication.***18**, 925-945 (2003) [[doi:10.1016/j.image.2003.08.009](https://doi.org/10.1016/j.image.2003.08.009)].
- [25] W. Zang, and R. Guo, "Enhancing the quality of service of video streaming over MANETs using MDC and FEC," *Proc. SPIE***8334**, Fourth International Conference on Digital Image Processing (ICDIP 2012), 83343T (2012) [[doi: 10.1117/12.966823](https://doi.org/10.1117/12.966823); <http://dx.doi.org/10.1117/12.966823>].
- [26] M. Kazemi, H. Sadeghi, and S. Shirmohammadi, "A Review of Multiple Description Coding Techniques for Error-Resilient Video Delivery," *Multimedia Systems.***20**, 283-309 (2014) [[doi:10.1007/s00530-013-0319-z](https://doi.org/10.1007/s00530-013-0319-z)].
- [27] C. Rafael Gonzalez, E. Richard. L Woods and E. Steven., "Digital Image Processing Using MATLAB", pp. 12-15, Pearson Prentice-Hall (2004).

- [28] Ahmed Ben Attallah, « Etude et Implantation d'Algorithmes de Compression d'Images dans un Environnement Mixte Matériel et Logiciel », thèse de doctorat en électronique, Université Bordeaux I (2007).
- [29] The Joint Photographic Experts Group (JPEG) committee: <https://jpeg.org/jpeg/index.html> (2017).
- [30] MPEG Committee Documentation, <http://www.chiariglione.org> (2017).
- [31] Iain E. Richardson., "The H.264 advanced video compression standard", 2nd edition, ISBN: 978-0-470-51692-8, John Wiley & Sons , (2010).
- [32] Philippe Bellaïche « Les secrets de l'image vidéo », 8ème édition, Eyrolles, ISBN 13 978-2-212-14212-9 (2015).
- [33] Le codage des images et de la couleur : <http://obligement.free.fr/articles/codagedesimages.php> (2018).
- [34] K R Rao, Do Nyeon Kim, and Jae Jeong Hwang, "Video coding standards , AVS China, H.264/MPEG-4 PART 10, HEVC, VP6, DIRAC and VC-1" ISBN 978-94-007-6742-3, Springer Netherlands (2014) [doi: [10.1007/978-94-007-6742-3](https://doi.org/10.1007/978-94-007-6742-3)]
- [35] J.Weiss, "Télévision: signal vidéo ", Supélec, campus de rennes (1998), www.rennes.supelec.fr/ren/perso/jweiss/tv/signal/tv_opt.pdf
- [36] R.Westwater, B.Furht and B.Furht, "Real-Time Video Compression: Techniques and Algorithms,"Kluwer Academic Publishers Boston, Dordrecht, London (1997), ISBN: 0792397878. [DOI :[10.1007/b102475](https://doi.org/10.1007/b102475)]
- [37] Iain E. G. Richardson,"H.264 and MPEG-4 Video Compression. Video Coding for Next-generation Multimedia", Wiley, pp.152-222 (2003)
- [38] WikiPedia, Télévision analogique, [shttps://fr.wikipedia.org/wiki/Télévision#En_analogique](https://fr.wikipedia.org/wiki/T%C3%A9l%C3%A9vision#En_analogique) (2018).
- [39] Q Yun. Shi and Sun Huifang, "Image and Video Compression for Multimedia Engineering, Fundamentals, Algorithms, and Standards," CRC Press, USA (2000).
- [40] Les définitions des TV : <https://www.cnetfrance.fr/produits/sd-hd-ultra-hd-4k-8k-comprendre-les-definitions-des-tv-39786402.htm> (2018)
- [41] F. Bossen et al., "HM reference software homepage," https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware (2018).
