

UNIVERSITÉ DU QUÉBEC

MÉMOIRE PRÉSENTÉ À
L'UNIVERSITÉ DU QUÉBEC À TROIS-RIVIÈRES

COMME EXIGENCE PARTIELLE
DE LA MAÎTRISE EN PHYSIQUE

PAR
PASCAL NSAME

MODÉLISATION D'UN SYSTÈME DE TRANSMISSION
D'IMAGES ANIMÉES SUR RÉSEAU SANS FIL

Hiver 97

Université du Québec à Trois-Rivières

Service de la bibliothèque

Avertissement

L'auteur de ce mémoire ou de cette thèse a autorisé l'Université du Québec à Trois-Rivières à diffuser, à des fins non lucratives, une copie de son mémoire ou de sa thèse.

Cette diffusion n'entraîne pas une renonciation de la part de l'auteur à ses droits de propriété intellectuelle, incluant le droit d'auteur, sur ce mémoire ou cette thèse. Notamment, la reproduction ou la publication de la totalité ou d'une partie importante de ce mémoire ou de cette thèse requiert son autorisation.

À mes parents M. Époté et L. Sunday Nfor

“Notre faim de connaître
est un feu toujours ardent.
Le vent de notre savoir
souffle et l’attise davantage.”

Kaydara: Réponse des initiés en Kestelout; Littérature africaine, Nathan, 1971

Remerciements

Rendons à César...

Je tiens à remercier:

- mon directeur de recherche Guy Bégin du département d'informatique à l'Université du Québec à Montréal qui, tout au long de la recherche, a su me soutenir sans douter de mes capacités de travail. Il m'a permis pendant l'élaboration du projet, d'acquérir le sens de la recherche et de la critique. Chaque rencontre avec lui était pour moi une source d'enthousiasme et de motivation.
- Lamiaa Msalka, pour toute l'attention et l'affection dont elle a su faire preuve pendant toute la durée de mes études. Ma reconnaissance est sans limite.
- Viviane et Joseph N., Michel S., Judith S., Dione et Emile M., pour le soutien familial inconditionnel et toujours présent.
- Mounir Boukadoum du département d'informatique à l'UQAM qui m'a apporté le soutien et l'encouragement nécessaire pour entreprendre et terminer mes études en Microélectronique et en Télécommunication à l' UQAM.
- Élie Boridy, J. C. Zanghi du département de Physique, Jacob Davidson, François Gros-d'Aillon, du département d'Informatique de l'UQAM, pour toute la patience dont ils ont su faire preuve pendant les rencontres interminables suscitées par mes études en maîtrise et au baccalauréat. Leur collaboration fut sans limites.
- Louis Machildon et le reste du département de Physique de l'Université du Québec à Trois-Rivières qui m'ont fourni l'encadrement nécessaire pour mener à bien mes études de deuxième cycle.

RÉSUMÉ

Le succès commercial obtenu avec l'introduction des radio-téléphones cellulaires mobiles et le développement des réseaux locaux sans-fil a généré un intérêt important pour la transmission des images vidéo via un canal sans-fil. Toutefois ce canal est aléatoire et non-stationnaire dans le temps. Donc l'effet d'une erreur de transmission varie de façon significative d'un bit à l'autre. Dans cette étude, nous avons modélisé et simulé un système de transmission d'images vidéo comprimées suivant la norme MPEG via un canal sans-fil intra-édifice. Pour développer une telle application, on doit considérer de nombreux facteurs: le codage efficace de la source d'information, l'allocation dynamique de la protection contre les erreurs de transmission, les évanouissements dans le canal et les paramètres vidéo qui affectent la qualité des images. L'utilisation d'une technique de correction d'erreurs sur les données vidéo comprimées est aussi étudié. Nous avons classifié les images vidéo comprimées en fonction de leur contenu. Les bits les plus importants sont transmis avec un niveau de protection plus élevé. Cette stratégie de protection adaptative a été réalisé au moyen des codes convolutionnels perforés à taux variable. La performance de la stratégie a été évaluée par simulation d'un canal à évanouissement en utilisant des signaux modulés par déplacement en quadrature de phase. L'évaluation de la performance d'ensemble est basée à la fois sur des critères de mesures objectifs de la qualité des images reproduites, l'erreur quadratique moyenne, le rapport signal sur bruit d'images à la réception et sur une évaluation subjective. Les résultats obtenus montrent qu'un critère unique n'existe pas, l'évaluation subjective et l'évaluation objective considérées séparément n'étant pas satisfaisantes. Nous avons démontré qu'avec une stratégie de protection adaptative, on obtient des images vidéo donc la qualité est comparable à celle d'images obtenues avec une stratégie de protection conventionnelle non-adaptative, mais en faisant un usage moindre de la redondance dans la transmission. Donc nous pouvons conclure que la stratégie de protection adaptative est plus efficace.

ABSTRACT

Due to widespread increase of Wireless-Local Area Network (W-LAN) and cellular radio subscribers, the demand for mobile image transmission over a multipath fading channel is increasing rapidly. However, the channel is generally very unreliable and produces burst errors due to the time-varying fading effect. Thus the consequence of a channel error varies significantly from one bit to another. In this work, we consider the design and simulation of an MPEG video transmission system over a wireless indoor channel. We address several problems that arise in developing such an application and provide possible solutions to them. These problems include: efficient source coding, dynamic allocation of channel protection against transmission errors, channel fading, and video related issues such as parameters that affect image quality. The effect of conventional Forward Error Correction (FEC) on highly data compressed moving images transmission is also studied. We classified the compressed moving images bits into classes of different importance, those bits in the more important class are transmitted with higher priority. We used the unequal error protection capabilities of convolutional codes belonging to the family of Rate-Compatible Punctured Convolutional Code (RC-PCC). The performance of these codes is simulated for a multipath fading channel with Quadrature Phase Shift Keying modulation (QPSK). The evaluation of the overall system performance was based on the computation of the Peak Signal-to-Noise Ratio (PSNR) and the Mean Square Error (MSE), and on subjective evaluation. The simulation results indicate that the evaluation of the overall image quality is more accurate when both a subjective and an objective measures are considered. Based on computer simulation, we have confirmed that, with limited channel coding redundancy the proposed scheme achieves a quality of image transmission comparable to conventional FEC. Thus the simulation results demonstrate that the unequal error protection technique used in the thesis is an attractive solution for the transmission of MPEG-like video over a time-varying fading channel.

TABLE DES MATIÈRES

REMERCIEMENTS	iii
RÉSUMÉ	iv
ABSTRACT	v
LISTE DES TABLEAUX	viii
LISTE DES FIGURES	viii
LISTE DES ABRÉVIATIONS	x
INTRODUCTION	1
Contributions et méthodologie	2
CHAPITRE 1: Système de communication vidéo intra-édifice	
1.1 Définition d'un système de communication vidéo.....	4
1.2 Composantes d'un système de communication vidéo	5
1.3 Étapes de conception du système de communication vidéo	9
1.4 Caractéristiques d'un système de communication vidéo	12
1.5 Applications d'un système de communication vidéo	13
CHAPITRE 2: Caractérisation, perception et représentation des images	
2.1 Système visuel humain	16
2.1.1 Colorimétrie, Photométrie et Radiométrie	18
2.2 Système d'acquisition et d'affichage	19
2.3 Modèle de l'information visuelle	22
2.3.1 Modèles déterministes	22
2.3.2 Modèles stochastiques	23
2.4 Filtrage dans les systèmes vidéo	24
2.4.1 Échantillonnage multi-dimensionnel spatial et temporel	26
2.4.2 Filtrage temporel	28
2.4.3 Filtrage spatial	30
2.4.4 Conversion des standards (structure d'échantillonnage)	31
2.4.5 Reconstruction d'un signal vidéo échantillonné	31
CHAPITRE 3: Analyse et caractérisation de l'encodeur et du décodeur MPEG-2	
3.1 L'Algorithme de compression MPEG-2	34
3.1.1 Format des données vidéo	35
3.1.2 Transformée en cosinus discrète (TCD)	39

3.1.3	Encodeur MPEG-2	42
3.1.4	Décodeur MPEG-2	47
3.2	Structure hiérarchique des données encodées MPEG-2	48
3.3	Techniques de masquage d'erreur du CODEC	51
3.3.1	Codage pyramidal	51
3.3.2	Prédiction temporelle robuste	51
3.3.3	Prédiction spatiale robuste	51

CHAPITRE 4: Modèle de canal et protection contre les erreurs de transmission

4.1	Phénomènes de propagation par chemins multiples	53
4.1.1	Caractérisation de la réponse impulsionnelle.....	56
4.1.2	Délai maximum et étalement de délai	59
4.1.3	Atténuations par perte de transmission	59
4.2	Performance des réseaux de campus	61
4.3	Codage de canal	63
4.3.1	Codes convolutionnels	66
4.3.2	Codes convolutionnels perforés	70
4.4	Hypothèses de simulation et classification des données	74
4.4.1	Stratégie de protection adaptative	76
4.5	Modulation	78
4.5.1	Calcul du gain de codage	80

CHAPITRE 5: Résultats et discussions

5.1	Paramètres du canal de communication	82
5.2	Paramètres de simulation et de codage	82
5.3	Mesures de la qualité des images décodées	87
5.4	Discussions	89
5.5	Améliorations possibles du système de communications vidéo	98

CONCLUSION.....100

RÉFÉRENCES.....103

APPENDICE

A. Modèles de simulation du système de communication108

Liste des Tableaux

Tableaux		Pages
3.1	Image en format CCIR 601 (60 Hz) et sous-échantillonnage des composantes de couleurs.....	36
3.2	Gain en complexité d'une TCD rapide.....	42
4.1	Quelques valeurs typiques de η	60
4.2	Métriques du décodeur de Viterbi à décisions pondérées.....	71
4.3	Répartition des bits d'information par rapport au bits de protection.....	78
4.4	Correspondance entre bits et symboles transmis	79
5.1	Caractéristiques de la trame I.....	84
5.2	Caractéristiques de la trame P.....	84
5.3	Caractéristiques de la première trame B.....	85
5.4	Caractéristiques de la deuxième trame B.....	85
5.5	Statistiques sur la séquence d'images originale.....	86
5.6	Qualité des images avant la transmission.....	91
5.7	Qualité des images après la transmission.....	92

Liste des Figures

Figures		Pages
1.1	Schéma simplifié du système de transmission d'images animées.....	5
1.2	Éléments du système de communication vidéo.....	6
1.3	Étapes de conception du système de communication vidéo proposé.....	10
1.4	Systèmes de communication sans-fil.....	12
1.5	Applications de la compression de signaux vidéo en fonction du débit.....	13
2.1	Réponse spatiale et temporelle du SVH.....	18
2.2	SVH par rapport à la photométrie.....	19

2.3	Acquisition et affichage.....	21
2.4	Correction Gamma.....	22
2.5	Échantillonnage d'un signal continu dans le temps.....	27
2.6	Spectre d'un signal continu 2-D et spectre du signal échantillonné.....	28
2.7	Filtrage réel par le SVH et filtrage rectangulaire.....	29
2.8	Spectre des composantes YUV d'un signal vidéo PAL.....	30
2.9	Balayage entrelacé et balayage progressif.....	32
3.1	Macrobloc en format 4:2:0(MPEG-1), 4:2:0(MPEG-2) et 4:2:2 (CCIR-601).....	36
3.2	Segmentation de l'image en champs, trames, tranches et macroblocs.....	38
3.3	Exemple de TCD, de quantification et de TCD inverse.....	41
3.4	Procédure d'encodage vidéo.....	43
3.5	Encodeur MPEG-2.....	45
3.6	Ordre de transmission et de réception des trames I, P et B.....	47
3.7	Décodeur MPEG-2.....	48
3.8	Structure hiérarchique des données encodées MPEG-2.....	49
4.1	Modèle mathématique du canal intra-édifice.....	53
4.2	Réponse impulsionnelle du canal pour un édifice de taille moyenne.....	54
4.3	Modèle géométrique de propagation radio dans un canal intra-édifice.....	55
4.4	Séquences des profils pour des points adjacents dans l'espace.....	56
4.5	Le modèle de réponse du canal.....	58
4.6	Classification de quelques codes de contrôle d'erreurs.....	64
4.7	Codeur par blocs et codeur convolutionnel.....	65
4.8	Codeur convolutionnel.....	66
4.9	Décodage de Viterbi pour $K = 3$ et $R = 1/3$	68
4.10	Seuils de décisions des décodeurs à décisions pondérées et dures	70
4.11	Exemples de patrons de perforation.....	73
4.12	Classification des données vidéo suivant un critère d'importance.....	75

4.13	Diagramme bloc du système de transmission.....	76
4.14	Illustration du principe de perforation.....	78
4.15	Diagramme d'un signal modulé par déplacement en quadrature de phase(QPSK).....	80
5.1	Performance de la stratégie de protection non-adaptative (canal intra-édifice - modèle de Saleh).....	89
5.2	PSNR vs Eb/No des trames I de la séquence "short" pour la composante Y.....	93
5.3	Images originales.....	94
5.4	Images décodées.....	94
5.5	Images décodées avec trop d'erreurs de transmission.....	95
5.6	Images décodées avec la stratégie de protection adaptative.....	97

Liste des abréviations

ADSL:	Asymmetric Digital Subscriber Loop
ATM:	Asynchronous Transfer Mode
CCD:	Charged Coupled Device
CCITT:	Comité consultatif international télégraphique et téléphonique
CCIR:	Comité consultatif international des radiocommunications
CD-I:	Compact Disc Interactif
CIE:	Commission internationale de l'éclairage
CODEC:	Codeur et décodeur
CPB:	Constrained Parameter Bitstream
EQM:	Erreur quadratique moyenne
FCC:	Federal Communication Commission
IES:	Interférence entre symboles
ISO:	International Standard Organisation

ITU:	International Telecommunication Union
LAN:	Local Area Network
MB:	MacroBlock
MPEG:	Motion Picture Expert Group
NTSC:	National Television System Committee
PAL:	Phase Alternate Line
PCS:	Personal Communication Systems
RNIS:	Réseau numérique à intégration de services
SECAM:	Sequential Color Amplitude Modulation
SIF:	Standard Interchange Format
SNR:	Signal to Noise Ratio
SPW:	Signal Processing Worksystem
SVH:	Système visuel humain
TCD:	Transformée en cosinus discrète
TRC:	Tube à rayon cathodique
TVHD:	Télévision à haute définition
YCbCr:	Luminance et chrominance numérisées
YUV:	Luminance (Y) et chrominance (U,V)
VOD:	Video on Demand
VHDL:	Very High Speed Integrated Circuits Hardware Description Language
VLSI:	Very Large Scale Integration
W-LAN:	Wireless LAN

Introduction

Un certain nombre de constatations sont à l'origine de notre travail. D'une part, les communications sans-fil bénéficient d'une croissance exponentielle et constituent un intérêt particulier pour la communauté scientifique. D'autre part, avec l'augmentation de la demande pour les communications personnelles et le développement des applications multimédia, il est de plus en plus probable que la transmission des images ou des séquences vidéo numérisées et comprimées via une liaison radio devienne une réalité.

Les systèmes de communication sans-fil couvrent une variété d'applications parmi lesquelles on retrouve la communication entre des personnes en mouvement à l'intérieur d'un bloc d'appartements, d'un hôpital, d'un édifice à bureaux, d'un supermarché ou d'un centre d'achats. Ces systèmes ont de nombreux avantages. En l'occurrence, ils permettent d'accommoder les utilisateurs gênés par le câblage conventionnel dans les édifices permettant ainsi une vraie mobilité. La communication sans-fil intra-édifice autorise aussi une certaine flexibilité dans la création et la modification des divers systèmes de communication existants dans les édifices. Ceci permet de réduire de façon importante le câblage dans les locaux nouveaux ou réaménagés. Les systèmes de communication sans-fil sont de plus en plus complexes. Une liaison sans-fil dans un environnement mobile ou intra-édifice est généralement affectée par les perturbations dues aux effets de propagation par trajets multiples donnant lieu aux évanouissements, à l'effet Doppler, à la diffraction, à la réflexion et à la diffusion des ondes électromagnétiques. Cette complexité est due aussi à la nature de l'information à transmettre, à l'environnement de propagation des signaux transmis et à la mobilité des utilisateurs potentiels. Dans notre cas, la représentation et la transmission de la vidéo numérique constituent un aspect clé dans la conception d'un système de transmission vidéo via un canal sans-fil.

La recherche dans le domaine du codage efficace des signaux vidéo, encore appelé codage de source, a été menée de façon intensive ces vingt dernières années. Ce domaine de recherche est suffisamment mature pour servir de base à l'élaboration de recommandations actuelles et à venir. Notamment, la recommandation ITU-T H.261 (International Telecommunication Union) anciennement CCIR H.261 (Comité Consultatif de Radiocommunication) pour les applications de téléconférence et la famille de recommandations ISO/IEC MPEG (Motion Picture Expert Group) pour la compression des images vidéo (MPEG-1) et de télévision (MPEG-2). Ces recommandations sont toutes basées sur la même architecture, à savoir un codage temporel compensé pour le mouvement et un codage spatial. Malgré le fait que beaucoup de travail reste à faire pour atteindre les limites fondamentales de performances, ces algorithmes de codage recommandés devraient rester à la base de plusieurs services pour plusieurs années encore (Dubois, 1992). Parmi ces services, on retrouve la vidéo sur demande, la télévision numérique à haute définition et les bases de données multimédia. Les images numérisées constituent une quantité volumineuse de données. Une compression de données est donc souhaitable avant la transmission. Ceci dans le but de réduire la bande passante du canal de communication. Sauf que, plus on réduit la redondance dans le signal à transmettre (compression élevée), plus l'information transmise est significative, ce qui rend l'effet d'une erreur de transmission plus importante.

Contributions et méthodologie

L'objectif principal de ce travail est d'adapter les techniques de protection de données à la transmission des images animées sur un réseau sans-fil. La stratégie de contrôle d'erreurs proposée est basée sur une correction automatique des erreurs de transmission plutôt que sur une retransmission en cas d'erreur. Une technique de retransmission en cas d'erreur (Yasushi, 1993) utilisée pour la transmission d'images fixes se prête plutôt mal au décodage en temps réel des séquences vidéo comprimées. Notre stratégie de protection adaptative et de contrôle d'erreurs se démarque des techniques conventionnelles de protection de données par le fait qu'elle tient compte de l'importance du

contenu de l'information à transmettre. Cette information est divisée en classes et une protection incrémentale est appliquée à chaque classe basée sur un critère de priorité. Nous disposons d'un codeur et d'un décodeur écrits en langage C sous le système d'exploitation UNIX et fonctionnant sur une station (poste de travail) *SUN* dans un environnement X-Windows. Le logiciel SPW (*Signal Processing Worksystem*) de Alta Group nous a permis de simuler un modèle de canal intra-édifice avec évanouissements auquel nous avons ajouté du bruit blanc gaussien. La modélisation et la simulation du système de communication a été faite avec le même logiciel (SPW).

Le premier chapitre de cette étude porte sur la définition, l'évaluation, les applications du système de communication proposé et les applications de la compression d'images vidéo. Nous tentons de mettre en lumière les différents éléments du système de codage et de transmission tout en insistant sur leurs rôles respectifs.

Le deuxième chapitre est consacré à la caractérisation, à la perception et à la représentation des images.

Le troisième chapitre est consacré à l'analyse et à la caractérisation du codeur de source basé sur le standard MPEG-2.

Le quatrième chapitre porte sur la caractérisation et la modélisation du canal de communication, ainsi que sur la théorie du codage de canal.

Le cinquième chapitre est consacré à la présentation et à l'analyse des résultats obtenus par simulation avec SPW.

CHAPITRE 1

Système de communication vidéo intra-édifice

Dans ce chapitre, nous tentons de mettre en lumière les différents éléments du système de communication vidéo intra-édifice tout en insistant sur le rôle de chacun des modules du système de codage et de transmission.

1.1 Définition d'un système de communication vidéo

Un système de communication effectue la transmission d'une information de façon analogique ou numérique. Cette information peut être: le son, des graphiques, des images, de la vidéo, du texte ou une combinaison de ces données. Les algorithmes de traitement et de compression d'images fixes ou animées permettent d'assurer, d'une part la transmission des images entre un émetteur et un ou plusieurs récepteurs, et d'autre part, le stockage de ces images sur un support magnétique ou optique en minimisant la largeur de bande et le volume. Ces algorithmes de traitement sont sujets à une contrainte principale qui est la qualité de l'information traitée et transmise. D'autres considérations comme le volume d'information à transmettre (largeur de bande), type de canal, fréquence d'émission et la complexité du codeur et du décodeur sont toutes aussi importantes dans la conception du système. La figure 1.1 présente le schéma simplifié d'un système de transmission d'images animées. L'aspect traitement en temps réel est important dans la mesure où les images transmises sont dynamiques. Cet aspect ne touche que les parties transmission et décodage des séquences d'images: la partie compression vidéo pouvant se faire en différé parce que trop complexe par rapport à la partie décompression vidéo.

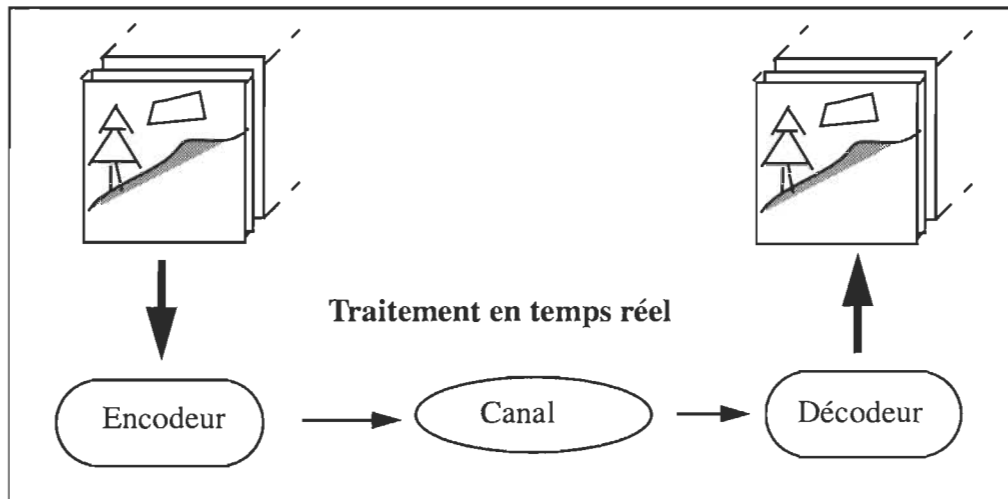


Figure 1.1 Schéma simplifié d'un système de transmission d'images animées

1.2 Composantes d'un système de communication vidéo

Dans cette section nous allons décrire brièvement les différents blocs du système de communication vidéo tels qu'illustrés à la figure 1.2.

- **L'acquisition** est faite par une caméra vidéo ou un numériseur.

- **Le pré-traitement** est un pré-filtrage ou une décomposition spectrale. Cette transformation permet de représenter le signal vidéo dans l'espace et dans le temps ou en fréquence, dans le but de limiter la perception de la distorsion dans l'image. Dans l'illustration suivante, $u(x,y,t)$ est le signal tri-dimensionnel représentant l'image et $U(x,y,t)$ est le signal reconstitué.

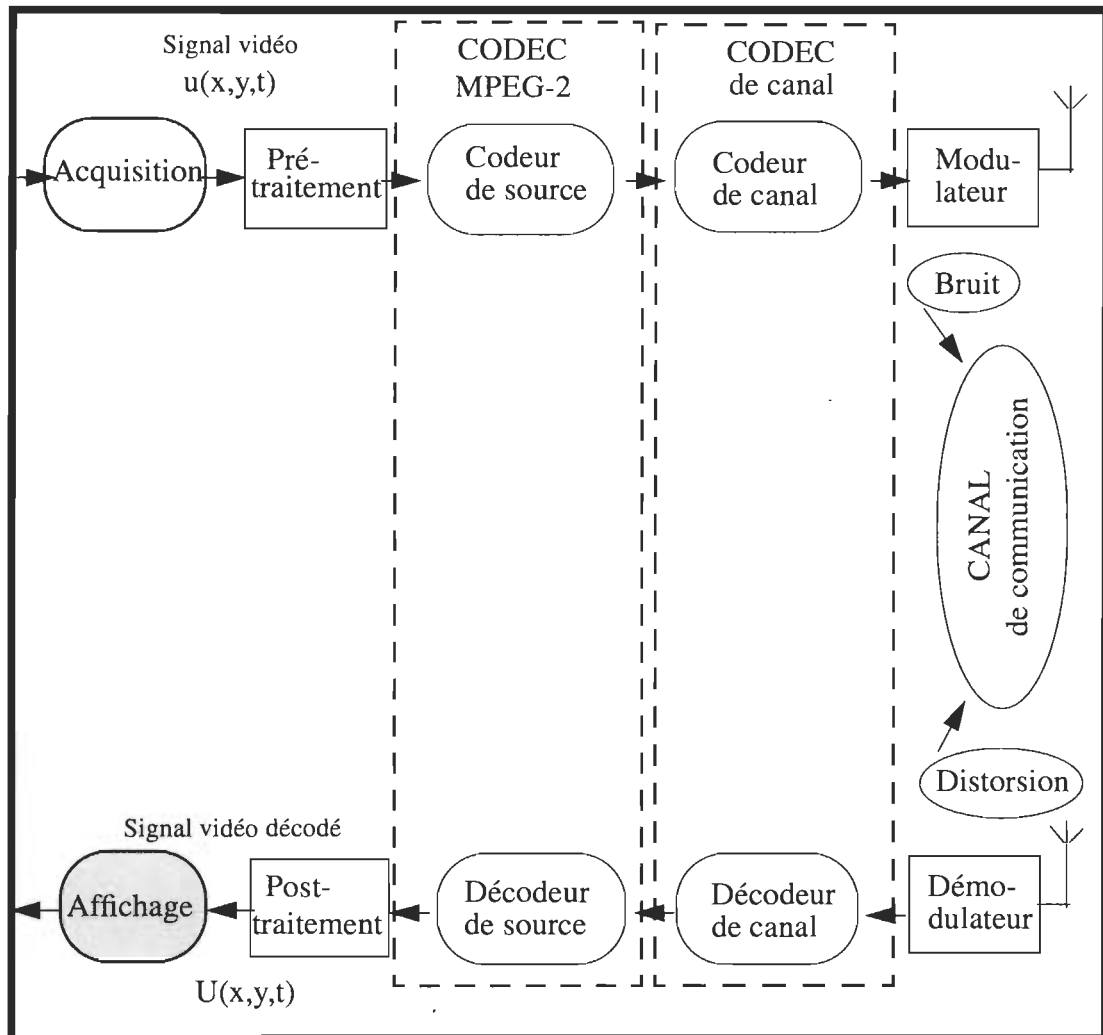


Figure 1.2 Éléments d'un système de communication vidéo

- **Le codeur de source** permet de coder les données à transmettre de façon efficace en faisant une compression dans le temps et dans l'espace. L'entrée du codeur de source peut être discrète ou continue, mais la sortie est toujours discrète. Dans notre cas, la transmission des données vidéo est faite sous un format numérique. Le processus de conversion d'un signal vidéo analogique à un signal vidéo numérique peut être intégré ou non au codeur de source. Toutes les sources d'information produisent une sortie

dont les valeurs sont aléatoires (si celles-ci étaient connues à l'avance, il ne serait pas nécessaire de les transmettre). Pour tenir compte de ce fait, la source d'information est caractérisée par un modèle statistique. Le codeur de source utilisé dans ce système inclut un quantificateur. La distorsion introduite par cette quantification provoque une perte de précision dans la transformation inverse qui a lieu dans le décodeur de source.

- **Le codeur de canal** introduit de façon contrôlée de la redondance dans la séquence binaire produite par le codeur de source. Cette redondance est utilisée par le décodeur de canal pour compenser les effets du bruit et des interférences de toutes sortes susceptibles d'affecter la qualité de la transmission. En général, le processus d'encodage met en correspondance K bits d'information à l'instant t avec un groupe de N symboles appelé mot de code. La quantité de redondance introduite par l'encodeur est mesurée par le rapport N/K encore appelé taux de codage.

- **Le modulateur** joue un rôle d'interface entre la sortie numérique du codeur de canal et l'entrée du canal de communication. La séquence binaire à la sortie du codeur de canal est envoyée au modulateur qui met en correspondance chaque symbole binaire avec un signal radio.

- **Le canal de communication** permet d'assurer la transmission de l'information d'un bout à l'autre du système. Supposons que les entrées et sorties du canal sont des ondes radio. Un bruit blanc gaussien s'ajoute au signal. Des distorsions α , autres que celles produites par le bruit additif, affectent le signal utile $x(t)$ de façon différente d'un canal à l'autre. $y(t)$ représente l'onde reçue.

$$y(t) = \alpha x(t) + n(t) \quad (1.1)$$

Dans le cadre de ce travail, nous allons nous attarder uniquement sur le canal avec évanouissement. Le chapitre 4 est consacré à la caractérisation du canal de communication utilisé dans un réseau sans-fil intra-édifice.

- **Le démodulateur** sert d'interface entre la sortie du canal et l'entrée numérique du décodeur de canal. L'onde radio dégradée reçue est réduite à un scalaire ou à un vecteur qui représente une estimation du symbole transmis (binaire ou multi-niveaux). Si on considère le processus de décision comme étant une forme de quantification, on observe que le décodage avec une décision dure correspond en fait à une quantification de la sortie du démodulateur. De façon générale, on peut considérer un démodulateur qui quantifie sur un nombre de niveaux $Q > 2$, comme un décodeur à décision pondéré, alors qu'un démodulateur qui transmet un "0" ou un "1" au bloc suivant dans la chaîne de réception correspond à une décision dure. La sortie du détecteur est envoyée au décodeur de canal.

- **Le décodeur de canal** exploite la redondance introduite par le codeur de canal pour corriger les erreurs de transmission introduites par le canal de communication. Le décodeur utilisé dans notre système est basé sur l'algorithme de Viterbi (voir chapitre 4).

- **Le décodeur de source** fait l'inverse de l'opération du codeur de source en faisant usage d'un critère de qualité. Ce critère permet de percevoir l'image animée transmise $u(x,y,t)$ comme étant égale à l'image animée reçue et reconstruite $U(x,y,t)$.

- **Le post-traitement** est un filtrage effectué afin d'améliorer la qualité de l'image et de compenser les imperfections du décodeur de canal et du décodeur de source.

