

Université des Sciences et Technologies de Lille

Ecole Doctorale des Sciences pour l'Ingénieur

Discipline : Microondes et Microtechnologies

## **Thèse**

pour l'obtention du grade de

**Docteur de l'Université des Sciences et Technologies de Lille**

présentée et soutenue par

**Karim HOUNI**

Le 7 Février 2008

## **Modélisation et Etude de la Transmission d'Information par Codes Graphiques**

Composition du Jury:

Président :

Olivier COLOT, Professeur à l'Université des Sciences et Technologies de Lille

Rapporteurs :

Jean-Marc BOUCHER, Professeur à Télécom Bretagne

Pierre PELLAT-FINET, Professeur à l'Université de Bretagne Sud

Examineur :

Didier RAMAGE, Responsable d'Unité R&D – Orange Labs

Directeur de thèse :

Yves DELIGNON, Professeur à Télécom Lille 1

Co-encadrant :

Wadih SAWAYA, Enseignant-Chercheur à Télécom Lille 1



*A mes parents,  
à mes frères.*



## Remerciements

J'adresse d'abord mes remerciements à Yves Delignon, mon directeur de thèse, ainsi qu'à Wadih Sawaya, mon co-encadrant, pour leur suivi et leur supervision.

Je tiens également à remercier Didier Ramage, mon directeur scientifique, pour ses conseils et son soutien, ainsi que ses nombreuses qualités humaines et professionnelles.

Je remercie mes rapporteurs, Jean-Marc Boucher et Pierre Pellat-Finet, pour leur relecture attentive de mon travail qui m'a permis de construire un exposé de thèse clair et concis.

Je remercie Olivier Colot d'avoir présidé mon jury de soutenance.

Je remercie France Télécom R&D, et plus particulièrement le site de Caen, de m'avoir accueilli dans ses murs et dans ses équipes. J'y ai croisé de nombreuses personnes pour lesquelles je souhaite étendre ces remerciements. J'emporte avec moi le souvenir de nombreux échanges et partages, tant personnels que professionnels.

Je remercie Christophe Seguinot, responsable du département Systèmes de Communication, de m'avoir permis d'accomplir ce travail de thèse dans les meilleures conditions possibles. Par là même, j'étends ces remerciements aux autres membres du département, à tout le personnel de Telecom Lille 1 et son directeur Bertrand Bonte pour la qualité de leur accueil.

Je souhaite également remercier mes compagnons de thèse, amis et doctorants du GRIF. Cette aventure aura été le prétexte d'une rencontre, et je souhaite le meilleur pour chacun d'entre eux.

Enfin, je remercie ma famille pour leur indéfectible soutien. Votre place fut centrale dans le cheminement et l'aboutissement, c'est donc tout naturellement que ce doctorat vous est dédié.



# Table des matières

## Chapitre 1

Introduction générale.....	11
1.1 Les codes graphiques.....	15
1.1.1 Codes-barres ou codes 1D.....	15
1.1.2 Codes 2D.....	17
1.1.3 Lecture des codes graphiques.....	19
1.1.4 Etat de l'art sur les codes graphiques.....	21
1.2 Problématique et démarche.....	24
1.3 Organisation du rapport.....	26

## Chapitre 2

Radiométrie et Formation d'Images.....	29
2.1 Radiométrie.....	30
2.1.1 Grandeurs fondamentales .....	31
2.1.1.1 Le flux .....	31
2.1.1.2 L'intensité .....	32
2.1.1.3 La luminance énergétique .....	34
2.1.1.4 l'Eclairement énergétique .....	36
2.1.1.5 L'Emittance énergétique .....	37
2.1.1.6 Récapitulatif.....	37
2.1.2 Les sources de rayonnement.....	38
2.1.2.1 Principes fondamentaux - terminologie.....	38
2.1.2.1.1 Exemples de source.....	40
2.1.2.1.2 Sources singulières et leurs propriétés.....	40
2.1.2.1.2.1 Source cohérente et source incohérente.....	40
2.1.2.1.2.2 Source ponctuelle isotrope.....	41
2.1.2.1.2.3 Source étendue homogène.....	42
2.2 Formation d'images.....	43
2.2.1 Optique géométrique.....	43
2.2.2 Géométrie dans l'espace.....	47
2.2.2.1 Repères et transformations.....	47
2.2.2.2 Profondeur de champ et flou.....	50
2.2.3 Optique ondulatoire.....	51
2.3 Radiométrie d'un système optique.....	55
2.4 Acquisition d'image.....	61
2.5 Conclusion.....	63

## Chapitre 3

Modélisation et analyse des transmissions par code-barres.....	65
3.1 Modélisation du canal de transmission.....	67
3.1.1 Formation d'image.....	67
3.1.2 Analyse radiométrique.....	69
3.1.3 Flou.....	69
3.1.4 Intégration.....	70

3.1.5 Equation du canal.....	70
3.1.5.1 Construction du canal par symbole.....	70
3.2 Mesure des performances du système.....	73
3.2.1 Théorie de l'information – Information Mutuelle Moyenne (IMM).....	73
3.2.2 Calcul de l'IMM.....	75
3.2.2.1 Entropie conditionnelle.....	75
3.2.2.2 Entropie .....	75
3.2.2.3 L'IMM comme mesure théorique de performances.....	80
3.3 Simulations et exploitation des résultats.....	82
3.3.1 Configuration du système.....	82
3.3.2 Profondeur de Champ Théorique.....	83
3.3.2.1 Caméra Longue Portée.....	84
3.3.2.2 Caméra compacte Longue Portée.....	85
3.3.2.3 Caméra Courte Portée.....	86
3.3.2.4 Caméras et seuils d'IMM, tableau récapitulatif.....	87
3.3.3 Effets d'un autofocus.....	89
3.3.4 Résolution spatiale théorique du système.....	96
3.3.4.1 IMM et RSB.....	99
3.3.4.2 IMM et RST.....	100
3.3.4.3 Diagrammes de résolution, codes-barres industriels.....	103
3.4 Conclusion.....	105
Chapitre 4	
Modélisation et analyse des transmissions par code 2D.....	109
4.1 Modélisation du canal de transmission.....	112
4.1.1 Formation d'image et acquisition.....	112
4.1.1 Equation du canal : construction.....	113
4.1.2 Equation du canal : écriture matricielle, caractéristiques statistiques.....	115
4.2 Information Mutuelle Moyenne.....	119
4.2.1 Entropie conditionnelle .....	120
4.2.2 Entropie .....	120
4.3 Calcul de l'IMM.....	121
4.3.1 Distribution du champ et modèle d'Ising.....	122
4.3.2 Estimation de la fonction de partition.....	122
4.3.3 Energie libre variationnelle et graphes.....	123
4.3.3.1 Graphes factoriels et Propagation de Croyance.....	123
4.3.3.2 Graphe de régions et Propagation de Croyance Généralisée.....	126
4.3.4 Algorithme d'estimation de l'IMM.....	133
4.4 Simulations et exploitation de résultats.....	134
4.4.1 Configuration du système.....	135
4.4.2 Comportement du canal.....	136
4.4.3 Evaluation de l'IMM du système LP.....	140
4.5 Conclusion.....	142
Conclusion Générale.....	145
Références bibliographiques.....	151





# **Chapitre 1**

## **Introduction générale**

Je me souviens de ma première fois au cinéma. Tout heureux de pouvoir assister au spectacle sur grand écran, je m'installais confortablement dans mon siège, tenant fermement entre mes doigts le ticket qui matérialisait l'unicité de l'instant et la magie des 90 minutes à venir. J'avoue avoir gardé cette habitude de conserver mon ticket de cinéma comme souvenir, mais cela va devenir de plus en plus difficile. Ce laissez-passer, écorné par les guichetiers des salles obscures, tend à progressivement céder sa place à des badges d'accès, voire à nos téléphones mobiles, nouveaux sésames du divertissement.

Et ce qui peut sembler relever de l'anecdote révèle en fait une tendance actuelle et massive du tissu économique et social de nos sociétés : il s'agit de la dématérialisation. La démocratisation des équipements numériques et des moyens de télécommunications a établi une trame sur laquelle se créent et s'appliquent de nouvelles méthodes de gestion des produits, de rationalisation des coûts et d'amélioration des échanges. Factures, billets ou identifiants, toutes sortes de documents sont numérisés, traités et échangés, simplifiant par l'informatique la gestion d'un immatériel qui se veut équivalent au réel.

Le cas du cinéma est un parfait exemple pour l'application de la dématérialisation. En effet, le billet acheté est un droit de passage symbolisé par un ticket dont la consommation est actée en en découpant une partie. Une autre réponse à ce cahier des charges fonctionnelles serait l'utilisation d'un badge radiofréquence à valider à un portique jouant le rôle de guichetier automatique, et dont l'achat donnerait le droit à un nombre de passages limité. On a ainsi changé de support physique de sorte que la place achetée est désormais une donnée numérique qui peut-être traitée automatiquement et surtout mise en relation avec une base de donnée, ou une infrastructure informatique, permettant l'introduction d'une pléiade de services à valeur ajoutée. L'apport de la dématérialisation est donc non seulement une simplification de l'usage, souvent par l'automatisation, mais aussi une ouverture vers des offres plus riches, plus cohérentes, plus communicantes.

Ce principe de droit pour un service ou pour un bien est extrêmement répandu. Du transport à la gestion d'accès en passant par le paiement, nous sommes de plus en plus amenés à nous identifier auprès de bornes et automates avant la délivrance de ce que nous escomptions. Dès lors, conjointement à la nécessité de concevoir des interfaces homme-machine (IHM) toujours plus évoluées et accessibles, se pose la question des communications

machine-machine (CMM) via un medium interposé. Comme dans le cas du ticket de cinéma, la dématérialisation n'implique pas forcément la disparition de tout support de l'information mais plutôt la modification du contenant et l'adaptation du contenu. Un badge d'accès est ainsi un possible vecteur de l'information à partir duquel s'établit un canal de communication avec le portique. Pour peu que le ticket numérique du porteur soit lu et compris comme valide, le champ lui sera laissé libre.

Pour les utilisateurs que nous sommes, la nouvelle dimension apportée par la dématérialisation de nos échanges transforme notre relation avec le monde. Si les principes de la consommation sont toujours les mêmes, c'est à dire basés sur les notions d'offre et de demande, tout tend à s'agencer et s'articuler autour de l'individu. Au quotidien, nous sommes poussés à être plus impliqués dans nos actes d'achat et de consommation, en traitant avec des machines qui estiment, pèsent, valident, et moins avec l'humain. Cet égotisme technologique nous donne l'impression, véridique ou pas, d'un monde à portée de main transcendée précisément par la dématérialisation, mais pas seulement. Le support de l'information numérique tient ici tout son rôle, ce afin que la CMM soit le lien évident d'une offre de masse vers une demande souhaitée l'être tout autant. Ce contenant idéal, cette clef polymorphe numérique est de fait toute désignée : le téléphone mobile. Ses capacités de stockage, de télécommunications, de communications locales et surtout sa présence dans toutes les poches le promeuvent d'office, tant auprès des fournisseurs de contenu que pour les usagers qui le voient comme un couteau suisse à la portée de chaque bourse. Au centre de toutes les attentions, il s'institutionnalise préfixe d'un grand pan des activités économiques actuelles pour lesquelles on invente le m-Business et ses implications, telles le m-Commerce ou le m-Ticketing. L'intérêt des investisseurs va ici de paire avec celui des industriels et surtout des scientifiques car avoir une unité de calcul et de communication dans plus de trois milliards de mains ouvre des possibilités dont l'appréhension est en soi un sujet de recherche.

Reprenons notre fil rouge du ticket de cinéma. Par le m-Business, l'acte d'achat ne se déroule plus exclusivement au guichet mais devient multimodal : via internet, un serveur vocal ou par SMS, le premier pas du quidam se veut le plus aisé possible. La finalité, qui nous intéresse ici, est d'avoir le "laissez-passer" dans le terminal. Si la première phase pouvait revêtir plusieurs formes il en est de même pour l'étape de consommation et le choix du vecteur d'information

est ici très important. En considérant que la validation est peu ou prou effectuée par un automate, les technologies permettant une CMM ne sont pas toutes sur un pied d'égalité. Celles disponibles dans les terminaux sont principalement l'infrarouge, le bluetooth et la communication radiofréquence locale (ou NFC). Si les débits et les portées varient de l'une à l'autre, il est un critère encore plus rédhibitoire : la disponibilité. Comment proposer un service innovant au plus grand nombre alors que tous ne sont pas dotés des dernières solutions de CMM, souvent synonymes des meilleures performances ? C'est cette question simple qui a amené la communauté scientifico-industrielle à effectuer le constat aussi génial que trivial : tous les terminaux ont un écran! L'idée n'est pas ici d'afficher sur l'écran du mobile un texte en clair correspondant au ticket et qui serait lu par un système de reconnaissance de caractères. Il s'agirait plutôt d'un cas intermédiaire où le laissez-passer numérique serait formaté graphiquement pour être affiché sur le mobile et lu (rapidement) par l'automate. La technologie répondant au mieux à ses critères est le code graphique, plus connu sous sa première version grand public qu'est le code-barres.

Les codes graphiques sont de véritables fichiers numériques qui transposent toute séquence binaire en points de couleurs, offrant ainsi un compromis entre complexité de lecture, occupation spatiale et capacité de stockage. De plus, lorsque l'on ne peut avoir son mobile sur soi pour présenter le pictogramme numérique sur son écran, on peut toujours l'imprimer chez soi ce qui est au fond l'unique méthode de dématérialisation sur papier en vue d'une CMM. Or l'édition de billets directement à domicile s'impose progressivement dans nos foyers majoritairement reliés à l'internet. Une technologie de plus d'un demi siècle, présente depuis trente ans sur tous les étals de nos supermarchés, est donc le vecteur d'information permissif d'un m-Business démocratisé. Ajoutons à cela qu'une majorité de terminaux mobiles sont dotés d'une caméra numérique les transformant en lecteurs potentiels. La boucle est bouclée.

Pour autant, l'utilisation de pictogrammes comme vecteurs d'information les met en concurrence avec des technologies de télécommunications telles que les NFC. Dans le cadre des CMM, un choix, ou au moins une bonne compréhension du système, impose de regarder le code graphique comme une problématique de télécoms afin de pouvoir évaluer ses performances en termes de débits et de portée à l'instar de ses concurrentes. Ce questionnement est peu présent dans la littérature car le code-barres et ses descendants ont

toujours été liés à leur utilisation historique, à savoir la logistique. De plus, étant intrinsèquement une image à traiter, la plupart des investigations ont été menées sous l'angle de l'imagerie et de son traitement. Ainsi, nos travaux s'inscrivent dans cette réflexion originale de l'utilisation de codes graphiques pour la transmission d'information. En nous basant sur l'optique et la théorie de l'information, nous caractérisons le système en un canal de communication à partir duquel sont évaluées les performances du système en termes de quantité d'information transmissible. La modélisation permet de surcroît l'étude de son comportement par rapport aux différents paramètres tels que le positionnement, la taille du code ou les spécifications du lecteur. Le pictogramme est ainsi considéré comme un mini émetteur, localisé dans l'espace, émettant une séquence binaire plus ou moins altérée lors de la transmission par affichage – lecture. Afin de mieux cerner cette technologie, attardons-nous quelques instants sur ses principes.

## **1.1 Les codes graphiques**

### **1.1.1 Codes-barres ou codes 1D**

L'introduction des codes graphiques remonte à la fin des années 50 avec le brevet de B. Silver et N.J. Woodland [1]. Les auteurs y présentent une solution d'identification automatique utilisant un motif bicolore formé de cercles concentriques ou de barres parallèles. Sa lecture s'effectue en analysant électroniquement les variations de la lumière réfléchie d'un faisceau projeté à sa surface. Il faudra ensuite attendre le milieu des années 70 pour voir la première lecture d'un code à barres dans un supermarché américain. Celui-ci était un code UPC ("universal product code", figure 1.1a/) issu des travaux d'IBM sur la création d'une symbologie, c'est à dire une méthode d'agencement des barres, qui se voulait universelle et suffisamment robuste pour la grande distribution [2]. De nombreuses évolutions ont émergé à partir de ce code et c'est aujourd'hui l'EAN13 (European Article Number) qui est le standard international (figure 1.1b/).

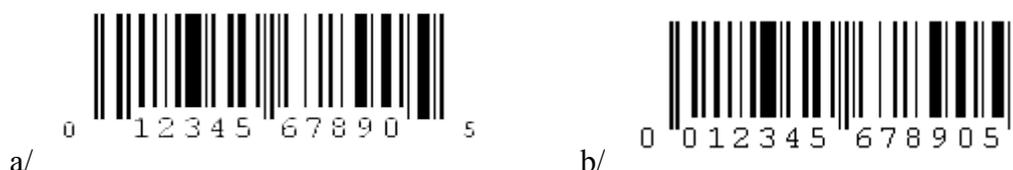


Figure 1.1 : a/ Code UPC - b/ Code EAN13

Bien qu'il y ait une infinité de manière d'agencer des barres pour former une séquence, l'UPC et ses descendants ont la particularité d'appartenir à une des deux grandes familles de codage graphique : les  $\delta$ -codes ([12], figure 1.2a/). Dans cette famille, les barres sont parallèles entre elles et d'égale largeur. Chacune de ces barres est alors appelée *module* et correspond à une information binaire. Dans la seconde famille, le code est composé de barres alternativement blanches et noires de largeur variable. Celle-ci peut prendre deux valeurs de sorte que l'information binaire est portée par la largeur des barres. Les codes 1D construits sur ce schéma sont appelés  $\omega$ -codes (figure 1.2b/).

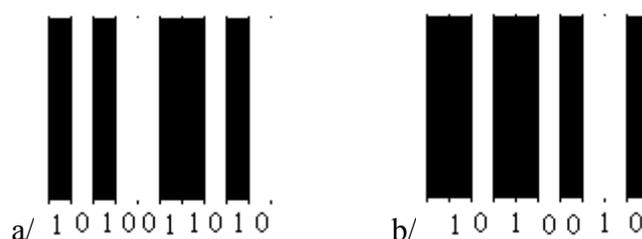


Figure 1.2 : a/  $\delta$ -code - b/  $\omega$ -code

A partir de ces deux groupes, des contraintes supplémentaires sont ajoutées afin de garantir une certaine robustesse des symboles lors d'une lecture perturbée (taches sur le code, effacements). Ces contraintes sont les suivantes :

- pour les  $\delta$ -codes, chaque caractère du message à coder est transposé en  $m$  barres dont  $n$  noires et  $n$  blanches. La notation de cette symbologie est alors  $\delta(m, n)$ .
- pour les  $\omega$ -codes, chaque caractère du message à coder est transposé en  $p$  barres dont  $q$  larges. Cette symbologie est notée  $\omega(p, q)$ .

Ces règles réduisent l'ensemble des combinaisons possibles de barres, rendant les séquences

associées à chaque caractère plus facilement différenciables les unes des autres, même en cas de perturbations telles qu'un étalement d'encre ou la disparition d'une barre. En se référant à la théorie de l'information, ces règles introduisent un codage canal [3] que nous qualifions d'intrinsèque car lié au seul aspect graphique du symbole. Un codage canal extrinsèque affecterait ainsi le message source avant sa transposition graphique.

La majorité des symbologies 1D découlent de cette codification : l'UPC et l'EAN13 sont un  $\delta(7,2)$ , le Code128 utilisé en logistique est un  $\delta(11,3)$ , le Code39 instauré par le département d'état américain à la défense est un  $\omega(9,3)$ . Néanmoins, malgré ces règles de codage évolué des codes à barres, il subsiste que chaque barre correspond à un "0" ou un "1". C'est pourquoi nous nous focaliserons dans cette étude sur ce codage binaire que l'on peut qualifier de fondamental, en laissant de côté tout type de codage canal intrinsèque. De plus, les  $\delta$ -codes étant la famille la plus présente dans l'industrie, nous nous bornerons à cette seule symbologie dans la suite des travaux.