- [42] X265 homepage: <https://bitbucket.org/multicoreware> (2018)
- [43] Libde265 homepage : <https://www.libde265.org> (2017)
- [44] D.Legros, J. Crinon and P.Georget, "Les effets des systèmes et des outils multimédias sur la cognition, l'apprentissage et l'enseignement," <https://edutice.archives-ouvertes.fr/edutice-00000351> (2018)
- [45] Christoph Meinel, Harald Sack, "Digital Communication," Springer-Verlag Berlin Heidelberg, ISBN 978-3-642-54331-9 (2014), [doi: [10.1007/978-3-642-54331-9](https://doi.org/10.1007/978-3-642-54331-9)]

- [46] H.263++ : H.263 Annex U, V, W and X. Compléments de la recommandation H.263 à l'UITT. (2000).
- [47] G. J. Sullivan et al., "Overview of the high efficiency video coding (HEVC) standard," IEEE Trans Circuits Syst. Video Technol. **22** (12), 1649–1668 (2012).
- [48] Vivienne Sze, Madhukar Budagavi and Gary J. Sullivan, "High Efficiency Video Coding (HEVC) Algorithms and Architectures," Springer International Publishing, ISSN 1558-9412, 2014, [doi: 10.1007/978-3-319-06895].
- [49] S. Percheron, and J. Vieron, "HEVC, the key to delivering an enhanced television viewing experience "Beyond HD," SMPTE 2013 Annual Technical Conference & Exhibition (2013).
- [50] Ph. Hanhart et al., "Subjective quality evaluation of the upcoming HEVC video compression standard," Proc. SPIE **8499**, Applications of Digital Image Processing XXXV, 84990V (2012) [doi:10.1117/12.946036].
- [51] ITU-T Home: Study groups: ITU-T Recommendations: ITU-T H.265 (V2) (10/2014)". ITU. 29-10-2014 (2014)
- [52] WikiPedia, https://en.wikipedia.org/wiki/High_Efficiency_Video_Coding#Versions (2018)
- [53] B. Bross et al., "High Efficiency Video Coding (HEVC) text specification draft 10," Joint Collaborative Team on Video Coding (JCT-VC), Geneva (Switzerland), Tech. Rep. JCTVC-L1003 (2013).
- [54] CODE: Sequoia Website on HEVC <https://codesequoia.wordpress.com/2012/10/28/hevc-ctu-cu-ctb-cb-pb-and-tb/> (2017).
- [55] Mathias Wien, "High Efficiency Video Coding, Coding Tools and Specification", Springer-Verlag Berlin Heidelberg, ISSN 1860-4862 (2015) [doi:10.1007/978-3-662-44276-0]
- [56] G.Pastuszak, and M. Trochimiuk, "Algorithm and architecture design of the motion estimation for the H.265/HEVC 4K-UHD encoder," J Real-Time Image Proc (2016) **12**: 517-529. <https://doi.org/10.1007/s11554-015-0516-4>
- [57] K. Misra, J. Zhao, and A. Segall, "Lightweight slicing for entropy coding," JCTVC-D070 (2011).
- [58] G.Fuldseth., B.jontegaard, and M. Budagavi, "CE10: Core transform design for HEVC," JCTVC-G495 (2011).
- [59] J. Han, A. Saxena, and K. Rose, "Towards jointly optimal spatial prediction and adaptive transform in video/image coding," in Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP), pp.726–729 (2010).
- [60] S.Harize, "Etude et implémentation d'un codeur vidéo sur un système embarqué de type FPGA: Applications aux transformations orthogonales du codeur H264/AVC," thèse de doctorat en Science, Université badji mokhtar –Annaba (2014).
- [61] J. Lainema, K. Ugur, and A. Hallapuro, "Single entropy coder for HEVC with a high throughput binarization mode," JCTVC-G569, (2011).

- [62] T.Mahsa et al, "HEVC, the new gold standard for video compression", IEEE consumer electronics magazine, (2012).