- **L'affichage** se fait sur un écran ou sur papier. La taille des images est un paramètre du système de communication.

1.3 Étapes de conception du système de communication vidéo

Les différentes étapes de conception du système de communication vidéo intra-édifice que nous proposons sont les suivantes: la spécification, l'organisation fonctionnelle, la conception logique, la simulation, l'optimisation, le développement des utilitaires de conversion de formats de données, l'expérimentation et le prototypage. Telle que présentée à la figure 1.3, la méthodologie suivie pour la modélisation de ce système de communication s'est faite en huit étapes. La neuvième étape, qui est la réalisation matérielle (prototype), dépasse le cadre de ce travail. L'objectif de ce mémoire étant de modéliser le système de communication vidéo via un canal sans-fil intra-édifice, cette dernière étape n'a donc pas été couverte.

Étape 1: Spécification du système et choix des paramètres de simulation en fonction du domaine d'application visé qui est dans notre cas, la transmission des images animées via un canal sans-fil intra-édifice.

Étape 2: Représentation, caractérisation et perception des images à transmettre. Cette étape est importante dans la mesure où elle permet d'évaluer l'effet de la distorsion dans l'image reconstruite. La caractérisation du canal de communication est faite à ce niveau aussi.

Étape 3: Organisation fonctionnelle et modularisation de l'architecture du système de communication vidéo en fonction des ressources matérielles et logicielles.

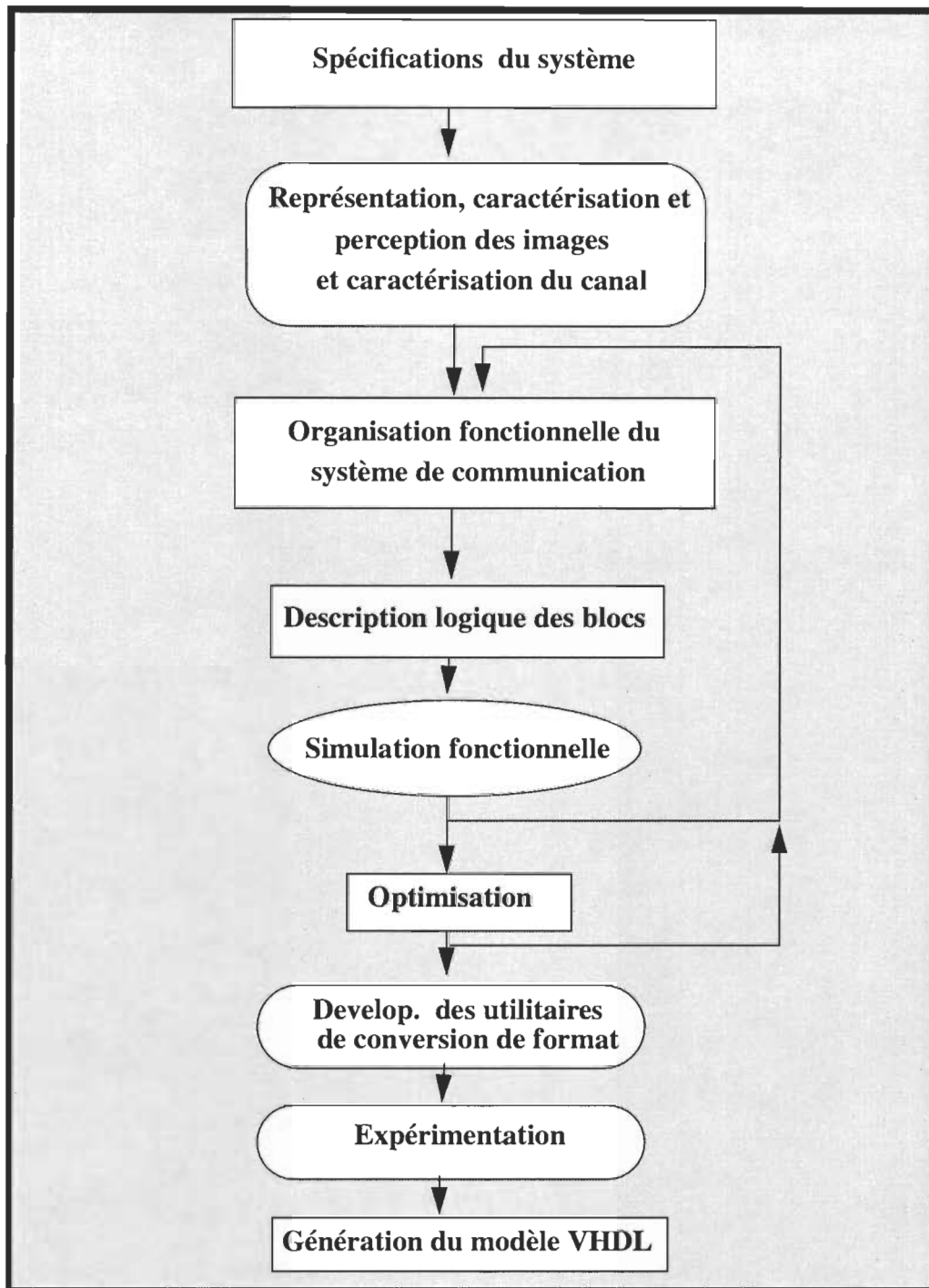


Figure 1.3 Étapes de conception du système de communication vidéo

Étape 4: Description logique des différents blocs (modules) qui composent le système de communication en utilisant les blocs élémentaires de la librairie de communication de SPW. Configuration des blocs en utilisant les paramètres de communication spécifiés à l'étape 1.

Étape 5: Simulation du système de communication avec les outils d'aide à la conception assistée par ordinateur et retour à l'étape 3 afin de corriger les erreurs de conception. La simulation permet de mesurer les performances fonctionnelles (taux d'erreurs, robustesse contre les distorsions produites par le canal intra-édifice, protection adaptative) du système avant l'étape de la fabrication ou de la réalisation.

Étape 6: Optimisation des paramètres pertinents. Une fois que nous sommes satisfait des résultats de la simulation fonctionnelle, les paramètres du système sont optimisés afin de respecter les spécifications du système de communication.

Étape 7: Développement des utilitaires qui permettent de convertir les données d'un outil à l'autre. Ceci permet l'intégration des différents modules en un système complet et fonctionnel.

Étape 8: À l'étape d'expérimentation, les différentes stratégies de protection sont simulées. Cette étape se termine avec l'analyse et l'interprétation des résultats obtenus par simulation.

Étape 9: Cette étape qui n'est pas couverte par ce travail permet de générer les fichiers en langage de description matérielle VHDL (Very High Speed Integrated Circuits Hardware Description Language) à partir du système conçu avec SPW illustré à l'appendice A. Par la suite, le modèle du système de communication serait encore simulé, cette fois-ci en VHDL à l'aide des outils d'aide à la conception de circuits inté-

grés tels que ceux de Mentor Graphics, Synopsys ou Cadence. Après le placement routage et la génération des masques, les circuits sont envoyés dans une fonderie de silicium pour la fabrication et le test.

1.4 Caractéristiques d'un système de communication vidéo

La transmission de la vidéo à travers un canal de capacité limitée nécessite l'encodage d'un signal tri-dimensionnel $u(x,y,t)$ en un signal uni-dimensionnel $v(t)$. Pour construire un tel système, il faut connaître: la largeur de bande et les propriétés statistiques de la source, la capacité du canal, le niveau de bruit, le taux d'erreurs acceptable à la réception et le critère de qualité. La figure 1.4 illustre quelques systèmes de communication sans-fil déjà existants de nos jours.

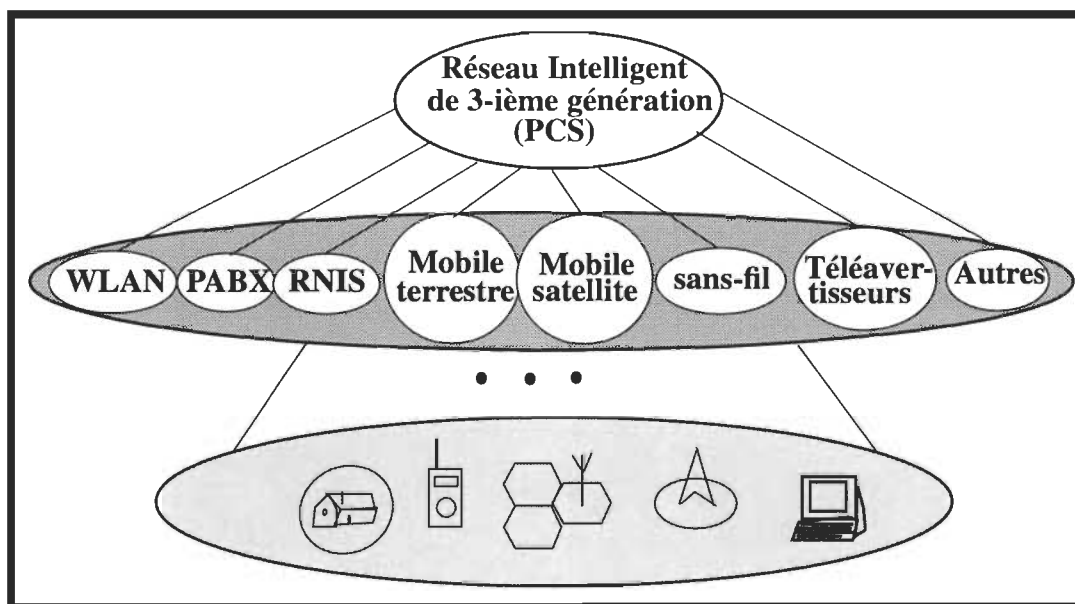


Figure 1.4 Systèmes de communication sans-fil

1.5 Applications d'un système de communication vidéo

Les applications d'un système de communication vidéo peuvent être regroupées en fonction du débit de transmission telle qu'illustré à la figure 1.5.

1- La télédiffusion

- la télévision numérique (*Digital Broadcasting System - DBS*)
- la télévision à haute définition (TVHD) terrestre

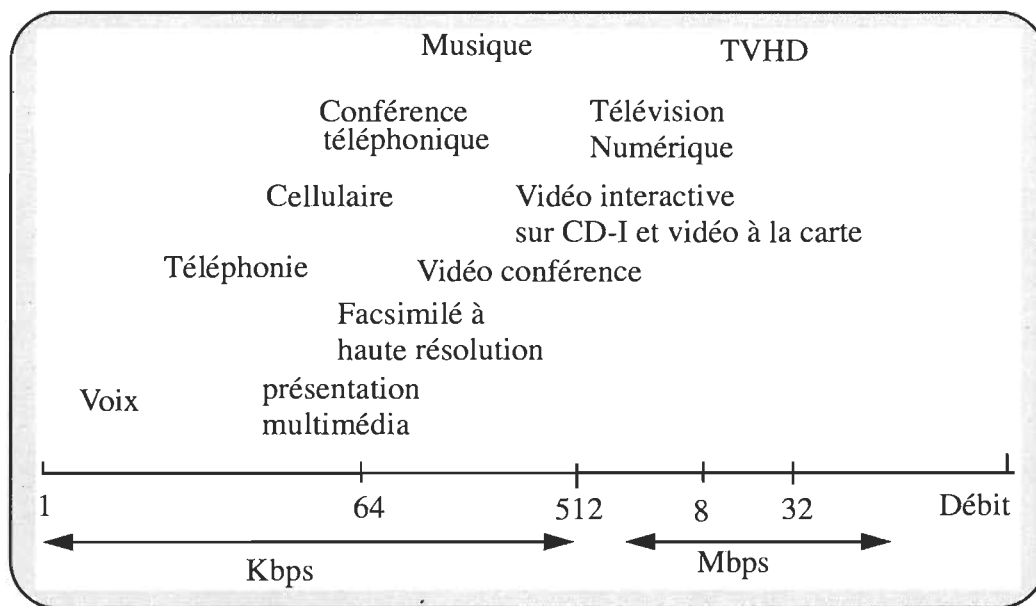


Figure 1.5 Applications de la compression de signaux vidéo en fonction du débit

2- Les communications bidirectionnelles

- la vidéoconférence
- le vidéophone

3- Les communications unidirectionnelles

- la télésurveillance

4- L'accès aux bases de données multimédia

- la vidéo sur demande par ligne téléphonique
- les serveurs vidéo sur réseau ATM

5- Les divertissements

- le magnétoscope numérique
- le disque laser numérique (CD-I de Philips, Photo-CD de Kodak)

Nous avons vu le rôle et l'importance de chacun des modules constituant le système de communication sans-fil intra-édifice. Dans le chapitre suivant, nous allons présenter plus en détail la théorie à la base du fonctionnement du codeur de source MPEG-2.

CHAPITRE 2

Caractérisation, perception et représentation des images

La compression de signal consiste à représenter un signal analogique de façon numérique, avec le moins de redondance possible. Dans notre cas, le signal vidéo est codé et transmis, puis reconstruit et perçu selon un critère ultime qui est la qualité. Celle-ci est généralement jugée ou mesurée par le système visuel humain (SVH). L'objectif dans la conception d'un codeur ou d'un algorithme de codage vidéo vise à minimiser la perception des distorsions (mesures subjectives). L'évaluation de cet objectif étant souvent problématique, on a souvent recours à l'optimisation d'un critère objectif qui est l'erreur quadratique moyenne entre le signal à l'entrée et à la sortie (mesure objective). Par conséquent, on introduira une notion de "codage perceptuel" (Jayant, 1993) basée sur un concept de masquage de la distorsion du signal comprimé.

Définitions

La qualité est un concept relatif qui peut vouloir dire l'absence de distorsion dans l'image, l'absence de bruit dans l'image ou l'intelligibilité de l'information dans l'image.

L'image est la lumière émise ou réfléchiée par des objets dans le champ visuel capté par les yeux. Les capteurs du système visuel humain logés dans la rétine sont constitués de cellules appelées cônes et bâtonnets. Les cônes sont adaptés à la vision diurne et en couleur alors que les bâtonnets sont adaptés à la vision nocturne.

La luminance est une valeur quantifiable ou une mesure objective de l'aspect visible de l'énergie radiante produisant la sensation de brillance. Elle est exprimée comme étant le quotient de l'intensité lumineuse d'une surface par l'aire apparente de cette surface pour un observateur lointain.

La brillance est une valeur qualitative associée à la luminance. Elle dépend de la luminance de l'objet par rapport à la luminance du fond de l'image (arrière-scène).

La résolution est la densité des échantillons d'une image donnée par unité de surface. Elle est une caractéristique importante de l'information visuelle. La qualité de reproduction d'une image est déterminée jusqu'à un certain point par la résolution. Une image à haute résolution (meilleure qualité) est exigeante pour un système de communication vidéo en terme d'espace occupé et de bande passante pendant la transmission.

La distorsion est une indication de la différence entre une information visuelle transmise et l'information reconstruite. Par exemple, si une image de télévision est transmise via un canal numérique, la distorsion pourrait être indiquée par le nombre relatif de téléspectateurs qui appellent pour se plaindre de la qualité des images transmises. Ou encore, dans une application où des images satellites sont utilisées pour mesurer la superficie d'un champ de blé, la distorsion pourrait être évaluée par le pourcentage de terrains mal classifiées. Donc la performance des systèmes de codage et de communication peut être évaluée sur la base de la distorsion.

2.1 Système visuel humain

Le système visuel humain présenté dans cette section est un modèle qui décrit le comportement de l'oeil humain de façon consistante avec sa physiologie. Un traitement plus détaillé du SVH est fait par M. H. Pirenne (1967) et G.S. Brindley (1970). Les

informations sur le SVH sont basées sur les études physiologiques des vertébrés tels que les chats et les singes qui ont un système visuel semblable à celui des humains.

Définition de la fréquence spatiale

Si peu de personnes ignorent ce qu'est une fréquence, beaucoup cependant pourraient ne pas percevoir ce que sont les fréquences spatiales. La capacité de l'oeil humain à reconnaître certains aspects d'une image dépend précisément de ce que l'on nomme fréquence spatiale. Lorsque deux lignes verticales d'une image sont si proches que l'oeil commence à avoir des difficultés à les percevoir séparément, elles définissent avec l'oeil un angle minimum de perception. La fréquence spatiale est proportionnelle à l'inverse de cet angle et elle est égale au nombre de lignes par unité de cet angle. Lorsque l'angle diminue, la fréquence augmente. L'oeil ne pouvant résoudre des lignes trop rapprochées, agit comme un filtre passe-bas. La fréquence spatiale dépend de la distance à laquelle sont vues les lignes ainsi que de la couleur.

La figure 2.1 présente le modèle du SVH ainsi qu'une illustration de sa réponse en fréquence spatiale et temporelle. Le SVH se comporte comme un filtre passe-bas (FPB) spatial ayant une fréquence de coupure d'environ 60 Cycles/degé illustré par la réponse en fréquence $|H(f)|$. Ainsi les fréquences hautes dans l'image ne sont pas visibles à l'oeil. Il n'est donc pas nécessaire de les transmettre. Le SVH agit aussi comme un filtre passe-bande (FPBande) temporel ayant une fréquence de coupure haute située entre 50 et 70 Cycles/degé et une réponse en fréquence $|V(f)|$, dépendant des conditions de vision.

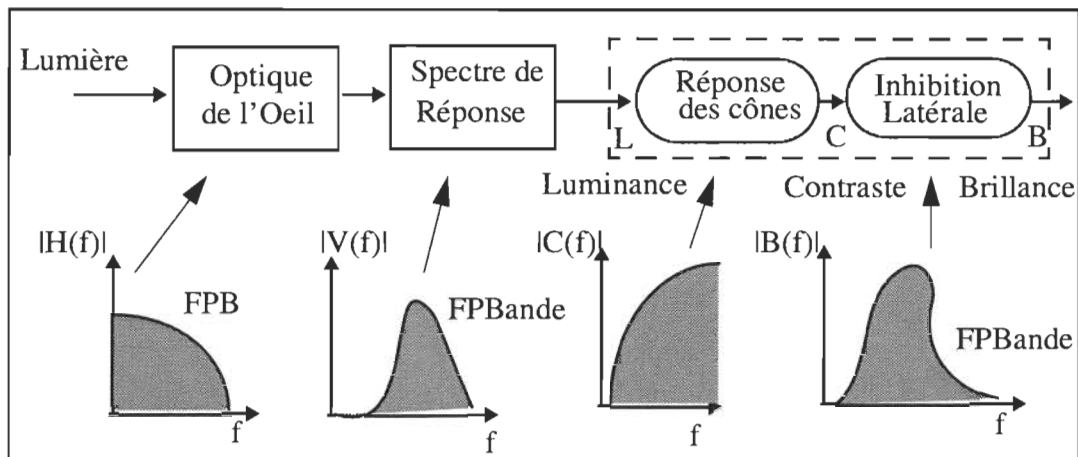


Figure 2.1 Réponse spatiale et temporelle du SVH

En général, l'effet du bruit et de la distorsion sur la qualité des images est moins visible pour les grandes valeurs de luminance que pour des valeurs moyennes et petites comme l'indique la réponse en fréquence de la luminance $|C(f)|$. En outre, les conditions de visibilité telles que la brillance de la scène et la lumière ambiante affectent grandement la perception de ces anomalies telle qu'indiquée par la réponse en fréquence $|B(f)|$ au voisinage des zones de transitions de couleur ou d'intensité dans l'image. Le bruit est moins visible dans les hautes et basses fréquences que dans les fréquences moyennes à cause du filtrage temporel passe-bande.

2.1.2 Colorimétrie, Photométrie et Radiométrie

La lumière est un attribut de toutes les perceptions ou sensations qui sont particulières à l'organe de la vision et qui se produisent par son intermédiaire. C'est un agent physique capable d'impressionner l'oeil, de rendre les choses visibles. La lumière visible est un rayonnement électromagnétique dont la longueur d'onde se situe entre 350 et 780 nm. On distingue trois descripteurs de couleurs:

- 1- la luminosité: l'intensité lumineuse

2- la dominance: la longueur d'onde dominante (ex: rouge, vert, bleue)

3- la pureté: la pureté d'excitation (gris -> pastel -> couleur pure)

Les stimulus de référence sont: Rouge (700 nm), Vert (546.1 nm) et Bleu (435.8 nm).

Le système visuel humain peut être schématisé comme étant le lien entre la colorimétrie, la radiométrie et la photométrie (voir figure 2.2) dans la perception des images naturelles. La radiométrie est la mesure des grandeurs relatives aux rayonnements d'ondes électromagnétiques. C'est-à-dire l'émission ou le transport d'énergie sous forme d'ondes électromagnétiques ou de particules.

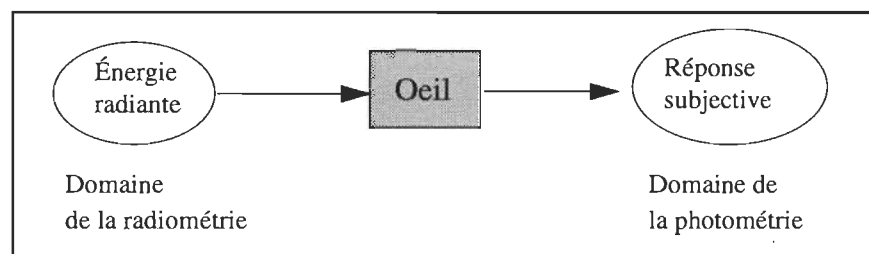


Figure 2.2 SVH par rapport à la photométrie et la radiométrie.

Le SVH perçoit un rayonnement grâce aux cônes et aux bâtonnets dans la rétine. Ce rayonnement perçu par le SVH est représenté dans un autre domaine appelé photométrie. Celle-ci est une mesure des grandeurs relatives aux rayonnements évaluées selon l'impression visuelle produite par ceux-ci et sur la base de certaines conventions. La photométrie tient compte des propriétés d'un observateur pour mesurer des grandeurs relatives aux rayonnements afin d'exprimer les quantités photométriques de façon standard.

Maxwell a démontré en 1855 en utilisant un projecteur que toutes les couleurs naturelles pouvaient être synthétisées en faisant un mélange approprié de trois couleurs primaires bien choisies. Aujourd'hui, ce domaine s'est développé et s'appelle la colorimétrie.

2.3 Système d'acquisition et d'affichage

Éléments du système d'acquisition

Dans cette section nous allons présenter le système typique d'acquisition et d'affichage des images. Chaque bloc du système est numéroté et suivi d'une brève description qui donne le rôle de chacun des éléments. La figure 2.3 illustre les éléments du système d'affichage.

- 1) Le système optique consiste en la projection de la scène sur un plan d'image.
- 2) Le détecteur de luminance (écran monochrome) ou des couleurs (rouge, vert, bleu: écran couleur) est constitué d'une surface photosensible à l'éclairement. Il est composé de trois filtres + un détecteur ou de trois détecteurs (RVB). Les lampes de type *Vidicon*, ou *Plumbicon* anciennement utilisées ont une haute qualité mais sont encombrantes et consomment beaucoup d'énergie. Les détecteurs de types CCD (*Coupled Charged Device*) produisent des images de bonne qualité, ont une petite dimension et consomment moins. Par conséquent, ils sont de plus en plus utilisés.
- 3) Le pré-filtrage passe-bas spatial et temporel est nécessaire pour limiter la bande passante du signal avant l'échantillonnage. Toutefois, le filtrage par le système optique (fonction de transfert de la lentille de la caméra) et par le détecteur (dimension et forme du *spot* mobile) est souvent suffisant.
- 4) L'échantillonnage 2-D ou 3-D se fait par balayage ou par transfert d'un format d'encodage à un autre (ex. NTSC, PAL, SÉCAM).

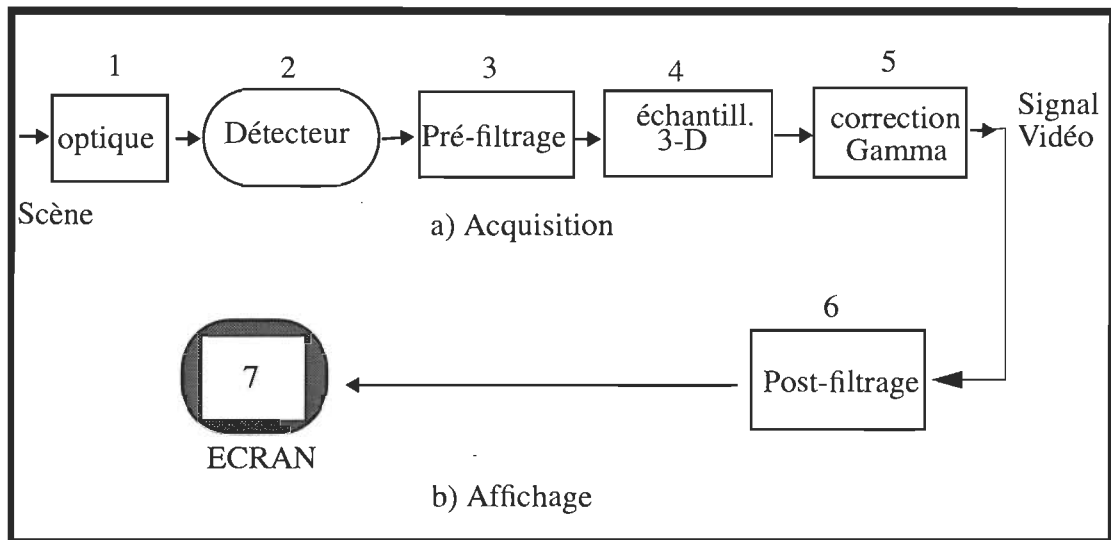


Figure 2.3 Acquisition et affichage

5) La correction Gamma permet d'égaliser la non-linéarité entre la caméra vidéo et l'écran d'affichage. La majorité des caméra vidéo et les écrans TRC (tube à rayon cathodique) ont une non-linéarité inhérente entre l'intensité de la lumière captée ou affichée et la tension du signal vidéo. Dans le but d'éviter d'intégrer un circuit de correction dans chacun des récepteurs vendus dans le monde entier, la correction Gamma est habituellement faite dans la caméra vidéo avant la transmission de l'image telle qu'illustré à la figure 2.4 La relation entre le voltage du signal vidéo et l'intensité de la lumière est définie par:

$$B_s = cv^\gamma + B_e \quad (2.1)$$

où B_s est l'intensité de la lumière, B_e est l'intensité de l'écran correspondant au noir, v est le voltage du signal vidéo, c est une constante de proportionnalité désignant un facteur de gain et γ est compris entre 1 et 3.

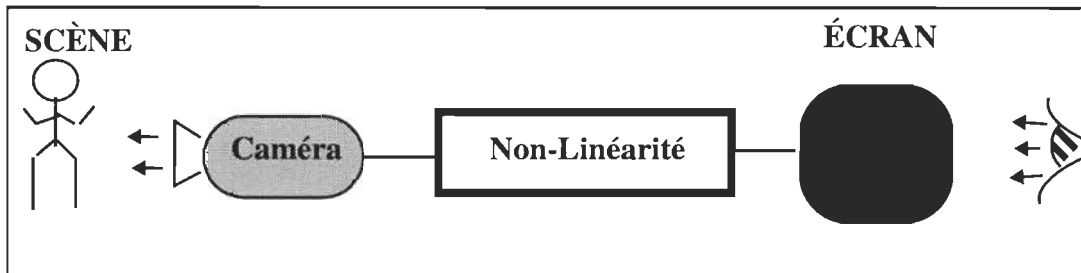


Figure 2.4 Correction Gamma

Si γ_c est le facteur Gamma de la caméra γ_e est celui de l'écran TRC et le voltage de la caméra est corrigé suivant la formule:

$$v_e = v_c^{\gamma_c/\gamma_e} \quad (2.2)$$

Des recherches antérieures ont démontrées que la qualité de l'image produite par la caméra est maximale si la brillance à l'affichage est égale à la brillance de la scène originale.

Éléments du système d'affichage

Le but d'un système d'affichage est de reproduire une image de telle sorte que la perception résultante soit la plus proche possible de la vraie image. La complexité d'un système d'affichage dépend de certaines propriétés de l'image telle que:

- 1- images monochromes ou en couleurs
- 2- images monoscopiques ou stéréoscopiques
- 3- images statiques ou dynamiques (en mouvement)

6) Le post-filtrage est fait par le SVH, puis par les phosphores de l'écran d'affichage.

Le comportement des phosphores de l'écran d'affichage est décrit dans le temps par

une réponse impulsionnelle exponentielle décroissante. Les phosphores sont utilisés dans l'écran pour visualiser l'intensité d'un pixel de l'image pendant la durée de la persistance rétinienne du SVH.

7) L'écran permet de visualiser l'image.

2.4 Modèle de l'information visuelle

Typiquement, un algorithme est développé pour une classe d'images. Chaque classe peut être définie par un modèle. On distingue les modèles déterministes et les modèles stochastiques.

2.4.1 Modèles déterministes

À partir d'observations, on pose certaines hypothèses qui décrivent des propriétés d'une classe d'images. Ces propriétés sont les suivantes: la largeur de bande de l'image (dépend du contenu de l'image), la forme typique des objets ou des blocs dans l'image, la linéarité locale des composantes de l'image (luminance ou chrominance, RVB) et les mouvements d'objets par blocs dans l'image.

2.4.2 Modèles stochastiques

Les phénomènes aléatoires qui existent dans la nature dépendent, en général, du temps. Par exemple, un signal audio transmis à travers le canal téléphonique est un signal aléatoire. De façon similaire, la sortie d'un codeur d'image est considérée être un signal aléatoire qui varie dans le temps. Ces hypothèses sont basées sur les propriétés statistiques d'une classe d'images donnée. Parmi ces modèles stochastiques on distingue: le modèle de covariance séparable et non-séparable, le modèle prédictif causal, semi-causal et non-causal et le modèle de probabilité conjointe de Markov. Cependant, dans la pratique, seuls les modèles déterministes sont utilisés, les autres étant difficile-

ment réalisables.

2.5 Filtrage dans les systèmes vidéo

Quelques définitions

Pixel: un pixel (ou *pel*) réfère à un “*picture element*” ou un échantillon d’image.

Débit: le débit est une mesure de la quantité d’information par unité de temps de l’image. Il peut être exprimé en bits par échantillon, en bits par pixel (bpp) ou en bits par seconde (bps).