### 1.1.2 Codes 2D

Les codes-barres ne portent pas d'information sur leur hauteur. Le message codé graphiquement est porté le long d'une ligne traversant le pictogramme sur sa longueur d'où l'appellation "code 1D". La dimension verticale participe à un meilleur confort de lecture car plus le symbole est haut, plus la ligne de lecture peut être oblique. Cette redondance constitue une forme de codage canal supplémentaire où l'information est répétée sur la hauteur, à l'instar d'un code à répétition. L'exploitation de la surface totale d'affichage mène naturellement à l'idée de "code 2D" qui prend principalement deux formes en pratique:

- l'empilement de codes 1D étroits, on parle de codes hybrides ou de codes 1D empilés
- une matrice de points bicolores, ce sont les codes matriciels ou codes 2D.

Dans le cadre de ce travail, les codes empilés sont appréhendés comme une suite de codes-barres. Si ce n'est une contrainte de positionnement, l'analyse du système global ne sera pas différente selon que l'on utilise l'un ou l'autre.

Les codes matriciels sont en revanche une nette évolution des identifiants graphiques. Historiquement, l'un des premiers codes 2D matriciels a été le DotCode de Philipps en 1986 ([4], Figure 1.3).

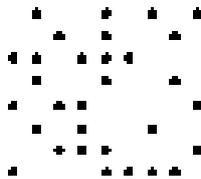


Figure 1.3 : Un DotCode

Il faudra attendre le milieu des années 90 pour voir l'apparition des symbologies 2D majeures, plus que jamais d'actualité avec l'essor du m-Commerce. Ces codes de première ligne sont au nombre de trois : le Datamatrix et le QR Code introduits en 1994, l'Aztec proposé en 1995 ([5,6,7], Figure 1.4).

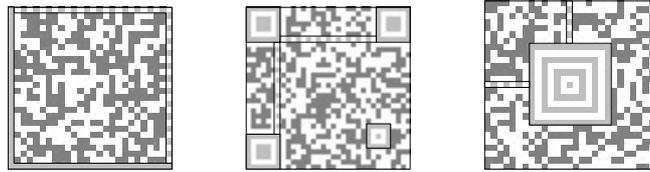


Figure 1.4 : "Une image vaut mieux que mille mots." codé en

a/ DataMatrix, b/ QR Code, c/ Aztec

Le champ d'application de ces codes matriciels ne se limitent pas à la logistique à l'instar de leur aîné. L'utilisation des deux dimensions offrent une capacité de stockage beaucoup plus grande que celle des code-barres ce qui fait du code 2D un véritable fichier numérique portable. Pour information, un DataMatrix peut transporter un peu plus d'un kilo-octets, un Aztec 1.8 Ko et un QR Code 2.8 Ko.

Bien que d'aspects différents, les trois symbologies et les codes matriciels en général possèdent la même structure. Ils sont composés de points bicolores appelés *modules* qui portent une information binaire. De plus, ils sont dotés d'un moyen de localisation qui est le plus souvent un motif prédéfini permettant de déterminer la position et l'orientation du code lors de la lecture. Enfin un ou plusieurs motifs indiquent la taille des modules ainsi que celle du code. Ce dernier élément de structuration du code s'apparente à une "horloge spatiale". Cette architecture des codes 2D est illustrée en figure 1.5.



*Figure 1.5 : Structure d'un code matriciel - en rouge la cible de localisation, en bleu l'horloge spatiale et en vert la zone de donnée*

Tout comme les code-barres, les codes 2D sont sujets à des perturbations qui peuvent altérer l'intégrité du message transporté. Taches, déchirures ou froissements, on adopte contre eux des mécanismes de codage canal qui permettent de détecter et voire corriger les erreurs introduites dans la zone de donnée. La position des modules et leur regroupement en octets est donc fonction de la protection que l'on veut (ou peut) appliquer au message à transmettre. Pour autant, le principe des codes matriciels est l'utilisation de points noirs et blancs codant des "0" et des "1" d'information agencés en matrice. Le codage canal vient se greffer sur cette correspondance séquence binaire – séquence graphique. Par conséquent, lorsque nous parlerons de codes 2D dans la suite de cette étude, il s'agira de matrices de points bicolores dépourvues d'un éventuel codage. De plus le code sera considéré comme localisé et de taille connue de sorte qu'il ne contienne que la zone de donnée.

### 1.1.3 Lecture des codes graphiques

En une ou deux dimensions, la récupération du message porté par un code graphique passe par la retransposition des modules en information binaire. Celle-ci s'opère sur les pictogrammes bicolores par la mesure des variations de contraste selon un rythme défini, d'où l'importance de l'horloge spatiale. Les lecteurs permettant cette acquisition de contraste se basent sur deux technologies : le laser et l'acquisition d'images.

Avec la première, un faisceau laser balaye un code-barres sur sa longueur, le faisceau réfléchi est récupéré par un capteur photosensible qui convertit l'énergie reçue en signal électrique proportionnel. Lors du balayage, le faisceau est plus réfléchi par les parties claires (idéalement blanches) que les sombres de sorte que le signal électrique issu du capteur est un

train de créneaux plus ou moins bruité et filtré. On l'aura compris, ces lecteurs lasers opèrent exclusivement sur les codes linéaires ou les codes empilés. Si la lecture de codes matriciels par ce procédé n'est certes pas impossible, en asservissant le balayage du faisceau à la forme du code, elle est très complexe à mettre en oeuvre et inexistante en pratique. Les lecteurs lasers ont émergé avec l'intérêt croissant pour les codes-barres. Cette solution était alors relativement peu coûteuse et présentait l'intérêt, toujours valide aujourd'hui, de permettre la lecture éloignée d'un code par les caractéristiques même du laser. Le développement des codes 1D, le faible coût des lecteurs et l'intérêt du secteur de la logistique ont établi un cercle vertueux qui a permis aux lecteurs laser de dominer cette fonction. Cependant, l'avènement des codes 2D et la démocratisation des capteurs d'image a modifié cette donne.

Avec l'acquisition d'images, la lumière provenant du code est acquise et numérisée par un capteur constitué d'éléments photosensibles (pixels) basés sur les technologies CCD [8,9] ou CMOS [9,10]. La lumière transmise l'est par réflexion d'une source éclairante (code sur papier) ou par le code lui-même (code sur écran), puis le flux lumineux est focalisé par un bloc optique intégré au lecteur. Le capteur a le plus souvent une structure régulière permettant la discrétisation du champ lumineux issu du bloc. On distingue principalement deux agencements, linéaire et matriciel, définissant deux types de lecteurs :

- les imageurs 1D dotés d'un capteur linéaire, c'est à dire composé d'une ligne de pixels contigus,
- les imageurs 2D dotés d'un capteur rectangulaire ou carré, où les pixels sont disposés côte à côte en matrice.

Les imageurs 1D sont dédiés à la lecture des codes-barres et des codes empilés. Ils émulent le balayage effectué par les lecteurs laser sans toutefois permettre une acquisition distante du code par la dispersion de l'énergie lumineuse. En revanche, leur coût et leur robustesse à l'usage les ont imposés comme de sérieux concurrents. Ils sont ainsi très présents dans la grande distribution. Les imageurs 2D sont quant à eux les plus polyvalents. En effectuant l'acquisition d'une surface complète, ils permettent l'acquisition de tout type de codes graphiques et surtout des codes 2D. Ces lecteurs sont donc dédiés à l'acquisition des codes matriciels.

De par ce dernier point nous laissons donc de côté la lecture laser qui, de plus, est une ancienne technologie pour laquelle de nombreux travaux ont été publiés [11]. Par conséquent,

nous considérerons dans notre étude que les lecteurs utilisés sont des imageurs : lorsque le code sera à barres il s'agira d'un lecteur linéaire, lorsque le code sera matriciel il s'agira d'un lecteur matriciel.

### 1.1.4 Etat de l'art sur les codes graphiques

L'activité de la communauté scientifique sur les codes graphiques s'articule autour de quatre axes principaux : la théorie de l'information, l'analyse des systèmes, le traitement d'images et les usages des codes graphiques. Nous présentons ci-après des travaux sélectionnés pour représenter ces quatre courants et illustrer l'activité sur ce sujet.

Dans le premier, l'objectif est la conception et l'analyse de symbologies dont les codes sont considérés comme des unités de stockage optique. Ainsi Pavlidis *et al.* établissent dans [12] les fondamentaux de la conception des codes 1D. Les caractéristiques des  $\delta$ -codes et de  $\omega$ -codes (déjà présentés plus haut) sont analysées, notamment concernant l'alphabet des mots-codes possibles définis par les contraintes  $\delta(m, n)$  et  $\omega(p, q)$ . Ceci permet aux auteurs de définir l'entropie des codes, grandeur théorique clef mesurant la quantité d'information d'une variable aléatoire ([13]). De plus, les auteurs modélisent l'étalement de l'encre lors de l'impression d'un pictogramme comme une distortion affectant la lecture. Ils déterminent ainsi l'impact de cette perturbation et du bruit. Poursuivant dans cette idée d'appliquer la théorie de l'information aux codes graphiques, Pavlidis *et al.* publient dans [14] l'extension des travaux sur les codes-barres aux codes empilés. Il en découlera une symbologie très utilisée aujourd'hui dans le contrôle d'accès : le PDF417 (figure 1.6). Pour information, le sigle PDF417 veut dire "Portable Data File  $\delta(17,4)$ ".



Figure 1.6 : Exemple de PDF417

Citons également dans cet axe de recherche Van Gils qui a publié dans [4] les caractéristiques du DotCode (Figure 1.7). La démarche de l'auteur pour la conception de ce code matriciel a

été de s'inscrire dans la théorie de l'information pour établir un codage source (compression des données) et un codage canal (protection des données) adaptés à un medium de transmission graphique. Enfin signalons que, bien que cette liste ne soit pas exhaustive, il est apparu dans nos recherches que peu de travaux ont été publiés dans cette thématique. Il est probable toutefois que l'activité des codes graphiques étant fortement liée à l'industrie, nombre de résultats et symbologies ont été brevetés sans que les études et les méthodologies aient été publiées ([15, 16, 17]).

L'analyse des systèmes regroupe les activités portant sur tout ou partie d'une chaîne de lecture de codes. Dans ce domaine, Tsi *et al.* analysent dans [18] la lecture de codes-barres par un capteur linéaire. Les auteurs déduisent des considérations d'optique géométrique et d'électronique l'étendue de fonctionnement du système qui est la distance code – lecteur permettant une récupération du message transporté par le pictogramme. Riyanto *et al.* [19], et plus récemment Vincenzi et Oliva [20] ont proposé une étude de l'éclairage du code afin d'en améliorer le contraste et donc sa lecture. Les premiers déterminent les positions optimales des sources éclairantes ainsi que leur nombre, les seconds établissent la forme de la source pour obtenir l'éclairage produisant le meilleur contraste possible. La lecture de codes graphiques considérée dans sa globalité est un système multiparamétré dont l'efficacité est entre autres fonction de l'éclairage, de l'électronique du lecteur et bien sûr du code.

Le traitement d'image autour des codes graphiques est de loin l'axe de recherche le plus fourni. Nous en avons extrait ci-après une portion représentative des tendances des activités. La plupart des travaux s'orientent vers l'amélioration de la lecture des codes par une meilleure robustesse aux perturbations ([21-24]), une plus grande rapidité de traitement ([25,26]) ou une meilleure localisation du code dans l'espace ([26-28]).

Avec le téléphone mobile, et surtout les photophones, chaque usager a en main un lecteur de code potentiel. Cependant ces terminaux n'ont pas été conçus pour cette fonction, d'où les travaux portant sur des algorithmes adaptés aux capacités limitées des téléphones ([29]) ainsi qu'aux conditions de prise de vue ([30]). Les fonctions communicantes des terminaux permettent de définir des stratégies de décodage décentralisé par l'envoi d'une image vers un serveur central ([31]). Enfin, les symbologies les plus usitées en 1D et 2D ont été conçues alors que les terminaux mobiles n'existaient pas. L'utilisation de codes graphiques dans ce nouveau cadre impose de réfléchir aux codes les mieux adaptés ([32,33]).

Enfin, les activités de recherche autour des usages des codes graphiques est une conséquence logique de l'analyse des systèmes et du traitement du signal au sens large, l'un et l'autre mettant en évidence les limites et les possibilités de la communication graphique. Dans le prolongement de leur héritage historique dans la logistique, Sriram et al. [34] traitent de l'utilisation de codes barres pour la gestion automatique de stocks. Yu et al [35] proposent quant à eux une implémentation d'un système de contrôle de ligne de production de circuits intégrés utilisant les codes 2D, et les codes graphiques sont également mis à profit dans les travaux de Li et al. [36] pour l'automatisation d'une banque du sang.

Mais l'évolution principale dans l'usage de codes graphiques vient avec le téléphone mobile. Gao et al. [37] proposent une inspection du processus de décodage des codes 2D en vue de leur utilisation en m-Commerce. Kato et Tan [33] traitent la question de la pertinence des principales symbologies dans cette application, et Adelman [38] propose une implémentation d'un lecteur de codes pour une solution d'aide à la recherche d'allergènes. Chaisatien et Akahori illustrent de leur côté [39] la création d'un lien entre monde réel et virtuel dans l'enseignement en présentant l'utilisation de codes 2D comme liens vers le web avec son photophone.

Ce pont vers le virtuel est justement très étudié dans la communauté de la réalité augmentée ([40-42]) car avec un photophone, les codes graphiques permettent d'identifier simplement et à bas coût les objets du monde réel. Dans le cas de l'aide à la personne, la capacité à reconnaître et lier les objets à du contenu enrichi avec son téléphone a permis de proposer des solutions intéressantes pour assister les personnes déficientes visuelles dans leurs déplacements [43, 44].

Enfin signalons l'utilisation originale des codes graphiques dans le tatouage d'image [45] : un code 2D transportant une signature numérique est introduit dans le plan fréquentiel de l'image à marquer pour l'identifier de manière unique et non modifiable.

La plus grande représentativité des deux derniers axes de recherche, le traitement d'images et les usages des codes graphiques, n'est pas un artifice de notre présentation. Elle révèle le réel intérêt de la communauté sur la manière de mieux lire et décoder, notamment avec des terminaux à faibles capacités de calcul, et sur les possibilités qu'offrent ces solutions lorsque mises dans toutes les mains avec nos mobiles et photophones. Toutefois, ce

formidable élan est à notre sens la conséquence des travaux portant sur la théorie de l'information et l'analyse des systèmes qui ont permis d'avoir un médium numérique performant. Pour autant, la chaîne de transmission code – lecteur n'était alors pas vue comme un canal de communication numérique, si ce n'est dans une certaine mesure avec le travail de Pavlidis et al. [12,14] qui insistent sur la conception de codes et l'impact de l'étalement d'encre. Nous présentons ainsi dans le prochain paragraphe notre démarche, liée aux communications numériques, pour la modélisation et l'étude de la transmission d'information par codes graphiques.

## **1.2 Problématique et démarche**

Nous considérons la lecture d'un code graphique par un imageur comme la réalisation d'une communication numérique où la transposition de l'information binaire en modules est l'analogie de la modulation. L'émetteur est alors un pictogramme, dessiné ou affiché sur son support, et le récepteur est une caméra numérique. A l'usage, le positionnement du code est le plus souvent libre devant la caméra impliquant a priori une grande variabilité des conditions de transmission. L'orientation télécommunications des investigations met ainsi en exergue l'idée force du questionnement : la quantité d'information transmissible. En effet, dès lors qu'il s'agit d'une technique de télécommunication, les critères d'intérêt tournent autour du débit et de ses variations par rapport aux conditions de transmission. La question centrale est ainsi l'évaluation des performances que l'on peut attendre d'un système à codes graphiques. En d'autres termes, quelle quantité d'information peut-on récupérer à l'issue de l'acquisition du code par l'imageur.

Le problème ainsi posé, la théorie de l'information découle naturellement comme cadre scientifique d'étude. Ce choix est appuyé par le souhait d'avoir un regard objectif sur les performances du système, objectivité apportée par ce cadre formel. Ainsi, nous cherchons à définir la relation entrée-sortie du système, déterminant l'équation du canal de transmission. Les performances sont évaluées par le calcul de l'information mutuelle moyenne (IMM) en se basant sur les caractéristiques statistiques du système.

L'équation du canal est construite en se basant sur la radiométrie et l'optique. Partant d'une séquence binaire, nous modélisons sa transposition graphique et l'acquisition par la caméra du

pictogramme généré. Conformément à l'usage courant, nous n'imposons pas de contraintes sur le positionnement du code, une caméra à focale fixe (cas le plus répandu) forme donc une image floue sur le capteur qui, de son côté, introduit un bruit lors de la quantification de l'éclairement à sa surface. Le flou est modélisé par un filtre linéaire paramétré par la position du code et les caractéristiques du bloc optique et s'appliquant à l'image du code déformé par la projection perspective. L'intégration de la lumière opérée par les pixels du capteur est modélisée par un filtre moyenneur à la sortie duquel s'ajoute un bruit blanc gaussien. Le système ainsi modélisé, l'information à la sortie de la caméra est liée à la séquence émise par une matrice caractéristique du canal synthétisant l'ensemble des distorsions. L'évaluation de l'information transmissible à travers le canal ainsi défini est alors différente selon que le système est en une dimension (code 1D et lecteur linéaire) ou en deux dimensions (code 2D et lecteur matriciel). Dans le premier cas, on met en évidence que l'équation du canal implique l'existence d'une mémoire qui est un processus markovien à partir duquel nous évaluons par simulations l'IMM. Dans le second, nous démontrons que l'information émise sachant celle reçue est un champ aléatoire de Gibbs, champ sur lequel nous adaptons les outils de la physique statistique pour l'estimation du flux d'information.

Cette démarche de modélisation du canal de transmission optique pour une évaluation de ses performances se rapproche de travaux précédemment publiés. Huck et al. modélisent dans [46] une chaîne d'acquisition en une suite de filtres modifiant le contenu spectral de l'image. En se basant sur la théorie de l'information, ils en déduisent un ensemble de critères permettant l'évaluation de la qualité d'une image destinée à être vue par l'œil. Cette méthode s'illustre dans d'autres publications de ces auteurs [47-49] et est à rapprocher des travaux de O'Sullivan et al. liant imagerie et théorie de l'information [50,51]. Plus proches du domaine du stockage optique des données, Chen et Siegel proposent dans [52] une méthode de calcul des bornes du flux d'information pour le modèle de canal de stockage optique TwoDOS [53].

Sur ce dernier point, la construction du modèle proposé dans nos travaux pour l'étude de la transmission d'information par codes graphiques a révélé qu'il n'y a, du point de vue formel, pas de différences avec le stockage optique. Les codes-barres et codes 2D soutiennent ainsi la comparaison avec les plus récentes unités de stockage de données, ce qui n'était pas évident de prime abord!

### **1.3 Organisation du rapport**

Le chapitre 2 est dédié à la présentation des outils de radiométrie et d'optique utilisés pour la construction du modèle. Les notions de flux, d'intensité et surtout de luminance sont introduites, ainsi que les caractéristiques des sources. Le processus de formation d'image est détaillé, impliquant les principes de l'optique géométrique. Ceux-ci mettent en évidence le problème du flou d'où découle la notion de profondeur de champ géométrique. Le flou est toutefois modélisé plus généralement avec l'optique de Fourier, et le système optique est assimilé à un système linéaire en puissance fonction des paramètres optiques de la caméra. Le bilan radiométrique d'un système optique est ensuite établi, déterminant l'éclairement produit sur le plan de projection. L'acquisition d'image est enfin modélisé par un filtre moyennneur traduisant le processus d'intégration des pixels.

Au chapitre 3, les systèmes de transmission par code-barres et lecteur linéaire sont modélisés à partir des éléments du chapitre précédent. L'équation du canal de communication est définie telle que les symboles reçus sont liés à l'information émise par une matrice synthétisant l'ensemble des distorsions. Nous mettons en évidence qu'un tel canal est à interférences entre symboles et dont la mémoire est markovienne. De cette propriété, une méthode de calcul de l'IMM est proposée pour la mesure des performances. De ce flux d'information sont déduits deux nouveaux critères : la profondeur de champ théorique et la résolution spatiale théorique que l'on met en perspective avec leur pendant géométrique.

Le chapitre 4 est consacré au systèmes de transmission par code 2D et imageur matriciel. De manière analogue au chapitre 3, l'équation du canal est construite à partir des éléments de radiométrie et d'optique introduits au chapitre 2. Il en découle une relation linéaire entre les symboles émis et reçus liés également par une matrice caractéristique du canal. L'analyse du modèle permet de mettre en évidence l'existence d'un champ aléatoire gibbsien. A partir de ses propriétés, nous démontrons que le flux d'information du système est proportionnel à l'énergie libre de ce champ, grandeur fondamentale de la physique statistique. Nous proposons ainsi une stratégie d'estimation de l'IMM dérivée de la propagation de croyance de Pearl [54].