- [63] C.-M. Fu, C.-Y. Chen, Y.-W. Huang, and S. Lei, "Sample adaptive offset for HEVC," in Proc.Int. Workshop on Multimedia Signal Processing (MMSP) (2011).
- [64] C.-Y. Tsai, C.-Y. Chen, C.-M. Fu, Y.-W. Huang, and S. Lei, "One-pass encoding algorithm for adaptive loop filter in high-efficiency video coding," in Proc. Visual Communications and Image Processing (VCIP) (2011).
- [65] Albanese A, Blomer J, Edmonds J, Luby M, Sudan M, "Priority encoding transmission. Information Theory," IEEE Transactions on **42** (6):1737-1744 (1996) [doi:10.1109/18.556670]
- [66] R.Puri, K.Ramchandran, "Multiple description source coding through forward error correction codes". Paper presented at the 33rd Asilomar Conf.Signals, Systems and Computers, CA (1999).
- [67] Guy Pujolle, Initiation aux réseaux, Editions Eyrolles 61, Bld Saint-Germain 75240 Paris cedex 05, www.editions-eyrolles.com, Éditeur délégué: Olivier Salvatori
- [68] TCP/IP, <http://www.frameip.com/osi/>, par Sylvain, Copyright (2011-2014)
- [69] Guy Pujolle, "Cours réseaux et télécoms Avec exercices corrigés," Avec la contribution de Olivier Salvatori, 3 e édition © Groupe Eyrolles,(2008).
- [70] Ph. Atelin, " Wi-Fi Réseaux sans fil 802.11 : Technologie - Déploiement – Sécurisation," 2 édition © Eni editions,(2008).
- [71] C.Loison, T.Ruocco, C.Rives, "routage multicast dans les réseaux véhiculaires (VANET)," Université d'AVIGNON, année 2012-2013, disponible sur: http://www.projetvanet.sitew.com/fs/Root/9kbf1-Rapport_final_VANET.pdf
- [72] Mobile Ad hoc Networking (MANET): Routing Protocol Performance Issues and Evaluation Considerations: <https://tools.ietf.org/html/rfc2501> (2016)
- [73] Su, GM., X, Su, Y, Bai, et al., "QoE in video streaming over wireless networks: perspectives and research challenges," Wireless Netw (2016) **22**: 1571. <https://doi.org/10.1007/s11276-015-1028-7>
- [74] ITU-T recommandation G.1080 "Quality of experience requirements for IPTV services," (2008)
- [75] S. Möller and A. Raake (eds.), "Quality of Experience", T-Labs Series in Telecommunication Services, Springer International Publishing (2014) [doi: 10.1007/978-3-319-02681-7].
- [76] M.Alreshoodil and J.Woods, "Survey on QoE/QoS correlation models fo rmultimedia services," International Journal of Distributed and Parallel Systems (IJDPS).**4**(3), (2013) [doi: 10.5121/ijdps.2013.4305 53].
- [77] P. Sylvain, "Le modèle OSI," (2003), disponible sur : <http://www.frameip.com/osi/> (2017),
- [78] Farouk Boumehrez, Radhia Brai, Noureddine Doghmane, Khaled Mansouri, "Quality of experience enhancement of high efficiency video coding video streaming in wireless packet networks using multiple description coding," Journal of Electronic Imaging **27**(1), 013028 (27 February 2018). <https://doi.org/10.1117/1.JEI.27.1.013028>

- [79] V. Menkovski, “Computational Inference and Control of Quality in Multimedia Services,” Springer Theses, [doi:10.1007/978-3-319-24792-2]
- [80] K. McCann et al., “High efficiency video coding (HEVC) test model 11 (HM11) encoder description,” ITU-T/ISO/IEC joint collaborative team on video coding (JCT-VC) document JCTVC-M1002 (2013).
- [81] P. Dymarski, S. Kula, and T. N. Huy, “QoS conditions for VoIP and VoD,” *Journal of Telecommunications and Information Technology*, 3, 29–37 (2011).