Signaux Vidéo

Un téléviseur génère les couleurs à partir d’un mélange additif de sources de couleurs Rouge, Vert et Bleue (RVB). Deux sources primaires peuvent être utilisées en principe, une source rouge-orange et une source vert-bleue. Mais l’image résultante n’est pas naturelle, ce qui explique pourquoi les systèmes à base de deux couleurs n’ont jamais eu de succès ni à la télévision, ni dans l’animation. Un des concepts spéciaux introduit par la commission internationale de l’éclairage (CIE), était l’isolation de la luminance (*Brightness*) de la chrominance (*Hue*). En utilisant le concept de la CIE comme référence, l’organisme NTSC (*National Television System Committee*) a adopté un système de transmission de signaux vidéo sous forme de signaux de luminance (Y) et de chrominance (U,V), plutôt que le format basé sur les composantes RVB des téléviseurs. Il existe en effet moins de corrélation entre les composantes Y, U et V qu’entre les composantes R, V et B. En Europe, deux standards ont été établis, le standard PAL (*Phase-alternate Line*) et le standard SECAM (*Sequential Color Amplitude Modulation*), tous les deux utilisent aussi l’espace de couleurs YUV. L’algorithme MPEG est formulé pour les trois composantes YUV. Les valeurs numériques des sources de couleurs RVB doivent être comprises entre 0 et 255 (des contributions égales de chaque composante représentent le blanc). Pour que leur conversion produise des valeurs de Y comprises entre 0 et 255 et de U et V comprises entre -128 et 128. Ces deux dernières

sont décalées de 128 afin d'obtenir des valeurs toujours positives. Le format YUV demande moins de bande passante à la transmission par rapport au format RVB. De plus, la composante Y est en fait un signal vidéo noir et blanc directement compatible avec une télévision monochrome. La matrice de conversion s'écrit comme suit:

$$\begin{bmatrix} Y \\ C_b \\ C_c \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.1687 & -0.3313 & 0.5 \\ 0.5 & -0.4187 & -0.0813 \end{bmatrix} \begin{bmatrix} R \\ V \\ B \end{bmatrix} \quad (2.3)$$

où l'espace YCbCr est l'équivalent numérique de l'espace YUV, et Y, Cb et Cr correspondent respectivement à la version numérisée des composantes Y, U et V à un facteur près. La matrice de conversion inverse est la suivante:

$$\begin{bmatrix} R \\ V \\ B \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1.402 \\ 1 & -0.34414 & -0.71414 \\ 0.5 & 1.772 & 0 \end{bmatrix} \begin{bmatrix} Y \\ C_b \\ C_c \end{bmatrix} \quad (2.4)$$

Le format CCIR-601 est une norme couramment utilisée pour la compression des signaux vidéo. Elle spécifie une conversion semblable à celle décrite ci-dessus mais à un facteur près. Notamment, les valeurs de la luminance sont limitées à l'intervalle [16 ... 235] et celles des composantes de chrominance sont limitées à l'intervalle [12 ... 240]. Les équations suivantes décrivent l'ajustement lors du passage d'un espace à l'autre:

$$Y' = \frac{219}{255}Y + 16 \quad (2.5)$$

$$C_b' = \frac{224}{255}C_b + 128$$

$$C_r' = \frac{224}{255}C_r + 128$$

2.5.1 Échantillonnage multi-dimensionnel spatial et temporel

La conversion de signaux continus dans le temps en signaux discrets dans le temps peut-être considérée comme un processus d'échantillonnage suivi d'une conversion des trains d'impulsions en séquence. Comme le montre la figure 2.5, le signal analogique $x(t)$ peut être reconstruit en soumettant le signal échantillonné $x_s(t)$ à un filtre passe-bas idéal avec une fréquence de coupure F_c , à condition que: $F_\omega \leq F_c \leq F_s - F_\omega$, où $F_s \geq 2F_\omega$. Cette condition est imposée par le critère de Nyquist qui s'énonce comme suit:

Un signal de bande passante limitée à F_ω rad/s est complètement déterminé par les valeurs instantanées de ses échantillons uniformément répartis dans le temps à la fréquence d'échantillonnage $F_s \geq 2F_\omega$ rad/s. La fréquence d'échantillonnage minimale $F_s = 2F_\omega$ est appelée fréquence de Nyquist (Nyquist, 1928).

Si les conditions du critère de Nyquist ne sont pas respectées, il y a recouvrement du spectre de fréquence. C'est-à-dire qu'il y a chevauchement des composantes du spectre de fréquence. Il est important de bien comprendre que le critère de Nyquist s'applique aussi dans d'autres contextes, tel que l'échantillonnage d'une image dans un espace à deux dimensions (2-D). Dans ce cas, le critère de Nyquist nous donne:

$$\frac{1}{\Delta x} \geq 2F_v, \quad \frac{1}{\Delta y} \geq 2F_h \quad (2.3)$$

où Δx et Δy sont les périodes d'échantillonnage spatiales verticales et horizontales respectivement. F_v et F_h sont les fréquences spatiales verticales et horizontales maximales de la transformée de l'image respectivement.

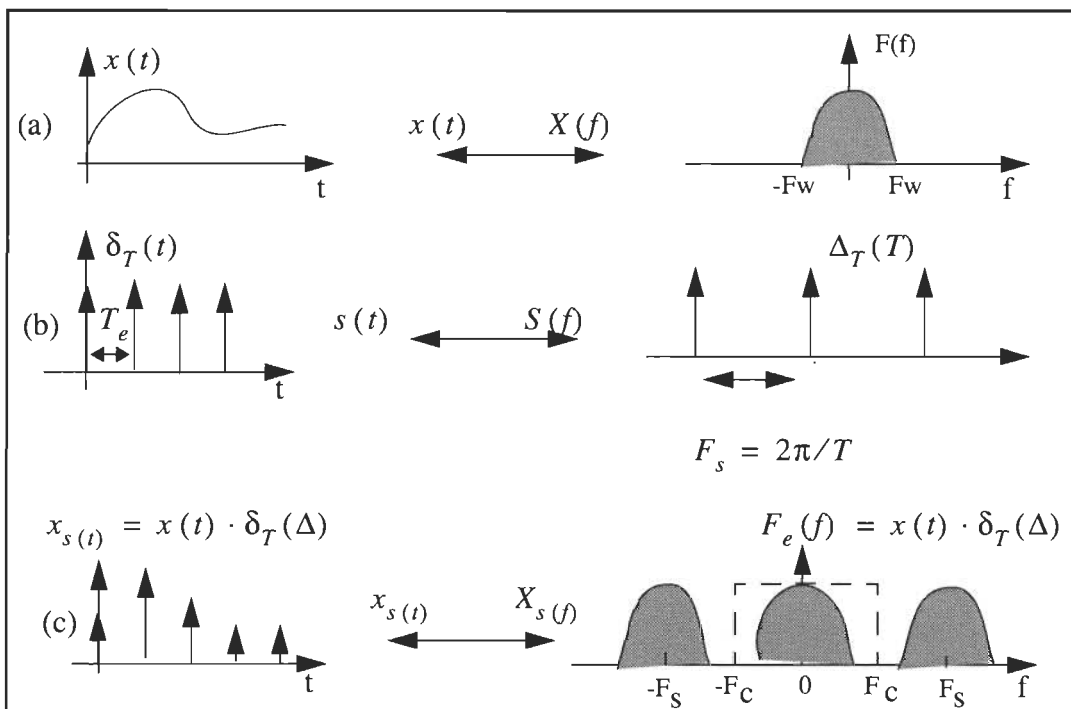


Figure 2.5 Échantillonnage d'un signal continu dans le temps, a) signal, b) train d'impulsions et c) signal échantillonné

Le critère de Nyquist s'applique également au spectre en trois dimensions (3-D) d'un signal vidéo. Les spectres obtenus après échantillonnage sont perçus avec une bande passante limitée par les filtrages inhérents au SVH. Pour la manipulation numérique d'images, la luminance (Y) ou les composantes de couleurs (U et V) doivent être:

1. échantillonnées. C'est-à-dire que la coordonnée continue devient la coordonnée discrète.
2. quantifiées. C'est-à-dire que l'amplitude continue devient l'amplitude discrète.

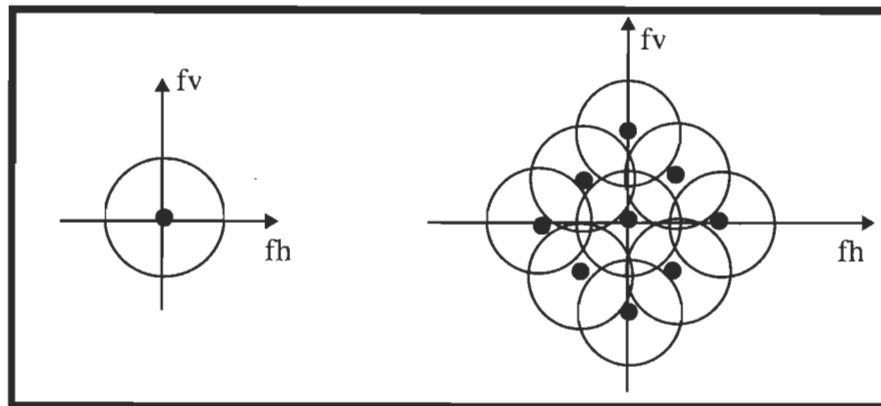


Figure 2.6 a) Spectre du signal continu 2-D. b) Spectre du signal échantillonné avec recouvrement de spectre (Éric Dubois, 1985)

Une image fixe est considérée comme une fonction à deux dimensions (positions horizontale et verticale). Par contre, une image dynamique ou animée est considérée comme une fonction à trois dimensions (positions horizontale, verticale et temporelle). La figure 2.6 illustre le spectre d'un signal vidéo avec l'effet du recouvrement de spectre dû à une fréquence d'échantillonnage inférieure à la fréquence de Nyquist.

2.5.2 Filtrage temporel

Le filtrage temporel est parfois réalisé avant la transmission des signaux vidéo afin de réduire le bruit aléatoire. On utilise souvent un filtre transversal pour la mise en forme des échantillons d'images à l'émetteur et un filtrage adaptatif (détection de corrélation) à la réception afin de réduire les effets du bruit. Les filtrages spatial et temporel peuvent tous les deux être utilisés pour séparer la luminance et la chrominance (séparation Y-C) dans un décodeur PAL (*Phase Alternate Line*) ou NTSC (*National Television System Committee*), ou encore, pour "étaier" le signal vidéo dans les applications qui nécessitent la détection du mouvement.

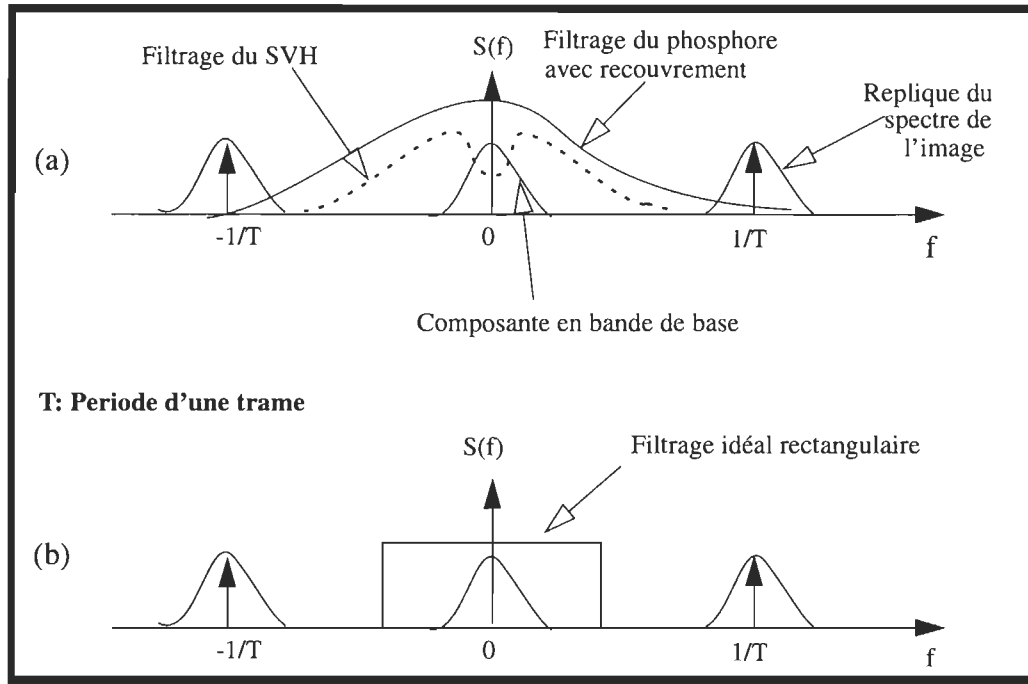


Figure 2.7 (a) Filtrage réel par le SVH. (b) Filtrage idéal rectangulaire

La figure 2.7 montre qu'en utilisant un filtre ayant une réponse en fréquence $H(f)$ rectangulaire (figure 2.7.b) plutôt qu'une réponse en fréquence (phosphores d'un écran TRC) ayant une pente d'environ de 6 dB par octave (figure 2.7.a), on pourrait réduire le taux d'échantillonnage temporel (nombre de trames par unité de temps) du signal vidéo. Ceci sans augmenter le recouvrement de spectre. La réponse en fréquence $H(f)$ d'un signal $h(t) = \text{sinc}(t)$ a une forme rectangulaire. En effet, la décomposition en série de Fourier du signal rectangulaire donne des valeurs alternativement négatives et positives avec une amplitude décroissante. En pratique, on obtient difficilement une réponse en fréquence rectangulaire avec un écran phosphorescent.

2.5.3 Filtrage spatial

Le signal de télévision est souvent séparé en deux composantes: le signal de luminance Y et le signal de chrominance C (voir chapitre 3). On utilise pour cela un filtrage simple de faible coût. Ce filtrage nécessite essentiellement un filtre passe-bande 1-D pour séparer l'information de chrominance de l'information de luminance et un filtre coupe-bande pour rejeter l'information de chrominance du canal réservé à la luminance. Cette approche simple provoque des anomalies qui se présentent sous forme de points mobiles sur les bords des couleurs et de perte de certaines composantes de la luminance Y . On peut réduire ces perturbations avec une meilleure résolution $Y-C$.

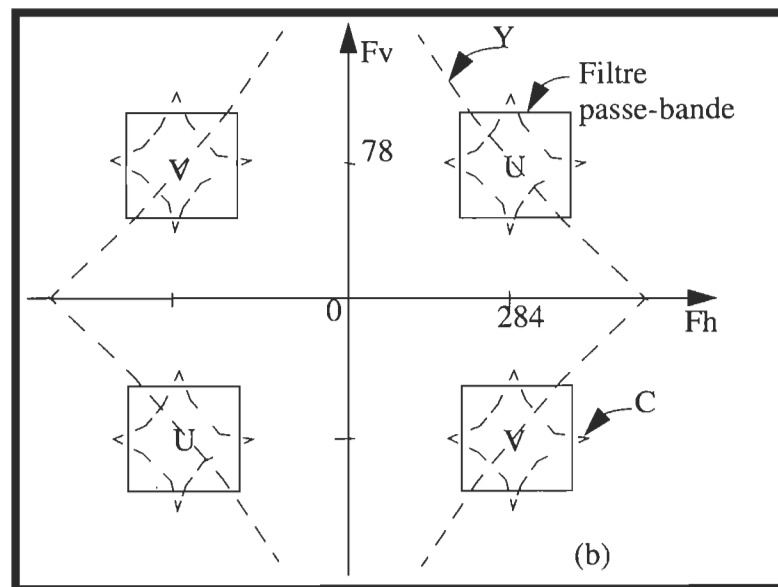


Figure 2.8 Spectre des composantes YUV d'un signal vidéo PAL.

Une solution consiste à employer un filtre passe-bande 2-D intra-zone. Le spectre du signal 2-D résultant est représenté par les fréquences horizontales (F_h) et verticales (F_v). La figure 2.8 représente les composantes numériques YUV du signal vidéo PAL.

en bande de base. Ce spectre est reproduit le long de l'axe des fréquences horizontales.

2.5.4 Conversion des standards (structure d'échantillonnage)

La multiplicité des standards entre les différents systèmes de télévision rend indispensable la conversion entre formats, notamment la conversion d'une séquence vidéo en format NTSC, avec 525 lignes à 60 Hz, à un format PAL, avec 625 lignes à 50 Hz et vice-versa.

Étapes de la conversion:

1. On effectue la démodulation d'un signal vidéo composite (ex. PAL, NTSC, SÉCAM).
2. On effectue éventuellement une transformation linéaire entre deux espaces de couleur (RVB, YUV).
3. On effectue une interpolation spatio-temporelle de la luminance et des chrominances.
4. Puis, enfin, on module de nouveau le signal composite dans un format approprié pour l'affichage.

2.5.5 Reconstruction d'un signal vidéo échantillonné

Le balayage est le déplacement du spot lumineux sur chaque ligne d'un écran de gauche à droite et de haut en bas. C'est un paramètre important dans un système de communication vidéo. On distingue deux types de balayages. Le balayage séquentiel et le balayage entrelacé. **Le balayage séquentiel ou progressif** est simple, adéquat pour les images fixes où le papillotement est absent. Comme exemples, on a le balayage des télécopieurs et écrans d'ordinateur. **Le balayage entrelacé** est plus complexe, une partie de l'image est prise à la fois et N trames forment une image. Le rapport d'échantillonnage entre les trames qui forment l'image est de N:1 (typiquement 2:1 trame paire

et impaire). Comme exemple, on a le balayage des images de télévision.

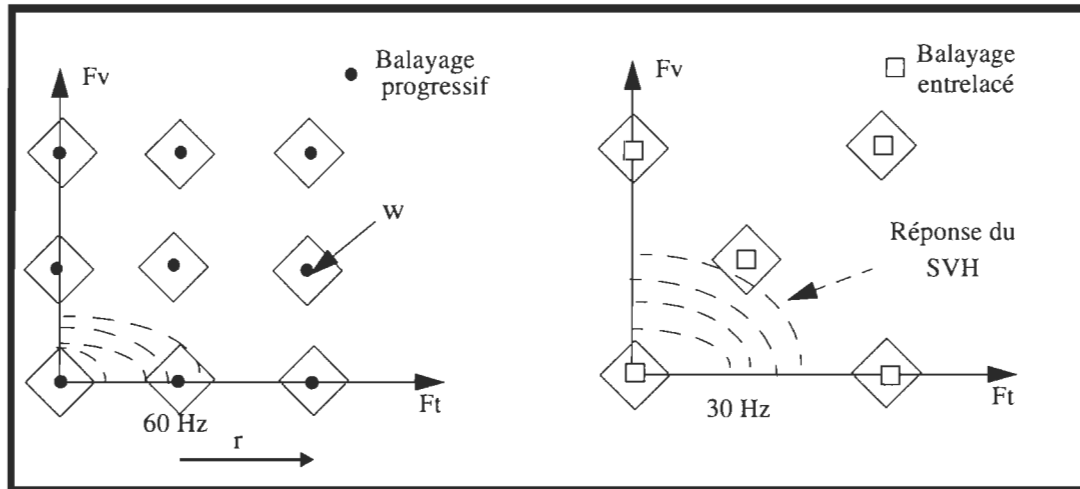


Figure 2.9 Balayage entrelacé et balayage progressif

Le choix du type de balayage pour différents écrans est borné par le système visuel humain. Si la luminance est petite, alors la bande passante du SVH est réduite alors que pour une grande luminance, la bande passante du système est élevée. Ainsi pour les écrans sombres, il est souhaitable d'avoir un entrelacement dont le rapport N:1 est grand et pour les écrans clairs, il est souhaitable d'avoir un entrelacement dont le rapport N:1 petit. La figure 2.9 montre la structure d'un signal vidéo pour les deux types de balayage. La reconstruction d'un signal échantillonné est toujours l'étape finale dans un système de communication vidéo. La reconstruction exacte d'un signal continu à partir de ses échantillons est possible si et seulement si on a pas de recouvrement de spectre. La reconstruction d'un signal vidéo consiste en la transformation du signal vidéo échantillonné en un signal vidéo analogique. Ce processus est complété par la réponse du SVH qui permet d'interpoler les échantillons (pixels) manquants afin de reproduire le plus parfaitement possible l'image sous forme analogique. F_t est la fréquence temporelle et F_v est la fréquence verticale.

Nous avons vu comment représenter les images dans ce chapitre. Nous avons également vu comment le système visuel humain perçoit l'information visuelle. Il a été démontré que lorsqu'on exploite les capacités du SVH (codage perceptuel) pendant l'acquisition et pendant l'affichage des images, on arrive à représenter ces images avec moins de redondance en faisant usage d'un critère de distorsion. Dans le chapitre suivant, nous allons décrire l'algorithme de compression d'images utilisé dans le système de communication vidéo sans-fil.

CHAPITRE 3

Analyse et caractérisation de l'encodeur et du décodeur MPEG-2

Dans un système de communication numérique, les messages produits par la source sont convertis en une séquence binaire. Idéalement, nous aimerions représenter ce message avec le moins de bits (symboles binaires de valeurs 0 ou 1) possible. En d'autres termes, nous voulons une redondance minimale ou nulle dans la séquence binaire résultante. Ce processus de conversion efficace d'une source analogique ou numérique en une séquence binaire est appelé codage de source ou compression de données. Dans les sections suivantes nous allons décrire l'algorithme MPEG-2.

3.1 L'Algorithme de compression MPEG-2

Les efforts de l'organisation de standardisation internationale (OSI) pour représenter les images numérisées ont aboutis à la création d'un standard appelé MPEG-2 sous la référence ISO/IEC JTC1/SC29/WG11-13818-2 en mars 1994. Subséquemment, l'organisme qui régit les lois sur les communications aux États-Unis a décidé d'adopter, sous le nom de la Grande Alliance (Chalapali, 1995), le standard MPEG-2 pour représenter les signaux vidéo qui seront utilisés pour la télévision à haute définition (TVHD) en Amérique du Nord. Le standard MPEG-2 définit les détails de l'algorithme de décodage, mais ne spécifie que partiellement l'encodage requis afin obtenir une qualité optimale d'images codées. En l'occurrence la façon d'estimer le vecteur de mouvement, ou de gérer la mémoire tampon afin de stabiliser le débit en sortie, ou encore la quantification adaptative des coefficients dans le but de maintenir de façon uniforme la qualité pendant la transmission. La première phase du standard, connue sous le nom de MPEG-1 (ISO/IEC JTC1/SC29/WG11 11172-2), est destinée aux applications de stoc-

kage des signaux vidéo sur des supports numériques tels que les disques compacts (CD-ROM, CD-I) avec un débit de 1,5 Mbps. La seconde phase, connue sous le nom de MPEG-2, utilise des techniques de codage avancées de la vidéo entrelacée ou progressive avec un débit de 1 à 20 Mbps. Le groupe de normalisation MPEG-2, conjointement avec le groupe 15 de l'union internationale des télécommunications (ITU-TS, anciennement CCITT), travaillent sur les techniques de codage vidéo adaptées au réseau ATM en plus du forum ATM qui travaille sur l'aspect transmission. Dans le cadre de ce mémoire, nous nous limiterons à la caractérisation du CODEC MPEG-2 utilisé dans notre modèle du système de communication vidéo.

3.1.1 Format des données vidéo

Quelques définitions:

Bloc: C'est la plus petite entité utilisée dans le codage spatial.

Macrobloc: C'est le plus petit commun multiple des blocs 8x8 de la luminance Y et des deux chrominances Cb, Cr (généralement égal à 16x16).

L'élément de base dans l'algorithme MPEG est un macrobloc; celui-ci regroupe toutes les composantes qui forment un élément de codage de l'image. L'essentiel de l'information est concentré dans la luminance. De plus, l'oeil humain est plus sensible à la luminance qu'aux deux chrominances. Les composantes YUV (ou leur version numérique) sont indépendantes (décorrélées) et peuvent donc être codées séparément sans risque de perte d'efficacité. Le tableau 3.2 illustre la relation entre la taille des images et les différents rapport d'échantillonnage pour le format CCIR 601. La réduction de la dimension des données vidéo après la conversion de RVB à YUV est de l'ordre d'un facteur de 4 ou 2, selon qu'on applique le format 4:2:0 ou le format 4:2:2. Le format 4:2:0 veut dire qu'on a 1 échantillons de couleurs pour quatre échantillons de luminance. Donc la résolution verticale et horizontale des composantes de couleurs est réduite de moitié par rapport à la résolution de la composante de luminance.

Tableau 3. 1: Image en format CCIR 601 (60 Hz) et sous-échantillonnage des composantes de couleurs

format	Y	Cb, Cr	Vertical	Horizontal
4:4:4	720 x 480	720 x 480	1:1	1:1
4:2:2	720 x 480	360 x 480	1:1	2:1
4:2:0	720 x 480	360 x 240	2:1	2:1
4:1:1	720 x 480	720 x 120	4:1	1:1
4:1:0	720 x 480	180 x 120	4:1	4:1

Le format 4:2:0 veut dire qu'on a un échantillon de couleurs pour quatre échantillons de luminance pour un total de six échantillons par macrobloc. Le format 4:2:2 veut dire qu'on a deux échantillons de couleurs pour quatre échantillons de luminance pour un total de huit échantillons par macrobloc. Le format 4:4:4 veut dire qu'on a quatre échantillons de couleurs pour quatre échantillons de luminance pour un total de douze échantillons par macrobloc. La réduction du taux de transmission des composantes YUV (par rapport à RVB) lorsqu'on utilise les formats 4:2:0 ou 4:2:2, est un avantage de codage certain. Par défaut, un bloc a une dimension de 8x8 pixels. La figure 3.1. illustre la structure des différents formats de macroblocs 4:2:0 pour MPEG-1, 4:2:0 pour MPEG-2 et 4:2:2 pour MPEG-2 (recommandation 601 du CCIR). Dans le format MPEG-1, les composantes de couleurs sont centrées autour des composantes de luminance afin de maintenir une certaine compatibilité avec le format CCIR 601. Dans les trois cas a, b et c de la figure 3.1, les deux composantes de couleurs sont superposées à la même position.

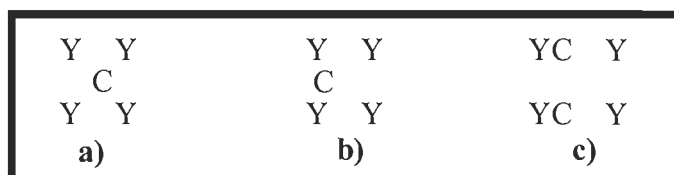


Figure 3.1 a) MPEG-1 (4:2:0), b) MPEG-2 (4:2:0), c) MPEG-2 (CCIR 601 4:2:2)

Le format vidéo MPEG restreint (*Constrained Parameter Bitstream, CBP*) a été conçu pour permettre à différents décodeurs de décoder une séquence vidéo avec une complexité normalisée. Certains paramètres du codeur ont des valeurs qui ont été limitées tout en étant destinées à une grande variété d'applications. Parmi ces paramètres on retrouve la complexité du traitement des données vidéo, la taille de la mémoire tampon et de la mémoire vive. La limite des paramètres a été choisie afin que le décodeur n'ait pas besoin de plus de 4 Mbits de mémoire vive. La dimension d'une image est indiquée de la façon suivante: 352x240x30, où 352 est le nombre de lignes dans l'image, 240 le nombre de colonnes et 30 le nombre d'images/secondes.

Résolution:	352x240 ou 352x288 (<i>SIF, Standard Interchange Format</i>)
Débit:	1,862 Mbps
Affichage:	30 Hz (nombre d'images par seconde)
Mémoire:	40 Kbits
Macroblocs:	330 (84480 pixels/images)

L'application principale du format restreint (CBP) est la vidéo numérique sur disque compact interactif (CD-I) et la vidéo sur ordinateur. Le décodeur TV numérique tombe dans une autre catégorie avec un taux d'échantillonnage plus élevé, défini par la recommandation CCIR-601 (*Broadcast rate*). Dans cette catégorie, la bande passante et la fréquence d'échantillonnage sont approximativement égales à 4 fois celles du format SIF (352x240 pixels).

Pixel: La norme MPEG définit un pixel comme étant le vecteur unitaire contenant un échantillon complet de la luminance et d'une fraction de chacune des deux chrominances (0,25 pour le format 4:1:1 donc $1 + 0.25 + 0.25 = 1.5$ bits/pixel).

La bande passante de l'image codée est calculée de la façon suivante:

$$\begin{aligned}
 B &= 352 \text{ (pixels/ligne)} \times 240 \text{ (lignes/image)} \\
 &\quad \times 30 \text{ (images/seconde)} \times 1,5 \text{ bits/pixel} \\
 &= 3,8 \text{ Mbps} \qquad (3.4)
 \end{aligned}$$

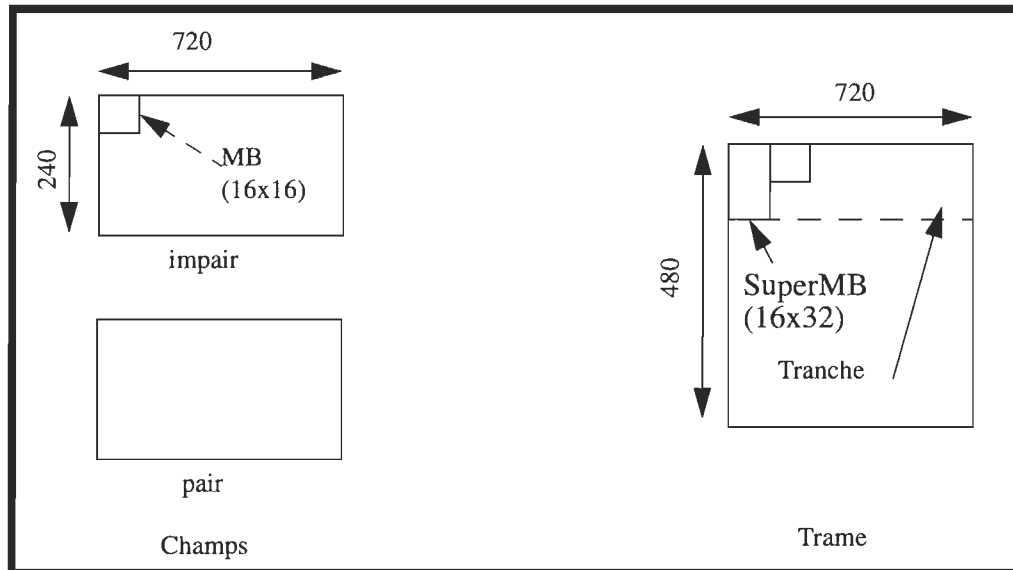


Figure 3.2 Segmentation de l'image en champ (vidéo entrelacée), trames (vidéo progressive), tranches et macroblocs.