Le chapitre 5 enfin conclut ce rapport en rappelant les principaux apports de ces travaux, et en en présentant les perspectives.





## **Chapitre 2**

# **Radiométrie et Formation d'Images**

L'utilisation d'un pictogramme tel que l'EAN13 ou le Datamatrix (figure 2.1) pour la transmission d'un numéro d'identification ou d'une URL revient à coder graphiquement une information binaire. Ce codage en une ou deux dimensions fait correspondre une zone sombre du pictogramme au "1" binaire et une zone claire au "0" (ou inversement). Cette modulation graphique constitue la base de la transmission. L'information transportée peut, elle, être d'une quelconque nature, compressée ou codée sans que cela ne modifie ce principe. Le pictogramme est ainsi un fichier binaire graphique.

Par conséquent nous identifions le système basé sur l'affichage du pictogramme et son acquisition par une caméra numérique comme un système de communication. L'étude du système proposée va nous amener à mesurer et quantifier la lumière, vecteur de l'information, et à déterminer les distorsions induites lors de son acheminement vers le récepteur. Pour l'un et l'autre de ces objectifs, nous introduisons dans ce chapitre les outils théoriques nécessaires à la poursuite de l'étude. Tout d'abord la radiométrie, d'où nous tirons les grandeurs fondamentales caractérisant le rayonnement des sources ainsi que les lois et principes qui les régissent [55, 56, 57, 58]. Ensuite nous traitons le processus de formation d'images et les distorsions que l'instrument optique induit entre l'objet et son image. La modélisation proposée tient compte de la projection perspective et du flou influant sur la distribution énergétique de la lumière projetée sur le capteur. Enfin, nous présentons un modèle d'acquisition de l'image basé sur des capteurs numériques de type CCD ou CMOS.



*Figure 2.1 : a/ EAN13 - b/ Datamatrix*

## **2.1 Radiométrie**

Quel que soit le support du pictogramme, l'émission d'information s'effectue en utilisant la lumière comme vecteur de communication . L'étude complète du système implique

donc la compréhension et la mesure du phénomène lumineux. La radiométrie est à ce titre la mesure énergétique d'un rayonnement. En parallèle, on définit la photométrie comme la radiométrie adaptée à l'oeil humain, alors considéré comme récepteur de référence. La lecture et l'identification d'un pictogramme (code-barres 1D ou code 2D) se réalise à l'aide de capteurs artificiels, nous nous positionnons donc pour la suite de ce rapport selon le point de vue de la radiométrie. Les notions fondamentales telles que le flux, la luminance et l'éclairement sont ainsi présentées, de même que les propriétés des sources de rayonnement. Ces grandeurs et leurs caractéristiques sont les éléments de base pour la construction du modèle d'un système de lecture de codes graphiques.

## 2.1.1 Grandeurs fondamentales

### 2.1.1.1 Le flux $F$

Le *flux* est l'énergie  $E(t)$  rayonnée par une source en une seconde. Il s'agit donc d'une mesure de puissance qui s'exprime en watts (figure 2.2).

$$F = \frac{1}{t_1 - t_0} \int_{t_0}^{t_1} E(t) dt \quad (2.1)$$

avec  $t_1 - t_0 = 1s$ .

Le flux est la grandeur de base de la radiométrie, celle dont toutes les autres découlent. Elle est d'une importance majeure pour notre étude car la réponse des capteurs d'images est fonction de l'énergie reçue sur leur surface.

En marge, notons que l'on lit parfois dans la littérature les expressions "flux énergétique", "flux radiant" ou encore "puissance radiante". Les qualificatifs "énergétique" ou "radiant" permettent de préciser que l'on se place dans le cadre de la radiométrie.

La notion de flux concerne l'ensemble du volume englobant la source. L'introduction d'une directivité mène à la définition de l'*intensité* d'une source.

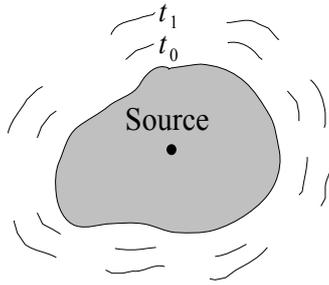


Figure 2.2 Flux d'une source rayonnant dans toutes les directions

### 2.1.1.2 L'intensité $I$

En préambule, rappelons que l'angle solide  $\Omega$  sous lequel est vu un objet depuis un point d'observation est le rapport entre l'aire  $S'$  de la calotte sphérique définie par le cône qui s'appuie sur le contour apparent de l'objet (et dont le centre est le point d'observation) et le carré du rayon  $r$  de cette calotte (figure 2.3). L'unité est le stéradian (sr).

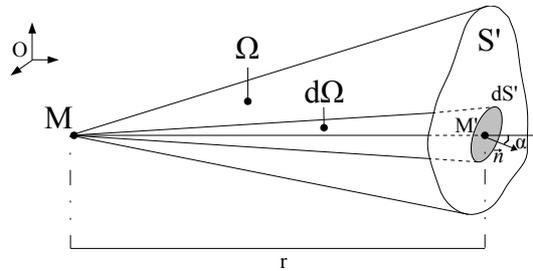


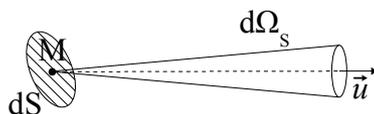
Figure 2.3: Angle Solide  $\Omega$

$$\Omega = \frac{S'}{r^2} \quad (2.2)$$

L'angle solide élémentaire est égal à :

$$d\Omega = \frac{dS' \cos \alpha}{r^2} \quad (2.3)$$

L'intensité  $I$  d'une source – également appelée intensité énergétique ou radiante – est la mesure du flux élémentaire  $dF$  émis par celle-ci dans une direction donnée et par unité d'angle solide (figure 2.4).



Source

Figure 2.4 : Intensité d'une source

Définie par l'équation (2.4), elle s'exprime en watts par stéradian ( $\text{W}\cdot\text{sr}^{-1}$ ).

$$I = \frac{dF}{d\Omega_s} \quad (2.4)$$

En photométrie on parle d'intensité lumineuse, exprimée en candela (cd).

Lorsque l'intensité est exprimée pour plusieurs, voire toutes les directions de l'espace, on obtient sa fonction indicatrice d'intensité  $I(\vec{u})$  (figure 2.5). Celle-ci renseigne sur l'enveloppe de la puissance rayonnée. Les sources telles que les lasers ont une indicatrice d'intensité allongée sur une seule direction tandis que les sources homogènes ont une indicatrice sphérique. Nous reviendrons un peu plus tard sur les caractéristiques des sources.

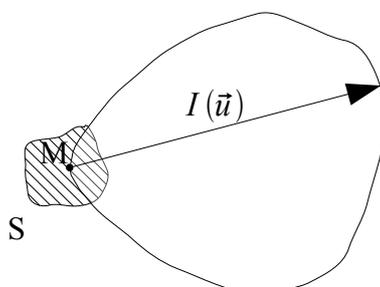


Figure 2.5: Indicatrice d'intensité

### 2.1.1.3 La luminance énergétique $L$

La luminance énergétique  $L$  est l'intensité d'une source par unité d'aire apparente. Elle est définie par l'équation (2.5) et son unité est le watt par stéradian par mètre carré ( $\text{W}\cdot\text{sr}^{-1}\cdot\text{m}^{-2}$ ).

$$L = \frac{dI}{dS \cos \alpha} \quad (2.5)$$

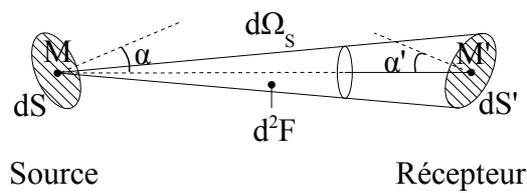


Figure 2.6 : Luminance énergétique

Ainsi, en se reportant à la figure 2.6, la luminance en un point  $M$  de la source rayonnante dans la direction d'un point  $M'$  lie le flux élémentaire  $d^2F$  émis par la surface élémentaire  $dS$  autour de  $M$ , et la surface élémentaire  $dS'$  autour de  $M'$ . On a :

$$L = \frac{d^2F}{d\Omega_s \cdot dS \cos \alpha} = \frac{d^2F \cdot r^2}{dS' \cos \alpha' dS \cos \alpha} \quad (2.6)$$

Le pendant photométrique de la luminance énergétique est la luminance visuelle exprimée en  $\text{cd}/\text{m}^2$ .

A partir de l'équation de luminance (2.6), on détermine le flux élémentaire  $d^2F$  porté par le pinceau allant de  $dS$  vers  $dS'$  (figure 2.7):

$$d^2F = L \frac{dS dS' \cos(\alpha) \cos(\alpha')}{r^2} = L d^2G \quad (2.7)$$

$d^2G$  est appelée *étendue géométrique* du pinceau reliant les deux surfaces  $dS$  et  $dS'$ . Elle modélise la portion de l'espace à travers laquelle le flux  $d^2F$  s'écoule de  $dS$  vers  $dS'$ .

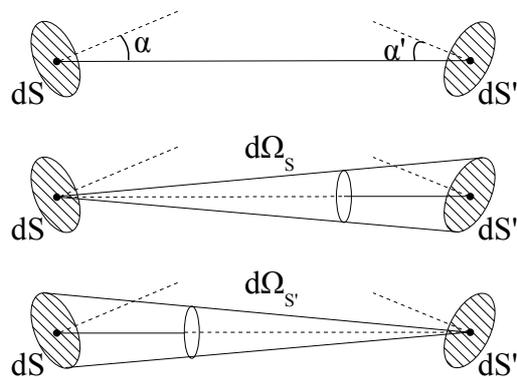


Figure 2.7 : Notations pour le calcul de l'étendue géométrique

La luminance, qu'elle soit énergétique ou visuelle, est une grandeur très utilisée car elle renseigne quant à la puissance émise par une unité de surface et dans une direction donnée. Dans le cas de la photométrie, elle est une grandeur objective (quantifiable) de l'excitation que la source va induire à l'oeil de l'observateur. La luminance complète donc la notion de luminosité, ou anciennement brillance, qui est une perception subjective du rayonnement d'une source. En effet, la clarté perçue d'un objet est relative à son environnement de sorte que pour une luminance constante, celui-ci paraîtra plus ou moins brillant suivant son voisinage (figure 2.8).

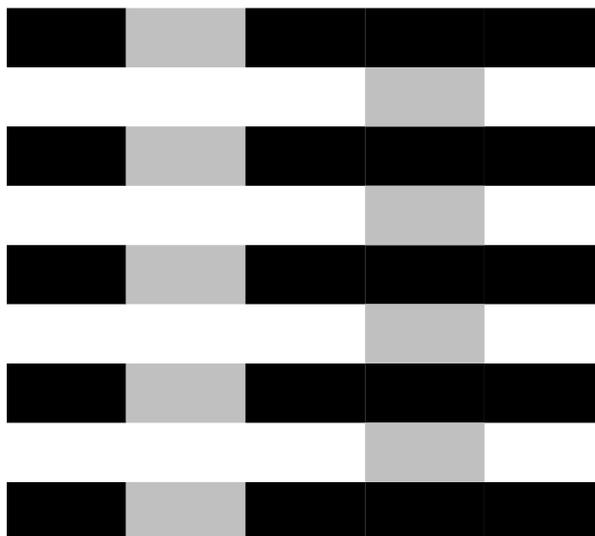


Figure 2.8 : Illusion d'optique de White, la luminance du gris est toujours la même

Après avoir défini l'outil radiométrique général caractérisant un rayonnement quelconque, nous introduisons deux grandeurs relatives à la distribution surfacique du flux: l'éclairement d'une surface réceptrice, et l'émittance d'une surface émettrice. Dans notre système, l'émittance sert à mesurer la puissance émise du code et l'éclairement celle atteignant le capteur.

#### 2.1.1.4 l'Eclairement énergétique $E_c$

L'éclairement énergétique  $E_c$  est le flux reçu par unité de surface réceptrice.

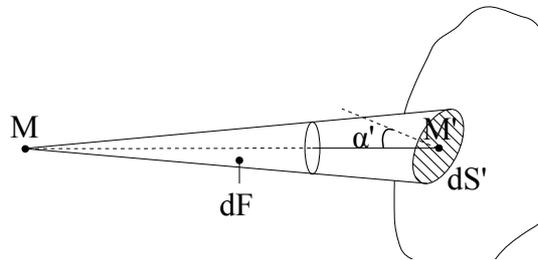


Figure 2.9 : Eclairement

Ainsi, d'après la figure 2.9, l'éclairement  $E_c$  au point  $M'$  est égal à :

$$E_c = \frac{dF}{dS'} \quad (2.8)$$

L'éclairement énergétique s'exprime en watts par mètre carré ( $W \cdot m^{-2}$ ). Il est appelé "éclairement" en photométrie, sans complément, et son unité est alors le lumen par mètre carré, ou le *lux* avec  $1 \text{ lm} \cdot m^{-2} = 1 \text{ lx}$ .

Signalons que l'éclairement énergétique peut également s'exprimer en fonction de l'intensité énergétique. Combinant (2.3), (2.4) et (2.8) nous obtenons :

$$E_c = \frac{I \cos(\alpha')}{r^2} \quad (2.9)$$

### 2.1.1.5 L'Émittance énergétique $E_m$

L'émittance (ou exitance) énergétique est la mesure du flux émis par unité de surface émettrice (figure 2.10):

$$E_m = \frac{dF}{dS} \quad (2.10)$$

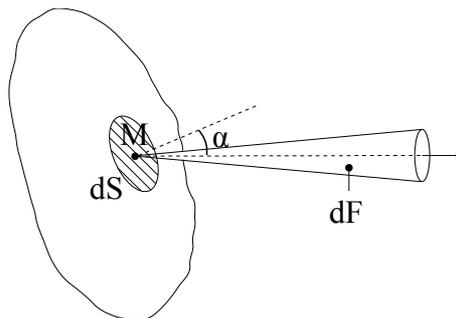


Figure 2.10 : Émittance

A l'instar de l'éclairement énergétique  $E_c$ , l'émittance s'exprime en  $\text{W.m}^{-2}$ . Notons qu'en photométrie on parle également d'exitance, exprimée cependant en lux.

### 2.1.1.6 Récapitulatif

Le tableau 2.1 regroupe l'ensemble des grandeurs radiométriques présentée dans ce paragraphe, leur pendant photométrique ainsi que leur unité. De plus, il présente la dénomination anglaise équivalente couramment rencontrée dans la littérature.

<b>Grandeur radiométrique</b>	<b>Symbole</b>	<b>Unité</b>	<b>Equivalent anglais</b>	<b>Equivalent photométrique</b>
Flux radiant	$F$	W	Radiant flux	Flux lumineux (lumen)
Intensité	$I$	W.sr <sup>-1</sup>	Intensity	Intensité lumineuse (cd)
Luminance énergétique	$L$	W.sr <sup>-1</sup> .m <sup>-2</sup>	Radiance	Luminance visuelle (cd.m <sup>-2</sup> )
Eclairement énergétique	$E_c$	W.m <sup>-2</sup>	Irradiance	Eclairement (lux)
Emittance énergétique	$E_m$	W.m <sup>-2</sup>	Radiant Emittance	Exitance (lux)

*Tableau 2.1: Grandeurs Radiométriques*

## **2.1.2 Les sources de rayonnement**

La lecture d'un code, qu'il soit mono ou bidimensionnel, implique que le pictogramme émette directement ou indirectement de la lumière. L'information graphique est donc systématiquement associée à un générateur de lumière caractérisé par les grandeurs radiométriques présentées précédemment. Afin de compléter l'étude énergétique du rayonnement menant à la construction du modèle du système de transmission, nous introduisons les sources de rayonnement et leurs propriétés, tout en restant dans le cadre de la lecture de codes. Une bonne compréhension des bases de la production de lumière et sa terminologie permettra au lecteur de mieux se figurer les sources réelles, leur spectre émis et leurs liens respectifs. Pour une approche plus exhaustive, nous invitons le lecteur à consulter les références [58, 59].

### **2.1.2.1 Principes fondamentaux - terminologie**

La production de lumière repose sur l'excitation des électrons d'un atome. De l'énergie est injectée au complexe noyau-électron d'une molécule, modifiant la configuration du nuage électronique. Son retour à un état d'équilibre implique l'évacuation de l'excès d'énergie qui est alors émis sous forme de rayonnement. Une source est ainsi un générateur de rayonnement. Lorsque celui-ci se situe dans les longueurs d'onde détectables par l'oeil, c'est à dire comprises

entre 400 et 700 nm, on parle de générateur de lumière. Toutefois il est rare qu'une source n'émette que dans le spectre visible et elle est bien souvent également générateur d'ondes infrarouges et de rayons UV.

La production du rayonnement s'effectue à partir de différentes formes d'énergie telles que la chaleur (les lampes à incandescence) ou l'électricité (les lampes à décharges ou néons).

De ces énergies excitatrices découlent deux modalités de source : un mode "chaud" lié à la température propre de l'objet rayonnant, et un mode "froid" englobant les cas où la température n'intervient pas. Le premier cas englobe les sources thermiques (incandescence). L'étude de leur rayonnement se réfère au modèle du *corps noir* introduit par Gustav Kirchoff dont l'émittance  $Em^0$  est donnée par la loi de Planck :

$$Em^0_{\lambda} = \frac{2\pi hc^2}{\lambda^5} \frac{1}{e^{\frac{hc}{\lambda k_B T}} - 1} \quad (2.11)$$

où  $c \approx 3 \cdot 10^8 \text{ m} \cdot \text{s}^{-1}$  est la vitesse de la lumière,  $k_B \approx 1.38 \cdot 10^{-23} \text{ J} \cdot \text{K}^{-1}$  est la constante de Boltzmann,  $T$  est la température exprimée en kelvin (K) et  $\lambda$  est la longueur d'onde en m. L'équation (2.11) donne une bonne approximation de l'émittance de sources thermiques réelles telles que le soleil ou les lampes classiques.

Dans le second cas (mode "froid") on utilise le terme luminescence. Les procédés d'excitation des électrons des atomes de la source sont alors multiples : réactions chimiques pour lesquelles on parle de chimiluminescence (bâtons lumineux de secours "lightsticks" utilisés en mer) et de bioluminescence chez les être vivants (lucioles), l'électricité qui donne l'électroluminescence (lampes néons, diodes électroluminescentes, etc.), la contrainte mécanique qui donne la triboluminescence, etc. En marge des ces variantes, précisons les termes "fluorescence" et "phosphorescence" souvent rencontrés dans la littérature ne désignent pas un procédé mais une durée. La fluorescence consiste en une courte émission après excitation alors que pour la phosphorescence, l'émission est persistante dans le temps.

### 2.1.2.1.1 Exemples de source

Une courte sélection de source lumineuse est donnée au tableau 2.2. Ces sources sont caractérisées par leur luminance énergétique.

Source	Luminance énergétique ( $\text{W}\cdot\text{m}^{-2}\cdot\text{sr}^{-1}$ )
Soleil (au niveau de la mer)	$1.3\cdot 10^7$
Ciel non couvert	100
Ciel couvert	$< 10$
Lune (pleine)	22
Lampe à incandescence (100 W)	$2\cdot 10^5$

Tableau 2.2 : Quelques sources et leur luminance

### 2.1.2.1.2 Sources singulières et leurs propriétés

En marge des principes généraux de la production de lumière, certaines sources se démarquent par la corrélation de leur spectre ou par la variation de leur rayonnement suivant sa direction. Nous présentons ici trois cas dont les propriétés seront utilisées dans la construction du modèle du canal, notamment pour le flou et lors de l'acquisition.

#### 2.1.2.1.2.1 Source cohérente et source incohérente

La nature ondulatoire du rayonnement lumineux permet de définir le concept de cohérence d'une source :

- si le rayonnement émis est constitué de paquets d'ondes quasi monochromatiques dont les phase sont corrélées, on parle de *cohérence temporelle*,
- si les vibrations émises par deux points de la source présentent une différence de phase qui ne dépend que des positions de ces points rayonnants, on parle de *cohérence spatiale*.