- [82] J.Frnda, M.Voznak, L. Sevcik “Prediction Model of Triple play services for QoS Assessment in IP Based Networks,” *Journal of networks*, 10(4) (2015).
- [83] S. Winkler, P. Mohandas, The evolution of video quality measurement: from PSNR to hybrid metrics, *IEEE Trans. Broadcast.* 54 (3), 660–668 (2008)
- [84] Z. Wang, A. Bovik, H. Sheikh, E. Simoncelli, Image quality assessment: from error visibility to structural similarity, *IEEE Trans. Image Process.* 13 (4) , 600–612 (2004)
- [85] T. De Pessemier et al., “Quantifying subjective quality evaluations for mobile video watching in a semi-living lab context,” *IEEE Trans. Broadcast.* 58 (4), 580-589 (2012).
- [86] Catellier et al., “Impact of mobile devices and usage location on perceived multimedia quality,” in *Fourth Int. Workshop on Quality of Multimedia Experience (QoMEX)*, 39-44 (2012).
- [87] Z. Wang et al., “Objective video presentation QoE predictor for smart adaptive video streaming,” *Proc. SPIE 9599*, 95990Y (2015).
- [88] Q.Irondi,. Wang, and C. Grecos, “Subjective evaluation of H.265/HEVC based dynamic adaptive video streaming over HTTP (HEVC-DASH),” *Proc. SPIE 9400*, 94000B (2015).
- [89] VQMT, homepage, <http://mmspg.epfl.ch/vqmt> (2016)
- [90] <http://www.comlab.uniroma3.it/retrieved.htm> (2017)
- [91] <http://vqa.como.polimi.it/sequences.htm>,” (2014).
- [92] Fliegel, “Qualinet multimedia databases v5.5,” Czech Technical University in Prague, Faculty of Electrical Engineering, Multimedia Technology Group, <http://dbq-wiki.multimediatech.cz/> (2014).
- [93] F. DeSimone et al., “Subjective quality assessment of H.264/AVC video streaming with packet losses,” *EURASIP J. Image Video Process*, 190431 (2011).
- [94] P. Topiwala et al. “Performance comparison of AV1, HEVC, and JVET video codecs on 360 (spherical) video,” *Proc. SPIE10396*, *Applications of Digital Image Processing XL*, 1039609 (2017) [doi: 10.1117/12.2272119;http://dx.doi.org/10.1117/12.2272119].
- [95] G. Esakki, “Dynamic Switching of GOP Configurations in High Efficiency Video Coding (HEVC) using Relational Databases for Multi-objective Optimization,” thesis of Science Computer Engineering, electrical and Computer Engineering (2014).
- [96] F. Bossen, “Common HM test conditions and software reference configurations,” Document JCTVC-H1100, JCT-VC, San Jose, California (2012).

- [97] M. Hussain and A. Hameed, “Adaptive video-aware forward error correction code allocation for reliable video transmission,” *Signal Image Video Process.* 12(1), 161–169 (2017).
- [98] C Lin et al., “Multiple description coding for stereoscopic videos with stagger frame order,” *IEEE Trans. Circuits Syst. Video Technol.* 25(6), 1016–1025 (2015).
- [99] “FFmpeg Developers. ffmpeg tool (Version be1d324) [Software],” <http://ffmpeg.org/> (2016).
- [100]FFmpeg [Wiki ubuntu-fr] - Documentation Ubuntu, <https://doc.ubuntu-fr.org/ffmpeg> (2017)
- [101] Network emulation with NetEM,”
<http://www.linuxfoundation.org/collaborate/workgroups/networking/netem> (2016).
- [102] A. Floris et al., “QoE Assessment of Multimedia Video Consumption on Tablet Devices,” GC'12 Workshop: Quality of Experience for Multimedia Communications, IEEE (2013) [doi: 10.1109/GLOCOMW.2012.6477775]