Le débit est le paramètre qui est le moins souvent respecté dans le format restreint. En effet, un encodeur qui utilise intelligemment les spécifications MPEG-2 produit des images codées en format SIF dont la qualité atteint la saturation à environ 2 Mbps. Par contre un encodeur qui produit un code ne contenant pas de trame prédites ou interpolées, requiert environ 4 Mbps pour atteindre une qualité équivalente.

Segmentation

Avant le codage MPEG-2, les données vidéo sont segmentées en blocs. La dimension d'une image de qualité standard *Main Level at Main Profile* ou *ML@MP* (suivant la

terminologie du standard MPEG-2) a une résolution de 720x480 à 25 Hz ou 30 Hz selon qu'on utilise la norme PAL ou NTSC. Une tranche est formée d'une ou plusieurs lignes complètes de macroblocs (MB). Une trame est constituée de tous les MB qui forment une image complète. Un champ est formé d'une partie des MB formant une image complète (ex. images entrelacées). La dimension typique d'un MB est de 16x16 ou 16x32 pour les super-macroblochs, voir figure 3.2. La dimension d'un champ (pair ou impair) est égale à la moitié de celle d'une trame.

3.1.2 Transformée en cosinus discrète (TCD)

La TCD prend les blocs de pixels d'une image et les représente dans un autre domaine, en l'occurrence, dans le domaine des fréquences. Le codage par transformée permet de décorréler les données vidéo (trouver une représentation avec une corrélation minimale entre les éléments) et de minimiser le nombre de coefficients (de la transformée) ayant une amplitude significative (*energy compaction*). La compression est possible parce que les coefficients obtenus de la TCD peuvent être représentés avec une certaine perte de précision (d'où la quantification) et certains coefficients (insignifiants) peuvent être éliminés par un sous-échantillonnage non-uniforme.

Le choix de la taille des blocs de 8x8 plutôt que 4x4, 16x16 ou 32x32 est le résultat de plusieurs expériences et simulations intensives. Des tests de performances confirment en effet ces choix (Netravali, 1991). Conceptuellement, la TCD sur une dimension (TCD-1D) est semblable à la transformée de Fourier d'une séquence infinie symétrique par rapport à l'axe des ordonnées. En effet, la transformée de Fourier d'une fonction purement réelle et paire ne contient que des valeurs réelles correspondants aux termes en cosinus. La TCD est l'équivalent de la transformée de Fourier d'une fonction paire. Pour une séquence donnée, la TCD est plus simple et plus rapide que la transformée de Fourier. La TCD sur deux dimensions (TCD-2D) peut-être obtenue directement ou en

combinant la TCD-1D sur les colonnes et la TCD-1D sur les lignes. Ceci est possible car la TCD est une transformée séparable. Une équation explicite de la TCD-2D pour un bloc de 8x8 peut être décrite en terme de pixels $f(i,j)$ dans le domaine spatial, et de coefficients dans le domaine fréquentiel $F(u,v)$ comme suit:

$$f(i, j) = \frac{1}{4} C_u C_v \sum_{u=0}^7 \sum_{v=0}^7 F(u, v) \cos\left((2i+1) u \frac{\pi}{16} \right) \cos\left((2j+1) v \frac{\pi}{16} \right)$$

$$\text{où } C_u = C_v = \frac{1}{\sqrt{2}} \quad \text{si } u, v = 0 \quad (3.5)$$

$$C_u = 1 \quad \text{et } C_v = 1 \quad \text{autrement .}$$

L'équation de la TCD inverse est la suivante: (3.6)

$$F(u, v) = \frac{1}{4} C_u C_v \sum_{u=0}^7 \sum_{v=0}^7 f(i, j) \cos\left((2i+1) u \frac{\pi}{16} \right) \cos\left((2j+1) v \frac{\pi}{16} \right)$$

La figure 3.3 montre l'effet des différentes transformations (TCD, quantification, quantification inverse et TCD inverse) sur les pixels d'un bloc de 8x8 d'une image quelconque. La matrice de quantification est déterminée en fonction de la classe d'images à transmettre. Elle est transmise au décodeur au début (entête) de la séquence comprimée. La comparaison entre le premier pixel en haut à gauche (139) de l'image originale et le premier pixel en haut à gauche (144) de l'image reconstruite montre une légère perte d'information (chrominance ou luminance). La méthode de transformation choisie par le CCITT est la TCD bi-dimensionnelle par blocs de 8x8 pixels.

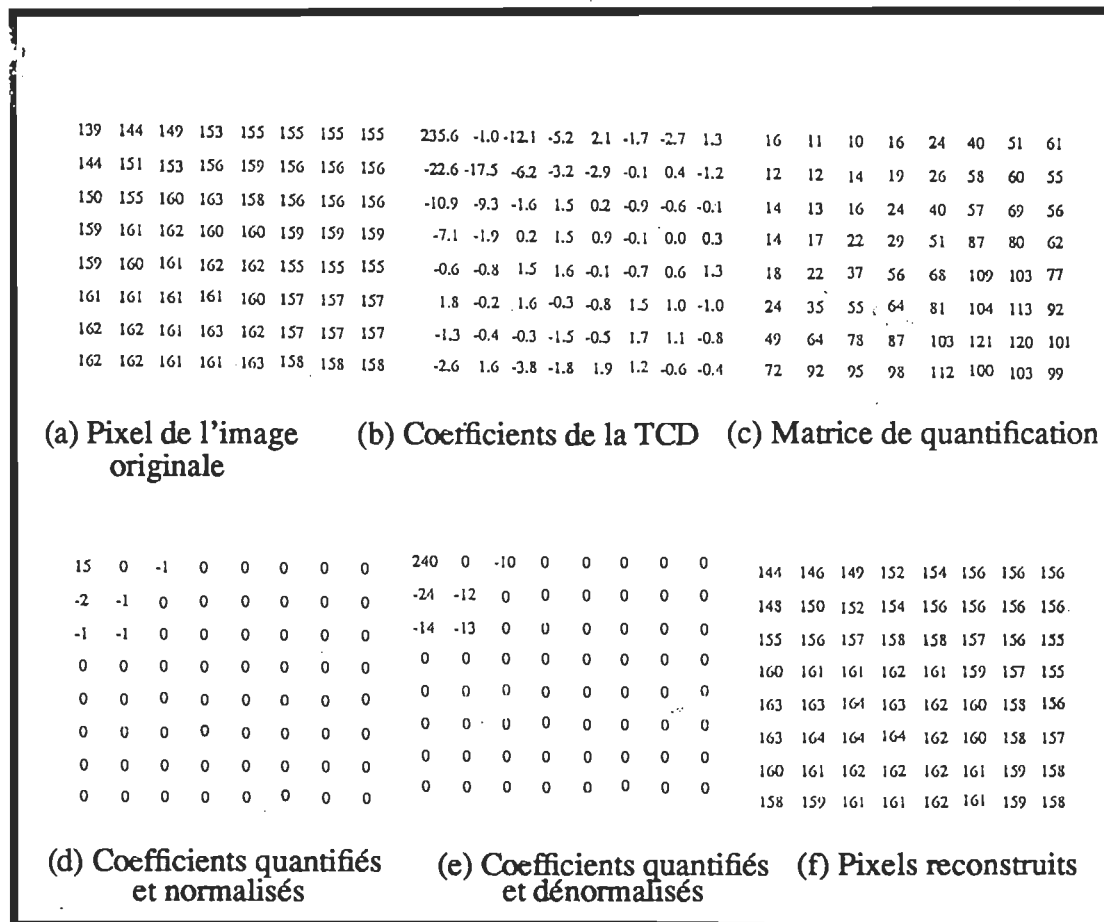


Fig. 3.3 Exemple de TCD, quantification et TCD inverse d'un bloc de 8x8 pixels

Les recherches sur la transformée en cosinus discrète ont été poussées au maximum ces dernières années. Pour les applications visées par la norme MPEG, la TCD est considérée dans l'ensemble de la communauté scientifique comme étant la technique de transformation la plus efficace, réalisable avec une complexité raisonnable. En pratique, la TCD est calculée suivant un algorithme rapide (MPEG Video Committee Draft, 1994). Le tableau 3.2 illustre le facteur de gain de l'algorithme rapide de calcul de la TCD.

Tableau 3.2: Gain en complexité d'une TCD rapide

N	TCD	TCD rapide	Facteur de gain
4	16	8	2
8	64	24	2.7
16	256	64	4
32	1024	160	6.4
64	4096	384	10.7
128	16386	896	18.3
256	65536	2048	32

3.1.3 Encodeur MPEG-2

Le codage MPEG est destiné à la compression d'images animées, il comprend le codage vidéo, le codage du son et celui des signaux de synchronisation. La résolution maximale de l'image est de 1920x1152 pixels avec un espace chromatique dans le rapport 4:2:0. Il est évident que lorsque les images de cette dimension sont codées, la compression n'est possible dans des délais raisonnables que si l'on fait usage d'ordinateurs extrêmement puissants. La recherche du minimum de redondance dans une séquence vidéo a conduit à déterminer deux techniques de compression, l'une intra-image (génère les trames I), l'autre inter-image (génère les trames P et B). La première consiste en le codage de chaque image indépendamment des autres et la seconde à coder les différences existant entre des trames successives (figure 3.4). La trame P est prédite à partir de la trame I. La trame B est une interpolation de la trame I et la trame P. L'interpolation se fait en calculant la moyenne entre deux blocs. Ceci a l'avantage de réduire le bruit dans l'image de moitié. Toutefois, l'interpolation a certains désavantages: notamment, le délai, la complexité et l'espace mémoire nécessaire pour garder les deux trames requises pour coder la trame B. Le codage inter-image utilise une estima-

tion de mouvement entre deux images consécutives. Il prend en compte le mouvement en recherchant le déplacement des blocs de pixels d'une image à l'autre.

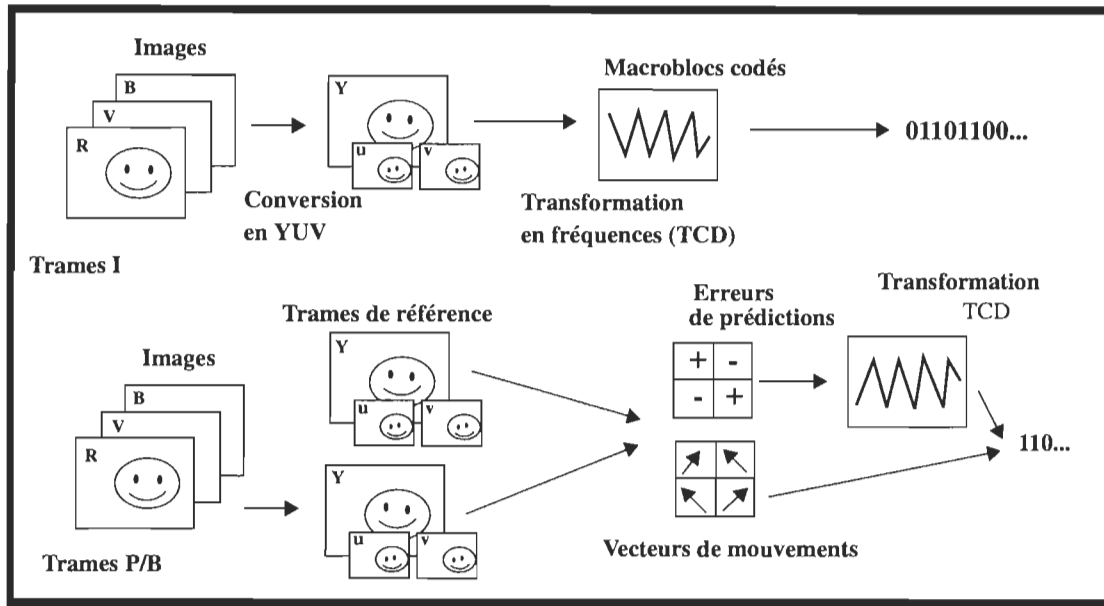


Figure 3.4 Procédure d'encodage vidéo

La différence entre les images s'effectue alors entre un bloc et le bloc correspondant dans l'image suivante déplacé du vecteur de mouvement. Le résultat de la transformée en cosinus discrète (TCD) sera ordonné dans une matrice de 64 coefficients afin que la valeur moyenne (coefficient DC) soit le premier en haut à gauche. C'est le cas du coefficient ayant la valeur 235.6 dans la figure 3.3 b). Le coefficient le moins significatif valant 0.4 en bas à droite de la figure 3.3 b). Le reste des coefficients seront classés par ordre d'importance relative par rapport à un critère de corrélation. Ceux de hautes fréquences (images avec beaucoup de détails) étant les moins importants, sont de plus en plus éloignés du coefficient continu (DC) dans la matrice résultante. Les fréquences verticales sont représentées par les coefficients sur les lignes de la matrice, et les fréquences horizontales sont représentées par les coefficients sur les colonnes de la matrice. Si la transformation inverse était appliquée après la TCD, l'image originale serait parfaitement reconstituée. Cette étape en elle-même ne contribue donc pas à la

compression. La compression commence par la quantification de l'image, qui consiste à diviser l'amplitude de chaque fréquence spatiale par un coefficient distinct (fig. 3.3.c). Ainsi le nombre de valeurs possibles après quantification est limité. C'est lors de cette opération que s'effectue la perte d'information dans la compression de l'image. Les blocs de pixels sont différenciés en blocs de luminance et de chrominance par le passage en format YUV. La compression peut être augmentée sur les blocs de chrominance par rapport aux blocs de luminance.

Mode de décisions

Pour augmenter l'efficacité du codage d'une image, chaque macrobloc est analysé et codé suivant un mode de codage spécifique (Coppisetti, 1993). Le choix du mode de codage dépend du type de trame (I, P ou B) et est basé sur la variance des coefficients de la TCD. Les macroblocs peuvent être codés suivant une des techniques suivantes: le codage prédictif compensé par le mouvement ou non et le codage intra-image ou inter-images (voir figure 3.5). La prédiction compensée par le mouvement est une technique qui consiste à anticiper le déplacement d'un objet à partir de la connaissance des images antérieures. Cette technique peut être implémentée de deux façons. La première utilise des équations récursives (calculs de gradients) pour déterminer le vecteur de déplacement dans l'image, et ne transmet que l'erreur de prédiction. Le décodeur effectue le même calcul pour reconstituer l'image. Comme dans tous les systèmes récursifs à prédiction, le résultat peut diverger en raison des erreurs de transmission. La seconde technique cherche à mettre en correspondance les blocs en mouvement dans l'image courante avec ceux dans l'image précédente. Les images sont découpées en blocs et la prédiction compensée par le mouvement s'opère en deux temps.

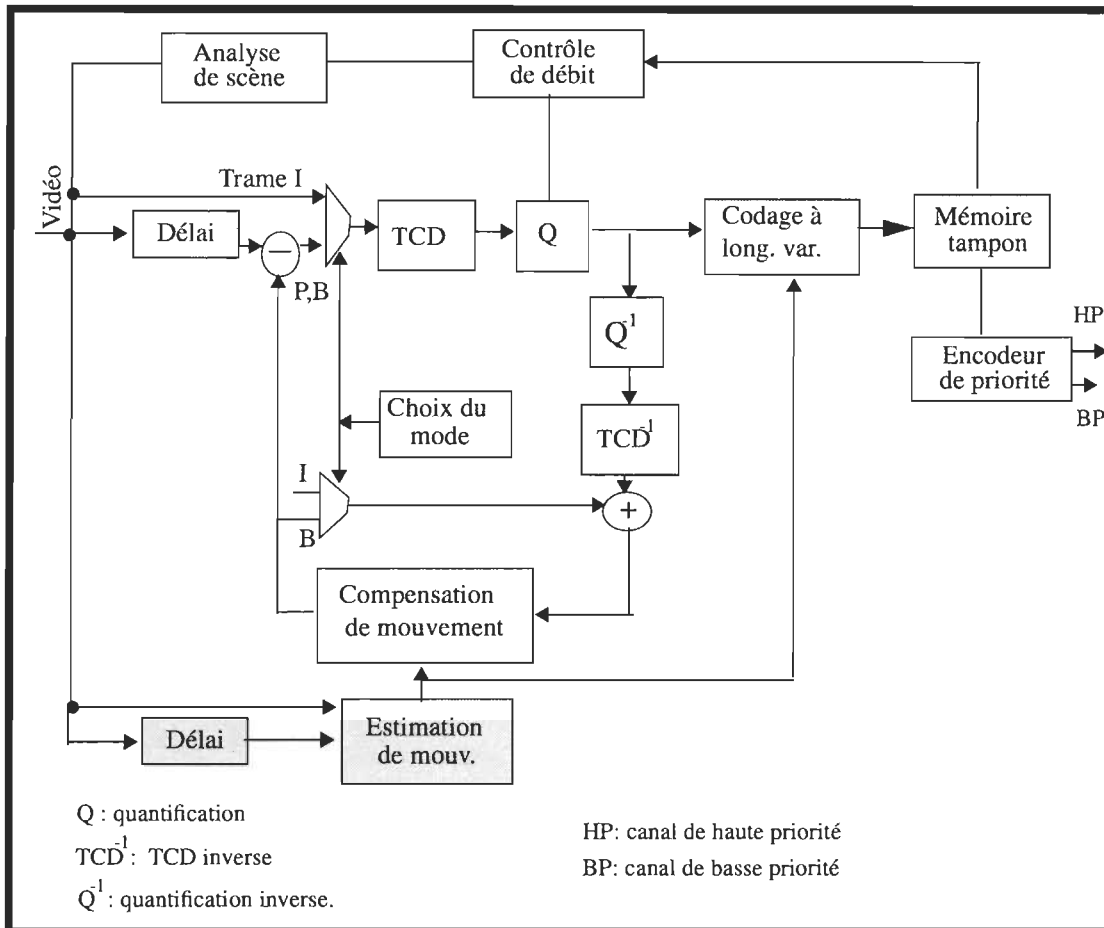


Figure 3.5 Encodeur MPEG-2

Dans un premier temps il y a, l'estimation du mouvement, c'est-à-dire la recherche dans l'image précédente d'une zone centrée sur une projection du bloc analysé. La longueur du vecteur de déplacement calculé est limitée à un maximum de 2047 (typiquement à un maximum de 15). Si cette longueur est plus petite que le déplacement réel d'un objet dans l'image, la compensation de mouvement devient inefficace et la qualité des images baisse. D'un autre côté, si cette longueur est trop importante, un nombre trop élevé de bits sera alloué au vecteur de mouvement pendant la transmission. Dans ce cas, la qualité des images est indirectement réduite aussi.

Dans un deuxième temps il y a, la compensation du mouvement, c'est-à-dire l'action de faire coïncider la position antérieure du bloc et la projection orthogonale du même bloc dans l'image courante en utilisant le vecteur de mouvement (ou de déplacement) calculé dans la première étape (estimation de mouvement). Les deux étapes se font dans le codeur, mais seule la compensation de mouvement se fait dans le décodeur. Plus robuste que la prédiction spatiale, la prédiction compensée par le mouvement nécessite la transmission du vecteur déplacement associé à chaque bloc, ce qui conduit à réserver une part du taux de transmission à cet effet. En revanche, la complexité du décodeur est moindre par rapport à celle du codeur.

Après la transformation et la quantification, le bloc de fréquences quantifiées peut contenir beaucoup de coefficients nuls (zéros). Pour obtenir une compression de donnée supplémentaire, un codage arithmétique est utilisé. Celui-ci consiste à coder une suite de coefficients nuls par un seul coefficient nul suivi du code indiquant le nombre de répétitions. Ainsi, au lieu de transmettre tous les coefficients nuls consécutifs obtenus après la TCD, un code plus court est transmis. Le nombre de coefficients nuls consécutifs est augmenté par un balayage des blocs de 8 X 8 pixels des basses fréquences vers les hautes fréquences. À la suite du codage des zéros, on effectue un codage à longueur variable qui consiste à choisir pour un ensemble de valeurs possibles (alphabet), un code tel que les valeurs les plus fréquentes soient codées avec peu de bits. Le codage de Huffman est le codage à longueur variable le plus couramment utilisé en compression d'images. Les données encodées sont par la suite transmises ou stockées. La figure 3.6 illustre la structure des différents types de trames codées et l'ordre de transmission par rapport à l'ordre de réception. Le nombre de trames de type B consécutives est variable. Si les trames de type B (trames interpolées de I et de P) ne sont pas utilisées (ce qui résulte en un délai de codage plus court et une complexité réduite), on a une séquence semblable à celle-ci: I P P P I P P P I P P P I.

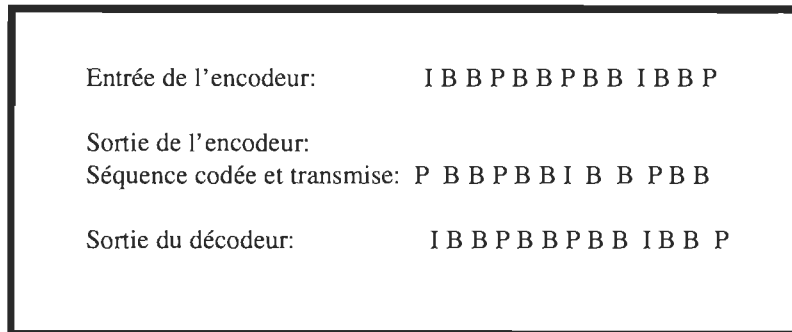


Figure 3.6 Ordre de transmission et de réception des trames I, P et B

Deux paramètres importants M et N sont utilisés pour indiquer la façon dont les images sont codées aussi bien dans la norme MPEG-2 que dans la norme MPEG-1. Un groupe d'images contient N trames dont la première est de type I et M trames de type P sont prédites de la première (la trame I). La longueur typique d'un groupe d'images est de 15 (avec $N=15$ et $M=3$).

3.1.4 Décodeur MPEG-2

Le décodeur est fondamentalement plus simple que l'encodeur. Il reçoit une séquence MPEG-2 comme entrée et produit des fichiers contenant des composantes Y, U et V. Le nombre de fichiers produits dépend de la fréquence d'affichage (25 ou 30 images/s) et de la longueur de la séquence d'images encodées (typiquement, une séquence vidéo de 5 secondes contient $30 \times 3 \times 5 = 450$ fichiers YUV). La figure 3.7 illustre le schéma bloc du décodeur MPEG-2. Lorsqu'un seul canal est disponible et qu'on ne fait pas usage du codage pyramidal ou du partitionnement des données, seul le canal de basse priorité (BP) est utilisé.

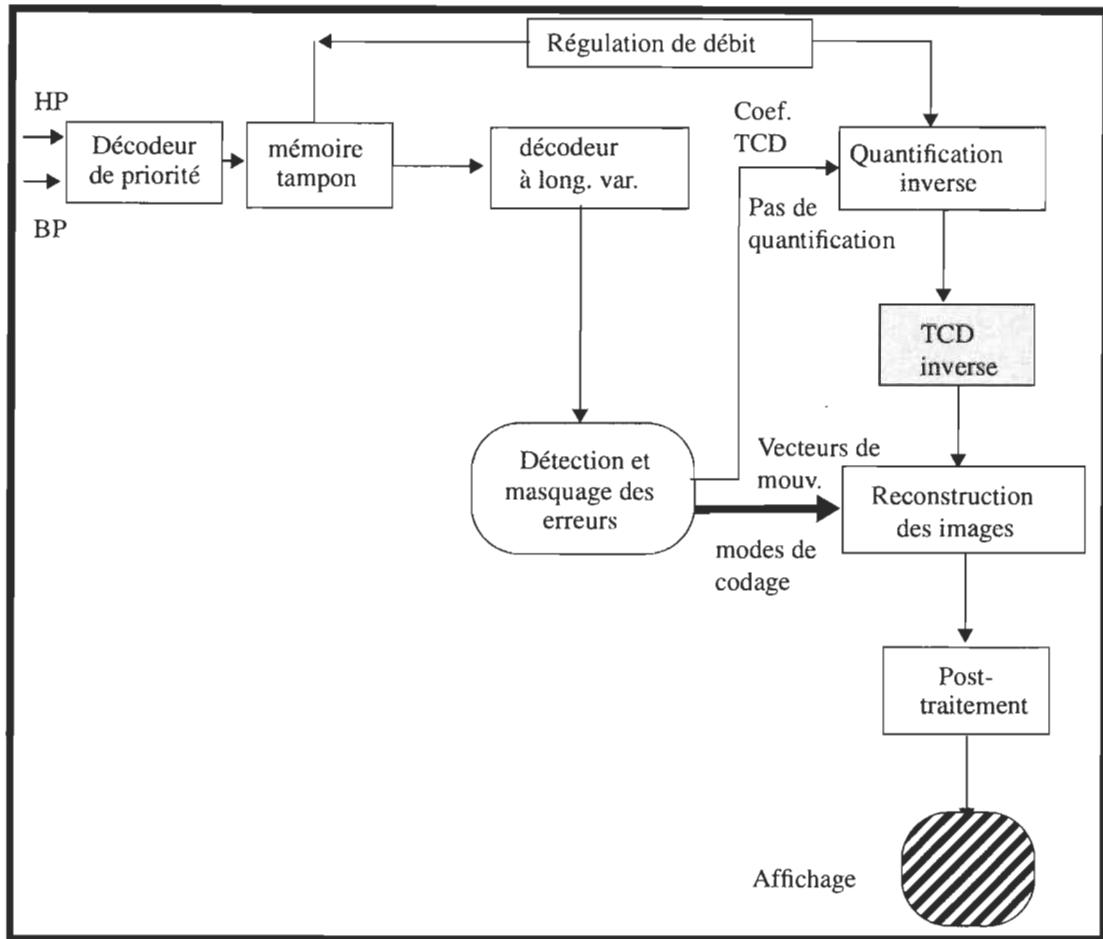


Figure 3.7 Le décodeur MPEG-2

3.2 Structure hiérarchique des données encodées MPEG-2

La structure des données encodées MPEG-2 est composée de plusieurs couches. Nous allons décrire brièvement les différentes couches afin de donner une vue d'ensemble de la hiérarchie de l'information (figure 3.8).

1. La première couche est celle de la séquence vidéo. Elle est très importante pour la signalisation de l'information. La résolution de l'image, la proportion de l'image (*aspect ratio*), le taux de transfert, les matrices de quantification utili-

sées (sinon la matrice par défaut) sont quelques paramètres que l'on retrouve dans cette couche. L'algorithme MPEG-2 est compatible avec l'algorithme MPEG-1.

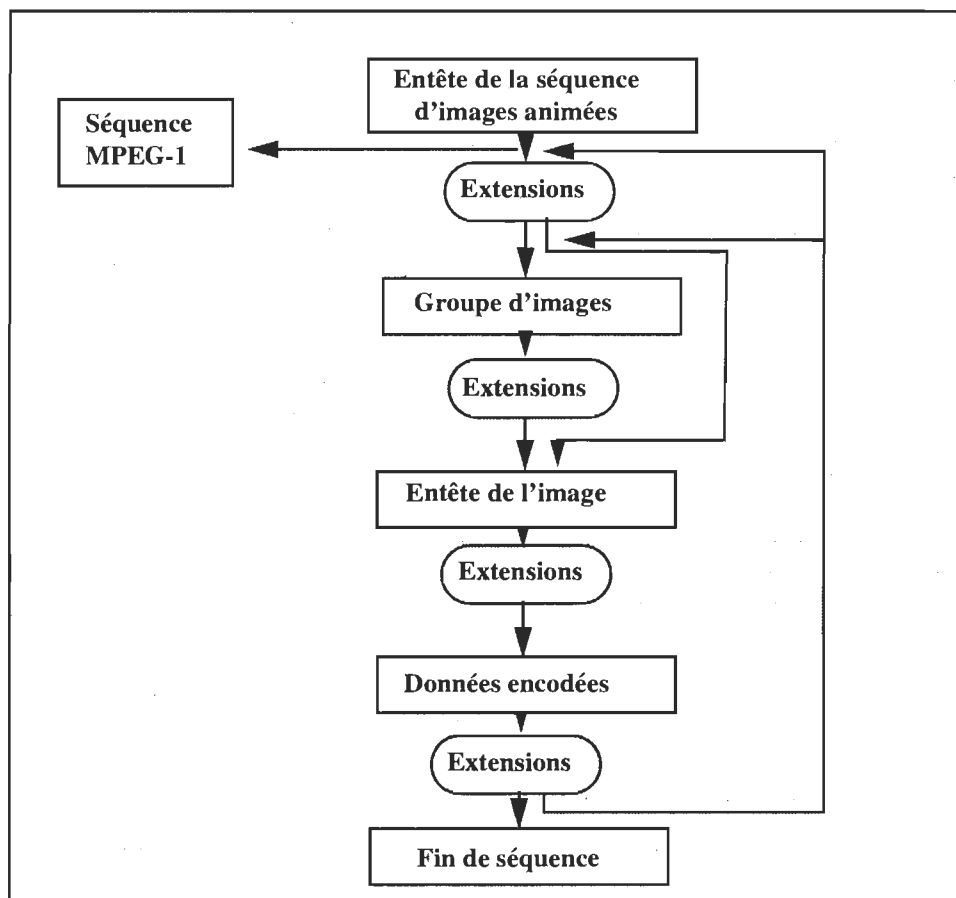


Figure 3.8 Structure hiérarchique des données encodées MPEG-2

2. La couche suivante est celle du groupe d'images ("*Group of Pictures*" ou GOP). Elle contient les informations relatives à la synchronisation et nécessaires à la reconstruction du signal vidéo.
3. La troisième couche est la couche d'images; elle décrit la position courante de la séquence vidéo (par rapport au GOP) ainsi que le type d'image (I, B ou P), le

niveau de la mémoire tampon, la référence temporelle et d'autres informations pertinentes pour le décodeur.

4. Dans la couche des tranches ("*slices layer*"), l'image est divisée en tranches de façon à réduire l'effet d'une erreur de transmission sur la qualité des images. Ces tranches sont placées les unes après les autres dans l'ordre de balayage de l'image. Elle consiste en une ou plusieurs lignes de macroblocs dont la description suit.
5. La couche des macroblocs est l'élément le plus petit qui décrit l'image dans le processus de codage. Le signal source est composé de la luminance Y et de deux chrominances U et V dans un rapport 4:1:1. Puisqu'on dispose de 4 fois plus de blocs de luminance que de blocs de chrominance, un macrobloc contient 4 blocs (8x8 pixels) de luminance et un bloc pour chaque chrominance pour un total de 6 blocs. Tous les blocs de chaque composante de l'image dans un macrobloc partagent les mêmes propriétés utilisées pour l'estimation et la compensation de mouvement. Pour cette raison, ils sont codés ensembles (le même vecteur de mouvement est appliqué sur tous les blocs d'un macrobloc). L'ordre de transmission commence avec les 4 blocs de luminance et puis les blocs de chrominance.
6. La couche des blocs représente les blocs 8x8 qui forment un macrobloc de l'image.