Le laser est l'exemple le plus représentatif d'une source cohérente (figure 2.11). Son mode d'émission stimulée génère une onde quasi monochromatique, et sa taille permet de l'apparenter à une source quasi ponctuelle.

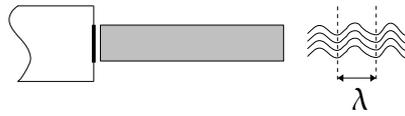


Figure 2.11: Cohérence spatiale et temporelle

Par opposition, un exemple de source *incohérente* est un objet large rayonnant une lumière blanche (figure 2.12). Le soleil, les lampes à incandescence sont de bons exemples de ce type de source et c'est dans ce cadre que nous inscrivons nos travaux.

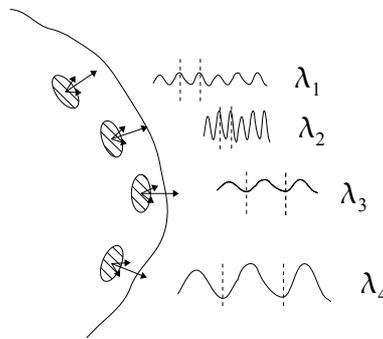


Figure 2.12 : Source incohérente

#### 2.1.2.1.2.2 Source ponctuelle isotrope

Une source ponctuelle dont la luminance ne varie pas selon la direction est dite *isotrope*. L'émittance d'une telle source est alors directement proportionnelle à sa luminance.

En effet :

$$Em = \frac{dF}{dS} = \int L \cos(\theta) d\Omega \quad (2.12)$$

Or  $L$  est constante quelle que soit l'orientation :

$$Em = L \int \cos(\theta) d\Omega \quad (2.13)$$

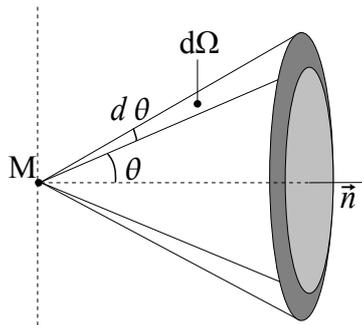


Figure 2.13 : Source ponctuelle - intégration sur l'angle solide

De plus, la mesure de l'angle solide illustré en figure 2.13 est donné par le résultat classique :

$$d\Omega = 2\pi \sin(\theta) d\theta \quad (2.14)$$

L'émittance dans une direction donnée est donc le résultat de l'intégration sur la demi-sphère :

$$Em = 2\pi L \int_0^{\frac{\pi}{2}} \cos(\theta) \sin(\theta) d\theta = 2\pi L \left[ \frac{\sin^2(\theta)}{2} \right]_0^{\frac{\pi}{2}} = \pi L \quad (2.15)$$

Une source isotrope vérifie donc l'égalité :  $Em = \pi L$  appelée loi de Lambert (la source est qualifiée de *lambertienne*).

### 2.1.2.1.2.3 Source étendue homogène

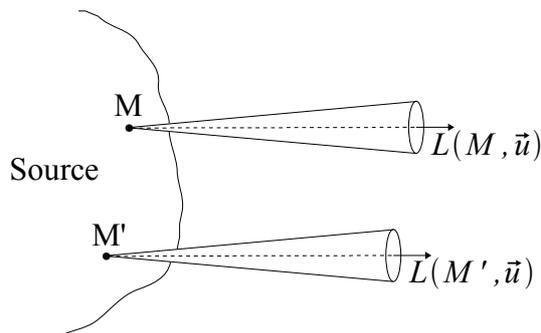


Figure 2.14 : Source étendue homogène

Une source étendue homogène est une source dont la luminance est la même quelle que soit le point émetteur (figure 2.14). On a :

$$L(M, \vec{u}) = L(M', \vec{u}) \quad (2.16)$$

## 2.2 Formation d'images

Durant la transmission, le pictogramme irradie dans l'espace un flux énergétique propageant son message binaire modulé graphiquement. Ce rayonnement, soit issu directement du pictogramme (lorsque le code est affiché par un écran), soit indirectement par réflexion d'une source éclairante (code imprimé sur papier), est capturé en partie par le système optique de la caméra formant une image sur le plan du capteur numérique. L'appréhension de ce processus de formation d'image est complexe dans le sens où elle nécessite plusieurs outils mathématiques pour pouvoir en définir le modèle et étudier la propagation de l'énergie lumineuse au travers du système optique. Nous employons pour cela l'optique géométrique ainsi que l'optique ondulatoire, notamment pour le phénomène de diffraction. Dans ce paragraphe, nous présentons les concepts et outils de la formation d'une image par un système optique analogue aux lecteurs de code. Dans une démarche analytique, nous passons en revue les transformations de l'espace pour la modélisation des problèmes de positionnement du système, l'optique géométrique pour la projection de la scène sur le capteur, et l'optique ondulatoire pour la modélisation du flou de défocalisation.

### 2.2.1 Optique géométrique

En optique géométrique, seule la direction de propagation de l'énergie lumineuse est considérée : le rayon lumineux. Cette simplification du rayonnement occulte sa nature ondulatoire mise en évidence par Maxwell, ne permettant pas d'expliquer par exemple le phénomène de diffraction ou les interférences. Elle permet toutefois de modéliser de manière simple les principaux processus optiques, à savoir la réflexion et la réfraction, ainsi que le concept même du couple {objet, image}. Le rayonnement lumineux ainsi résumé à la propagation de son énergie permet l'étude des systèmes optiques classiques par simple construction géométrique, d'où son nom.

Deux hypothèses fondamentales en constituent la base :

- la propagation de la lumière est rectiligne dans un milieu homogène et isotrope, caractérisé par un indice dit de réfraction  $n=c/v$  constant, où  $c$  est la vitesse de la lumière dans le vide et  $v$  est la vitesse de la lumière dans le milieu considéré,

- le principe du retour inverse de la lumière, en d'autres termes la direction de propagation prévaut sur son sens.

A partir de ces fondamentaux, on établit le comportement du rayon lumineux lors de la rencontre de différents milieux et les conditions de la formation d'une image. Pour le détail de ces développements, nous orientons le lecteur vers [57, 60] et ne présentons ici que les résultats utilisés pour l'étude. Nous considérons notamment le cas d'un milieu d'indice de réfraction  $n$  de faible épaisseur plongé dans l'air, et dont les délimitations sont des calottes sphériques. Cet élément optique est une *lentille mince* caractérisé par un centre optique  $O$ , une ouverture optique  $D$  et une distance focale image  $f$  (figure 2.15).

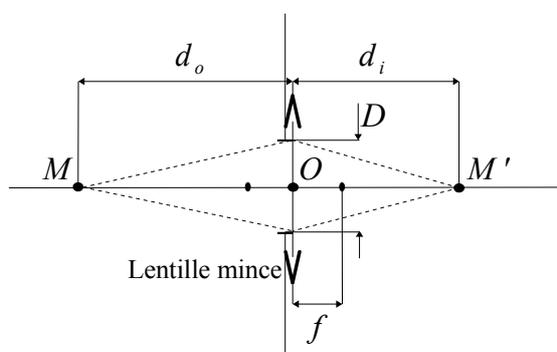


Figure 2.15 : Schéma d'une lentille mince

On a alors la relation :

$$\frac{1}{d_i} - \frac{1}{d_o} = \frac{1}{f} \quad (2.17)$$

avec les distances algébriques objet  $d_o = \overline{OM}$  et image  $d_i = \overline{OM'}$ . Notons que  $f > 0$  pour les systèmes convergents.

La formule (2.17) est la *loi des lentilles minces* et les distances objet et image définissent le *grandissement transversal* de la lentille par :

$$G_t = \frac{d_i}{d_o} \quad (2.18)$$

L'objectif d'une caméra est un système optique composé de nombreux éléments alignés sur le même axe optique (le système est dit centré). Il est caractérisé par sa distance focale et son diaphragme, et sa principale fonction est de canaliser le flux radiant de la scène observée pour être projeté sur la zone photosensible (film ou capteur numérique).

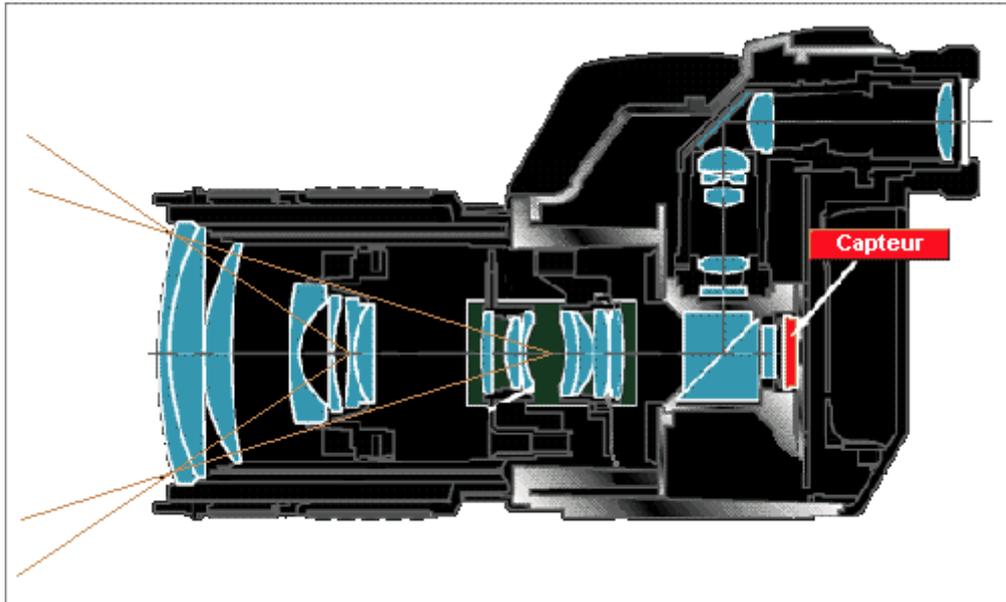


Figure 2.16 : Caméra numérique en vue de coupe (© Arnaud Frich)

En utilisant les principes de l'optique géométrique précédemment introduits, le processus de formation d'image d'une caméra est modélisé en assimilant son objectif à une lentille mince dont les caractéristiques (focale, ouverture, centre optique) sont tirées de ses spécifications. Cette modélisation, classique en vision par ordinateur, rend compte de la principale déformation induite qu'est la projection perspective. Elle ne conserve ni les distances entre les points, ni les rapports de distances. En se reportant à la figure 2.17, on a :

$$\frac{\overline{AC}}{\overline{BC}} \neq \frac{\overline{A'C'}}{\overline{B'C'}} \quad (2.19)$$

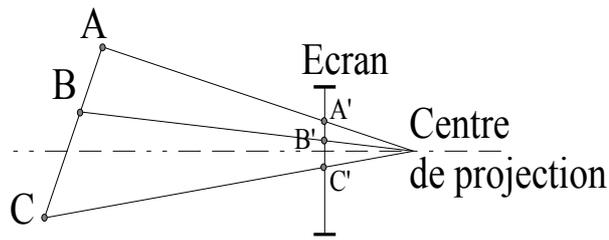


Figure 2.17 : La projection perspective et les rapports de distances

Cette projection fait apparaître les objets plus petits au fur et à mesure qu'ils s'éloignent du centre optique. Le grandissement transversal de la lentille équivalente tend donc à diminuer avec cet éloignement et l'on a de manière schématisée la situation suivante (figure 2.18 a et b):

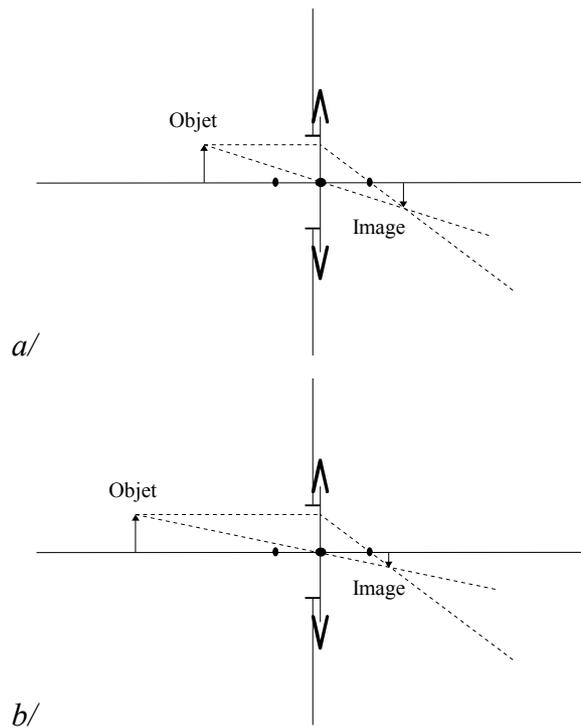


Figure 2.18 : Eloignement et rétrécissement de l'image

Conséquemment à cette modélisation, la connaissance de la position du point objet implique celle de son conjugué. Nous introduisons ainsi au prochain paragraphe les notions de

géométrie dans l'espace, et principalement le formalisme des *coordonnées homogènes* permettant le calcul matriciel des transformations telles la translation, la rotation ou la projection.

## 2.2.2 Géométrie dans l'espace

### 2.2.2.1 Repères et transformations

Le positionnement relatif du code graphique et de la caméra est effectué par l'intermédiaire de deux repères orthonormés dans l'espace tridimensionnel  $E$  (figure 2.19). Le premier est le *repère objet*  $R(O, \vec{x}, \vec{y}, \vec{z})$  relatif au pictogramme observé. Tout point  $M$  de l'espace a pour coordonnées dans  $R$  le vecteur  $\mathbf{M} = [x, y, z]^T$ . Le second repère, lié à la caméra, est le *repère image*  $R'(O', \vec{x}', \vec{y}', \vec{z}')$ . Tout point  $M'$  de l'espace a pour coordonnées dans  $R'$  le vecteur  $\mathbf{M}' = [x', y', z']^T$ . L'axe  $[O' \vec{z}')$  est confondu avec l'axe optique de la caméra.

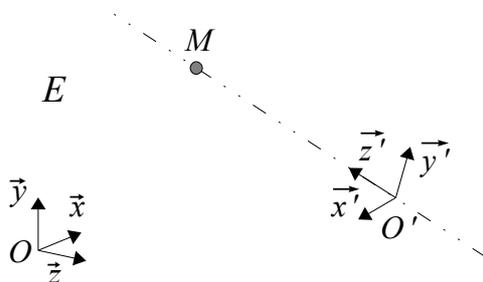


Figure 2.19 : Repères monde et image

L'expression des coordonnées d'un point du repère objet dans le repère image s'effectue par translations et rotations. L'usage courant des codes veut qu'il n'y ait pas de contraintes imposées à la position du pictogramme par rapport à la caméra, si ce n'est la nécessité d'être "visible". Cette liberté implique de pouvoir traiter dans la modélisation du système les déplacements et changements d'échelle éventuels, ainsi que la projection perspective effectuée par l'imageur.

A ces fins, nous adoptons un outil classique de l'infographie : les coordonnées homogènes

[55]. Il s'agit en pratique d'une représentation des points d'un espace de dimension  $d$  par un vecteur à  $d+1$  composantes. Ainsi un point  $M[x, y, z]^T$  de  $E$  a pour coordonnées homogènes le vecteur  $[x, y, z, w]^T$  avec l'une des quatre coordonnées non nulles. Par convention, on prend  $w=1$ . Pour la justification de cette définition et les implications théoriques d'une telle représentation, nous orientons le lecteur vers [55, 60, 61]. Le point important est que l'utilisation des coordonnées homogènes permet de transposer les transformations affines de l'espace (translation, rotation, homothétie) et les projections en multiplications matricielles.

En effet, soit le point  $M$  de coordonnées homogènes  $M^{(h)}=[x, y, z, 1]^T$ . L'image  $M'{}^{(h)}$  de  $M^{(h)}$  par la translation de vecteur  $T=[T_x, T_y, T_z]^T$  (figure 2.20) a pour coordonnées homogènes :

$$M'{}^{(h)} = \begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & T_x \\ 0 & 1 & 0 & T_y \\ 0 & 0 & 1 & T_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = T^{(h)} \cdot M^{(h)} \quad (2.20)$$

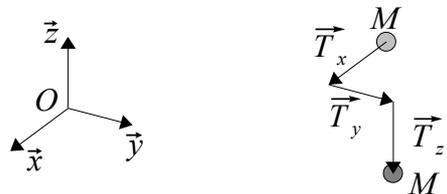


Figure 2.20 : Translation de  $M$  vers  $M'$

L'application du formalisme de coordonnées homogènes revient donc à définir à partir du vecteur  $T$  la matrice  $T^{(h)}$ .

Soient les rotations d'angles  $\alpha$ ,  $\beta$  et  $\gamma$  dans  $E$  (figure 2.21).

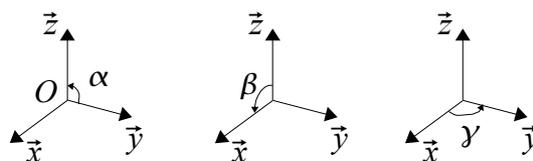


Figure 2.21 : Rotations dans  $E$

On définit respectivement les matrices :  $\mathbf{R}_\alpha = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha & 0 \\ 0 & \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ ,

$$\mathbf{R}_\beta^{(h)} = \begin{bmatrix} \cos \beta & 0 & \sin \beta & 0 \\ 0 & 1 & 0 & 0 \\ -\sin \beta & 0 & \cos \beta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \text{ et } \mathbf{R}_\gamma^{(h)} = \begin{bmatrix} \cos \gamma & -\sin \gamma & 0 & 0 \\ \sin \gamma & \cos \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Soit l'homothétie de facteurs d'échelle  $\mathbf{S} = [S_x, S_y, S_z]^T$  (figure 2.22).

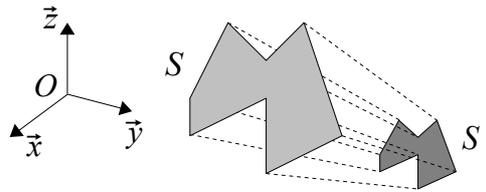


Figure 2.22 : Homothétie

Son équivalent en coordonnées homogènes est défini par la matrice :

$$\mathbf{S}^{(h)} = \begin{bmatrix} S_x & 0 & 0 & 0 \\ 0 & S_y & 0 & 0 \\ 0 & 0 & S_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Enfin, nous avons modélisé la déformation géométrique opérée par l'imageur par une projection perspective. Appliquée au point  $M$ , cette opération est transposée en coordonnées homogènes par la multiplication du vecteur  $\mathbf{M}^{(h)}$  par une matrice  $\mathbf{P}^{(h)}$  égale à :

$$\mathbf{P}^{(h)} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1/f & 0 \end{bmatrix}$$

où  $f$  est la distance focale de l'imageur. Le vecteur des coordonnées du projeté  $\mathbf{M}^{(h)}$  de

$M^{(h)}$  est obtenu après normalisation du résultat du produit  $P^{(h)} \cdot M^{(h)}$  par sa quatrième composante égale à  $z/f$ .

### 2.2.2.2 Profondeur de champ et flou

En plus de la déformation géométrique induite par la projection perspective, la modélisation de la caméra met en évidence un problème de convergence des rayons émergents lorsque la position du plan image et la focale sont constantes. En effet, par l'équation (2.17), il n'existe qu'une position possible du plan objet pour un plan image donné et toute autre disposition donne un étalement de l'information lumineuse traduisant l'inadéquation focalisation, c'est le flou. On définit alors des plages de variation du plan objet (ou image) pour un étalement acceptable en regard de la technologie d'acquisition utilisée. Ces plages sont respectivement la *profondeur de champ* et la *profondeur de focus* (figures 2.23 a/ et b/).