La séquence binaire Mpeg-2 (*bitstream*)

Les données vidéo encodées MPEG -2 commencent avec un code de synchronisation (*Start code*) qui indique le début de la séquence. Celui-ci est formé d'un code préfixe de 23 bits de valeur 0 suivi d'un bit de valeur 1 et d'un code identificateur de séquence. La valeur de ce code est un entier de 8 bits qui indique le type de séquence (ex. 000001B8). L'entête complète de la séquence après le code de synchronisation peut

avoir une longueur totale d'environ 1500 bits si toutes les options sont utilisées.

3.3 Techniques de masquage d'erreur du CODEC

Les techniques de dissimulation d'erreurs du CODEC se résument en 3 catégories. La première catégorie comprend les techniques de codage temporel robuste, la deuxième comprend les techniques de codage spatial robuste et la dernière les techniques de codage en couches. Toutes ces techniques permettent de masquer les erreurs de transmission lorsqu'elles sont détectées et la limitation de l'influence de l'erreur dans l'espace (à l'intérieur d'une image) et dans le temps (d'une image à l'autre).

3.3.1 Codage pyramidal

La multiplicité des formats vidéo, des dimensions d'images, des résolutions et des applications rend inévitable l'utilisation du codage pyramidal spécifié dans le standard MPEG-2. La stratégie de protection adaptative que nous proposons au chapitre 4 s'adapte bien au codage pyramidal.

3.3.2 Prédiction temporelle robuste

La prédiction temporelle robuste comprend les techniques de substitution et de compensation de mouvement sur les macroblocs intra en cas d'erreurs de transmission. Elle est efficace pour des applications à haute qualité avec une dégradation progressive en présence d'erreurs de transmission.

3.3.3 Prédiction spatiale robuste

La prédiction spatiale robuste comprend la prédiction du coefficient DC et quelques coefficients AC d'un bloc perdu pendant la transmission à partir des blocs précédents reçus sans erreurs. Cette méthode permet aussi un codage multi-résolution des services vidéo interconnectés. Cette technique permet la compatibilité du standard MPEG-2 avec les autres tels que MPEG-1, H.261.

Nous avons présenté l'algorithme MPEG-2 utilisé pour compresser une séquence d'images vidéo dans un format normalisé. L'utilisation d'un tel algorithme dans un système de transmission vidéo via un canal sans-fil pose quelques problèmes. Dans un tel canal, en l'absence de protection de la voie numérique, que vaut un système de communication vidéo? Dès que le taux d'erreurs dans le canal de transmission dépasse un niveau critique (par ex. 1 symbole sur 1000) comment le système réagit-il? Dans une séquence constituée d'une image de référence codée et de plusieurs images prédites ou interpolées, que se passe-t-il si on perd l'image de référence pendant la transmission? Ou encore, qu'arrive-t-il, si on perd la matrice de quantification nécessaire pour décoder toute la séquence d'images à la réception? À l'opposé, la perte d'un coefficient de la transformée en cosinus discrète (codage spatial) dans une image donnée est-elle significative? Y a-t-il dans le contenu de l'information à transmettre, des données plus importantes que d'autres? En d'autres termes, doit-on accorder le même niveau de protection à tous les bits (symboles binaires de valeurs 0 ou 1) de la séquence générée par le codeur MPEG-2? Nous allons d'abord présenter la modélisation des caractéristiques physiques du canal de communication intra-édifice dans le chapitre suivant, avant de tenter de répondre à ces questions.

CHAPITRE 4

Modèle de canal et protection contre les erreurs de transmission

Les caractéristiques physiques du canal sont d'une importance capitale dans l'analyse et la conception de ce système. En règle générale, ces caractéristiques affectent de façon significative la conception de chaque module du système. Dans ce chapitre nous traiterons particulièrement du canal de communication sans-fil intra-édifice.

4.1 Phénomènes de propagation par chemins multiples

La réponse impulsionnelle est un moyen efficace pour caractériser un système linéaire, puisque la sortie du système peut-être obtenue en faisant une convolution du signal à l'entrée par la réponse impulsionnelle du système. Le canal de communication intra-édifice est aléatoire et non-stationnaire dans le temps. Pour chaque point de l'espace tri-dimensionnel, le canal se comporte comme un filtre linéaire variable dans le temps avec une réponse impulsionnelle donnée par l'équation à la figure 4.1.

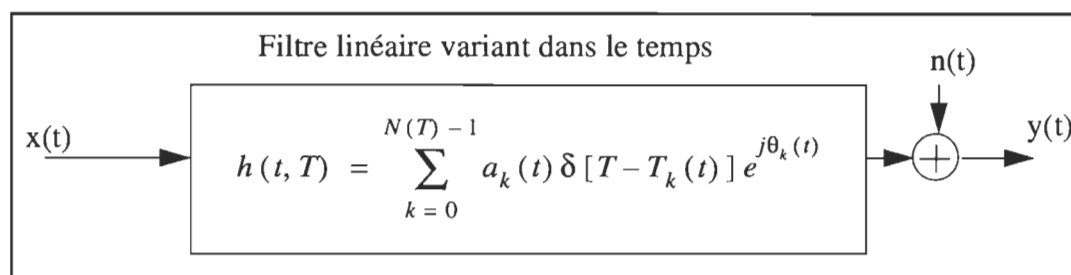


Figure 4.1. Modèle mathématique du canal intra-édifice

où t et T sont respectivement le temps d'observation et le temps d'application de l'impulsion, $N(T)$ est le nombre de chemins de propagation possibles, $\{a_k(t)\}$, $\{T_k(t)\}$ et $\{\theta_k(t)\}$ sont les paramètres aléatoires variant dans le temps et représentant respectivement, l'amplitude, le temps d'arrivée et la phase. $\delta(\cdot)$ est l'impulsion de Dirac. Le canal est complètement caractérisé par ces paramètres. La réponse impulsionnelle du canal pour un édifice de taille moyenne est illustré à la figure 4.2 et correspond au modèle à large bande. Ce modèle a l'avantage d'être assez général pour être utilisé afin d'obtenir la réponse du canal pour n'importe quel signal $s(t)$ en faisant une simple convolution de $s(t)$ et de $h(t)$ puis en additionnant le bruit $n(t)$. Comme le montre la figure 4.2 a), l'onde qui se propage en ligne droite est atténuée de -50 dB après 200 ns. Par contre l'onde qui a subi plusieurs réflexions est atténuée de -50 dB après plus de 400 ns.

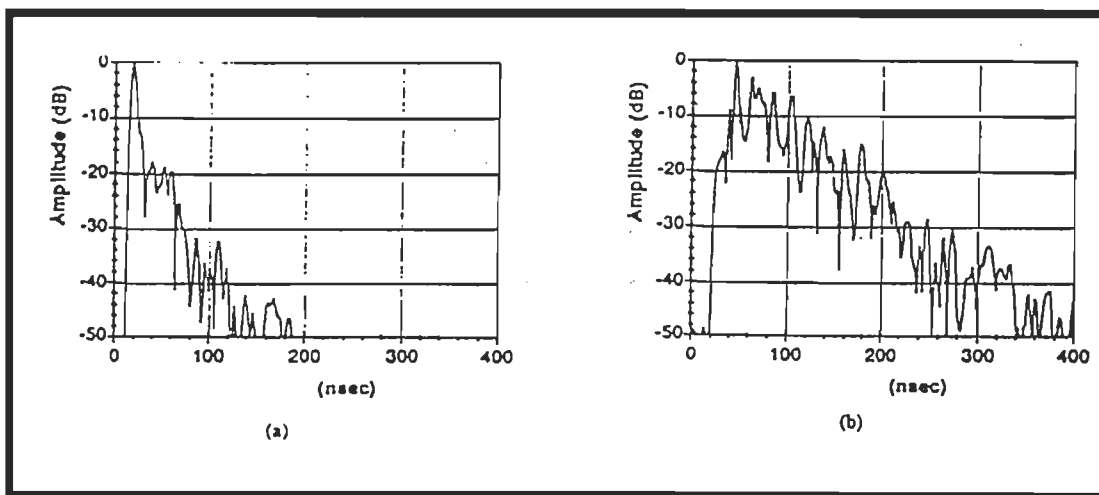


Figure 4.2 Réponse impulsionnelle du canal pour un édifice de taille moyenne. La séparation des antennes est de 5 m. (a) Propagation en ligne droite (LOS). (b) Propagation avec chemins multiples.

L'approche de la réponse impulsionnelle décrite dans la section précédente est basée sur le modèle géométrique de la figure 4.3. Le signal transmis par la base arrive aux récepteurs portatifs via une ou plusieurs ondes principales. Ces ondes principales, consistent en un rayonnement en ligne directe et plusieurs rayonnements diffractés et réfléchis par les structures telles que les murs, le plafond, le sol, etc. L'onde directe peut éventuellement être atténuée au point où elle devient indétectable. Les ondes principales arrivent à proximité du récepteur portatif de façon aléatoire, puis elles se dispersent dans l'environnement du portatif à cause des équipements avoisinants. Les signaux résultants pour chaque rayonnement arrivent avec des délais presque identiques et sont affectés par la même atténuation, mais ont des phases différentes dues aux différences de longueurs dans les chemins. Les différents chemins sont additionnés en fonction de leur temps d'arrivée, leur amplitude, leur phase et l'enveloppe complexe résultante est captée par le portatif. Le nombre de chemins distincts mesurés pendant une expérience à un instant donné et à un point donné dans l'espace dépend de la forme, de la structure de l'édifice et de la résolution des mesures expérimentales.

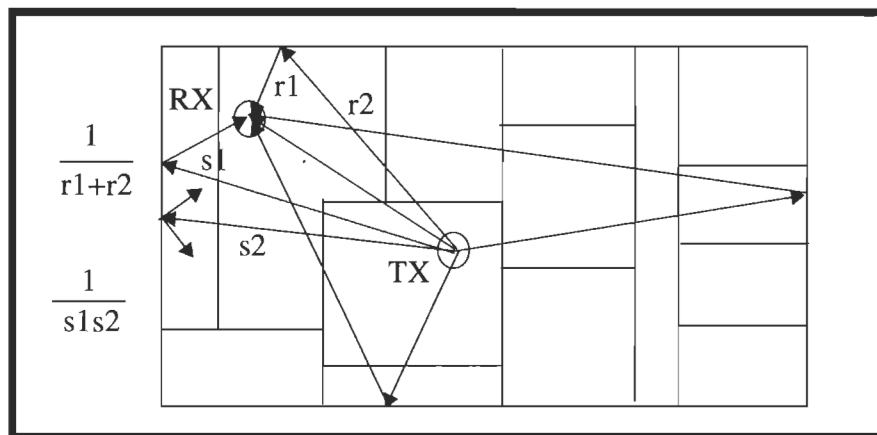


Figure 4.3 Modèle géométrique de propagation radio dans un canal intra-édifice.

Les profils de réponse impulsionnelle collectés pour deux portatifs sur deux sites 1 et 2 illustrés sur la figure 4.4 sont normalement très différents à cause des différences dans la structure le long du trajet base-portatif et des différences dans l'environnement local des récepteurs portatifs.

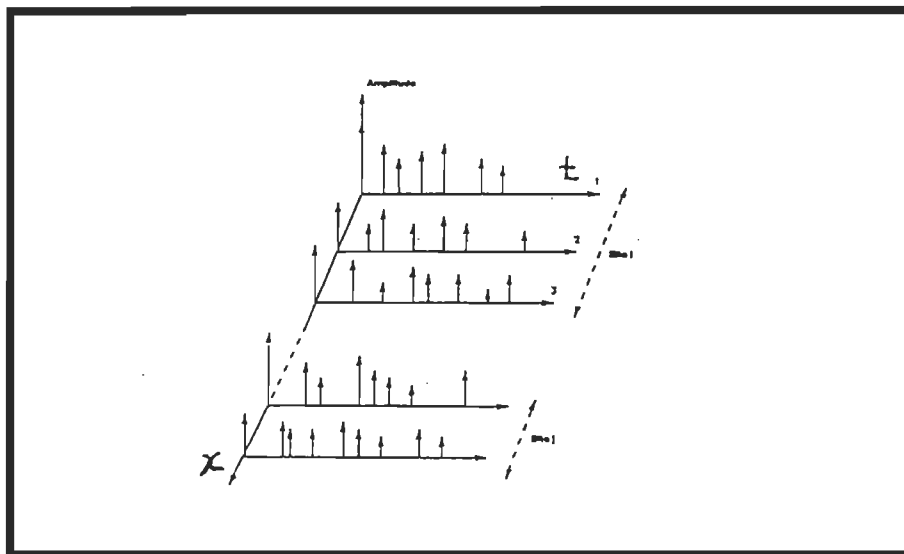


Figure 4.4 Séquences des profils pour des points adjacents dans l'espace

4.1.1 Caractérisation de la réponse impulsionnelle

Nous avons utilisé la représentation discrète de la réponse impulsionnelle (Saleh, 1987) décrite par l'équation suivante:

$$h(t) = \sum_k \beta_k e^{j\theta_k} \delta(t - T_k) \quad (4.1)$$

où β_k est le gain en amplitude (réel positif) de chaque rayonnement, T_k est le délai de propagation, θ_k le décalage de phase et k l'indice représentant le chemin. En principe k varie de 0 à ∞ . $\delta(\cdot)$ est l'impulsion de Dirac. À cause du mouvement des person-

nes et des équipements dans l'édifice, les paramètres β_k , T_k et θ_k sont des variables aléatoires variant dans le temps. Toutefois, puisque leur taux de variation est faible, ces variables aléatoires peuvent être considérées comme invariantes dans le temps. Soit $x(t)$ le signal à transmettre:

$$x(t) = A(t) e^{j(\omega t + \phi)} \quad (4.2)$$

où $A(t)$ est la forme du signal, ω la fréquence angulaire et ϕ la phase du signal.

Le signal reçu $y(t)$ est la convolution du signal $x(t)$ par $h(t)$:

$$y(t) = \sum_k \beta_k A(t - T_k) e^{j[\omega(t - T_k) + \phi + \theta_k]} \quad (4.3)$$

Si on considère que les ondes arrivent en groupes (*clusters*), le temps d'arrivée de ces groupes forme un processus de Poisson avec un taux d'arrivée Λ fixe. Dans une groupe, les ondes arrivent aussi suivant un processus de Poisson avec un taux d'arrivée également fixe. Typiquement un groupe est constituée de plusieurs ondes. Soit T_n le temps d'arrivée de le n-ième groupe, $n = 0, 1, 2, \dots$, et soit T_{kn} le temps d'arrivée de la k-ième onde de le n-ième groupe, $k = 0, 1, 2, \dots$. Par définition $T_{0n} = 0$ et $T_0 = 0$ donc T_{kn} et T_n sont décrits par une fonction inter-arrivée ayant une densité de probabilité indépendante exponentielle:

$$p(T_n | T_{n-1}) = \Lambda e^{-\Lambda(T_n - T_{n-1})}, \quad n > 0 \quad (4.4)$$

$$p(T_{kn} | T_{(k-1)n}) = \Lambda e^{-\Lambda(T_{kn} - T_{(k-1)n})}, \quad n > 0$$

Soit β_{kn} le gain en amplitude de la k-ième onde dans le n-ième groupe et sa phase θ_{kn} .

La réponse impulsionnelle complexe du canal sera notée par:

$$h(t) = \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} \beta_k e^{j\theta_{kn}} \delta(t - T_n - T_{kn}) \quad (4.5)$$

La figure 4.5 représente schématiquement le modèle de canal intra-édifice. La puissance moyenne des groupes d'ondes est donnée par:

$$\overline{\beta_{kn}^2} \equiv \overline{\beta^2(T_n, T_{kn})} \quad (4.6)$$

$$\overline{\beta_{kn}^2} \equiv \overline{\beta^2(0, 0)} e^{-T_n/\Gamma} e^{-T_{kn}/\gamma}$$

où $\overline{\beta^2(0, 0)} \equiv \overline{\beta_{00}^2}$ est la puissance moyenne de la première onde dans le premier

groupe d'ondes. $\overline{\beta_{kn}^2}$ est la puissance moyenne de la k-ième onde dans le n-ième groupe.

Γ et γ sont respectivement la constante de délai de puissance du groupe et celle de l'onde. Pratiquement, la sommation sur les n s'arrête lorsque $e^{-T_n/\Gamma}$ est très petit par

rapport à 1 et la sommation sur les n s'arrête lorsque $e^{-T_{kn}/\gamma}$ est très petit par rapport à

1. Typiquement, un groupe d'ondes dure environ 200 ns.

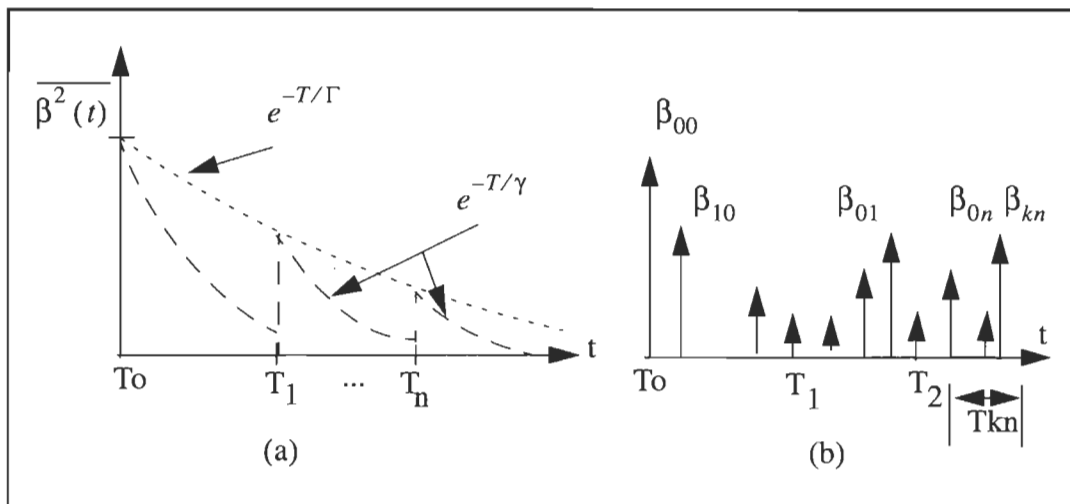


Figure 4.5 Le modèle de réponse de canal. (a) Puissance moyenne des groupes d'ondes. (b) Réponse impulsionnelle

4.1.2 Délai maximum et étalement de délai

La moyenne quadratique du moment de la puissance du profil de délai, encore appelé étalement de délai, T_{rms} est une représentation numérique (quantification grossière), d'un profil de réponse impulsionnelle et est donnée par la formule suivante:

$$T_{rms} = \left[\frac{\sum_k (t_k - T_m - t_A)^2 a_k^2}{\sum_k a_k^2} \right]^{1/2} \quad (4.7)$$

t_A est le temps d'arrivée du premier chemin dans un profil, t_k le temps d'arrivée du profil, a_k l'amplitude des ondes et T_m est le délai moyen défini comme:

$$T_m = \left[\frac{\sum_k (t_k - t_A) a_k^2}{\sum_k a_k^2} \right] \quad (4.8)$$

T_{rms} est une bonne indication de l'étalement de délai. Elle donne l'indication d'une éventuelle interférence entre symboles. Elle varie proportionnellement avec l'écho (par rapport au rayonnement en ligne droite) et le délai. Il a été démontré (Jakes, 1974) que la performance des systèmes de communication opérant dans un environnement intra-édifice était très sensible à la valeur de T_{rms} . Une valeur typique moyenne de T_{rms} est située entre 20 et 50 ns pour les édifices à bureaux de taille moyenne et petite (Saleh, 1987), entre 30 et 300 ns pour un environnement industriel, et moins de 100 ns pour la plupart des édifices universitaires (Huang, 1992).

4.1.3 Atténuations par pertes de transmission

La réponse impulsionnelle est utile pour caractériser le canal intra-édifice au niveau microscopique. Les pertes de transmission décrivent le canal intra-édifice au niveau macroscopique. La perte de transmission dans un environnement intra-édifice est essentielle pour déterminer la couverture radio du système de communication et

l'emplacement optimal de l'antenne de la base. Le tracé des diagrammes de contour de propagation tri-dimensionnels basés sur les plans et la structure des matériaux de construction de l'édifice est un travail énorme (Hashemi, 1993). Une connaissance détaillée des modèles efficaces de pertes de transmission est requise. L'environnement de propagation est complexe, et de surcroît, il est largement connu qu'un modèle décrivant les pertes de transmission de façon efficace n'existe pas encore.

Tableau 4.1: Quelques valeurs typiques de η

Propagation en ligne droite dans un entrepôt (Rappaport, 1989)	1.5 - 1.8
Propagation avec obstruction dans un entrepôt (Rappaport, 1989)	2.4 - 2.8
Grandes surfaces (ex. supermarchés, salles de bureaux (Seidel, 1992)	2
Propagation à 37.2 GHz et 21.6 GHz respectivement dans un édifice universitaire (Kalivas, 1992)	1.1 - 1.2

Une revue de la littérature sur les mesures de propagation intra-édifice indique qu'il existe quatre modèles distincts qui décrivent plus ou moins bien le canal intra-édifice. Nous allons nous limiter à la description du premier modèle, les autres étant des variantes.

La puissance du signal $P(d)$ reçue varie de façon proportionnelle à l'inverse de la distance à la puissance η entre les antennes.

$$P(d) = P_o d^{-\eta} \quad (4.9)$$

d est la distance qui sépare l'antenne du récepteur de l'antenne du transmetteur. P_o est la puissance reçue pour $d = 1 \text{ m}$. Elle dépend de la puissance du transmetteur, de la fréquence d'émission, de la hauteur et du gain des antennes, etc. Le paramètre η dépend de l'environnement. La courbe d'atténuation de transmission résultante correspond à une ligne droite avec une pente de $10\eta \log d$. Des valeurs typiques de η sont répertoriées dans le tableau 4.1.

4.2 Performance des réseaux de campus

L'analyse de performance des systèmes de communication sans-fil intra-édifice semble complexe en raison des conditions de propagation et des traitements sophistiqués requis à l'émetteur et au récepteur. Le concepteur du système de communication peut adopter l'une des approches suivantes:

- (1) - l'analyse de performance utilisant un modèle mathématique basé sur des considérations réalistes. Cette approche est pratiquement impossible à réaliser parce que trop complexe.
- (2) - l'analyse de performance utilisant un modèle mathématique basé sur une théorie simplifiée.
- (3) - l'analyse de performance utilisant un modèle statistique dont on se sert pour simuler la chaîne de transmission.

Dans le cadre de ce travail, nous utiliserons la troisième approche. Le modèle de canal sera simulé avec le logiciel SPW.

Performance des systèmes sans-fil intra-édifice

La réponse impulsionnelle décrite plus haut peut être utilisée pour caractériser complètement le canal. Cependant, une revue de la littérature et les efforts de modélisation de

certaines chercheurs (Hashemi, 1993) montrent que:

- (1) - Le nombre de chemins N pour un profil de réponse impulsionnelle varie de façon aléatoire. De plus, la valeur moyenne de N est différente d'un édifice à l'autre. La séquence des variables $\{a_k\}$, $\{t_k\}$, $\{\theta_k\}$ pour chaque point dans l'espace est une séquence aléatoire.
- (2) - Les chemins adjacents associés à un même profil de réponse impulsionnelle ne sont pas indépendants. Une distribution standard de Poisson est inadéquate pour décrire les temps d'arrivée de ces séquences. Les amplitudes adjacentes sont susceptibles d'avoir une corrélation dans l'évanouissement, si on fait des mesures avec une grande précision, puisqu'un certain nombre d'obstacles qui produisent ces évanouissements sont identiques.
- (3) - Les points du profil de réponse impulsionnelle qui sont rapprochés dans l'espace sont corrélés puisque la structure du canal ne change pas de façon appréciable sur une courte distance. La corrélation spatiale gouverne les amplitudes, les temps d'arrivées et les phases, ainsi que la variance et la moyenne des amplitudes. Il peut y avoir de petits changements dans les statistiques du canal et des variations globales importantes dues aux effets d'occlusions et à la non-stationnarité spatiale.
- (4) - La plupart du temps, les pertes de transmission sont très importantes dans la communication sans-fil intra-édifice. La dynamique de changement est appréciable sur une courte distance. Des règles simples de pertes de transmission peuvent être appliquées de façon satisfaisante pour décrire le comportement du canal mobile, mais pas pour le canal intra-édifice.

- (5) - Les paramètres du canal sont très dépendants de la forme, de la taille et de la structure des matériaux utilisés pour construire l'édifice. Des variations en fonction de la fréquence de l'onde transmise sont aussi possibles.

En général, le canal intra-édifice est non-stationnaire dans le temps. Les variations dans le temps sont dues au mouvement des personnes et des équipements autour des antennes émettrice et réceptrice. Un modèle de canal intra-édifice réaliste devrait prendre en considération tous ces facteurs. Mieux encore, il devrait dériver les valeurs de ces paramètres à partir des mesures expérimentales sur le terrain plutôt que d'une théorie simplifiée. La source d'information et le canal étant caractérisés, les données à transmettre doivent être adaptées au type de canal de communication que nous allons utiliser pour simuler le système de communication vidéo. C'est le rôle du codage de canal.

4.3 Codage de canal

Le codage de canal peut-être considéré comme l'adaptation du signal codé par la source au canal de communication. C'est un processus d'adaptation dans le sens où il tente de rendre les performances du système de transmission indépendantes des séquences de données transmises. En d'autres termes, le code de transmission doit être transparent à n'importe quelle séquence de données. Par la suite, un transcodage qui est une partie de la modulation permet de convertir les code binaires en des symboles de transmission possédant une densité spectrale de puissance mieux adaptée au canal de communication. Les erreurs de transmission peuvent apparaître de façon aléatoire, ou bien en rafale (lorsqu'une simple impulsion de bruit affecte une série de symboles transmis consécutivement). Des codes ont été réalisés spécialement pour chacun de ces cas. Par exemple, les codes Fire (Fire, 1959) et RS (Reed et Solomon, 1960) sont très connus pour leur capacité de corriger les erreurs en rafale. Par contre, d'autres codes comme les codes BCH, permettent la correction d'erreurs multiples aléatoirement réparties. Fondamentalement, un code correcteur d'erreurs aléatoires corrige toute

combinaison de x erreurs sur une séquence d'une longueur donnée si chaque symbole est indépendamment affecté par le bruit. Les codes de correction de x erreurs aléatoires peuvent également servir à corriger les erreurs en rafale, grâce à une technique de codage par multiplexage et démultiplexage connue sous le nom de code à blocs entrelacés.

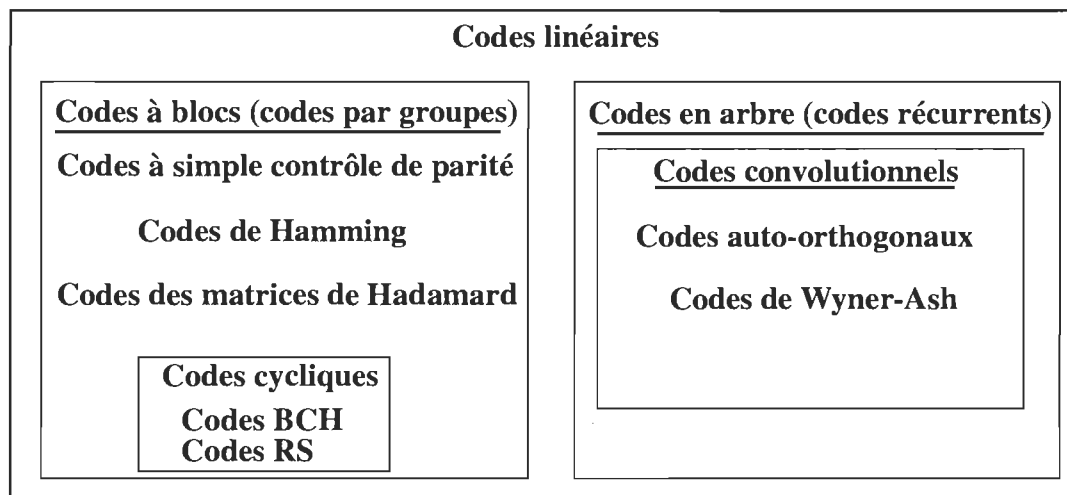


Figure 4.6 Classification de quelques codes de contrôle d'erreurs

Qu'il s'agisse d'erreurs aléatoires ou d'erreurs en rafale, les codes de contrôle d'erreurs sont habituellement classés en deux catégories: les codes à blocs et les codes récurrents. Mais, la plupart appartiennent au domaine plus général des codes linéaires. Tenant compte de cette classification, la figure 4.6 présente quelques uns des codes les plus connus. Nous avons utilisé les codes convolutionnels dans la conception de ce système de communication vidéo et nous allons donc décrire ces codes convolutionnels dans les sections suivantes de ce chapitre. Le lecteur qui serait intéressé par une étude mathématique approfondie d'une variété de codes de contrôle d'erreurs pourra se référer aux expériences de Michelson et Levesque (1985) ou à celle de Lin et Costello (1983).

Dans sa forme la plus simple, la protection contre les erreurs nécessite un code de détection d'erreurs (habituellement un simple contrôle de parité). Lorsqu'une erreur est détectée, le processus est arrêté ou, mieux encore, le message est ré-émis. Si la ré-émission est impossible, ou si elle représente un gaspillage, il est possible de faire en sorte que le récepteur évalue l'erreur, une fois qu'il l'a détectée, et remplace l'échantillon erroné par une valeur estimée.

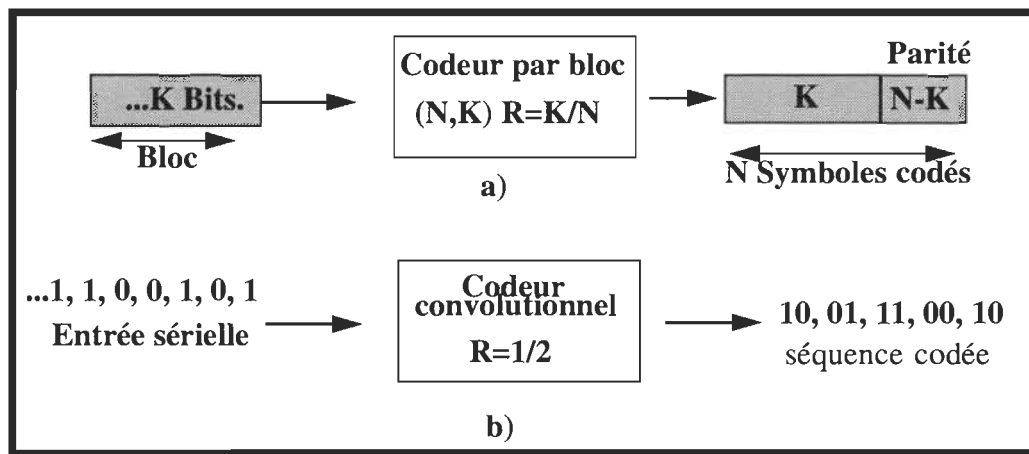


Figure 4.7 a) Codeur par bloc. b) Codeur convolutionnel.

Les techniques de contrôle d'erreurs se regroupent en trois grandes catégories:

- le codage correcteur d'erreurs (CCE) dans lequel on distingue deux autres types de codage: le codage par bloc illustré à la figure 4.7 a) et le codage convolutionnel, illustré à la figure 4.7 b)
- la détection d'erreurs et la correction par retransmission (CR)
- le codage hybride qui combine les deux techniques précédentes (CCE/CR).

Nous nous intéressons au codage correcteur d'erreurs utilisant les codes convolutionnels.

4.3.1 Codes convolutionnels

Le lecteur qui n'est pas familier avec les codes convolutionnels est référé à l'ouvrage de S. Lin et D. J. Costello (1983). Le taux de codage se réfère au rapport entre le débit des bits d'information et le débit des symboles. Les codes convolutionnels dominés par des codes de taux faible $R=1/N$ $\{1/2, 1/3, 1/4\}$ ont des gains de codage élevés. De plus les décodeurs sont pratiques. Lorsque le débit augmente, la largeur de bande disponible diminue. Toutefois, on note un certain intérêt vers les systèmes de codage correcteur d'erreurs ayant des taux de codage élevés $R=k/n$ $\{2/3, 3/4, 7/8\}$. Ces codes ont les caractéristiques suivantes: les bons codes $R=k/n$ sont peu nombreux et mal connus, les techniques de codage et décodage des codes de taux $R=1/n$ sont difficiles à appliquer à des codes de taux $R=k/n$ et la complexité du décodeur augmente rapidement lorsque le taux $R=k/n$ augmente. Avec un codeur de taux $1/n$, à 1 bit d'entrée on associe n symboles. La longueur du registre à décalage ou encore le nombre total d'éléments de mémoire du codeur est appelé la longueur de contrainte L (voir fig. 4.8).

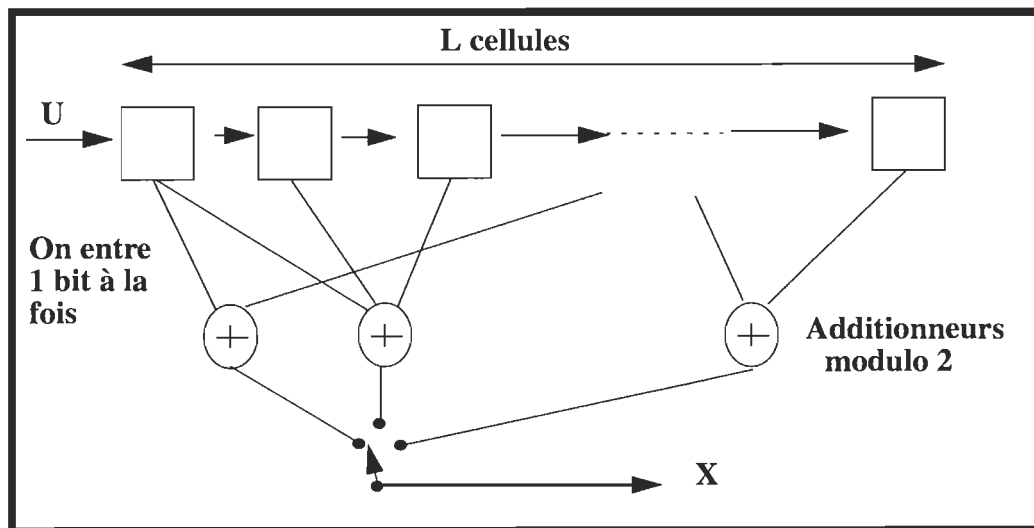


Figure 4.8 Le codeur convolutionnel

Les connexions entre cellules caractérisent le code obtenu. Une séquence de L bits d'information est suivie d'une queue de $(K-1)$ zéros.

$$\text{Le taux de codage } R = \frac{L}{L+K-1} \approx \frac{1}{N} \text{ bits/symboles si } L > K. \quad (5.10)$$

Vecteurs de connexions

$$G_j = (g_{1j}, g_{2j}, g_{3j}) \quad j = 1, 2, 3, \dots, \quad (5.11)$$

$g_{ij} = 1$ si le i -ième additionneur est considéré à la j -ième cellule, sinon $g_{ij} = 0$.

On a k vecteurs de dimension n , les vecteurs de connexions sont exprimés en octal, par exemple, $G_1 = 111$, $G_2 = 101$, $G_3 = 011$, $G_4 = 011$. Chaque symbole de sortie dépend du bit d'entrée présent et des $(K-1)$ précédents. Donc la séquence de sortie est une combinaison linéaire des entrées présentes et passées (fig.4.8). La mémoire du code est égale à $K-1$.

Décodeur de Viterbi

Le décodage est un ensemble de procédures de recherche d'un chemin dans un graphe (arbre ou treillis). Les recherches le long d'un ou plusieurs chemins se regroupent en trois catégories: selon la largeur (*Breadth-first*); selon la métrique (*Metric-first*, ex. décodage séquentiel, algorithme de Fano) et selon la profondeur (*Depth-first*, ex. décodage séquentiel, algorithme à pile). Nous allons décrire brièvement la procédure selon la largeur car elle est implémentée dans l'algorithme de Viterbi que nous utilisons dans notre système de communication vidéo. Elle est caractérisée par:

- l'exploration du treillis le long de plusieurs chemins, tous de même longueur
- l'élimination des chemins selon un critère de rejet (pas de recherche en arrière).

Puisque la recherche dans le treillis (décodage de Viterbi) consiste à examiner systématiquement tous les chemins distincts (selon l'algorithme de programmation dynami-

(décodage séquentiel) consiste à explorer seulement le chemin qui semble le plus vraisemblable (fraction de l'arbre); de ce fait elle est sous-optimale.

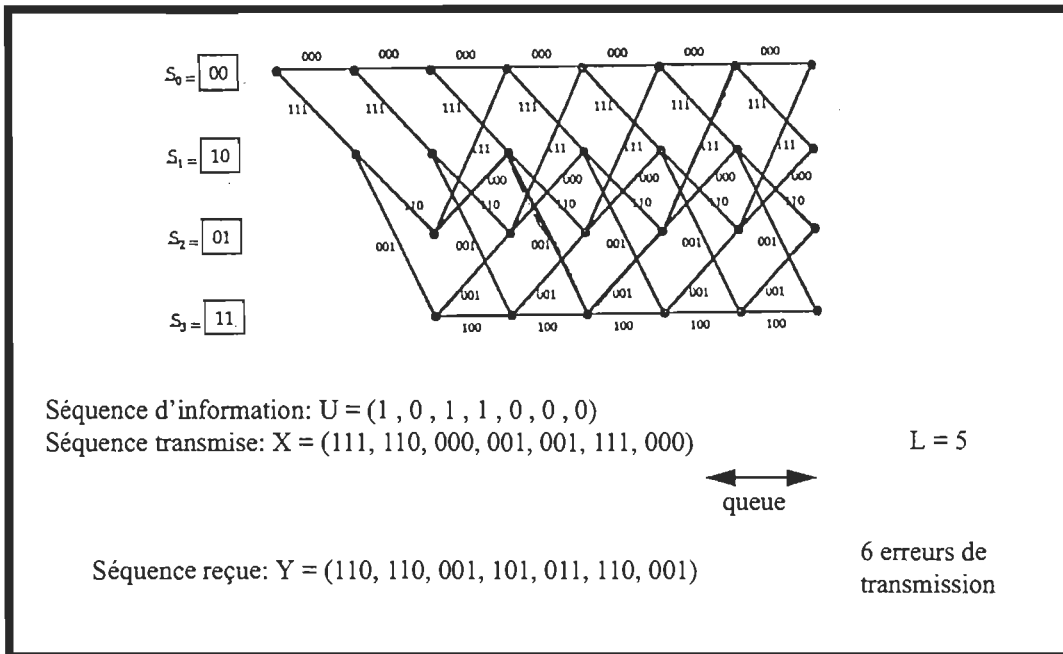


Figure 4.9 Décodage de Viterbi pour $K=3$ et $R=1/3$

La figure 4.9 illustre graphiquement le décodage de Viterbi pour une longueur de contrainte $K=3$ et un taux de codage $R=1/3$. Une description détaillée de l'algorithme de Viterbi est faite par G. C. Clark (1981) et S. Lin (1983).

Les caractéristiques du décodeur de Viterbi sont les suivantes:

- pour un code de taux $1/N$, on a 2^{K-1} calculs par bit décodé;
- la probabilité d'erreur est $P(\epsilon) < Le^{-K(R_o/R)}$ (bornes d'erreurs sur l'ensemble des codes);
- en pratique, en raison de la complexité, le décodeur usuel a une longueur de contrainte $K \leq 9$;

contrainte $K \leq 9$;

- les opérations sont identiques d'un niveau du treillis à un autre;
- une parallélisation des opérations est possible;
- le décodage est rapide (plusieurs mégabits par seconde);
- le gain de codage est important, > 5 dB
- il peut être utilisé aussi bien dans un canal à quantification dure (BSC) que dans un canal à quantification pondérée. Celui-ci apporte un gain supplémentaire de 2 dB. Le décodage de Viterbi est largement utilisé dans les communications terrestres et par satellites.

Le décodeur de Viterbi implanté dans SPW utilise une métrique cumulative pour chaque branches. Une métrique de valeur élevée indique une probabilité élevée d'avoir le bon chemin.

L'utilisation du décodeur de Viterbi

À chaque noeud, on procède comme suit:

- Addition des nouvelles métriques de branche (distances) et calcul de la métrique totale. Comparaison des métriques totales des chemins convergents et sélection du chemin de métrique maximale.
- Mise à jour de l'historique du chemin.

Pour des raisons pratiques, le quantificateur qui convertit les signaux reçus en signaux numériques traitables par le décodeur a normalement des niveaux de quantifications (ou seuils de décision) uniformément espacés. L'intervalle optimum varie en fonction du rapport signal-sur-bruit. Toutefois, la plage de variation du rapport signal-sur-bruit dans notre cas est telle que l'espacement entre les seuils de décision est constant.

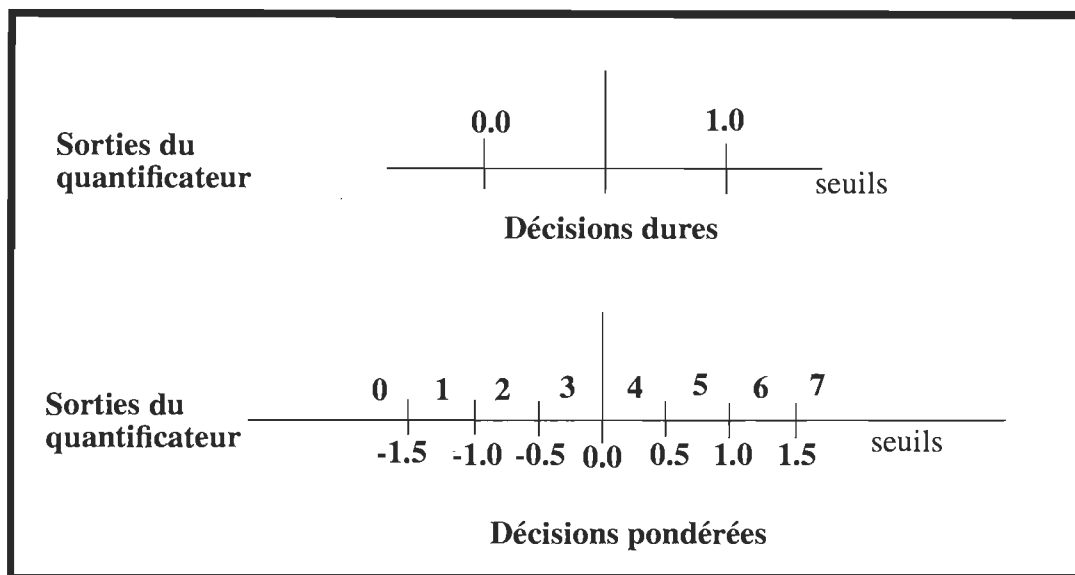


Figure 4.10 Seuils de décisions des décodeurs à décisions pondérées et dures

La figure 4.10 montre les seuils de décisions du décodeur à décisions pondérées par rapport au décodeur à décisions dures. Dans le tableau 4.2, Q est le nombre de niveau de quantification utilisé pour le décodeur à décisions pondérées. $M(A/B)$ est la métrique de transition à utiliser quand le niveau de quantification A est reçu alors que B a été transmis (B est “0” ou “1”).

4.3.2 Codes convolutionnels perforés

Le concept de perforation est utilisé afin d’obtenir des codes convolutionnels avec plusieurs niveaux de protection en utilisant le même décodeur. L’avantage principal des codes convolutionnels perforés est qu’ils peuvent être décodés par l’algorithme de Viterbi et que seulement un décodeur est requis, même si le taux de codage (donc le niveau de protection) change plusieurs fois à l’intérieur d’une séquence d’images codée et transmise.

Tableau 4.2: Métriques du décodeur de Viterbi à décisions pondérées

Formats			Exemple		
Q	M(1/0)	M(1/1)	8	7.0	0.0
Seuil_1	M(2/0)	M(2/1)	-1.5	6.0	1.0
Seuil_2	M(3/0)	M(3/1)	-1.0	5.0	2.0
			-0.5	4.0	3.0
·	·	·	0.0	3.0	4.0
·	·	·	0.5	2.0	5.0
			1.0	1.0	6.0
Seuil_Q	M(Q/0)	M(Q/1)	1.5	0.0	7.0

Une famille de codes convolutionnels perforés est construite à partir d'un code origine de taux $R = 1/n$, et de mémoire M ayant une matrice génératrice G . La perforation permet d'obtenir des codes de taux élevé $R = P/(P+l)$ avec $l = 2, 4, \dots, (n-1)P$ à partir d'un code origine de taux faible $R = 1/n$ en éliminant de façon périodique des symboles de sortie du codeur de taux faible. Comme autre avantage des codes convolutionnels perforés, on a le maintien de la structure du code origine, le décodage du code $R = P/(P+l)$ est effectué comme celui du code origine $R = 1/n$ et enfin le décodage de codes de taux variable est possible. Proposés à l'origine pour le décodage de Viterbi (Clark, 1979), (Yasuda, 1983, 1987), ces codes ont été adaptés pour le décodage séquentiel (Haccoun, Bégin, 1986, 1987).

Matrice de perforation

$$a(l) = \begin{bmatrix} 1 & 1 & \dots & 0 & \dots & 0 \\ 0 & 1 & \dots & 1 & \dots & 1 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 1 & \dots & 0 & \dots & 1 \end{bmatrix}$$

Dans la matrice de perforation, 1 indique que le symbole est conservé et 0 indique que le symbole a été éliminé. Un même codeur de taux faible $R = 1/n$ peut générer plusieurs codes de taux $R = P/(P+l)$ par changement de la matrice de perforation, d'où le codage à taux variable. Le transmetteur et le récepteur doivent connaître l'ordre de transmission des paquets ainsi que les différents patrons de perforation afin de garder une bonne synchronisation des données.

Exemples:

$$\text{Code origine: } R = \frac{1}{N} = 1/3 \Rightarrow P_1 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{Code perforé: } R = 2/3$$

$$\text{Code origine: } R = \frac{1}{N} = 1/2 \Rightarrow P_2 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \quad \text{Code perforé: } R = 2/3$$

Patrons de perforation pour taux compatibles

Une restriction dans les patrons de perforation permet que tous les symboles du code de taux élevé soient utilisés dans le code de taux faible. Ceci permet de transmettre les données avec une redondance (ou une protection) incrémentale. Les codes résultants sont presque aussi bons que les meilleurs codes convolutionnels généralement connus (Hagenauer, 1988) pour les mêmes taux. De plus, pour l'ensemble des codes d'une famille

donnée, nous avons utilisé le même codeur et le même décodeur de Viterbi; seule la règle de perforation (patron) change d'un taux à l'autre.

Exemple: Code origine $R = 1/N = 1/2$

Codes perforés $R = 4/5; 4/6; 4/7; 4/8$

$P_1 = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$	$P_1 = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & \boxed{1} & 0 & 1 \end{bmatrix}$	$P_1 = \begin{bmatrix} 1 & 1 & 1 & \boxed{1} \\ 1 & 0 & 0 & 1 \end{bmatrix}$	$P_1 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & \boxed{1} & 1 \end{bmatrix}$
$R = 4/5$	$R = 4/6$	$R = 4/7$	$R = 4/8$

Figure 4.11 Exemples de patrons de perforation

Tous les symboles codés d'un code de taux élevé donné sont nécessaires pour les codes de taux plus faible. Cette technique est utile dans les systèmes de communication avec retransmission qui exploite la combinaison de codes.

Critères justifiant le choix de bons codes perforés

Une borne supérieure sur la probabilité d'erreur de séquence de l'algorithme de Viterbi est donnée par:

$$P_E \leq \frac{1}{P} \sum_{d=d_{free}}^{\infty} a_d P_d \quad (4.12)$$

P_d est la probabilité qu'un mauvais chemin à la distance d soit choisi plutôt que le chemin correct. Si on utilise le décodeur de viterbi, les critères d'optimalité sont une distance libre d_{free} élevée, un nombre faible de chemins a_d , $d \geq d_{free}$ et un coefficient de pondération de l'erreur sur l'information sur tous les chemins, c_d , faible. Le spectre de distances a_d et c_d , qui doit donc être le plus petit possible, dépend du code choisi

donc de n , de la mémoire M , de la période P et des matrices génératrices G (Hagenauer, 1988).

Pour la plupart des stratégies de transmission, le taux de transmission des symboles dans le canal est fixe à cause de la modulation et des paramètres des différents filtres utilisés; donc, changer le taux de codage veut dire changer le taux de transmission des bits d'information. Dans ce cas, une mémoire tampon est requise afin de permettre le passage d'un taux de transmission à l'autre sans problèmes.

Un désavantage des codes convolutionnels perforés de taux élevés est que les événements d'erreurs sont parfois très longs dans le treillis. Ceci nécessite un délai de décision plus long. Dans le processus de décodage, l'algorithme de Viterbi utilise le même treillis pour tous les codes, ce qui facilite la réalisation matérielle (VLSI) du décodeur. En effet, le décodeur des codes convolutionnels perforés requiert que seuls la mémoire, les conditions du canal et les informations sur l'état du canal de communication soient des paramètres configurables. Nous avons présenté le codeur de canal. L'étape suivante consiste en la modulation des symboles et leur transmission à travers le canal de communication intra-édifice.

4.4 Hypothèses de simulation et classification des données

Pour plusieurs codeurs de source, l'effet des erreurs de transmission varie de façon significative d'un bit à l'autre (Hagenauer, 1990). Pour faire un meilleur usage de la redondance du codage de canal, une protection inégale des données est requise. Dans plusieurs situations, l'état du canal est très variable dans le temps. Cette variation temporelle peut-être évaluée à partir de la puissance du signal reçu. De plus, à cause de la structure des images encodées MPEG-2, certaines données sont plus importantes que d'autres à l'intérieur d'une séquence d'images. Toutes ces raisons font que le taux de transmission et le taux de codage doivent être adaptés dynamiquement aux conditions

du canal de transmission afin de maintenir une qualité uniforme des images reconstruites. La non-stationnarité du canal et la nature de l'information à transmettre sont les facteurs qui justifient la nécessité du concept d'adaptabilité utilisé pour la modélisation du système de transmission. Nous avons fixé la largeur de bande du canal à 2 MHz. L'étalement du délai (environ 50 ns pour un édifice universitaire) n'est pas considéré comme un problème majeur à cette fréquence.

La figure 4.12 illustre la méthode de classification des séquences d'images encodées MPEG-2. Pour des raisons de simplicité, nous avons considéré deux classes de données pour nos simulations. Toutefois, le même raisonnement s'applique pour N classes. Les données encodées en format MPEG-2 ont une structure qui permet en effet de les regrouper en deux classes ayant chacune une priorité associée à son contenu. La première classe comprend l'entête des séquences vidéo qui est d'environ 500 bits. Cette entête contient: les codes de synchronisation, les tables de quantification, les informations sur le type de profil et le type de codage pyramidal utilisé, les vecteurs de mouvement, les tables de décodage de codes à longueur variable. Toutes ces informations ont une importance primordiale pour le décodeur MPEG-2. Pour cette raison, nous avons donné à cette classe la priorité la plus élevée (protection la plus grande contre les erreurs).

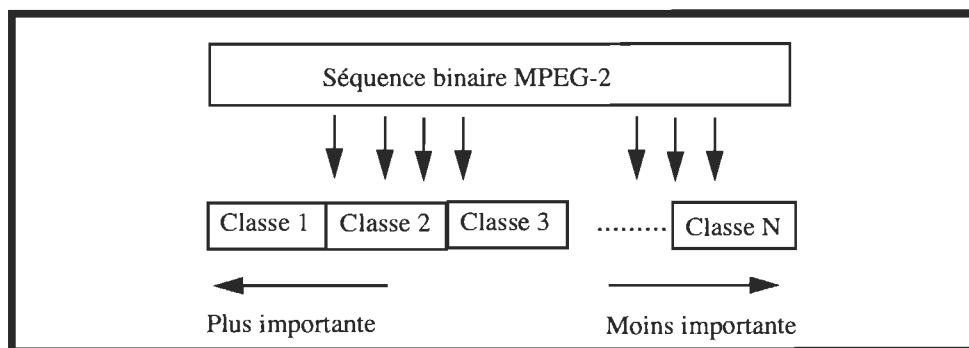


Figure 4.12 Classification des données vidéo suivant un critère d'importance

La deuxième classe comprend le reste des données vidéo encodées MPEG-2, en l'occurrence les coefficients de la TCD des macroblocs de chaque trame. Puisque l'entête se trouve toujours au début de la séquence d'images, cette classe est délimitée par le code de fin d'entête et le code de fin de séquence. Dans ce cas, les trames I, P et B bénéficient du même niveau de protection; seules les données représentant l'entête de la séquence sont prioritaires.

4.4.1 Stratégie de protection adaptative

La figure 4.13 montre le schéma-bloc du système de transmission vidéo sans-fil. Les images animées de la séquence ont été converties en format YUV et sauvegardées dans trois fichiers distincts ayant chacun un composant Y, U ou V du signal vidéo composite. Les trois fichiers représentant la séquence sont envoyés au codeur MPEG-2 au point A sur la figure 4.13. Le codeur MPEG-2 produit un code binaire représentant la séquence à transmettre avec un taux de transmission $D_b = 1$ Mbps.

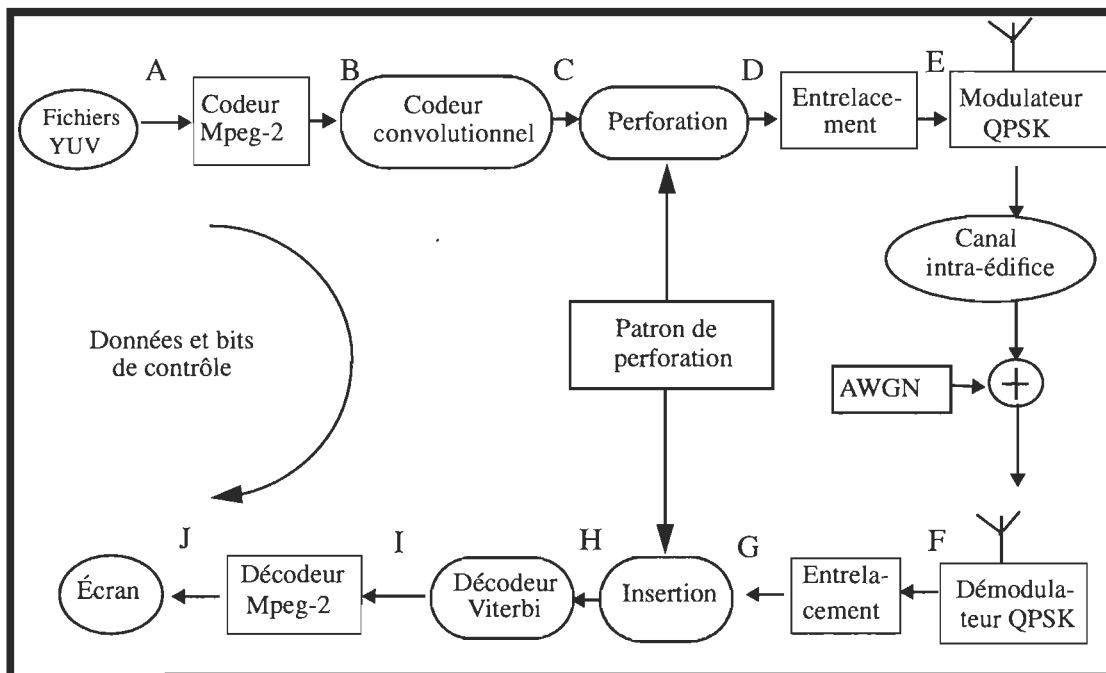


Figure 4.13 Diagramme bloc du système de transmission

À ce niveau les données ne sont pas encore protégées contre les erreurs. Par la suite, ces données sont envoyées au codeur convolusionnel au point B. Le codeur convolusionnel produit une séquence codée avec un taux de transmission $D_c = 2D_b = 2$ Mbps, car le taux de codage du codeur convolusionnel est de $R = 1/2$. Le codage convolusionnel a comme inconvénient de doubler la bande passante du canal de communication lorsqu'on utilise un taux de $1/2$. La perforation permet de diminuer la bande passante du canal tout en accordant aux données une protection moins importante mais suffisante. Au point C, la sortie du codeur convolusionnel est envoyée au perforateur. C'est à ce niveau que le contrôle adaptatif d'erreurs est fait. La première classe, la plus importante, est allouée à l'entête des données encodées MPEG-2. La deuxième classe est allouée aux coefficients de la TCD quantifiés et codés. La délimitation des deux classes de données est basée sur une estimation de la longueur de séquence correspondant à l'entête des données encodées MPEG-2. Cette estimation est faite avant la transmission proprement dite. Nous avons utilisé les taux de perforation $R_1 = 8/14$ (correspond à 100 bits d'information plus 75 bits de protection) pour la première et $R_2 = 8/12$ (correspond à 100 bits d'information plus 50 bits de protection) pour la deuxième classe. Les patrons (Hagenauer, 1988) de perforation correspondants sont:

$$a(1) = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 \end{bmatrix}, a(2) = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}$$

Les mots de codes perforés sont par la suite entrelacés par blocs. L'entrelacement et le désentrelacement sont généralement utilisés pour décorréler ou étaler un groupe de symboles en erreurs dont le nombre dépasse la capacité de correction du décodeur, permettant ainsi d'exploiter au maximum les capacités de correction des codes convolusionnels.

Au point H de la figure 4.14 les symboles codés transmis et démodulés sont envoyés au bloc d'insertion. Ce bloc se charge d'introduire des symboles factices à la place des symboles perforés en utilisant le même patron de perforation que le perforateur. Le décodeur de Viterbi se charge par la suite de décoder les séquences reçues tel qu'illustré à la figure 4.13. Enfin le décodeur MPEG-2 se charge de convertir la séquence binaire décodée en un format qui est affichable sur l'écran utilisé. La synchronisation est essentielle pour l'alignement du patron de perforation de l'encodeur et l'insertion des symboles factices du décodeur. Le calcul du taux d'erreurs lorsqu'on utilise des codes perforés nécessite que l'on recueille les données sur une période suffisamment longue pour que les statistiques obtenues soient significatives.

Tableau 4.3: Répartition des bits d'informations par rapport au bits de protection

Bits d'info.	Bits de protection	Taux de codage	Total de bits
1000	1000	8/16	2000
1143	857	8/14	2000
1333	667	8/12	2000
1000	0	1	1000

Le tableau 4.4 illustre la répartition des bits d'information par rapport aux bits de redondance pour trois stratégies de codage que nous avons implémentées.

4.5 Modulation

Dans un système de communication numérique, la modulation consiste en la conversion d'une séquence binaire en une ou plusieurs ondes électromagnétiques dont un ou plusieurs paramètres sont modifiés en fonction des valeurs binaires de la séquence. Afin de faire un meilleur usage de la bande passante du canal de communication disponible, il est parfois utile de faire une modulation multi-niveaux. Il existe plusieurs types de modulation numériques (BPSK, DPSK, 8PSK, 16PSK, 16QAM, 32QAM,

etc...). Les symboles de modulation sont déterminés soit par la variation de la phase, de l'amplitude ou de la fréquence d'un signal de base. Nous avons utilisé la modulation en quadrature de phase (QPSK) d'un signal de porteuse de 800 MHz. Le signal message $x(t)$ contenant l'information codée à transmettre (images animées codées) est appelé le signal modulant et a un débit de 2 MHz. Le niveau de modulation $m = 1, 2, \dots$ est le nombre de bits par symbole transmis (voir figure 4.15 pour $m = 2$). Ainsi, $M = 2^m = 2^2 = 4$ phases représentent les quatre symboles du signal à deux niveaux pour la modulation QPSK. La constellation des points est obtenue en utilisant le code Gray. Le signal numérique à transmettre est décrit par la formule suivante:

$$\alpha_k = e^{jd_{k2}(\pi/2 - d_{k1}(\pi/4))} e^{-j\pi/4}, \quad k = 1, 2, \dots; \quad (4.13)$$

$d_{k1}, d_{k2} \in \{-1, 1\}$ sont les symboles à transmettre (voir tableau 4.5).

Soit Z_k la distorsion complexe introduite par le canal avec évanouissement et n_k le bruit blanc gaussien complexe. Le signal discret reçu s'exprime comme suit:

$$r_k = Z_k \alpha_k + n_k$$

Le tableau 4.5 permet de faire la correspondance entre les bits et les symboles transmis.