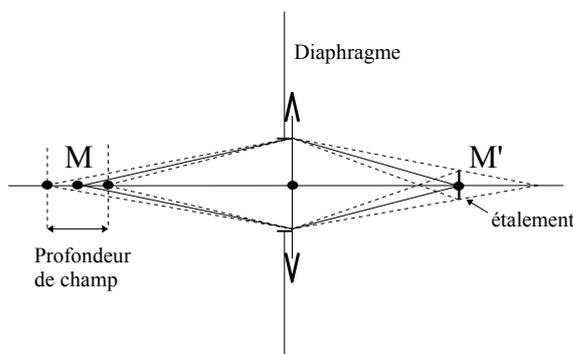


Figure 2.23a : Profondeur de champ

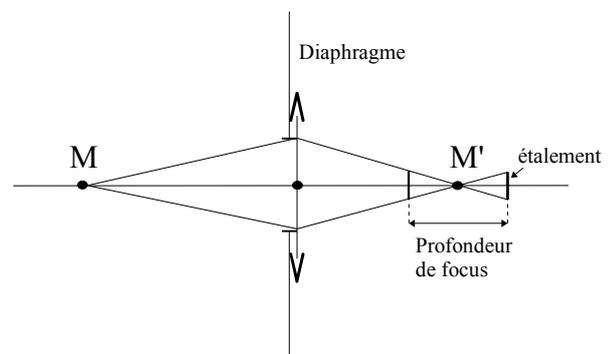


Figure 2.23b : Profondeur de focus

La position du plan capteur de l'imageur est le plus souvent non modifiable de sorte que la profondeur de focus est moins utilisée.

Analytiquement, la profondeur de champ que nous qualifions ici de géométrique (PdCG), revient à évaluer les positions limites du plan objet pour lesquelles on obtient un certain étalement  $e$  acceptable. Cet étalement correspond le plus souvent à la taille minimale d'un point image que la zone photosensible peut reproduire. Avec les caméras numériques cela correspond à la taille d'un pixel. Nous notons  $z_{pp}$  la position sur l'axe optique du plan le plus

proche du centre  $O$ , et  $z_{pe}$  celle du plan éloigné.

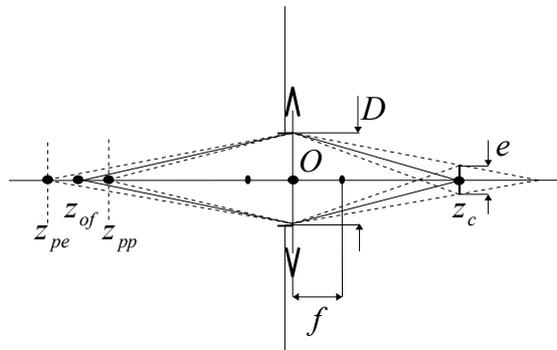


Figure 2.24 : Notations pour le calcul de la profondeur de champ géométrique

En se reportant à la figure 2.24 et en utilisant le classique théorème de Thalès il vient :

$$z_{pp} = \left( \frac{1}{f} + \frac{e-D}{D z_c} \right)^{-1} \quad (2.21)$$

et

$$z_{pe} = \left( \frac{1}{f} - \frac{e+D}{D z_c} \right)^{-1} \quad (2.22)$$

En pratique, on constate que malgré un bon positionnement du plan objet suivant (2.17) pour une focalisation sur le plan du capteur, il subsiste un étalement dont l'ampleur varie suivant la taille du diaphragme de l'imageur. Ce flou non prédit par l'optique géométrique est dû à la diffraction, c'est à dire à la modification par le diaphragme du front d'onde incident. Ce phénomène inhérent au fait que la lumière est une onde électromagnétique nécessite l'introduction de nouveaux concepts et outils : l'optique ondulatoire.

### 2.2.3 Optique ondulatoire

Durant les cinq derniers siècles, de nombreux scientifiques illustres émirent l'hypothèse que la lumière était une onde se propageant de proche en proche à l'instar du son ou des vaguelettes produites en jetant un caillou à la surface de l'eau. Descartes, Hooke, Huygens, Euler et bien d'autres édifièrent chacun leur théorie intégrant l'hypothèse ondulatoire de la lumière, mais c'est J.C. Maxwell qui, en 1873 dans [62], apporte la

modélisation et le formalisme le plus probant. En effet, en constatant que les ondes électromagnétiques se propagent à la même vitesse que la lumière (dont la mesure la plus précise avait été effectuée par Foucault en 1862), Maxwell émet l'hypothèse que la lumière est elle-même une onde électromagnétique, répondant donc aux mêmes lois.

*"The agreement of the results seems to show that light and magnetism are affections of the same substance, and that light is an electromagnetic disturbance propagated through the field according to electromagnetic laws." [63]*

A ce stade, l'étude du phénomène lumineux est une étude rigoureuse et mathématique. Celle-ci, sous couvert de certaines simplifications (introduites par Kirchoff, Rayleigh et Fresnel), permet l'interprétation du comportement de l'onde lors de sa rencontre avec une ouverture (telle un diaphragme), de l'impact de la cohérence de la source ou encore de la convergence des lentilles. Pour le détail de ces développements, fondamentaux en optique mais dont seuls les résultats sont utilisés dans ce mémoire, nous invitons le lecteur à consulter [64].

Nous retenons ici le cas particulier des lentilles en éclairage incohérent, qui correspond à la modélisation de notre système optique (paragraphe 2.2.1). Un des principaux résultats est que le système optique est linéaire en puissance, c'est à dire par rapport au flux radiant  $F$  (paragraphe 2.1.1.1). Il est caractérisé par une fonction de transfert optique  $FTO(f_{x'}, f_{y'})$  (ou son pendant spatial, la réponse impulsionnelle optique  $RIO(x', y')$ ), et la puissance du champ lumineux formé est la convolution de la puissance radiante projetée sur le plan image par  $RIO$ . On montre que  $FTO$  dépend de la fonction de forme  $D(x', y')$  du diaphragme du système, également appelée "fonction pupille" ([64]). Tout système optique est donc limité en résolution par sa tâche de diffraction qui est fonction du diaphragme du système.

Prenons le cas d'un bloc optique muni d'un diaphragme carré de largeur  $2l_D$  (figure 2.25).

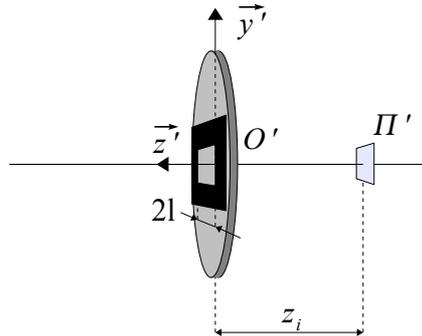


Figure 2.25 : Schéma du système optique

Sa fonction de transfert optique est donnée par :

$$FTO(f_{x'}, f_{y'}) = \Lambda\left(\frac{f_{x'}}{2f_o}\right) \Lambda\left(\frac{f_{y'}}{2f_o}\right) \quad (2.23)$$

où  $\Lambda$  est la fonction triangle,  $f_o = l_D / (\lambda z'_i)$  est la fréquence de coupure optique,  $\lambda$  la longueur d'onde et  $z'_i$  est la position du plan image sur l'axe optique.

Un second résultat de l'analyse ondulatoire de systèmes optiques important pour notre étude est la généralisation de la fonction de forme du diaphragme. Cette généralisation permet de modéliser une déformation du front d'onde en incluant un terme de phase. Ce déphasage entre les ondes émanant du diaphragme est une représentation mathématique élégante des éventuelles aberrations du système. On définit ainsi le diaphragme généralisé par la fonction complexe:

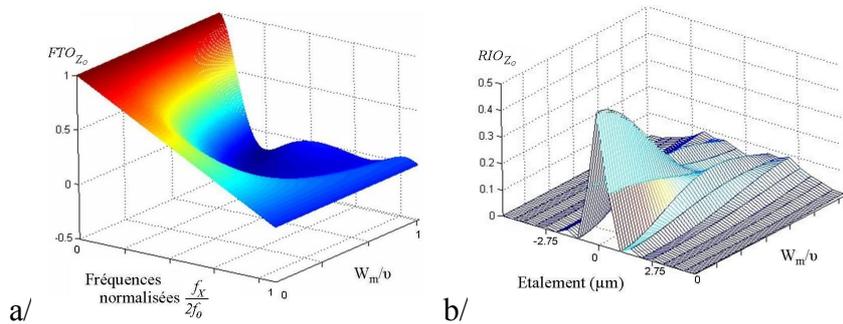
$$D_{\text{gén}}(x', y') = D(x', y') e^{jkW(x', y')} \quad (2.24)$$

où  $k = 2\pi/\lambda$  et  $W(x', y')$  mesure les éventuelles différences entre les trajets des fronts d'onde émanant du diaphragme.

Pour un système optique à lentilles, un mauvais positionnement du plan objet par rapport à la distance focale revient à un décalage du point de convergence du front d'onde. Dans le cas du système optique précédent à diaphragme carré, une telle erreur de focalisation a un impact direct sur la fonction de transfert (2.23). On montre que celle-ci se développe en :

$$FTO(f_{x'}, f_{y'}) = \Lambda\left(\frac{f_{x'}}{2f_o}\right) \Lambda\left(\frac{f_{y'}}{2f_o}\right) \cdot \text{sinc}\left[\frac{8W_m}{\lambda}\left(\frac{f_{x'}}{2f_o}\right)\left(1 - \frac{|f_{x'}|}{2f_o}\right)\right] \text{sinc}\left[\frac{8W_m}{\lambda}\left(\frac{f_{y'}}{2f_o}\right)\left(1 - \frac{|f_{y'}|}{2f_o}\right)\right] \quad (2.25)$$

avec  $W_m = 1/2(1/z'_a - 1/z'_i)l^2$  la fonction de décalage,  $z'_a$  la position théorique du point image formé et  $z'_i$  la position du plan image sur l'axe optique.



Figures 2.26 : a/ FTO et b/ RIO d'un système diaphragme carré

Nous avons tracé en figure 2.26 les variations de la fonction de transfert par rapport à  $W_m/\lambda$ . La FTO étant symétrique, nous n'avons reporté que les valeurs pour les fréquences spatiales horizontales.

En l'absence d'erreur de focalisation, soit  $W_m/\lambda = 0$ , on constate que la réponse fréquentielle diminue linéairement pour les hautes fréquences. Il s'agit du flou de diffraction, non prédit par l'optique géométrique mais clairement illustré avec l'approche ondulatoire. Lorsque la défocalisation augmente la bande passante du filtre diminue, et à partir d'un décalage par longueur d'onde  $(W_m/\lambda) > 0.5$ , la FTO prend des valeurs négatives traduisant des inversions de contraste que l'on observe en pratique pour les images fortement défocalisées (figure 2.27).

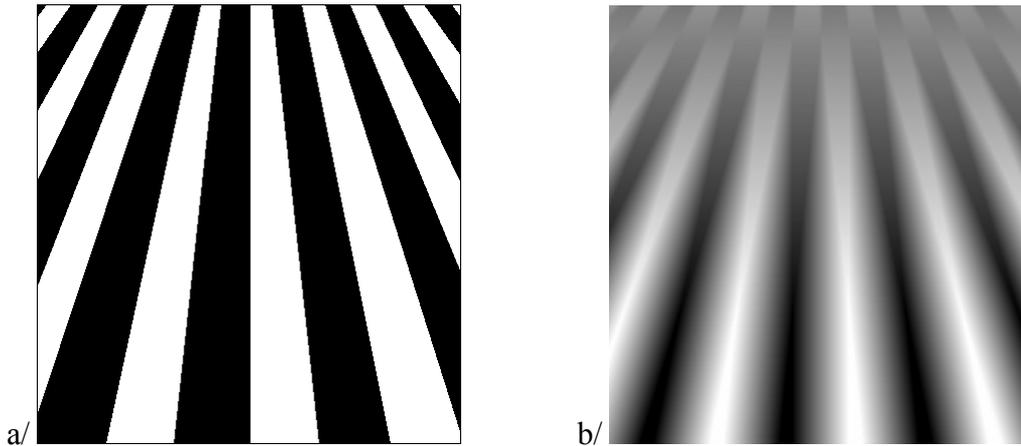


Figure 2.27 : Inversion de contraste d'une mire (a/) avec une défocalisation progressive (b/)

Pour les défocalisations très fortes, le système ne laisse passer que les très basses fréquences, de sorte que la fonction de transfert peut se réduire à :

$$FTO(f_{x'}, f_{y'}) \approx \text{sinc} \left[ \frac{8W_m}{\lambda} \left( \frac{f_{x'}}{2f_o} \right) \right] \text{sinc} \left[ \frac{8W_m}{\lambda} \left( \frac{f_{y'}}{2f_o} \right) \right] \quad (2.26)$$

La réponse impulsionnelle correspondant à cette  $FTO$  réduite est un signal carré, c'est à dire précisément la fonction de forme du diaphragme du système (à un facteur d'échelle près). Ce résultat concorde avec les prédictions de l'optique géométrique qui stipule que la réponse du système est la projection de son diaphragme sur le plan image [64]. De par sa généralité, l'optique ondulatoire englobe et complète l'optique géométrique.

### 2.3 Radiométrie d'un système optique

La formation d'une image par un système optique est ainsi modélisée par le filtrage de la puissance du champ lumineux capturé par un noyau de convolution qui dépend de la forme du diaphragme et des conditions d'acquisition. Les grandeurs radiométriques (la luminance énergétique, l'éclairement, ou encore l'émittance) présentées au paragraphe 2.1 sont toutes proportionnelles au flux radiant qui est homogène à une puissance. La linéarité de l'opérateur de convolution nous permet donc d'effectuer également des calculs radiométriques de l'image

formée, et de mettre en relation les grandeurs radiométriques de la scène observée et de son image.

Considérons dans un premier temps l'impact d'un système optique sur la luminance énergétique d'un objet. Soit le schéma de la figure 2.28, nous considérons que les dioptries constituant le bloc optique n'induisent pas de réflexions parasites et qu'il n'y a pas d'absorption de l'énergie lumineuse par les éléments optiques du bloc. Dès lors, le flux du faisceau entrant est égal au flux émergent du bloc : il y a conservation du flux radiant.

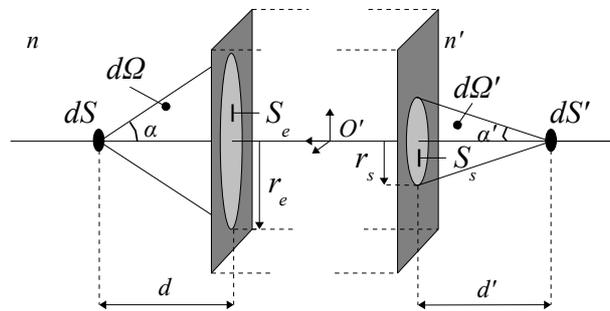


Figure 2.28: Bloc optique et conservation du flux

L'étendue géométrique du faisceau entrant est :

$$d^2 G = dS \frac{S_e}{d^2} = dS d \Omega \quad (2.27)$$

Le flux entrant est :

$$d^2 F = L d^2 G \quad (2.28)$$

où  $L$  est la luminance énergétique de la surface élémentaire objet  $dS$ .

L'étendue géométrique du faisceau sortant est :

$$d^2 G' = dS' \frac{S_s}{d'^2} = dS' d \Omega' \quad (2.29)$$

où  $dS'$  est l'image de  $dS$ .

Le flux sortant est :

$$d^2 F' = L' d^2 G' \quad (2.30)$$

On suppose que le stigmatisme du système est conservé sur les plans perpendiculaires à l'axe optique (c'est la *condition d'aplanétisme*). On a alors l'égalité suivante (relation des sinus d'Abbe) :

$$n d \sin \alpha = n' d' \sin \alpha' \quad (2.31)$$

En considérant que les milieux d'entrée et sortie ont même indice de réfraction on obtient :

$$d \sin \alpha = d' \sin \alpha' \quad (2.32)$$

Détaillons l'expression des angles solides des faisceaux entrant et sortants. On a :

$$d \Omega = \frac{S}{d^2} = \frac{\pi r_e^2}{d^2} = \pi \tan^2 \alpha \quad (2.33)$$

de même

$$d \Omega' = \frac{S'}{d'^2} = \frac{\pi r_s^2}{d'^2} = \pi \tan^2 \alpha' \quad (2.34)$$

Pour  $\alpha$  et  $\alpha'$  petits on peut effectuer les approximations suivantes:  $\tan^2 \alpha \approx \sin^2 \alpha$  et  $\tan^2 \alpha' \approx \sin^2 \alpha'$ .

Donc

$$d \Omega \approx \pi \sin^2 \alpha \quad (2.35a) \text{ et } d \Omega' \approx \pi \sin^2 \alpha' \quad (2.35b)$$

et par conséquent, on peut écrire

$$\frac{\sin^2 \alpha}{\sin^2 \alpha'} = \frac{d \Omega}{d \Omega'} \quad (2.36)$$

Ainsi, d'après (2.32)

$$\frac{d \Omega}{d \Omega'} = \frac{d'^2}{d^2} \quad (2.37)$$

Or le carré du grandissement est égal au rapport des surfaces élémentaires conjuguées :

$$\left(\frac{d'}{d}\right)^2 = \frac{dS'}{dS} \quad (2.38)$$

Donc :

$$\frac{d\Omega}{d\Omega'} = \frac{dS'}{dS} \quad (2.39)$$

ou bien

$$d\Omega dS = d\Omega' dS' \quad (2.40)$$

par conséquent

$$d^2G = d^2G' \quad (2.41)$$

Il y a conservation des étendues géométriques.

Or, il y a également conservation des flux radiants. En se reportant à (2.28) et (2.30) on en déduit qu'il y a également conservation des luminances énergétiques.

$$L = L' \quad (2.42)$$

Dans le cadre de la modélisation d'une caméra par une lentille (paragraphe 2.2.2), étudions le lien entre la luminance énergétique de l'objet observé et l'éclairement énergétique de son image. Soit le schéma de la figure 2.29, le point  $M$  entouré d'un élément de surface  $dS$  a pour image  $M'$  entouré de  $dS'$ .  $M$  émet la luminance  $L$ , l'éclairement en  $M'$  est  $E_{c'}$ .

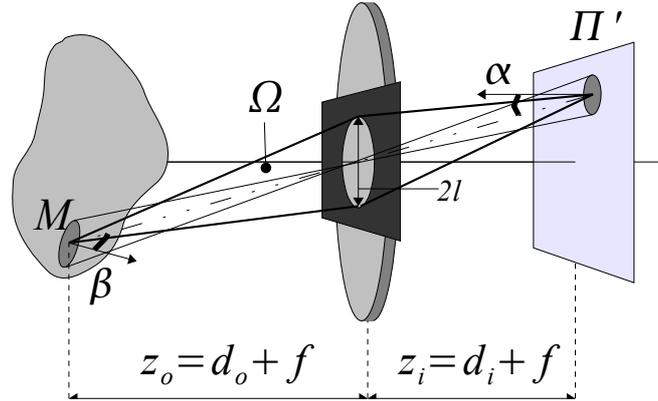


Figure 2.29 : Variation de l'éclairement dans un système optique

En se reportant à la figure 2.29, l'angle solide élémentaire  $d\Omega$  sous lequel le centre optique  $O'$  voit l'élément de surface  $dS$  (ou  $dS'$ ) est égal à :

$$d\Omega = \frac{dS' \cos \alpha}{(z_i / \cos \alpha)^2} = \frac{dS \cos \beta}{(z_o / \cos \alpha)^2} \quad (2.43)$$

ainsi

$$\frac{dS}{dS'} = \frac{\cos \alpha}{\cos \beta} \left( \frac{z_o}{z_i} \right)^2 \quad (2.44).$$

La surface apparente de la lentille délimitée par le diaphragme circulaire de diamètre  $2l_D$  est  $\pi l_D^2$ . L'angle solide  $\Omega$  par lequel le point  $M$  observe cette surface est :

$$\Omega = \pi \frac{l_D^2 \cos \alpha}{(z_o / \cos \alpha)^2} = \pi \left( \frac{l_D}{z_o} \right)^2 \cos^3 \alpha \quad (2.45)$$

Le flux énergétique  $dF$  émis par  $dS$  et capté par le système peut s'exprimer en fonction de la luminance  $L$  :

$$dF = L \Omega dS \cos \beta = \pi \left( \frac{l_D}{z_o} \right)^2 L dS \cos^3 \alpha \cos \beta \quad (2.46)$$

Le système optique n'introduisant ni réflexion parasite, ni absorption, on a conservation du

flux. L'éclairement est par conséquent donné par :

$$Ec' = \frac{dF}{dS'} = \pi \left( \frac{l_D}{z_o} \right)^2 L \frac{dS}{dS'} \cos^3 \alpha \cos \beta \quad (2.47)$$

en utilisant l'équation (2.44) on obtient l'égalité :

$$Ec' = \left[ \pi \left( \frac{l_D}{z_i} \right)^2 \cos^4 \alpha \right] L \quad (2.48)$$

Dans le cas d'un diaphragme carré de largeur  $2l_D$  on obtient un éclairement égal à:

$$Ec' = \left[ 4 \left( \frac{l_D}{z_i} \right)^2 \cos^4 \alpha \right] L \quad (2.49)$$

Ainsi, l'éclairement d'une image formée par un système optique est proportionnelle à la luminance de l'objet observé. Surtout, elle dépend de l'ouverture du diaphragme du système et est inversement proportionnelle à la position du point image. En se reportant à la figure 2.29 on pose  $z_i = d_i + f$ , pour un objet suffisamment éloigné on peut ainsi poser l'approximation  $z_i \approx f$ . Soit la grandeur  $f^{nb} = f/2l_D$ ,  $f^{nb}$  est le *nombre d'ouverture* du système (il s'agit du *f-number* de la littérature anglosaxonne). L'éclairement du plan image est donc inversement proportionnel au nombre d'ouverture du système. Notons que  $f^{nb}$  est un paramètre fondamental des systèmes optiques utilisés en imagerie. Il est, comme nous venons de le voir, un indicateur du niveau d'éclairement d'une acquisition, mais également un indicateur du comportement fréquentiel du système optique (cf paragraphe 2.2.3). Dans la suite de ce mémoire, nous adopterons la notation indiquée  $f_i^{nb}$  pour la valeur effective du nombre d'ouverture (pour  $f^{nb} = 2$  nous noterons  $f_2^{nb}$ ).