Tableau 4.4: Correspondance entre bits et symboles transmis

$d_{k1}, d_{k2} \in \{-1, 1\}$	Angle	Bits
1,1	0	00
1,-1	90	01
-1,1	180	11
-1,-1	-90	10

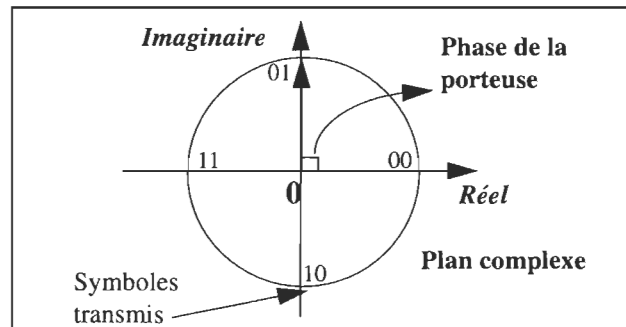


Figure 4.15 Diagramme d'un signal modulé par déplacement en quadrature de phase dans le plan complexe

4.5.1 Calcul du gain de codage

Le système de transmission de base est composé du codeur de canal (codeur convolusionnel), du modulateur QPSK, du démodulateur QPSK ($M=4$) et du décodeur de canal. Soit D le taux d'information fourni par le codeur de source. Le modulateur reçoit des bits codés du codeur de canal à un taux $D_c = D \cdot n/k$ en bits/s et produit des signaux à un taux $D_s = D_c / (\log M)$ en signaux/s. La méthode traditionnelle pour évaluer l'efficacité d'un système de transmission numérique est le rapport de l'énergie par bit E_b ou par symboles E_s sur la densité spectrale de puissance du bruit N_o .

$$\frac{E_b}{N_0} = \left(\frac{n}{k} \right) \frac{E_s}{N_0} \quad (4.14)$$

$\frac{n}{k}$ est le taux de codage convolutionnel.

Le rapport signal sur bruit peut donc se calculer ainsi:

$$SNR = \frac{E_b}{N_o} = \frac{E_s D_c}{N_o D} \quad (4.15)$$

Le gain de codage est la différence entre le SNR requis avec codage et le SNR requis sans codage pour un taux d'erreurs (par exemple 10^{-6}) donné. Il est défini par:

$$G(dB) = \left(\frac{E_b}{N_o} \right)_{noncode} - \left(\frac{E_b}{N_o} \right)_{code} \quad (4.16)$$

Le pouvoir correcteur global du système de codage dépend de la performance du démodulateur et du décodeur.

Nous avons présenté le canal de communication et le codage de canal requis pour le système de communication vidéo sans-fil intra-édifice. Dans le chapitre suivant, il est question d'énoncer les hypothèses de simulation et de les valider par nos résultats.

CHAPITRE 5

Résultats et discussions

Cette partie nous permettra de préciser les spécifications du système de communication vidéo et de valider nos hypothèses par simulation avec SPW.

5.1 Paramètres du canal de communication

Le modèle statistique d'un canal radio sans-fil intra-édifice caractérisé par une propagation par trajets multiples a été décrit au chapitre précédent. L'entrée du bloc représentant le canal est un signal complexe. Les sorties du bloc sont: la réponse du canal, la réponse impulsionnelle du canal, le délai moyen et l'étalement de délai. Quelques paramètres pertinents du canal ont été choisis de la façon suivante: la distance entre les deux antennes (transmetteur et récepteur) a été fixée à 1 m. L'intervalle de temps d'arrivée entre les groupes d'ondes a été fixé à 300 ns. La période de décroissance des groupes d'ondes à 60 ns et l'intervalle de temps moyen entre les ondes à l'intérieur d'un groupe d'ondes à 5 ns. La période de décroissance des ondes est fixée à 20 ns. Le facteur d'atténuation par perte de transmission a été fixé à 2 (pour un édifice universitaire). Elle doit être fixée à 1 pour simuler un canal de transmission sans obstacle. Le gain des antennes à une distance de 1 m est fixé à 1.

5.2 Paramètres de simulation et de codage

Taux de codage convolutionnel: $R_0 = 1/2$

Taux de codage après perforation: $R_1 = 8/14$ et $R_2 = 8/12$

Taux de transmission: $D = 1$ Mbps

Canal de communication: Intra-édifice sans-fil (modèle de Saleh)

Distance entre émetteur et récepteur (Tx-Rx): 1 m

Décodeur de Viterbi à décisions pondérées (quantification à 8 niveaux).

Vecteurs générateurs (octal): $G_{11} = 554$, $G_{12} = 744$

Facteur d'atténuation: $\eta = 2.0$

Codage MPEG I, 8x8 TCD, 1 bit/pixel, 30 trames/s, NTSC, 352x240 pixels

La matrice de macroblocs (MB) du tableau 5.1 présente les caractéristiques de la trame I dans la séquence d'images originale avant la transmission à travers le canal intra-édifice. Comme nous l'avons vu au chapitre 4, la trame I est celle qui contient le plus d'informations pertinentes (100% des blocs sont codés). Au total 330 MB représentent l'image encodée. La liste suivante spécifie les différentes techniques de codage utilisées dans le processus d'encodage MPEG avec leur code dans la matrice de MB.

Premier caractère du code (ex: I)

- **S**: MB abandonnés ("skipped"), MB dont l'erreur de prédiction est nulle. Ils ne sont pas transmis.
- **I**: MB codés Intra
- **F**: MB prédits en avant pour une séquence progressive
- **0**: MB prédits en avant sans compensation de mouvement
- **f**: MB prédits en avant pour une séquence d'images entrelacée
- **B**: MB prédits en arrière pour une séquence d'images progressive
- **b**: MB prédits en arrière pour une séquence d'images entrelacée
- **D**: MB interpolés pour une séquence d'images progressive
- **d**: MB interpolés pour une séquence d'images entrelacée

Deuxième caractère du code (ex: Q)

- **espace**: MB codés sans changement du pas de quantification par rapport au MB précédent
- **Q**: MB codés avec changement du pas de quantification
- **N**: MB non-codés (prédits seulement)

Tableau 5.1: Caractéristiques de la trame I

IQ	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	IQ	IQ	I
I	I	I	I	I	I	I	I	IQ	I	I	I	I	I	I	I	I	I	I	I	IQ	I
I	I	I	IQ	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	IQ	I
IQ	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	IQ	I	I	I
IQ	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	IQ	I	I	I	I
I	IQ	I	I	I	I	I	I	I	I	I	I	IQ	IQ	I	IQ	I	I	I	I	IQ	IQ
I	I	I	I	I	I	I	I	IQ	I	I	IQ	I	I	I	IQ	I	I	I	I	I	IQ
I	IQ	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
I	IQ	I	I	IQ	IQ	I	I	I	I	I	I	I	I	I	I	I	IQ	I	IQ	I	I
IQ	I	I	I	I	I	I	I	I	I	IQ	I	IQ	I	I	I	I	I	I	I	IQ	I
I	I	I	I	I	I	IQ	I	I	I	I	I	I	I	I	I	I	IQ	I	IQ	IQ	IQ

La matrice de MB du tableau 5.2 illustre les caractéristiques de la trame P de la séquence d'images originale. On remarque qu'environ 20 % des MB sont prédits en avant (codes 0 et F).

Tableau 5.2: Caractéristiques de la trame P

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0Q
0	0	0Q	0	0	0	0	0	IQ	IQ	0	0	0	0Q	0Q	0	0	0	0	0	0Q	0
0Q	0	0	0Q	0Q	0	0	0Q	I	0	0	0	0	0	0	0	0	0	0	0Q	F	
0	0	0	0Q	0	0	0	0	I	I	0	0	0	0	0	0	0	0	0	0Q	F	
0Q	0	0	0	0	0	0	0	I	I	0	0	0	0Q	SN	0	0	0	0	0	FQ	F
0Q	0	0	0	0	0	0	0	0	0	0Q	0Q	0	0	0	0Q	SN	0	I	F	FQ	F
0	0Q	0	0	0	0	0	0	0	0	0	0	0	0	0	0Q	0	I	FQ	F	F	F
0Q	0	0	0	0	0	0	0	0	0	0	0	0	F	I	F	FQ	FQ	FQ	F	F	F
FQ	0Q	0	0Q	0	0	0	0	0	I	I	I	FQ	FQ	F	FQ	F	F	F	F	FQ	0Q
0	0	0	0	0	0	0	F	F	FQ	F	FQ	FQ	FQ	FQ	FQ	F	F	F	0Q	0	0Q
0	0Q	0Q	0	0	SN	0	F	FQ	F	FQ	FQ	F	F	F	0	0Q	0	0	0	0Q	I
0	0	0	0	0	0	0	0Q	0	0	0	0	0	0	0	0	0	0	0	SN	0	IQ
F	FQ	0	FQ	0Q	0Q	0	0	0	0	0	0	0	0	0	0	SN	0Q	0	SN	SN	IQ
FN	SN	SN	SN	SN	SN	0Q	F	F	F	FQ	0	0Q	0Q	0Q	0	0Q	0Q	0	I	I	I
FN	SN	SN	SN	SN	SN	0Q	0	F	0	0	I	0	F	F	F	F	FQ	0	FQ	I	I

Les MB qui représentent les objets qui ont bougé entre deux images consécutives sont illustrés par le code F. Le code Q illustre les zones où il y a contraste dans la luminance et/ou la chrominance. La matrice de MB du tableau 5.3 illustre les caractéristiques de la trame B de la séquence d'images originale 'short'. On remarque qu'environ 15.2% des MB sont prédits en avant (code F) avec compensation de mouvement, 7% de MB sont prédits en arrière (code B) et 76.7% sont interpolés (code D).

Tableau 5.3: Caractéristiques de la première trame B

BN	DN	SN	SN	SN	SN	SN	BN	DN	SN	BN	DN	SN	SN	SN	SN	SN	SN	BN	DN	DQ	DN	
DN	BN	SN	D	B	DN	SN	SN	FQ	F	DN	D	SN	DQ	SN	SN	SN	SN	SN	SN	BQ	DN	
DN	SN	SN	SN	SN	BN	SN	DN	IQ	I	DN	SN	SN	SN	SN	SN	D	SN	SN	SN	DQ	DN	
DN	SN	SN	DQ	SN	SN	SN	SN	IQ	I	DN	SN	SN	SN	SN	SN	SN	SN	SN	D	DQ	D	
DN	SN	DN	DQ	SN	SN	SN	SN	B	B	DN	D	D	SN	FN	DN	SN	SN	SN	DQ	DN	DN	
DN	DN	DN	SN	SN	SN	SN	SN	DQ	SN	SN	SN	SN	SN	SN	SN	FN	DN	FQ	DQ	BN	DN	
D	SN	DQ	SN	SN	SN	SN	SN	D	D	D	BN	D	D	D	SN	DQ	F	DQ	DN	DN	DN	
DN	SN	DN	DQ	SN	D	SN	D	SN	SN	SN	D	B	D	D	F	D	FN	SN	DN	DN	FQ	
DN	SN	SN	SN	SN	DQ	SN	D	B	D	D	F	DQ	DQ	D	DQ	DN	DN	FN	F	SN	DN	
DQ	D	SN	SN	D	D	D	F	FQ	F	D	DN	FN	DN	DN	DQ	FQ	F	D	F	D	DQ	
DN	SN	SN	DQ	DQ	FQ	D	F	D	D	F	D	DQ	D	DQ	DN	DQ	SN	SN	SN	SN	DQ	
DN	SN	D	D	SN	SN	D	DN	SN	SN	DQ	SN	SN	SN	SN	SN	SN	SN	BN	DN	FN	DN	D
FN	FQ	F	DN	SN	FQ	F	DN	SN	D	D	SN	SN	SN	SN	SN	FN	D	D	FN	SN	F	
FN	FN	BN	BN	FN	FN	DQ	DN	D	D	BN	DN	DQ	D	D	SN	D	SN	DQ	B	DQ	FQ	
FN	FN	FN	DN	FN	DN	DN	FN	FN	FN	DN	FN	FN	DN	DN	DQ	D	BQ	D	BQ	FQ	F	

Tableau 5.4: Caractéristiques de la deuxième trame B

BN	DN	SN	SN	SN	D	SN	BN	DN	SN	BN	DN	SN	SN	SN	SN	DQ	SN	BN	DN	DQ	DN
DN	BN	SN	D	B	DN	SN	SN	DQ	D	DN	D	SN	DQ	SN	SN	SN	SN	SN	SN	BQ	DN
DN	SN	SN	SN	SN	BN	SN	DQ	BQ	BQ	DN	SN	SN	SN	D	SN	SN	SN	DQ	BN	DQ	DN
DN	SN	SN	DQ	D	SN	SN	SN	BQ	B	DN	SN	D	SN	SN	SN	D	SN	SN	D	DQ	D
DN	SN	SN	DQ	SN	SN	SN	SN	B	B	DN	D	D	SN	FN	DN	SN	SN	SN	DQ	FQ	BN
DN	DQ	SN	D	SN	D	D	SN	D	SN	SN	SN	SN	D	D	SN	FN	DN	DQ	DQ	DN	D
DN	D	DQ	SN	SN	SN	D	SN	D	SN	D	SN	SN	SN	D	D	D	DQ	DQ	DN	D	DN
DN	SN	DN	DQ	SN	SN	SN	D	SN	SN	SN	D	BN	D	F	D	D	DN	FN	DN	DN	DQ
DN	SN	SN	SN	SN	DQ	D	D	B	B	B	D	DQ	D	D	DQ	D	DN	DN	D	DQ	DN
DQ	D	D	D	D	SN	BQ	DQ	D	D	D	DQ	DN	DN	DQ	DQ	DQ	DQ	D	DQ	DN	DQ
DN	SN	SN	DQ	DQ	FQ	D	D	F	D	DQ	D	D	D	B	DN	D	SN	SN	D	SN	BQ
DN	SN	D	D	SN	SN	D	DN	SN	SN	DQ	FN	DN	SN	SN	SN	SN	BN	DN	FN	D	BQ
FN	FQ	F	DN	SN	SN	DQ	SN	SN	SN	D	SN	SN	SN	SN	SN	FN	D	D	FN	SN	FQ
BN	DN	FN	BN	FN	FN	DQ	DN	D	D	BN	DN	DN	DQ	DQ	SN	D	SN	SN	B	B	DQ
FN	FN	FN	FN	FN	FN	DN	DN	DN	DN	DN	DN	DN	DN	DN	DQ	D	BQ	D	BQ	D	D

On constate aussi que 36.1% des MB sont abandonnés. La majorité des MB sont interpolés, 76.7% pour la première trame B et 82.1% pour la deuxième trame B (voir tableau 5.4). Le tableau 5.5 résume les caractéristiques de la séquence d'images originale avant la transmission à travers le canal intra-édifice. Un MB est constitué de 6 blocs de 8x8 pixels (recommandation 4:1:1) répartis de la façon suivante, quatre blocs de luminance (Y), 2 blocs de chrominance (U et V). Tous les pourcentages sont calculés par rapport à un nombre total de 330 MB ou 1980 blocs. Certains MB (voir liste de code à la page 85) ne sont pas représentés dans le tableau 5 mais comptent dans le calcul des pourcentages.

Tableau 5.5: Statistique de la séquence d'images originale

Types de MB	trame I	trame P	trame B	trame B	trame I
# de MB intra codés	330 (100 %)	22 (6.7 %)	4 (1.2 %)	0 (0 %)	330 (100 %)
# de Blocs codés	1980 (100 %)	1061 (53.6 %)	237 (12.0 %)	259 (13.1 %)	1980 (100 %)
# de MB non-codés	0 (0 %)	2 (0.6 %)	90 (27.3 %)	85 (25.8 %)	0 (0 %)
# de MB abandonnés ("skipped")	0 (0 %)	17 (5.2 %)	119 (36.1 %)	108 (32.7 %)	0 (0 %)
# de MB prédits en avant	0 (0 %)	66 (20 %)	50 (15.2 %)	25 (7.6 %)	0 (0 %)
# de MB prédits en arrière	0 (0 %)	0 (0 %)	23 (7 %)	34 (10.3 %)	0 (0 %)
# de MB interpolés	0 (0 %)	0 (0 %)	253 (76.7 %)	271 (82.1 %)	0 (0 %)
taille de la trame (bits)	94824	38304	10313	10313	83374
taux moyen d'activité	160.5	143.7	140.3	139.8	154.0
Pas de quantification moyen	17.6	22.8	25.7	26.3	20.9

La trame I est utilisée comme image de référence pendant le codage des trames P et B. Le nombre de blocs codés varie d'une trame à l'autre car le pas de quantification (256 niveaux) des coefficients de la TCD à l'intérieur d'un macrobloc change dynamiquement. Il est déterminé par l'algorithme qui régularise le débit de transmission (par conséquent la qualité des images transmises aussi). L'allocation des bits est basée sur deux mesures: l'estimation du niveau de la mémoire tampon pour le N-ième macrobloc et le taux moyen d'activité. Au départ, l'image est considérée comme ayant une distribution uniforme de bits. L'activité locale est estimée comme étant la variance minimum entre les quatre blocs de luminance de la trame courante et les quatre blocs de luminance correspondants de l'image originale. La mesure de l'activité locale et du niveau de la mémoire tampon permet de déterminer la valeur finale du pas de quantification qui sera utilisé pour le bloc suivant dans le processus de codage. Le processus de décision du type de codage (compensé par le mouvement ou non, prédiction ou non, interpolation ou non) dépend de la mesure de la qualité (PSNR) relativement à un facteur de coût (en bits).

5.3 Mesures de la qualité des images décodées

Ce sujet préoccupe les spécialistes depuis les débuts de la télévision. Idéalement, on souhaiterait disposer d'une méthode de mesure objective, c'est-à-dire, une façon d'obtenir une valeur numérique caractérisant la qualité de l'image et ceci sans évaluation subjective. En dépit des résultats objectifs pour certaines dégradations homogènes et bien définies (par exemple le bruit), un tel outil reste hors de portée. Une évaluation subjective (observation sur un écran d'ordinateur) de la qualité de l'image est donc nécessaire. L'examen en laboratoire permet certes d'identifier les défauts, mais il doit être complété par une méthode de mesure de qualité rigoureusement codifiée. Trois méthodes principales sont utilisées. Suivant la méthode du double stimulus, deux images A et B sont présentées à l'observateur. L'une est la référence, l'autre est à évaluer. L'observateur qui ignore laquelle est la référence, évalue la qualité de A et B suivant une échelle continue bornée de 1 à 5. Une seconde méthode dite par comparaison pré-

sente deux images. L'image de référence est présentée en premier. Puis l'image à analyser est évaluée en terme de niveau de dégradation par rapport à l'image de référence selon une échelle quantifié de "non perceptible" à "image inutilisable". Enfin, une troisième méthode proposée récemment en télévision, permet à chaque observateur de se constituer une échelle arbitraire de qualité à condition qu'elle reste cohérente d'une image à l'autre. La qualité et le taux de compression sont des paramètres qui varient de façon significative avec les caractéristiques et le contenu des images (scènes). L'unité de mesure généralement utilisée est le bit/pixel, c'est-à-dire le nombre total de bits dans l'espace comprimé incluant toutes les composantes de couleurs, divisé par le nombre d'échantillons d'images.

- **0.25-0.5 bits/pixel:** qualité bonne à moyenne, suffisante pour certaines applications.
- **0.5-0.75 bits/pixel:** qualité bonne à très bonne, suffisante pour la plupart des applications.
- **0.75-2.0 bits/pixel:** impossible à distinguer de l'original, qualité professionnelle, suffisante pour les applications les plus exigeantes.

Comme méthode d'évaluation objective de la qualité des images, nous avons utilisé le rapport signal-sur-bruit (*Peak Signal-to-Noise Ratio*). La variation du rapport signal sur bruit d'une image monochrome avec 256 niveaux de gris pour une erreur intentionnelle sur un bit est décrite par la formule suivante (Netravali, 1989):

$$PSNR = 10 \log \frac{255^2}{EQM^2} \text{ (dB)} \quad (5.1)$$

$$EQM = \frac{1}{L_x L_y} \sum_{y=0}^{L_y} \sum_{x=0}^{L_x} (S_{x,y} - S'_{x,y})^2 \quad (5.2)$$

où EQM est l'erreur quadratique moyenne, L_x est la largeur de l'image et L_y la hauteur de l'image. $S_{x,y}$ et $S'_{x,y}$ sont respectivement les pixels de l'image originale et de l'image reconstruite. Nous nous intéressons aux valeurs relatives plutôt qu'aux valeurs absolues du PSNR.

5.4 Discussions

La performance du système de transmission est d'abord évaluée vs le rapport d'énergie par bit sur densité de puissance du bruit (E_b/N_o). La figure 5.1 montre le taux d'erreurs du système de transmission obtenu par simulation avec SPW. Quatre stratégies de protection sont comparées. La stratégie de protection avec taux de codage $R_2 = 8/16$ (protection maximale), la stratégie de protection avec taux de codage $R_1 = 8/14$, la stratégie de protection avec taux de codage $R_0 = 8/12$ et la stratégie sans codage de canal. Pour le même E_b/N_o , la performance du système de communication pour un taux de codage $R_1 = 8/14$ est inférieure à la performance du système pour un taux de codage convolutionnel $R=1/2$. Cela est normal car deux symboles sur 16 symboles transmis sont perforés, ce qui voudrait donc dire que les données ont été transmises avec une protection moindre.

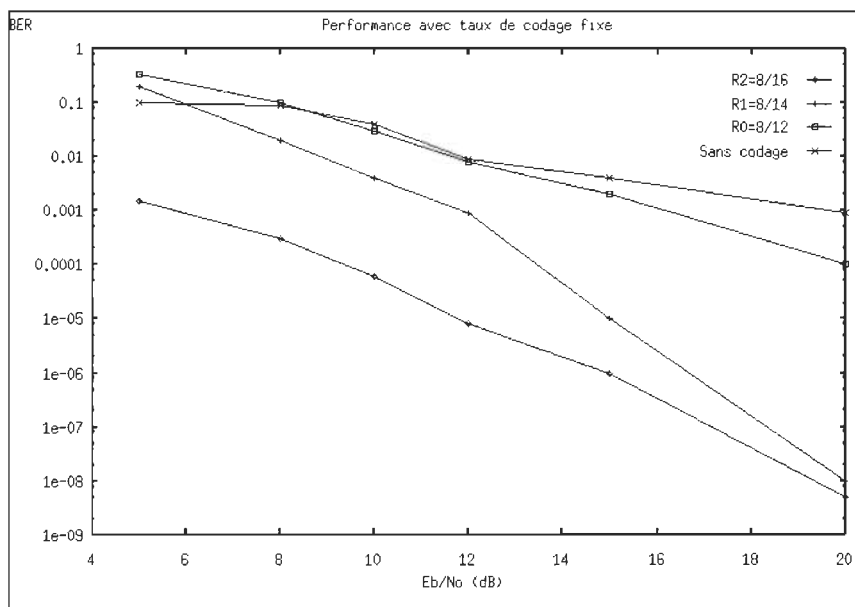


Figure 5.1 Performance de la stratégie de protection non-adaptative (canal intra-édifice - modèle de Saleh)

Le même raisonnement s'applique lorsqu'on utilise un taux $R_0 = 8/12$. Le codage convolutionnel permet de répartir l'énergie de n bits d'informations sur nk symboles avant la transmission ($k > 1$ et dépend du taux de codage). Ce désavantage est compensé par la capacité de correction du code convolutionnel. Les courbes de la figure 5.1 montrent qu'en dessous de 6 dB, les performances du système de transmission sont meilleures (taux d'erreurs plus bas) lorsqu'on utilise la stratégie de protection sans codage de canal. Ce résultat s'explique par le fait qu'il y a trop d'erreur de transmission lorsque le rapport E_b/N_o est inférieur à 6 dB, le décodeur de Viterbi fonctionne alors à sa capacité maximale (saturation). L'avantage présumé de la capacité de correction du code convolutionnel ne suffit pas à compenser la dilution d'énergie causé par le codeur convolutionnel. Par contre, lorsque le rapport E_b/N_o augmente, le rendement du décodeur s'améliore et l'effet de la correction d'erreurs devient important au point de compenser complètement la dilution d'énergie inhérente au codage convolutionnel. Ces résultats indiquent que seuls les taux de codage $R_1 = 8/14$ et $R_2 = 8/16$ seraient considérés dans une stratégie de protection adaptative. En effet, d'après les courbes de la figure 5.1, la performance du système de transmission avec taux de codage $R_0 = 8/12$ est médiocre (taux d'erreurs trop élevés pour ce type d'application). Nous pouvons aussi constater qu'un système de communication vidéo qui ne fait pas usage d'un codage de canal n'est pas envisageable à cause du nombre élevé d'erreurs de transmission (environ 1 symbole en erreur sur 100 transmis).

Le tableau 5.6 donne les différentes valeurs du PSNR pour chaque composante de la séquence d'images originales ('short') que nous avons utilisée pour évaluer la performance du système de transmission et de codage vidéo. Les composantes U et V ne contiennent que l'information de couleurs (chrominances). Pour un écran monochrome, seule la composante Y (luminance) est considérée.

Tableau 5.6: Qualité des images avant la transmission

Ordre d'affichage	Trames	YUV	EQM	PSNR
1	I	Y	88.5	28.7 dB
		U	7.57	39.3 dB
		V	9.73	38.3 dB
4	P	Y	81.4	29 dB
		U	6.34	40 dB
		V	9.53	38.3 dB
2	B	Y	71.7	29.6 dB
		U	6.22	40.2 dB
		V	8.58	38.8 dB
3	B	Y	71.3	29.6 dB
		U	6.18	40.2 dB
		V	7.91	39.1 dB
5	I	Y	106	27.9 dB
		U	7.62	39.3 dB
		V	9.8	38.2 dB

Le tableau 5.7 donne les statistiques sur la séquence d'images reconstruites. On constate une diminution du PSNR qui indique que les images obtenues ont une qualité inférieure à celle des images originales du tableau 5.6.

Tableau 5.7: Qualité des images après la transmission ($E_b/N_0=11$ dB)

Ordre d'affichage	Trames	YUV	EQM	PSNR
1	I	Y	88.5	28.7 dB
		U	7.59	39.3 dB
		V	9.78	38.2 dB
4	P	Y	878	18.7 dB
		U	545	21.6 dB
		V	735	19.5 dB
2	B	Y	3.01e3	13.3 db
		U	1.76e3	15.7 dB
		V	2.79e3	13.7 dB
3	B	Y	703	19.7 dB
		U	90.5	28.6 dB
		V	158	26.2 dB
5	I	Y	574	20.5 dB
		U	93.3	28.4 dB
		V	189	25.4 dB

Nous avons calculé le PSNR des images reconstruites pour plusieurs valeurs de E_b/N_0 . D'après les résultats obtenus par simulation avec SPW et présentés à la figure 5.2, au-dessus de 12 dB la performance globale du système de transmission et de codage vidéo pour la stratégie de protection avec taux de codage $R_1 = 8/14$ est presque identique à celle de la stratégie de protection avec taux de codage $R_2 = 8/16$. Ces résultats montrent l'avantage certain d'utiliser la stratégie de protection avec taux de codage $R_1 = 8/14$ aussi souvent que possible, car la performance résultante est satisfaisante bien que la redondance introduite par le codeur de canal soit moindre.

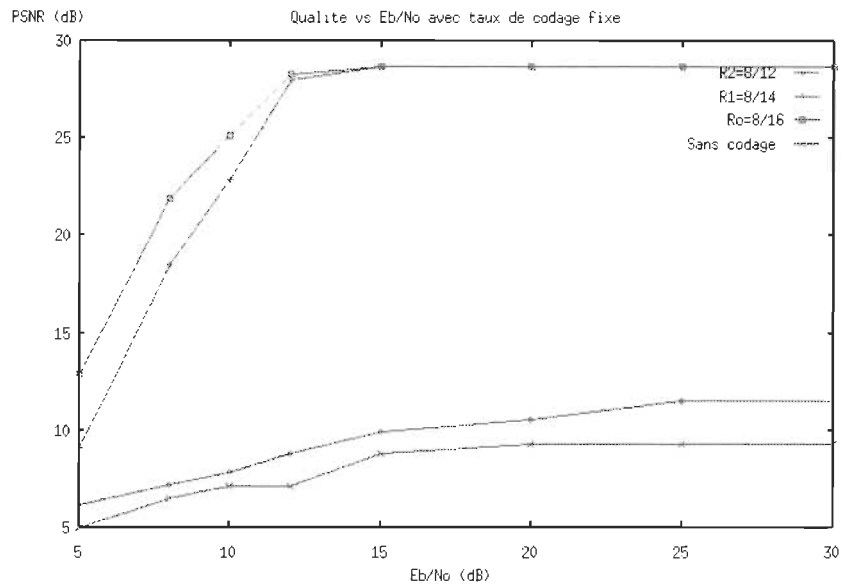


Figure 5.2 PSNR vs Eb/No des trames I 'short', pour la composante Y

Il est à noter que pour obtenir chacun des points des courbes de la figure 5.1 et de la figure 5.2, il a fallu environ dix heures de simulation avec SPW sur une station Sun-Sparc 2. Dans le premier cas, il fallait transmettre assez de bits pour calculer le taux d'erreur (ex: 10^{-5}), dans le deuxième cas, il fallait transmettre assez de bits pour analyser et calculer les statistiques de toutes les trames (IPBBI) de la séquence d'images 'short'.

Dans les figures 5.3 et 5.4, nous comparons les images de la séquence 'short' avant et après la transmission pour la stratégie de protection avec taux de codage $R1 = 8/14$ et avec $Eb/No = 12$ dB. Les blocs sombres dans les trames reconstruites P et B sont dus aux erreurs non corrigées par le décodeur de Viterbi.

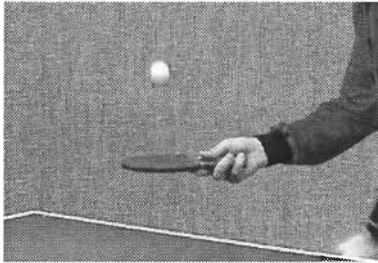


Figure 5.3.a Image originale (trame I)

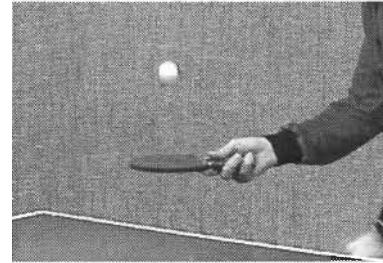


Figure 5.4.a Image décodée (trame I)
PSNR= 29.0 dB, BER = 0.00041, R1 = 8/14

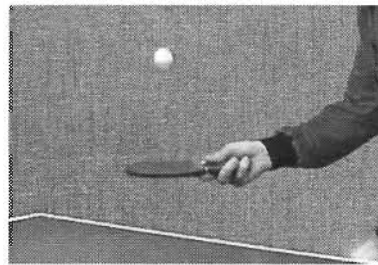


Figure 5.3.b Image originale (trame P)

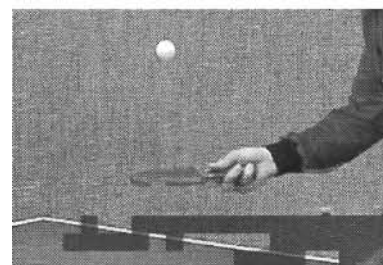


Figure 5.4.b Image décodée (trame P)
PSNR= 28.1 dB, BER = 0.00041, R1 = 8/14

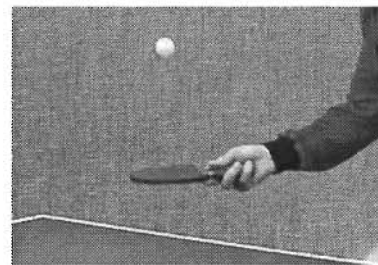


Figure 5.3.c Image originale (trame B)

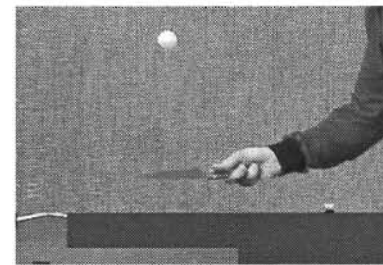


Figure 5.4.c Image décodée (trame B)
PSNR= 28.0 dB, BER = 0.00041, R1 = 8/14

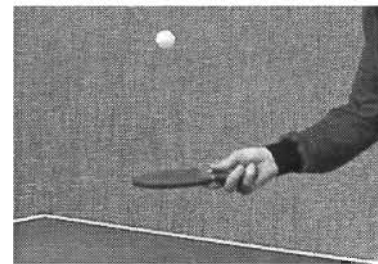


Figure 5.3.d Image originale (trame B)

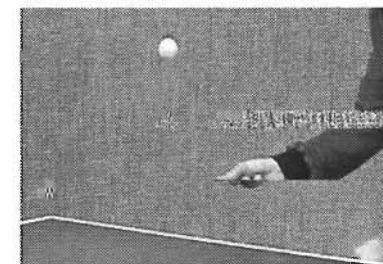


Figure 5.4.d Image décodée (trame B)
PSNR= 27.7 dB, BER = 0.0004, R1 = 8/14

On constate avec surprise que la différence dans la mesure objective (0.2 dB) et la différence dans la mesure subjective ne correspondent pas du tout. Ceci démontre que l'évaluation objective seule ne suffirait pas comme critère de mesure dans l'évaluation de la qualité (performance globale du système) des images reconstruites. On constate aussi que la distorsion dans les trames reconstruites P et B se retrouve concentrée au voisinage des zones où la texture dans l'image change (moins de corrélation entre pixels voisins le long de la table de tennis). De plus, l'effet des erreurs de transmission sur les composantes chromatiques (couleurs) est visible sur papier sous forme de zones d'ombres (nous avons utilisé une imprimante laser monochrome pour l'impression des images). Toutefois, le tableau 5.3 nous donne une mesure objective (PSNR) de la distorsion sur les composantes de couleurs des images transmises. L'observation de l'image 5.4.b sur un écran couleur montre que seules les composantes de couleurs (U et V) ont une distorsion visible sous forme de zones sombres ou grises superposées à la luminance. Dans les images 5.4.c et 5.4.d, l'information de luminance (Y) est aussi perdue au niveau de la petite raquette et ainsi qu'une partie de la main. On remarque pour ces deux dernières images que le décodeur MPEG a compensé la distorsion de la luminance seulement (car Y a un poids plus élevé que U et V) en remplaçant le macrobloc affecté par le macrobloc de la trame précédente encore en mémoire.

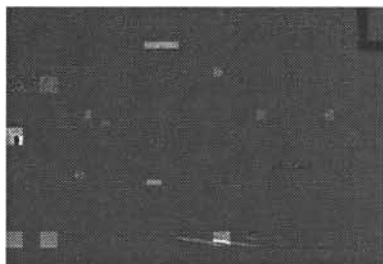


Figure 5.5.a Image décodée (trame I) avec trop d'erreurs
PSNR= 7.2 dB, BER = 0.02, R1 = 8/14

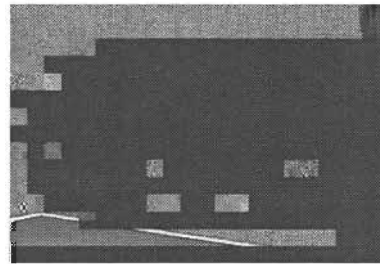


Figure 5.5.b Image décodée (P) avec trop d'erreurs
PSNR= 8.7 dB, BER = 0.02, R1 = 8/14

Les figures 5.5 montrent les cas extrêmes où l'image est complètement détruite, ceci lorsque les erreurs de transmission affectent l'entête des données vidéo contenant les informations vitales pour le décodage du groupe d'images. On voit que la qualité est grandement affectée. Les deux décodeurs (source et canal) sont saturés. Le canal a été sans pitié pour cette expérimentation. Globalement, on constate qu'il y a un "effet de seuil" dans la qualité lorsqu'on transmet avec une E_b/N_0 aux environs de 11 dB. La distorsion est très importante en-dessous de 10 dB et moins importante au-dessus de 12 dB.

Performance du système avec protection adaptative des données

Pour la trame ayant une sensibilité à l'erreur plus élevée (entête de la séquence vidéo), le taux de codage d'origine $R_2 = 8/16$ est appliqué parce que la protection maximale est requise dans ce cas. Les figures 5.6 montrent la performance du système de transmission lors des changements dynamiques de niveaux de protection pendant la transmission d'un taux de $8/16$ à un taux de $8/14$. L'évaluation objective de la qualité (pour $E_b/N_0 = 12$ dB) montre un PSNR proche du PSNR de l'image originale avant la transmission malgré un taux d'erreurs global assez élevé ($BER=0.00037$). On a choisi 12 dB pour faire ressortir l'effet de changer la stratégie de protection. Encore là, on remarque que seules les composantes de couleurs (U et V) ont une distorsion visible sous forme de zones sombres ou grises superposées à la composante de luminance. L'information de luminance a été perdue seulement au niveau de la petite balle blanche dans la figure 5.6.d. La compensation par le décodeur MPEG 2 a été encore faite avec le macrobloc à la même position dans la trame précédente (le mur dans l'image). Ces résultats démontrent la souplesse et la robustesse de la stratégie de protection adaptative, du système de codage et de transmission d'images animées proposé. Globalement, l'effet de la distorsion est moins perceptible lorsque la séquence d'images reconstruite est décodée en temps réel à 30 images/s.

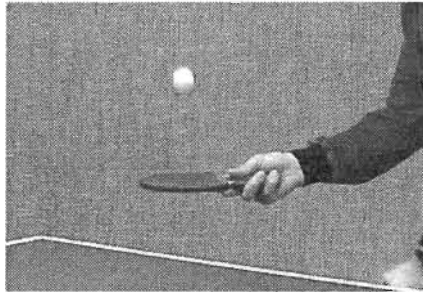


Figure 5.6.a Image décodée (trame I) avec la stratégie de protection adaptative
PSNR= 28.6 dB , BER= 0.00001 , R2 = 8/16, Eb/No = 12 dB

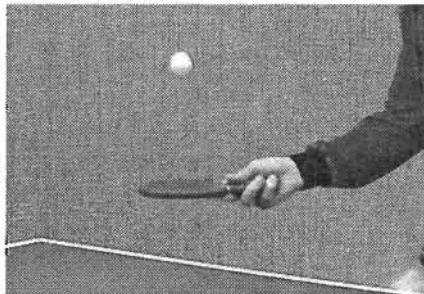


Figure 5.6.b Image décodée (trame P) avec la stratégie de protection adaptative
PSNR= 28.8 dB , BER= 0.00001 , R2 = 8/16, Eb/No = 12 dB

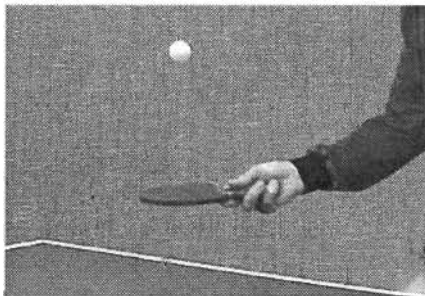


Figure 5.6.c Image décodée (trame B) avec la stratégie de protection adaptative
PSNR= 28.2 dB , BER= 0.0003, R1 = 8/14, Eb/No = 12 dB

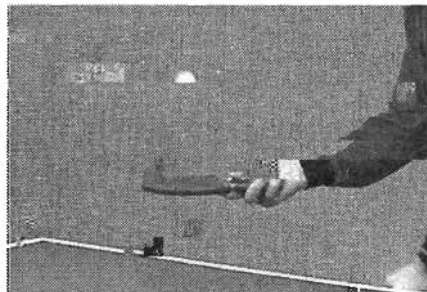


Figure 5.6.d Image décodée (trame B) avec la stratégie de protection adaptative
PSNR= 29 dB , BER = 0.00037 , R1 = 8/14, Eb/No = 12 dB

Le système visuel humain compense la détérioration lorsque les images sont évaluées subjectivement à l'écran. Les erreurs de transmissions sont encore moins visibles dans ce cas là.

Robustesse contre les erreurs de transmission

Si on perd une trame I pendant la transmission, la performance du système se détériore de façon significative et les trames P et B subséquentes ne sont plus décodables. Le décodeur MPEG détecte une perte de bits lorsque l'une des situations suivantes se présente:

- le code de synchronisation n'a pas été trouvé;
- le vecteur de mouvement transmis pointe dans une région en dehors de la zone de recherche locale de l'image;
- la somme des codes zéros à longueur variable et l'ordre courant des coefficients de la TCD n'est pas compris entre 0 et 63.

Lorsqu'une perte d'un paquet a été détectée malgré la protection supplémentaire de notre stratégie de codage adaptative, le décodeur MPEG remplace le macrobloc affecté par le même macrobloc dans la trame P ou I précédente encore en mémoire. Cette technique est dite de dégradation progressive. Les résultats de simulation de la figure 5.4.c et de la figure 5.6.d le démontrent clairement.

5.5 Améliorations possibles du système de communications vidéo

En plus des techniques de compensation que nous avons utilisées pour combattre l'effet de l'atténuation, du bruit, de l'effet Doppler et des distorsions de toutes sortes dans le canal de communication, à savoir l'entrelacement, la protection adaptative et l'égalisation adaptative, d'autres techniques pourraient être combinées aux précédentes afin de diminuer le taux d'erreurs dans le canal de transmission. On parle alors de diversité d'antennes, d'angle et de fréquence. La diversité dans sa forme la plus simple repose sur l'observation suivant laquelle, la même information est transmise continuellement via

plusieurs trajets indépendants. De cette façon, la probabilité qu'un évanouissement profond affecte exactement les mêmes symboles sur deux trajets au même moment reste faible. Dans notre cas, la diversité d'antennes semble plus adéquate. En effet, la plupart des systèmes de communication qui utilisent la diversité d'antennes disposent d'une antenne d'émission et de deux antennes réceptrices (ou plus). Toutefois, l'utilisation de deux antennes d'émission reste toujours possible, à condition de ne pas émettre le signal utile avec la même porteuse. De plus, les deux antennes d'émission ne doivent pas avoir la même polarisation, sinon les deux signaux transmis ne seront pas séparables dans les différents récepteurs. Le modèle de canal avec diversité d'antenne n'étant pas disponible dans les bibliothèques de SPW, et la construction d'un tel modèle SPW n'entrant pas dans les objectifs du présent travail, cet aspect n'a pas été évalué.

CONCLUSION

Ce mémoire décrit l'architecture et le modèle d'un système de communication vidéo permettant de transmettre des images animées et comprimées via un canal de communication sans-fil, à l'intérieur d'un édifice administratif, scolaire ou commercial. Le canal de communication du système est basé sur un modèle statistique simple qui cadre avec les mesures de propagation dans un canal sans-fil obtenues dans un édifice de taille moyenne (Saleh, 1987). Ce modèle statistique semble utilisable pour d'autres édifices. Ce mémoire nous a aussi permis de présenter la conception d'un système de communication vidéo sans-fil de troisième génération. La stratégie de protection adaptative des données vidéo encodées MPEG que nous proposons fait usage d'un concept de codage convolutionnel avec perforation adaptative. Les caractéristiques des données à transmettre ont été présentées en détail, le modèle de canal de transmission a été présenté, une stratégie de protection adaptative contre les erreurs a été élaborée et testée et le lien entre tous les éléments du système a été fait de façon satisfaisante.

Le logiciel SPW nous a permis de simuler et d'analyser le comportement du canal de communication. Nous avons examiné l'effet du bruit, des interférences, de l'atténuation et d'autres perturbations qui affectent le signal utile. Afin de compenser l'effet de ces perturbations sur le signal vidéo transmis, nous avons utilisé un égalisateur adaptatif implanté avec l'algorithme de Viterbi. Son rôle est de compenser en partie ou complètement l'effet de l'interférence entre symboles, du bruit et des évanouissements du signal utile transmis. L'évaluation de la performance globale du système de transmission et de codage vidéo montre qu'on observe un "effet de mur" lorsque le rapport E_b/N_0 varie entre 10 dB et 12 dB. Le décodeur de source (MPEG) semble mieux mitiger l'effet d'une erreur de transmission que le décodeur de canal (Viterbi) dans cet intervalle. En effet, pour ce type d'application, la nature propre des données (ces images sont plus sensibles à l'erreur de transmission lorsqu'elles sont comprimées) poussent à leurs limites les capacités de correction du codage de canal du moins en ce qui con-

cerne la dégradation progressive des images dans la séquence vidéo. Par contre, l'utilisation de la stratégie de protection adaptative que nous proposons permet de mieux adapter le système de codage et de transmission aux besoins de la transmission d'images animées encodées.

L'aspect qui semble très important dans l'évaluation de la performance des systèmes de communication qui sont basés sur la transmission d'une information visuelle (ex. images, vidéo ou télévision) est l'évaluation subjective de la qualité des images reconstruites contrairement au cas typique de transmission de données ou un critère de mesure objectif suffit. Les différentes méthodes d'évaluation de la qualité des images ont leurs avantages et leurs inconvénients. Dans le cas des images fixes, la méthodologie des essais subjectifs est satisfaisante. Le cas des images animées (notre cas) est plus difficile à traiter en particulier avec des codes à réduction de débit. L'un des problèmes les plus complexes est de définir les séquences d'essai. Il est nécessaire de représenter des types de mouvements dans l'image à transmettre qui sont critiques pour les codecs, ce qui est plus compliqué que la définition des configurations spatiales pures propres aux images fixes. En outre, le mode de présentation et les conditions de visibilité pour les observateurs ne sont pas directement transposables des images fixes aux images animées. Les systèmes de vidéoconférence soulèvent un problème différent. Non seulement il n'existe pas d'image de référence pour la comparaison au niveau du récepteur, mais une image dans ce contexte ne constitue cette fois que l'un des éléments de la qualité globale du système. Cet aspect du problème reste encore peu abordé. Les résultats obtenus au chapitre 5 sur l'évaluation de la qualité des images reconstruites ont démontré qu'un critère ultime de mesure n'existe pas. L'évaluation subjective et l'évaluation objective considérées séparément ne seraient pas satisfaisantes. Les deux mesures semblent nécessaires.

De plus, la variété des algorithmes de codage augmente avec les types de dégradations. Si la perception de ces dégradations était profondément comprise, la qualité d'une image produite par un codeur pourrait être vérifiée en faisant des mesures objectives de cer-

tains paramètres du signal vidéo. La transmission d'une information visuelle (images fixes ou animées) à un taux minimum, et avec une qualité acceptable, nécessite la compréhension et l'exploitation de connaissances réparties sur plusieurs disciplines de la science. Idéalement, il faudrait avoir une bonne appréciation des mécanismes d'acquisition et de représentation du SVH, de la théorie sur les signaux multi-dimensionnels, du codage numérique et des dispositifs d'acquisition et d'affichage de l'information.

Références

1. A. A. M. Saleh et R. A. Valenzuela “*A Statistical Model for Indoor Multipath Propagation*”, IEEE Journal on Selected Area in Communication, Vol. SAC-5, No. 2, February 1987.
2. A. Fiorentini et M. T. Zoli, “*Detection of a Target Surimposed to a Step Pattern of Illumination: Effects of a Just-Perceptible Illumination Step*”, Atti. G. Rondi, 1967, pp. 207-217.
3. A. Jain, “*Fundamentals of Digital Image Processing Information and System*,” Sciences Series, Prentice Hall, 1989.
4. A. N. Netravali et B. G. Haskell, “*Digital Pictures Representation and Compression, Application of Communication Theory*,” Plenum Press, 1989.
5. B. Sklar, “*Defining, Designing, and Evaluating Digital Communication Systems*”, IEEE communication Magazine, November 1993.
6. C. C. Huang et R. Khayata, “*Delay Spreads and Channel Dynamics Measurements at ISM Bands*”, in proc. IEEE Int. Conf. Commun., ICC ‘92, Chicago, Ill., June 14-17, 1992, pp. 1222-1226.
7. E. Dubois, “*The Sampling and Reconstruction of Time Varying Imagery with Application in Video System*”, Proceeding IEEE, Vol. 73, April 1985, pp. 562-522.
8. P. Fire, “*A class of Multiple-Error Correcting Binary Codes for Non-independent Errors*”, report RSL-E-2, Sylvania Reconnaissance System Lab., Mountain View, CA.
9. G. Bégin, “*An Adaptative Transmission Scheme for Meteor-Burst Communication*”, 1993 International Symposium on Information Theory, San Antonio, Texas, Jan. 1993, pp. 386.

10. G. Bégin, "***Variable-Rate Coding for Meteor Burst Communications***", IEEE International Conference on Selected Topics in Wireless Communications, Vancouver, B.C., June 1992, pp. 84-87.
11. G. Bégin et P. Nsame, "***Rate-Adaptive Type II ARQ Protocols for Meteor-Burst Communication using Punctured Convolutional Codes***", Canadian Workshop on Information Theory, Manoir du Lac Delage, Québec, May 1995.
12. G. C. Clark, Jr. et J. B. Cain, "***Error-Correction Coding for Digital Communication***", New York: Plenum, 1981.
13. G. S. Brindley, "***Physiology of the Retina and Visual Pathway***", Williams and Wilkins, 1970.
14. H. L. Bertoni, W. Honcharenko, L. R. Marciel, et H. H. Xia, "***UHF Propagation Prediction for Wireless Personal Communication***", Proceedings of IEEE, vol. 82, no. 9, Sept. 1994.
15. H. Hashemi, "***The Indoor Radio Propagation Channel***", Proceedings of the IEEE, vol. 81., no. 7, July 1993.
16. H. Hashemi, "***Impulse Response Modeling of Indoor Radio Propagation Channels***", IEEE J. Selected Areas Commun., Sept. 1993.
17. H. Hashemi, "***Simulation of the Urban Radio Propagation Channel***", IEEE Trans. Vehicular Techn., vol. VT-28, Aug. 1979, pp. 213-334.
18. H. Suzuki, "***A statistical Model fo Urban Radio Propagation***", IEEE Trans. Commun. vol. COM-25, July 1977, pp. 673-680.
19. W. C. Jake, "***Microwave Mobile Communications***", N.Y., J. Wiley, 1974.

20. J. Hagenauer et E. Lutz, "*Foward Error Correction Coding for Fading Compensation in Mobile Satellite Channel*", IEEE Journal on Selected Area in Communication, Vol. SAC-5, No. 2, February 1987.
21. J. Hagenauer, "*Rate Compatible Punctured Convolutional Codes and their Applications*", IEEE trans. Commun., Vol. COM-36, april 1988, pp. 389-400.
22. J.G. Proakis, *Digital Communication*, New York, MacGraw-Hill, 2nd edition, 1989.
23. K. Chalapali, et al., "*The Grand Alliance System for US HDTV*", Proceedings of the IEEE, Vol. 83, No. 2, February 1995.
24. S. Lin et D. J. Costello, *Error control coding: Fundamentals and Application*, Prentice-Hall, 1983.
25. L. C-I. Chuang, "*The Effects of Time Delay Spread on Portable Radio Communication Channels with Digital Modulation*", IEEE Trans. Selected Areas Commun., 1987, pp. 879-889.
26. K. Masahisa et al. "*Adaptation of the MPEG Video-Coding Algorithm to Network Applications*", IEEE Trans. Circ. Syst. Video Techn., Vol. 3, No. 4, August 1993.
27. M. H. Pirenne, *Vision and the Eye*, Science Paperbacks, Chapman and Hall, 1967.
28. A. M. Michelson et A. H. Levesque, *Error-control Techniques for Digital Communication*, Wisley, 1985.

29. MPEG Video Committee Draft, ISO/IEC/JTC1/SC2/WG11 13818-2, March 1994.
30. N. Coppisetti et al. "***Performance Analysis of the Advanced Digital HDTV Video Coding System***", IEEE Transaction on Consumer Electronics, Vol. 39, No. 4, Nov. 1993.
31. H. Nyquist, "***Certain Topics in Telegraph Transmission Theory***", AIEE, 47, 1928, pp. 617-644.
32. P. Pancha et al., "***MPEG Coding for Variable Bit Rate Video Transmission***", IEEE Communication Magazine, May 1994, pp. 54-66.
33. QUALCOMM Application Notes AN1401-29, "***Using Punctured Code Techniques with the Q1401 Viterbi Decoder***".
34. T. S. Rappaport et S. Y. Seidel "***Statistical Channel Impulse Response Models for Factory an Open Plan Radio Communication***", IEEE trans. Commun., Vol. 39, No. 5, May 1991.
35. I. S. Reed et G. Solomon, "***Polynomial Codes over Certain Finite Fields***", Journal Soc. ind. Appl. 8, 1960, pp. 300-304.
36. S. Stein, "***Fading Channel Issues in Systems Engineerings***", IEEE Journal on Selected Area in Communication, Vol. SAC-5, No. 2, February 1987.
37. S. Howard et K. Pathlavan, "***Measurement and Analysis of the Indoor Radio Channel in the Frequency Domain***", IEEE Trans. on Instrumentation and Measurement, IM-39, Oct. 1990, pp. 751-755.

38. S. Y. Seidel et T. S. Rappaport, "**914 MHz Path Loss Prediction Models for Wireless Communication in Multifloored Buildings**," IEEE Trans. Antennas Propagation vol. 40, no 2, Feb. 1992, pp. 207-217.
39. T. S. Rappaport et C. D. McGillem, "**UHF Fading in Factories**," IEEE J. Selected Areas Communications, vol. 7, no. 1, Jan. 1989, pp. 40-48.
40. T. S. Rappaport, S. Y. Seidel, et R. Singh, "**900 MHz Multipath Propagation Measurements for US Digital Cellular Radio Telephone**", IEEE Trans. Vehicular Techn., vol. VT-39, no. 2, May 1990, pp. 132-139.
41. O. K. Victor et al., "**Personal Communication Systems (PCS)**", Proceeding of the IEEE, Vol. 83, No. 9, Sept. 1995.
42. A. J. Viterbi, "**Convolutional Codes and their Performances in Communication System**", IEEE Trans. Commun. Techn., COM-19, (5,2), 1971, pp. 751-772.
43. Y. Koyama et S. Yoshida, "**Still Image Transmission Over a Fading Channel using ARQ Error Control**", ISITA '94, Sydney, Nov. 1994.

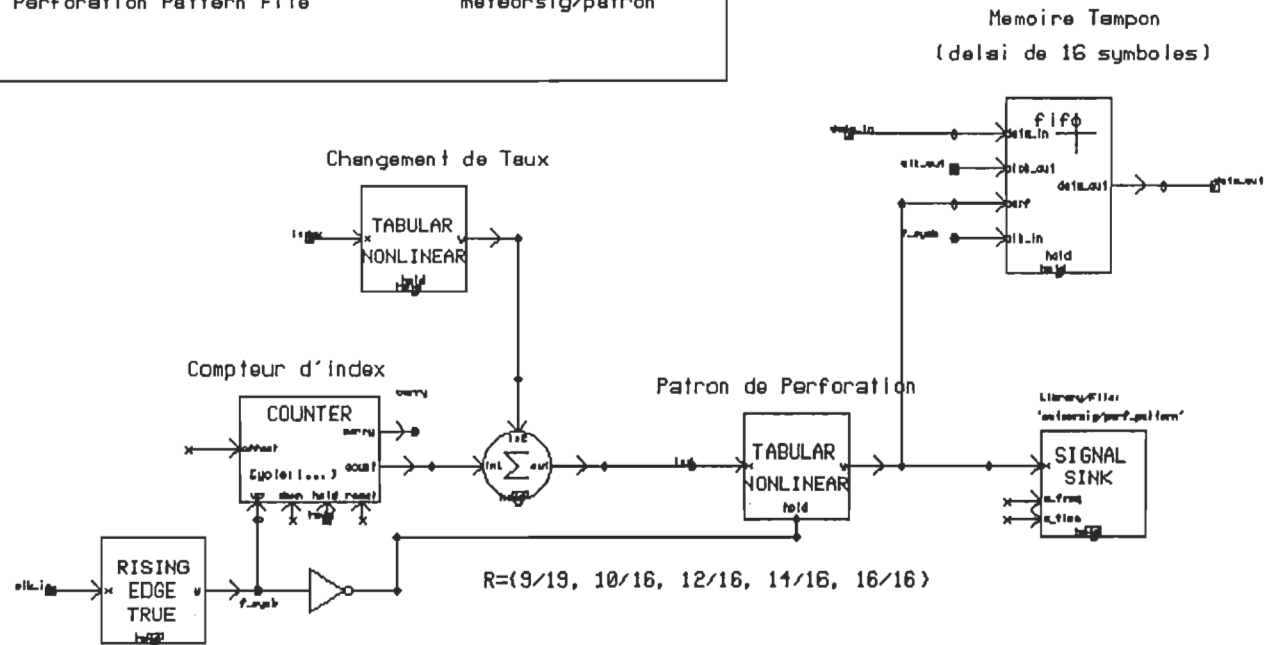
APPENDICE A

Modèles de simulation du système de communication

Puncture State Machine

Systems Parameters

Pattern Period	16
Sampling Frequency	16.0
Rate index file	'meteorsig/rate_index'
Main Clock Frequency	'meteorsig/ratio'
Rate File	'meteorsig/rate_table'
Perforation Pattern File	'meteorsig/patron'



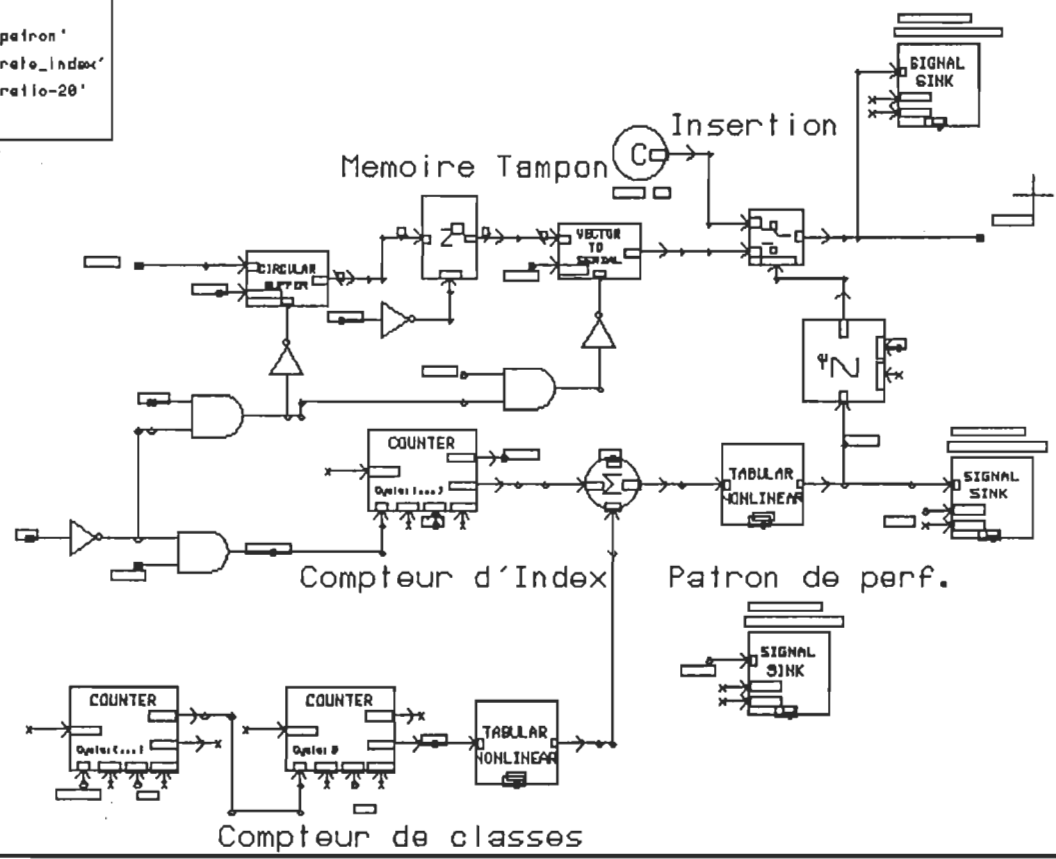
EDIT Affinity Connector

or tool configuration model. [0500035b]
 Model : illegal operator/operand combination in line 1
 Verifying SCHEMATIC model.

Insertion State Machine

Systems Parameters

- Temps d'ack (synbl) 1
- Paquet length 1000
- Punature Periode (P) 16
- Sampling Frequency 80.0e3
- Punature File Path 'metearsig/patron'
- Rate index file 'metearsig/rate_index'
- Clack Ratio File 'metearsig/ratio-20'



EDIT Affinity Connector

or tool configuration model. [0500035b]
 Model : illegal operator/operand combination in line 1
 Verifying SCHEMATIC model.