Enfin,  $Ec'$  est proportionnelle au facteur  $\cos^4 \alpha$  : l'éclairement chute donc au fur et à mesure que le point objet s'écarte de l'axe optique. Dans la littérature cette variation est connue sous l'appellation "loi du cosinus 4". En imagerie cette loi se traduit par un obscurcissement aux angles des photographies (figure 2.30) qui ne doit pas être confondu avec le *vignettage*. Ce dernier produit un obscurcissement similaire mais dû à l'obstruction

d'une partie des rayons lumineux par les tailles différentes des éléments optiques de l'objectif.

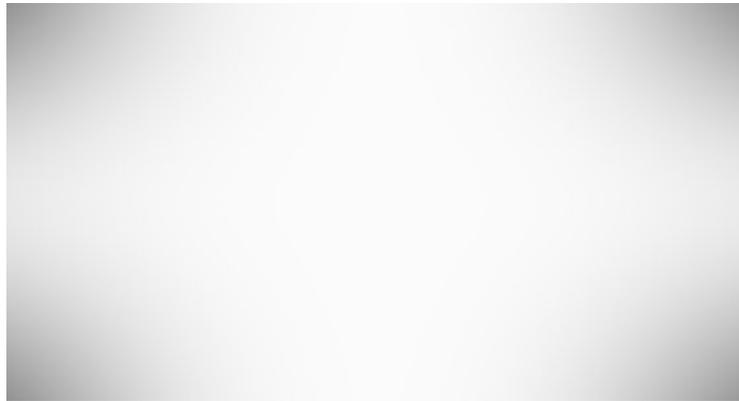


Figure 2.30 : Effets de la loi du cosinus 4 sur une image blanche

Finalement, concernant l'impact fréquentiel du système optique, celui-ci est comme nous l'avons rappelé en début de paragraphe linéaire en puissance. Par conséquent, l'éclairement du plan image tenant compte de l'impact de sa fonction de transfert optique est donné par :

$$Ec''(\mathbf{M}') = RIO(\mathbf{M}') ** Ec'(\mathbf{M}') \text{ dans le plan spatial (2.50)}$$

$$TF[Ec''(\mathbf{M}')] = FTO(\mathbf{M}') \cdot TF[Ec'(\mathbf{M}')] \text{ dans le plan fréquentiel (2.51)}$$

## 2.4 Acquisition d'image

Afin de permettre la visualisation et le traitement de l'image formée, une zone photosensible est placée en aval du bloc optique. Cette partie du système assurant l'acquisition de la scène observée est basée principalement sur deux approches :

- la première que nous qualifions d'analogique, où l'image de la scène observée est projetée sur un film recouvert d'une émulsion : le film argentique. Le champ lumineux vient modifier localement la structure moléculaire du support y laissant une empreinte pérenne.
- la seconde, dite numérique, où l'image de la scène observée vient recouvrir un matériau semi-conducteur : le capteur (ou détecteur) d'image. Le flux énergétique y génère localement un courant électrique qui est exploité afin d'en déduire une grandeur

physique représentative de l'énergie reçue.

Dans cette étude, nous nous plaçons dans le deuxième cas. Le processus permettant la génération d'un courant à partir d'un champ lumineux est l'effet photoélectrique découvert par H.R. Hertz en fin du XIX<sup>ème</sup> siècle [65]. Les détecteurs employés en imagerie de nos jours sont fondamentalement basés sur le même principe : la zone photosensible est une mosaïque dont chaque élément (le pixel) délivre un signal proportionnel à l'énergie lumineuse frappant sa surface. Cette conversion opérée par chaque pixel est une intégration temporelle : le détecteur est donc sensible au flux énergétique. On caractérise ainsi ses performances par sa sensibilité  $\kappa$  de sorte que le signal délivré  $s$  et le flux énergétique  $F$  sont liés par l'égalité  $s = \kappa F$ . La sensibilité est exprimé en ampères par watt (A/W) ou en volts par watt (V/W) suivant la nature du signal.

Par l'intégration spatiale qu'opèrent les pixels, le flux peut être directement déduit de l'éclairement à la surface de la mosaïque. En effet dans le cas de pixels carrés, on modélise le processus d'intégration par un filtre linéaire dont la réponse impulsionnelle  $RIC$  est :

$$RIC(x', y', z') = \frac{1}{\rho^2} \cdot 1_{\left[\frac{-\rho}{2}, \frac{\rho}{2}\right] \times \left[\frac{-\rho}{2}, \frac{\rho}{2}\right]}(x', y') \cdot \delta(z' - z_c) \quad (2.52)$$

où  $\rho$  est la largeur d'un pixel,  $1_{\left[\frac{-\rho}{2}, \frac{\rho}{2}\right] \times \left[\frac{-\rho}{2}, \frac{\rho}{2}\right]}(x', y')$  est la fonction unitaire sur l'étendue d'un pixel,  $\delta(z' - z_c)$  est l'impulsion de Dirac et  $z_c$  est la position du centre du plan capteur sur l'axe optique, ce dernier étant colinéaire à sa normale.

Ainsi le signal  $s_{i,j}$  délivré par le pixel  $p_{i,j}$  situé en ligne  $i$  et colonne  $j$  de la mosaïque est proportionnel au résultat de l'opération de convolution :

$$s_{i,j} \propto [RIC(x', y', z') ** Ec'(M')]_{M'=M'_{i,j}} \quad (2.53)$$

où  $M'_{i,j}$  est le centre du pixel  $p_{i,j}$  (figure 2.31).

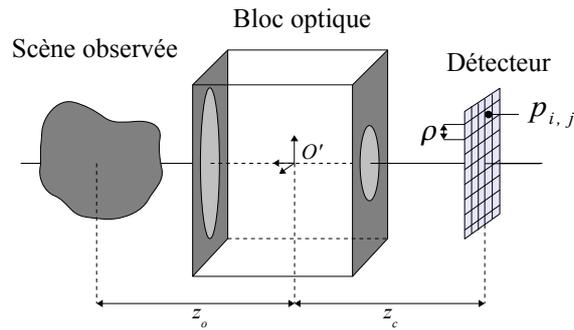


Figure 2.31: Disposition des éléments du système optique

Pour la suite de cette étude, les  $s_{i,j}$  sont considérés comme étant les niveaux de gris élémentaires délivrés à la sortie du système de transmission.

## 2.5 Conclusion

En préambule à la construction du modèle d'analyse du système de transmission, nous avons introduit dans ce chapitre les différentes notions d'optique et de vision par ordinateur nécessaires. Leur présentation ainsi que l'enchaînement de ces éléments ont été orientés pour cette étude et en constituent désormais le socle sur lequel nous allons nous appuyer. Ainsi, le chapitre suivant est consacré à l'étude du système de transmission 1D, c'est à dire lorsque l'émetteur est un code-barres et le récepteur une caméra linéaire. Le comportement du flux irradié par le code dans l'espace et à travers la caméra y est modélisé, nous amenant à décrire le canal de communication par une équation liant les symboles reçus à ceux émis ainsi qu'aux paramètres du système. Les propriétés de cette écriture multiparamétrée du canal nous permettent l'établissement d'une stratégie pour l'évaluation des performances globales.

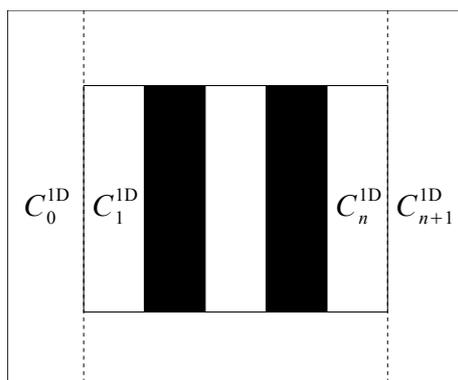


## **Chapitre 3**

### **Modélisation et analyse des transmissions par code-barres**

La majorité des produits et biens de consommation est identifiée au niveau des chaînes de production et de distribution par un code-barres. Cette utilisation standardisée d'un protocole de communication graphique entre les différents acteurs de l'industrie est à la base de leur logistique, ceci depuis les années soixante [66]. Dans ce cadre, une symbologie émerge de l'ensemble des codes-barres et symboles graphiques développés à ce jour : l'EAN13. Ce code est surtout représentatif d'une des deux grandes familles de codes à barres, les  $\delta$ -codes [12], la rendant de fait majoritaire. Elle se caractérise par  $n$  barres d'égale largeur, noires ou blanches, entourées d'une zone blanche constituant une marge ou "zone neutre".

Nous notons  $C_i^{1D}$ ,  $i \in \{1, 2, \dots, n\}$ , les barres du code et distinguons les marges situées de part et d'autre du code en les notant  $C_0^{1D}$  pour la partie gauche et  $C_{n+1}^{1D}$  pour la partie droite (figure 3.1).



*Figure 3.1 : Nomenclature du code-barres*

Cette séquence  $C_{0:n+1}^{1D}$  est l'équivalent graphique d'une séquence binaire notée  $X_{0:n+1}$ , la séquence d'entrée du système de transmission, et l'on note  $C^{1D}(X_i) = C_i^{1D}$ . En nous basant sur les principes optiques et radiométriques (chapitre 2), nous développons dans ce qui suit un modèle traduisant l'évolution de  $X_{0:n+1}$  au cours de l'acquisition du code-barres par un lecteur 1D. Cette modélisation nous permet de lier formellement l'entrée et la sortie du système en vue de son évaluation.

### 3.1 Modélisation du canal de transmission

Le système de transmission est constitué d'un code-barres  $C^{1D}$ , dont le support est plan et noté  $\Pi$  (figure 3.2). Le centre de  $C^{1D}$  est le centre  $O$  du repère objet  $R(O, \vec{x}, \vec{y}, \vec{z})$ . Le code est présenté devant une caméra 1D munie d'un diaphragme de fonction de forme  $D$  et de longueur focale  $f$ . Le bloc optique de la caméra est modélisé par une lentille mince (cf paragraphe 2.2.1) dont le centre optique est le centre  $O'$  du repère image  $R'(O', \vec{x}', \vec{y}', \vec{z}')$  et l'axe optique est l'axe  $[O'\vec{z}')$ . Le capteur linéaire est situé sur un plan  $\Pi'$  perpendiculaire à l'axe optique et coupant celui-ci en  $[0,0, -z_c]^T$  dans le repère image.

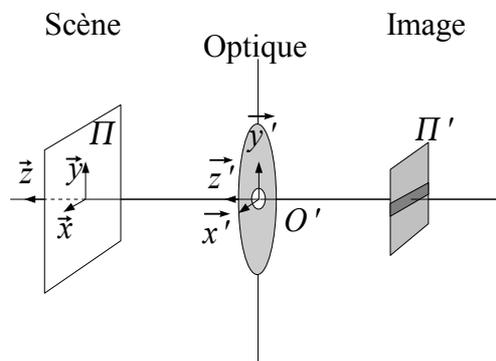


Figure 3.2 : Schéma du système 1D

#### 3.1.1 Formation d'image

L'image formée par le bloc optique de la caméra suit l'équation (2.17) du modèle proposé au paragraphe 2.2.1. Tout couple conjugué {objet; image} est lié à la distance focale du bloc de sorte que la projection bidimensionnelle de la scène 3D observée est géométriquement déformée (figure 3.3). Cette déformation, dont l'amplitude varie suivant la configuration, est une projection perspective [67].

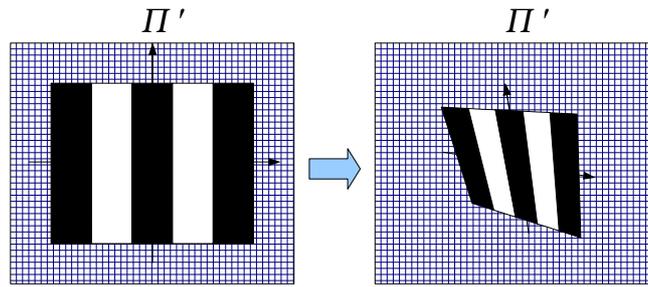


Figure 3.3 : Effets de la projection perspective

Le processus de formation d'image est ainsi une application de  $\mathbb{R}^3$  vers  $\mathbb{R}^2$  qui fait correspondre au point  $M_1=[x_1,y_1,z_1]^T$  de l'espace objet le point  $M'_1=[x'_1,y'_1,z'_1]^T$  du plan image. On note  $Proj()$  cette application et l'on a :

$$M'_1=Proj(M_1) \quad (3.1)$$

Un objet s'éloignant du centre optique occupe une surface de plus en plus petite sur le plan de projection. Dans notre système, cette distorsion géométrique implique que les barres initialement d'égale largeur ont leurs conjuguées de largeur variable. Cette déformation est d'autant plus importante qu'aucune contrainte n'est imposée quant à la position du plan objet par rapport à la caméra si ce n'est le fait que le code doit être capturé dans sa globalité. Dès lors, ce libre positionnement est formalisé par la définition de trois translations  $\{\vec{T}_x; \vec{T}_y; \vec{T}_z\}$  et trois rotations d'angle  $\{\alpha; \beta; \gamma\}$  (cf paragraphe 2.2.2). Par convention (figure 3.4) , nous lions les repères  $R$  et  $R'$  de sorte qu'en l'absence de translation et rotation leur centre soit confondus et les axes  $[O\vec{z})$  et  $[O\vec{z}')$  soient opposés:

$$O=O' \text{ et } \{\vec{x}=-\vec{x}'; \vec{y}=\vec{y}'; \vec{z}=-\vec{z}'\}$$

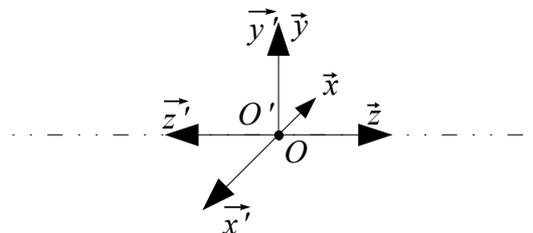


Figure 3.4 : Positionnement par défaut des repères

Sachant la position du point objet  $M_1$  , nous déterminons la position de son conjugué  $M'_1$

en utilisant la notation en coordonnées homogènes (cf paragraphe 2.2.2.1). Ainsi, les coordonnées cartésiennes du point image sont tirées de l'équation :

$$\mathbf{M}'_1^{(h)} = \mathbf{P}^{(h)} \cdot \mathbf{R}_y^{(h)} \cdot \mathbf{R}_\beta^{(h)} \cdot \mathbf{R}_\alpha^{(h)} \cdot \mathbf{T}^{(h)} \cdot \mathbf{M}_1^{(h)} \quad (3.2)$$

où les matrices de transformation sont détaillées au paragraphe 2.2.2.1, et  $R'$  est le repère de référence.

### 3.1.2 Analyse radiométrique

Le code  $C^{1D}$  est une source de rayonnement (primaire ou secondaire). Elle émet en tout point  $\mathbf{M} \in C^{1D}$  une luminance  $L(\mathbf{M})$  supposée constante quelle que soit l'orientation (source lambertienne). Nous considérons que le bloc optique de la caméra 1D n'est sujet ni à absorptions, ni à réflexions parasites. L'éclairement des points images du code est donc donné par l'équation (2.48) :

$$Ec'(\mathbf{M}') = \left[ \frac{S_D}{z'^2} \cos^4 \alpha_M \right] L(\mathbf{M}) \quad (3.3)$$

où  $S_D$  est la surface d'ouverture du diaphragme,  $\mathbf{M}' = Proj(\mathbf{M})$  et  $z'$  est la position en valeur absolue du point image sur l'axe optique.

### 3.1.3 Flou

Tout bloc optique utilisé en éclairage incohérent peut être modélisé par un système linéaire en puissance, caractérisé dans l'espace des fréquences spatiales par sa fonction de transfert optique  $FTO(f_{x'}, f_{y'})$  et par sa réponse impulsionnelle optique  $RIO(x', y')$  dans l'espace. Cette fonction de transfert rend compte du phénomène de diffraction et de l'éventuel flou dû à un positionnement hors du plan de mise au point de l'objet observé (paragraphe 2.2.3).

Dans notre étude, la caméra a une distance focale fixe et la position du capteur ne change pas par rapport au centre optique. L'éclairement de l'image floue  $Ec''(\mathbf{M}')$  sur le plan capteur

est ainsi donné par l'équation de convolution (2.50) :

$$Ec''(\mathbf{M}') = RIO(\mathbf{M}') ** Ec'(\mathbf{M}')$$

### 3.1.4 Intégration

Le flux radiant projeté par le bloc optique est numérisé par le détecteur linéaire placé à la distance  $z_c$  du centre optique. Il est constitué d'une ligne de  $N_p$  pixels intégrant l'énergie lumineuse incidente sur leur surface. Dans le cas de pixels carrés, nous modélisons cette intégration par un filtrage à moyenne glissante (cf section 2.4) dont la réponse impulsionnelle est donnée par l'équation (2.52) que nous rappelons ici :

$$RIC(x', y', z') = \frac{1}{\rho^2} \cdot 1_{\left[\frac{-\rho}{2}, \frac{\rho}{2}\right] \times \left[\frac{-\rho}{2}, \frac{\rho}{2}\right]}(x', y') \cdot \delta(z' - z_c)$$

Le signal délivré par le  $j^{\text{ème}}$  pixel  $p_{1,j}$  du détecteur est obtenu à partir de l'opération de convolution:

$$s_{1,j} \propto \left[ RIC_{p_{1,j}} ** Ec''(\mathbf{M}') \right]_{\mathbf{M}' \in p_{1,j}} \quad (3.4)$$

où  $RIC(x', y', z')$  associé à  $p_{1,j}$  est noté  $RIC_{p_{1,j}}$  pour alléger l'écriture.

### 3.1.5 Equation du canal

#### 3.1.5.1 Construction du canal par symbole

La séquence des  $s_{1,1:N_p}$  est considérée comme l'information utile récupérée à l'issue de la transmission des  $X_{0:n+1}$ . Les éléments de cette séquence de sortie du système sont obtenus à partir des convolutions successives avec les réponses impulsionnelles du bloc optique et des pixels du capteur :

$$s_{1,j} = \left[ RIC_{p_{1,j}} ** RIO(\mathbf{M}') ** Ec'(\mathbf{M}') \right]_{\mathbf{M}' \in p_{1,j}} \quad (3.5)$$

Les barres du code sont considérées comme des émetteurs lambertiens: la luminance

énergétique de chacune des barres du code  $L(\mathbf{M})_{\mathbf{M} \in C_i^{1D}}$  est constante et l'on pose  $L(\mathbf{M})_{\mathbf{M} \in C_i^{1D}} = X_i$  avec  $X_i = \{0; 1\}$ . Dès lors, en se basant sur (2.48) et (3.5), l'expression du niveau de gris par pixel  $s_{1,j}$  peut être développée en segmentant l'opération de convolution afin de faire ressortir la participation de chacune des barres  $C_i^{1D}$  et leur image  $C'_i{}^{1D}$ . En posant  $\mathbf{M}' = Proj(\mathbf{M})$ , on obtient la somme discrète suivante :

$$s_{1,j} = \sum_{k=0}^{n+1} \left( \kappa \cdot \frac{S_D}{z_c^2} \cdot \left[ \cos^4 \alpha_M \cdot 1_{C_k^{1D}}(\mathbf{M}') ** RIC_{p_{1,j}} ** RIO(\mathbf{M}') \right]_{\mathbf{M}' \in p_{1,j}} \cdot X_k \right) \quad (3.6)$$

où  $\kappa$  est un facteur d'échelle lié aux caractéristiques du détecteur (sensibilité, temps d'intégration) et  $S_D$  est la surface du diaphragme du bloc optique.

A l'issu du processus d'intégration par les pixels du détecteur, on modélise les perturbations parasites inhérentes aux imperfections des composants du système par un bruit. Nous considérons ce bruit comme additif, blanc et gaussien (ABG) et de variance  $\sigma_p^2$ .

A l'issu de la transmission des  $X_{0:n+1}$ , on considère que chaque symbole reçu est la moyenne de l'information utile des pixels englobant chacune des barres projetées. L'ensemble forme la séquence de sortie du système  $Y_{1:n}$ . La moyenne des niveaux de gris  $\bar{s}_{1,k}$  délivrés par les pixels englobant le projeté de la  $k^{\text{ème}}$  barre est donnée par :

$$\bar{s}_{1,k} = \sum_{j=1}^{N_p} g_{k,j} \cdot s_{1,j} \quad (3.7)$$

où  $g_{k,j}$  est l'aire normalisée de recouvrement du  $j^{\text{ème}}$  pixel  $p_{1,j}$  par la  $k^{\text{ème}}$  barre image  $C'_k{}^{1D}$ .

Les  $g_{k,j}$  vérifient les relations :

- $0 \leq g_{k,j} \leq 1$
- $\sum_{j=1}^{N_p} g_{k,j} = 1$

Concrètement le calcul des  $g_{k,j}$  s'effectuent tel qu'illustré sur l'exemple suivant (figure 3.5):

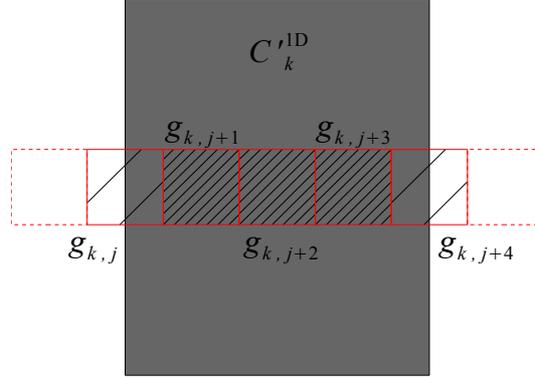


Figure 3.5 : Calcul des  $g_{k,j}$

avec  $g_{k,\{j,j+4\}}=0.5/5$  et  $g_{k,\{j+1,j+2,j+3\}}=1/5$ .

Développons (3.7) :

$$\begin{aligned}
\bar{s}_{1,k} &= \sum_{j=1}^{N_p} g_{k,j} \cdot s_{1,j} \\
&= \sum_{j=1}^{N_p} \left[ g_{k,j} \cdot \left[ \sum_{k_0=0}^{n+1} \left( \kappa \cdot \frac{S_D}{z_c^2} \cdot \left[ \cos^4 \alpha_M \cdot 1_{C_{k_0}^{1D}}(\mathbf{M}') ** RIC_{p_{1,j}} ** RIO(\mathbf{M}') \right]_{\mathbf{M}' \in p_{1,j}} \cdot X_{k_0} \right) \right] \right] \\
&= \sum_{k_0=0}^{n+1} \left[ \left[ \sum_{j=1}^{N_p} \left( \kappa \cdot \frac{S_D}{z_c^2} \cdot \left[ \cos^4 \alpha_M \cdot 1_{C_{k_0}^{1D}}(\mathbf{M}') ** RIC_{p_{1,j}} ** RIO(\mathbf{M}') \right]_{\mathbf{M}' \in p_{1,j}} \cdot g_{k,j} \right) \right] \cdot X_{k_0} \right] \\
\bar{s}_{1,k} &= \sum_{k_0=0}^{n+1} h_{k,k_0} \cdot X_{k_0} \quad (3.8)
\end{aligned}$$

Les symboles reçus étant une moyenne sur les pixels intégrant, le bruit  $B_k$  associé à chaque symbole est ABG et de variance  $\sigma^2$  égale au rapport entre  $\sigma_p^2$  et le nombre de pixels intégrant chaque barre.

L'équation du canal liant le symbole reçu  $Y_{1,k}$  aux  $\{X_{k_0}\}$  symboles émis se déduit ainsi de (3.8) et l'on a :

$$Y_{1,k} = \sum_{k_0=0}^{n+1} h_{k,k_0} \cdot X_{k_0} + B_k \quad (3.9)$$

## 3.2 Mesure des performances du système

### 3.2.1 Théorie de l'information – Information Mutuelle Moyenne (IMM)

L'utilisation de pictogrammes conjointement à un moyen de les lire tel qu'une caméra est une instanciation singulière d'un système de communication numérique. Une séquence binaire  $X_{0:n+1}$  est modulée graphiquement et transmise à travers un canal de communication que nous avons modélisé par un canal à mémoire et à bruit ABG. La récupération de l'information binaire transmise s'effectue à partir de la séquence de niveaux de gris reçus  $Y_{1:n}$ . Dans le domaine des communications numériques, l'étude des systèmes de communication se base sur le socle de la Théorie de l'Information, initiée et développée par C.E. Shannon [68]. C'est dans ce cadre théorique que nous plaçons notre étude.

D'une manière générale, soient les séquences  $X_{1:n}=\{X_1, \dots, X_n\}$  et  $Y_{1:m}=\{Y_1, \dots, Y_m\}$  ( $m$  et  $n$  non nécessairement égaux) deux ensembles de variables aléatoires. Chaque variable aléatoire  $X_k$  et  $Y_l$ ,  $k \in [1, n]$  et  $l \in [1, m]$ , prend respectivement ses valeurs dans un alphabet fini  $\Omega$  et dans l'ensemble des nombres réels  $\mathbb{R}$ . L'entropie d'une variable aléatoire discrète  $A$  est par définition égale à :

$$H(A) = - \sum_{\{a_k\}} P(A=a_k) \log P(A=a_k) \quad (3.10)$$

avec  $P(A=a_k)$  la probabilité que  $A=a_k$  et où les  $\{a_k\}$  sont l'ensemble des réalisations possibles de la variable aléatoire  $A$ .

L'entropie mesure une incertitude moyenne, c'est à dire l'information apportée par un événement. En effet, si une occurrence apparaît avec certitude, cette réalisation n'amène aucune information au récepteur. A l'opposé, une réalisation à priori peu probable sera plus informative.

Dans le cas où le logarithme est en base 2, l'entropie s'exprime en bits. Dans ce cas, une variable aléatoire  $X$  binaire et équiprobable a une entropie égale à 1 bit.

L'entropie d'une séquence aléatoire  $X_{1:n}$  est par définition égale à :

$$H(X_{1:n}) = - \sum_{\{x_{1:n}\}} P_{X_{1:n}}(x_{1:n}) \log_2 P_{X_{1:n}}(x_{1:n}) \quad (3.11)$$

De plus, on définit l'entropie de  $Y$  conditionnellement à  $X$  par :

$$H(Y|X) = - \sum_{\{y_j\}} \sum_{\{x_i\}} P_{X_i, Y_j}(x_i, y_j) \log_2 P_{Y_j|X_i}(y_j|x_i) \quad (3.12)$$

Dans le cas de séquences aléatoires, l'entropie conditionnelle est donnée par :

$$H(Y_{1:m}|X_{1:n}) = - \sum_{\{y_{1:m}\}} \sum_{\{x_{1:n}\}} P_{X_{1:n}, Y_{1:m}}(x_{1:n}, y_{1:m}) \log_2 P_{Y_{1:m}|X_{1:n}}(y_{1:m}|x_{1:n}) \quad (3.13)$$

Losque l'on traite des variables aléatoires continues, l'entropie est dite *différentielle* et l'on a :

$$H(Z) = - \int_{-\infty}^{\infty} p(z) \log_2 [p(z)] dz \quad (3.14)$$

où  $Z$  est une variable aléatoire continue de densité de probabilité  $p(z)$ .

En se basant sur ce concept d'entropie, on définit l'information mutuelle moyenne (IMM) de deux séquences aléatoires  $X_{1:n}$  et  $Y_{1:m}$  par :

$$IMM(X_{1:n}; Y_{1:m}) = H(Y_{1:m}) - H(Y_{1:m}|X_{1:n}) \quad (3.15)$$

Mathématiquement parlant, l'IMM mesure une dépendance statistique entre les deux séquences. En effet, la différence d'entropies revient à retrancher à la quantité d'information portée par  $Y_{1:m}$  l'information apportée par cette même séquence sachant  $X_{1:n}$ . Si le processus  $X$  n'apporte pas d'information quant au processus  $Y$ , l'IMM est nulle. Le concept de l'entropie étant basé sur la notion d'incertitude, ce cas revient à dire que  $X_{1:n}$  et  $Y_{1:m}$  sont, d'un point de vue probabiliste, indépendants.

Nous avons proposé au paragraphe 3.1.5 une modélisation du système de transmission par code-barres en formulant l'équation du canal de transmission. Désormais, les processus  $X$  et  $Y$  sont respectivement le message binaire transmis par le code et les niveaux de gris reçus en sortie du système (le message binaire est également une séquence n'ayant que deux niveaux de gris extrêmes). L'IMM constitue donc un indicateur quantitatif de l'impact du canal sur la transmission. C'est une mesure de flux d'information, dont le calcul revient à l'évaluation des performances du système par rapport à la modélisation proposée. Nous

qualifions cette évaluation d'objective, car elle s'inscrit dans le cadre rigoureux de la théorie de l'information et permet d'aller au delà d'une appréciation qualitative des performances.

### 3.2.2 Calcul de l'IMM

En se basant sur (3.15), L'IMM s'obtient par le calcul de l'entropie  $H(Y_{1:n})$  et de l'entropie conditionnelle  $H(Y_{1:n}|X_{0:n+1})$ .

L'équation du canal est pour rappel :

$$Y_{1,k} = \sum_{k_0=0}^{n+1} h_{k,k_0} \cdot X_{k_0} + B_k \quad (3.16)$$

#### 3.2.2.1 Entropie conditionnelle

En se référant à l'équation (3.16), le calcul de  $H(Y_{1:n}|X_{0:n+1})$  est totalement déterminé par les caractéristiques du bruit. Or l'entropie différentielle d'un processus suivant une loi normale de variance  $\sigma^2$  est égale à  $\frac{1}{2} \log_2(2\pi e \sigma^2)$  ([13] p225, [3] p107). On a alors directement le résultat :

$$H(Y_{1:n}|X_{0:n+1}) = \frac{n}{2} \log_2(2\pi e \sigma^2) \quad (3.17)$$

#### 3.2.2.2 Entropie $H(Y_{1:n})$

Le calcul de l'entropie de la sortie du système est quant à lui moins trivial. Notre approche est basée sur la simulation du canal et l'utilisation de ses propriétés. Ces caractéristiques singulières du canal sont mises en évidence en réécrivant l'équation du canal sous forme matricielle.

Soient les vecteurs  $\mathbf{X} = [X_0, \dots, X_{n+1}]^T$  et  $\mathbf{Y} = [Y_1, \dots, Y_n]^T$ , séquences d'entrée et sortie du système. L'équation complète du canal, c'est à dire pour l'ensemble des symboles émis et reçus, peut s'écrire :

$$\mathbf{Y} = \mathbf{A} \cdot \mathbf{X} + \mathbf{B} \quad (3.18)$$

avec  $\mathbf{B}=[B_1, \dots, B_n]^T$  le vecteur bruit et  $\mathbf{A}$  est une matrice rectangulaire de taille  $n \times (n+2)$  et dont les éléments sont tirés de (3.8) avec  $A_{k, k_0} = h_{k, k_0}$ .

On montre ainsi que le canal est markovien, de mémoire finie  $L$ , définie par le couple d'équations:

$$\begin{aligned} Y_k &= \mathbf{V}_k^T \mathbf{F}_k + B_k \\ \mathbf{F}_k &= \mathbf{D} \mathbf{F}_{k-1} + \mathbf{E}_k \end{aligned} \quad (3.19)$$

- $L$  longueur maximale telle que, pour  $|k_0 - k| > L$  on a  $A_{k, k_0} \rightarrow 0$  avec  $0 \leq k \leq n+1$ , et décomposable en  $L_k^g$  et  $L_k^d$  tels que pour  $0 \leq k \leq n+1$ ,  $L = L_k^g + L_k^d + 1$
- $\mathbf{V}_k = [A_{k, \min(k+L_k^d, n+1)}, \dots, A_{k, \max(k-L_k^g, 0)}]^T$
- $\mathbf{E}_k = [X_{\min(k+L, n+1)}, 0, \dots, 0]^T$
- $\mathbf{F}_k = [X_{\min(k+L_k^d, n+1)}, \dots, X_k, \dots, X_{\max(k-L_k^g, 0)}]^T$
- $\mathbf{D} = \begin{bmatrix} 0 & \dots & 0 \\ 1 & \ddots & \\ 0 & \ddots & \\ \vdots & \ddots & \\ 0 & \dots & 0 & 1 & 0 \end{bmatrix}$

et avec l'initialisation  $\mathbf{F}_{-1} = [X_{L-1}, \dots, X_0, 0, \dots, 0]^T$ .

L'estimation de l'entropie  $H(Y_{1:n})$  est le calcul d'une espérance mathématique (paragraphe 3.2.1). Afin de contourner la difficulté du calcul direct, nous procédons à une estimation de cette espérance par la méthode de Monte Carlo [69]. Par rapport à la modélisation proposée du canal de transmission, cette méthode empirique revient à évaluer la probabilité d'apparition de  $N_s$  séquences de sortie simulées :  $\{p(Y_{1:n} = y_{1:n}^{(s)}), s \in \{1, \dots, N_s\}\}$ . Les réalisations de  $Y_{1:n}$  sont tirées de l'équation du canal (3.18).

L'estimée de l'entropie est alors donnée par l'égalité :

$$\hat{H}(Y_{1:n}) = \frac{-1}{N_s} \sum_{s=1}^{N_s} \log_2 [p(Y_{1:n} = y_{1:n}^{(s)})] \quad (3.20)$$

Nous avons ainsi déplacé le problème du calcul de l'entropie vers un problème de calcul de

probabilité d'apparition de  $N_s$  séquences simulées  $Y_{1:n} = y_{1:n}^{(s)}$ .

Cette nouvelle problématique est traitée en utilisant la propriété du modèle de canal mise en évidence avec l'écriture matricielle : le caractère markovien de sa mémoire  $\mathbf{F}_{-1:n+1}$ . Cette singularité nous permet de proposer une solution récursive au calcul de  $p(Y_{1:n} = y_{1:n}^{(s)})$  basée sur les travaux de Bahl et al. [70] et Baum et al. [71].

En se reportant à l'écriture du canal sous la forme du couple d'équation (3.19), nous avons l'égalité :

$$Y_k = \mathbf{V}_k^T \mathbf{F}_k + B_k$$

Soit  $Z_k = \mathbf{V}_k^T \mathbf{F}_k$ , on a alors trivialement :  $Y_k = Z_k + B_k$ . A la position  $k$ , le système est à l'état  $\mathbf{F}_k$ . Une manière élégante de représenter les états du système et son évolution est un graphe en treillis [70, 72]. Dans le cas d'une entrée binaire et d'une mémoire de taille  $L=1$ , le graphe peut être représenté comme en figure 3.6 :

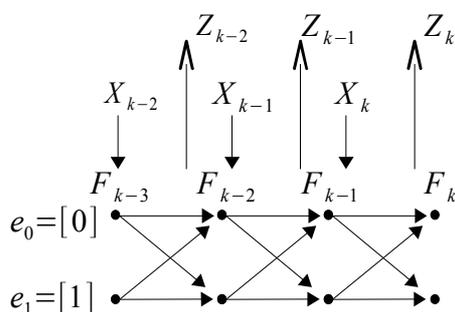


Figure 3.6 : Treillis d'un système à entrée binaire et de mémoire  $L=1$

où les  $\mathbf{e} = \{e_0; e_1\}$  sont les états que peut prendre la mémoire du système.

Soit la métrique associée aux noeuds du treillis  $\alpha'_k(\mathbf{e}) = P(\mathbf{F}_k = \mathbf{e} | Y_{1:k})$  avec  $k \leq n$ .

Développons  $\alpha'_k(\mathbf{e})$  :

$$\alpha'_k(\mathbf{e}) = P(\mathbf{F}_k = \mathbf{e} | Y_{1:k}) = \frac{P(\mathbf{F}_k = \mathbf{e}, Y_{1:k})}{P(Y_{1:k})}$$

Nous introduisons une marginalisation sur les états à l'étape  $k-1$  :

$$\alpha'_k(\mathbf{e}) = \frac{\sum_{\{\mathbf{e}'\}} P(\mathbf{F}_k = \mathbf{e}, \mathbf{F}_{k-1} = \mathbf{e}', Y_{1:k-1}, Y_k)}{P(Y_{1:k})}$$

que l'on décompose en :

$$\alpha'_k(\mathbf{e}) = \frac{\sum_{\{\mathbf{e}'\}} P(\mathbf{F}_{k-1} = \mathbf{e}', Y_{1:k-1}) P(\mathbf{F}_k = \mathbf{e}, Y_k | \mathbf{F}_{k-1} = \mathbf{e}', Y_{1:k-1})}{P(Y_{1:k})}$$

sachant  $\mathbf{F}_{k-1}, Y_{1:k-1}$  n'amène pas d'information sur  $Y_k$  et  $\mathbf{F}_k$  de sorte que :

$$\alpha'_k(\mathbf{e}) = \frac{\sum_{\{\mathbf{e}'\}} P(\mathbf{F}_{k-1} = \mathbf{e}' | Y_{1:k-1}) P(Y_{1:k-1}) P(\mathbf{F}_k = \mathbf{e}, Y_k | \mathbf{F}_{k-1} = \mathbf{e}')}{P(Y_k | Y_{1:k-1}) P(Y_{1:k-1})}$$

or  $P(\mathbf{F}_{k-1} = \mathbf{e}' | Y_{1:k-1}) = \alpha'_{k-1}(\mathbf{e}')$ , d'où la récurrence

$$\alpha'_k(\mathbf{e}) = \frac{\sum_{\{\mathbf{e}'\}} \alpha'_{k-1}(\mathbf{e}') P(\mathbf{F}_k = \mathbf{e}, Y_k | \mathbf{F}_{k-1} = \mathbf{e}')}{P(Y_k | Y_{1:k-1})}$$

Finalement, on a :

$$\alpha'_k(\mathbf{e}) = \lambda_k \sum_{\{\mathbf{e}'\}} \alpha'_{k-1}(\mathbf{e}') \gamma_k(\mathbf{e}', \mathbf{e}) \quad (3.21)$$

avec

$$\lambda_k = \frac{1}{P(Y_k | Y_{1:k-1})} \quad (3.22)$$

La métrique  $\alpha'_k(\mathbf{e})$  peut donc être calculée par récurrence. De plus étant homogène à une probabilité, la sommation à l'étape  $k$  des  $\alpha'_k(\mathbf{e})$  sur l'ensemble des sommets du treillis, soit

l'ensemble des états  $\{\mathbf{e}\}$  possibles de  $\mathbf{F}_k$ , est égale à 1 :

$$\begin{aligned} \sum_{\{\mathbf{e}\}} \alpha'_k(\mathbf{e}) &= \sum_{\{\mathbf{e}\}} \lambda_k \sum_{\{\mathbf{e}'\}} \alpha'_{k-1}(\mathbf{e}') \gamma_k(\mathbf{e}', \mathbf{e}) \\ &= \lambda_k \sum_{\{\mathbf{e}\}} \sum_{\{\mathbf{e}'\}} \alpha'_{k-1}(\mathbf{e}') \gamma_k(\mathbf{e}', \mathbf{e}) = 1 \end{aligned}$$

Donc  $\lambda_k$  est une constante de normalisation par étape du treillis définie par:

$$\lambda_k = \frac{1}{\sum_{\mathbf{e}} \sum_{\mathbf{e}'} \alpha'_{k-1}(\mathbf{e}') \gamma_k(\mathbf{e}', \mathbf{e})} \quad (3.23)$$

La métrique  $\gamma_k$  est associée aux transitions du treillis. On a :

$$\begin{aligned} \gamma_k(\mathbf{e}', \mathbf{e}) &= P(\mathbf{F}_k = \mathbf{e}, Y_k | \mathbf{F}_{k-1} = \mathbf{e}') \\ \gamma_k(\mathbf{e}', \mathbf{e}) &= P(Y_k | \mathbf{F}_{k-1} = \mathbf{e}', \mathbf{F}_k = \mathbf{e}) P(\mathbf{F}_k = \mathbf{e} | \mathbf{F}_{k-1} = \mathbf{e}') \end{aligned} \quad (3.24)$$

Or la connaissance de  $\mathbf{F}_{k-1}$  dans  $P(Y_k | \mathbf{F}_{k-1} = \mathbf{e}', \mathbf{F}_k = \mathbf{e})$  n'amène pas d'information sur  $Y_k$  sachant  $\mathbf{F}_k$ . On a donc finalement :

$$\gamma_k(\mathbf{e}', \mathbf{e}) = P(Y_k | \mathbf{F}_k = \mathbf{e}) P(\mathbf{F}_k = \mathbf{e} | \mathbf{F}_{k-1} = \mathbf{e}') \quad (3.25)$$

La probabilité  $P(Y_k | \mathbf{F}_k)$  est totalement déterminée par les caractéristiques du bruit du système. Pour un canal à bruit ABG de puissance  $N_0 = 2\sigma^2$  on a :

$$P(Y_k | \mathbf{F}_k) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y_k - z_k)^2}{2\sigma^2}} \quad (3.26)$$

La probabilité  $P(\mathbf{F}_k = \mathbf{e} | \mathbf{F}_{k-1} = \mathbf{e}')$  est la probabilité de transition du treillis du système.

Dans le cas d'une source binaire et équirobable,  $P(\mathbf{F}_k = \mathbf{e} | \mathbf{F}_{k-1} = \mathbf{e}') = 1/2$  si la transition est autorisée, 0 sinon.

Le parcours complet du treillis nous donne la séquence  $\lambda_k$ , avec  $1 \leq k \leq n$ . Or on sait que :

$$P(Y_{1:n}) = P(Y_1) P(Y_2 | Y_1) \cdots P(Y_n | Y_{1:n-1}) \quad (3.27)$$

donc 
$$P(Y_{1:n}) = \prod_{k=1}^n \frac{1}{\lambda_k} .$$

Finalement, en réutilisant l'équation (3.20), l'estimée de l'entropie de la sortie du système est

calculée en construisant et parcourant un treillis pour chacune des séquences simulées  $Y_{1:n} = y_{1:n}^{(s)}$ . Ainsi, en notant  $\{\lambda_k^{(s)}\}$  la séquence associée à chaque simulation, on a :

$$\hat{H}(Y_{1:n}) = \frac{-1}{N_s} \sum_{s=1}^{N_s} \log_2 \left[ \prod_{k=1}^n \frac{1}{\lambda_k^{(s)}} \right] = \frac{1}{N_s} \sum_{s=1}^{N_s} \sum_{k=1}^n \log_2 \lambda_k^{(s)} \quad (3.28)$$

Dans le cadre de notre modélisation du canal de transmission, le flux d'information binaire le traversant est donc estimé par l'équation :

$$IMM(X_{0:n+1}; Y_{1:n}) = \frac{1}{N_s} \sum_{s=1}^{N_s} \sum_{k=1}^n \log_2 \lambda_k^{(s)} - \frac{n}{2} \log_2(2\pi e \sigma^2) \quad (3.29)$$

L'algorithme utilisé est décrit de manière synthétique au tableau 3.1.

<p><b>Positionner</b> <math>C^{1D}</math></p> <p><b>Calculer</b> la matrice de canal : <math>\mathcal{A}</math></p> <p><b>Pour</b> numéroSimulation = {1 à <math>N_s</math>} <b>faire</b></p> <p style="padding-left: 40px;"><b>Simuler</b> <math>Y_{1:n}^{(numéroSimulation)}</math></p> <p style="padding-left: 80px;"><b>Pour</b> numéroModule = {1 à <math>n</math>} <b>faire</b></p> <p style="padding-left: 120px;"><b>Calculer</b> les <math>\alpha_{numéroModule}^{(numéroSimulation)}</math> et <math>\gamma_{numéroModule}^{(numéroSimulation)}</math></p> <p style="padding-left: 120px;"><b>En Dédire</b> les <math>\lambda_{numéroModule}^{(numéroSimulation)}</math></p> <p><b>En Dédire</b> <math>IMM(n, N_s, \sigma^2)</math></p> <p style="text-align: center;"><i>Tableau 3.1 : Algorithme de calcul du flux d'information</i></p>
--

### 3.2.2.3 L'IMM comme mesure théorique de performances

Nous proposons d'utiliser l'IMM pour mesurer les performances d'un système de transmission par code-barres. Notre objectif est d'étudier les variations de cette quantité d'information pour différentes valeurs des paramètres du système. Le code-barres étant positionné librement devant la caméra, les variations de l'IMM dépendent de la configuration globale du système modélisé en un canal de transmission régi par l'équation (3.18). Dans ce modèle, le canal apparaît comme un canal à mémoire et à bruit ABG. L'ensemble des

distorsions du canal sont synthétisées dans la matrice caractéristique  $A$ , où le flou ainsi que le processus d'intégration du capteur créent une interdépendance entre les signaux délivrés par les pixels. La position et l'orientation du code induisent une variation de la puissance rayonnante reçue.

L'IMM est certes avant tout une mesure théorique de la quantité d'information pouvant être acheminée sans erreur à travers un canal, ceci en se basant sur les lois de probabilités régissant les processus d'entrée et de sortie. Toutefois la modélisation proposée crée un lien entre cette grandeur théorique, les caractéristiques du système et les positions de son émetteur et récepteur de sorte que l'on peut établir une cartographie de l'espace indéxée sur les niveaux d'IMM du système. Notons qu'en toute rigueur ce niveau d'IMM est alors fonction de la symbologie utilisée.

De plus, un important résultat de la théorie de l'information est que le bruit limite l'IMM, mais n'influe pas sur les performances d'une transmission. En d'autres termes, pour peu que le quantité d'information émise soit inférieure ou égale à l'IMM, le message peut être totalement récupéré avec un codage adéquat ce quel que soit le niveau de bruit. Insistons cependant sur le fait que cette assertion classique de la théorie présuppose que le message aura été codé de manière appropriée pour sa récupération complète mais sans préciser quel est ce codage! L'IMM apparaît donc comme une quantité maximale et pour une matrice de canal  $A$  donnée, sa valeur décroît avec l'augmentation de la puissance du bruit. Le niveau de bruit restreint donc directement la zone de l'espace garantissant un seuil d'IMM pour une transmission efficace.

Ainsi, de par ces deux importantes remarques, nous utilisons l'IMM comme critère objectif pour l'estimation de ce que nous définissons comme la *profondeur de champ théorique* (PdCT). En photographie classique, la profondeur de champ géométrique est estimée en se basant sur la géométrie et les caractéristiques de l'appareil (distance focale, ouverture du diaphragme et taille du cercle de confusion, cf §2.2.2.4). Les calculs géométriques ne considèrent toutefois pas le bruit mais seulement une approximation du flou. La PdCT se base quant à elle sur la matrice de distorsion du canal ainsi que sur la puissance du bruit. Comme nous l'illustrons dans la suite, cette approche élargit le plus souvent les capacités des systèmes à code-barres au delà des prédictions classiques de la photographie.

### 3.3 Simulations et exploitation des résultats

Nous présentons dans ce paragraphe l'application des outils et concepts introduits jusqu'ici, illustrant ainsi l'intérêt de la démarche. L'IMM est évaluée pour différentes configurations de codes et de caméras. Le tracé de ses variations en fonction du positionnement et de l'orientation du pictogramme englobe l'ensemble des distorsions induites par le canal. De plus, nous évaluons le rapport signal sur bruit moyen (RSB) des transmissions. Ce rapport exprimé en décibels (dB) est défini par :

$$RSB = \frac{E_s}{N_0} \quad (3.30)$$

où

- $E_s$  est l'énergie moyenne du symbole binaire reçu, avec, en se reportant à l'équation du canal (3.18),

$$E_s = \frac{1}{n} \sum_{k=1}^n \frac{\sum_{\{x_{k0}\}} \left[ \sum_{k_0=1}^{n+2} A_{k,k_0} x_{k_0} \right]^2}{2^{n+2}} \quad (3.31)$$

- $N_0^{(s)}$  est la puissance par symbole binaire reçu du bruit du canal.

#### 3.3.1 Configuration du système

La caméra simulée est dotée d'un diaphragme carré de côté égal à  $2l_D$  et a pour distance focale  $f$ . Le capteur est formé de pixels contigus carrés de  $\rho = 5\mu\text{m}$  de côté, et est placé à la distance  $z_c$  du centre optique.

Nous optons pour trois configurations du bloc optique (figure 3.7):

- un bloc que nous qualifions de Longue Portée (LP) avec une focale  $f = 7.6\text{mm}$ , un nombre d'ouverture  $f_8^{nb}$  et  $z_c = 7.8\text{mm}$ . La position sur l'axe optique du plan objet correctement focalisé est donc à  $z_{of} = 296.4\text{mm}$
- un bloc compact Longue Portée (cLP), avec  $f = 1.492\text{mm}$ ,  $f_{1.4}^{nb}$ ,  $z_c = 1.5\text{mm}$  et  $z_{of} = 296.4\text{mm}$  également
- un bloc Courte Portée (CP) avec  $f = 7.5\text{mm}$ ,  $f_8^{nb}$ ,  $z_c = 9\text{mm}$  et  $z_{of} = 45\text{mm}$ .

Le choix des paramètres a pour premier objectif d'être réaliste en regard des standards de l'industrie [66]. Elles nous permettent surtout, par le calcul du flux, d'évaluer plus en détail les effets des distorsions du canal qui ne sont pas les mêmes suivant le bloc optique.

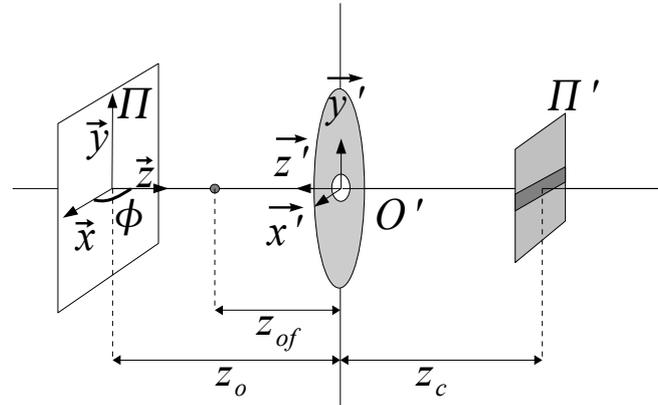


Figure 3.7 : Configuration et notations

Nous simulons un code  $C^{1D}$  de 95 barres d'égale largeur  $\delta=1\text{mm}$ . Le calcul de l'IMM est effectué en se basant sur la modélisation par barre du canal (équation (3.9)). Pour chaque bloc optique la position focalisée est prise comme point de référence, le flux radiant intégré sur le plan du capteur y est alors considéré comme nominal (normalisé par rapport au flux reçu lorsque le code est en  $z_{of}$ ). La source des symboles  $X_k$  associés aux barres  $C_k^{1D}$  est une source binaire et équiprobable. Le bruit enfin est simulé à différentes puissances  $N_0$  par pixel, afin d'avoir une plage suffisamment large de RSB (variant en pratique d'environ -10dB à 35 dB).

### 3.3.2 Profondeur de Champ Théorique

En nous référant au schéma de la figure 3.7, le code  $C^{1D}$  est déplacé le long de l'axe optique de sorte que la distance  $z_o$  varie entre  $[z_{of}; 2z_{of}]$ . Nous modifions également son orientation tel que  $\phi=\{0; 0.1\pi; 0.2\pi\}$ .

### 3.3.2.1 Caméra Longue Portée

Les variations de l'IMM sont tracées par rapport à  $z_o$  et  $\phi$  en figure 3.8 .

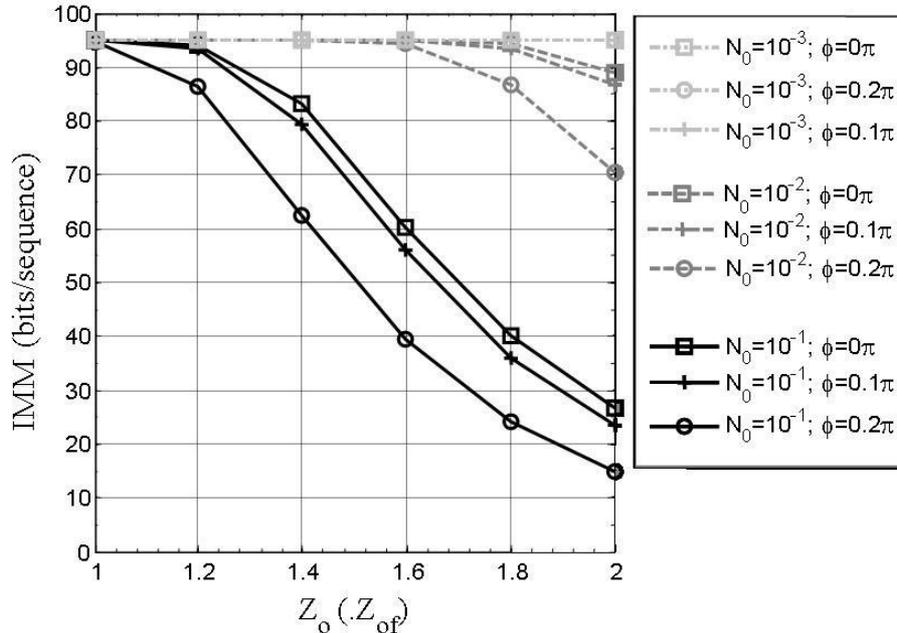


Figure 3.8 : Caméra LP - IMM par  $Z_o$

Pour les puissances de bruit faibles, les pertes d'IMM sont de moins d'1%, ceci même dans des régions situées bien après le point extrême de la profondeur de champ géométrique à  $z_{pe} = 370\text{mm} \approx 1.25 z_{of}$  du centre optique. Ces résultats illustrent le fait que la totalité de l'information transmise peut être récupérée dans un canal sans bruit dès lors que la matrice de distorsion du canal est connue, et ceci même pour de fortes distorsions (donc un flou important). Ceci renforce et valide l'intérêt de la définition d'une profondeur de champ théorique pour la lecture de code-barres. Dans le cas d'un canal plus fortement bruité, l'IMM est toutefois nettement diminué, ayant un impact sur l'étendue de la PdCT. Ainsi si l'on fixe un seuil pour l'IMM de 90%, le point à l'extrémité la plus éloignée de la PdCT est au delà de  $2 z_{of}$  pour  $N_0=10^{-3}$ , positionné à  $1.8 z_{of}$  pour  $N_0=10^{-2}$ , et proche de  $1.2 z_{of}$  pour  $N_0=10^{-1}$  quelle que soit l'orientation du code.

Lorsque  $C^{1D}$  est face à la caméra, les pertes de l'IMM sont de moins d'1% pour les bruits faibles. Les performances du système en termes de flexibilité par rapport à l'éloignement sont donc très bonnes lorsque le capteur, principal support du bruit, est de bonne qualité.

### 3.3.2.2 Caméra compacte Longue Portée

Considérons désormais la caméra cLP. La configuration de son bloc optique par rapport au cas LP nous permet d'étudier l'influence de modifications des caractéristiques optiques sur la PdCT. En photographie, l'impact du nombre d'ouverture sur la profondeur de champ est un résultat classique couramment utilisé par les photographes amateurs ou professionnels. Plus le nombre d'ouverture est petit, plus la profondeur de champ est grande. Cette modification du  $f^{nb}$  s'opère en jouant sur la taille du diaphragme et/ou sur la distance focale. La position prédite par la géométrie du point extrême de la profondeur de champ est de  $z_{pe} = 15.2m \approx 50 z_{of}$ . En se reportant à la figure 3.9, l'évaluation de l'IMM indique en revanche que la PdCT est (beaucoup) plus courte. En effet, pour un seuil de récupération à 90%, le point extrême se situe à  $2 z_{of}$  pour  $N_0 = 10^{-3}$ , à  $1.35 z_{of}$  pour  $N_0 = 10^{-2}$  et est situé en amont de  $z_{of}$  pour  $N_0 = 10^{-1}$ .

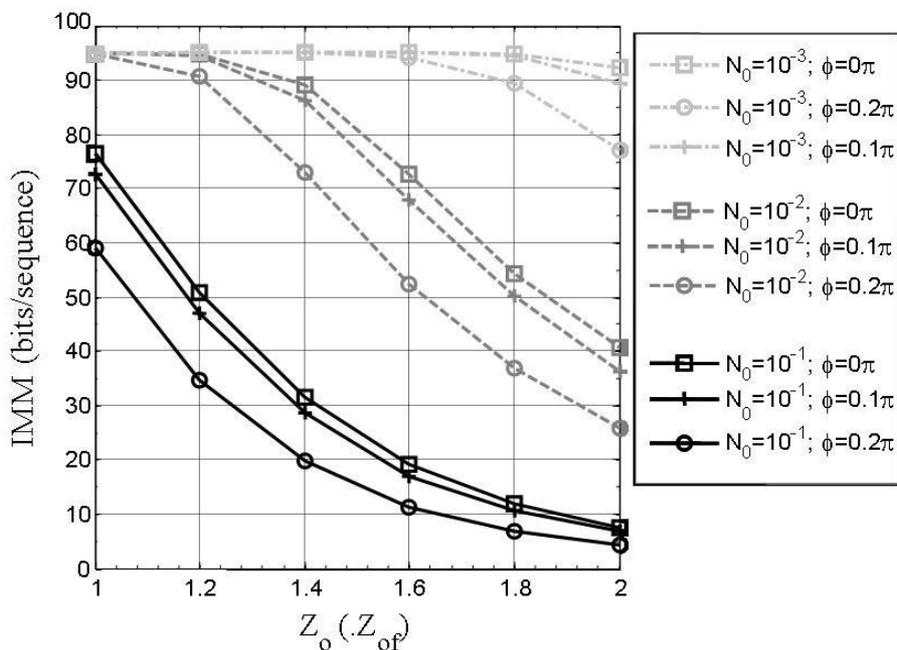


Figure 3.9 : Caméra cLP - IMM par  $Z_o$

La prise en compte du bruit du système amène donc un éclairage important sur la zone

d'utilisation optimale du système. De plus, le positionnement aléatoire des frontières des barres du code par rapport aux pixels intégrant influe sur les coefficients de  $A$ . Cette désynchronisation spatiale est d'autant plus importante que le code s'éloigne de la caméra et que son grandissement (paragraphe 2.2.1) est faible.

### 3.3.2.3 Caméra Courte Portée

Dans le cas du bloc CP, la position focalisée objet est à  $z_{of}=45\text{mm}$  du centre optique. La profondeur de champ géométrique ne s'étend pas quant à elle au delà de  $z_{pe}=50\text{mm}\approx 1.1 z_{of}$ . La PdCT du système sous un bruit de faible puissance  $N_0\leq 10^{-2}$  atteint les  $2 z_{of}$  avec une récupération totale de la séquence émise. Pour  $N_0=10^{-1}$  on peut récupérer au moins 95% du message émis jusqu'à  $1.8 z_{of}$  et pour  $-0.2\pi\leq\phi\leq 0.2\pi$ . Les variations de l'IMM par rapport à la position, à l'orientation et à la puissance du bruit illustrent clairement un excellent niveau de récupération de l'information transmise et une faible emprise du bruit (figure 3.10).

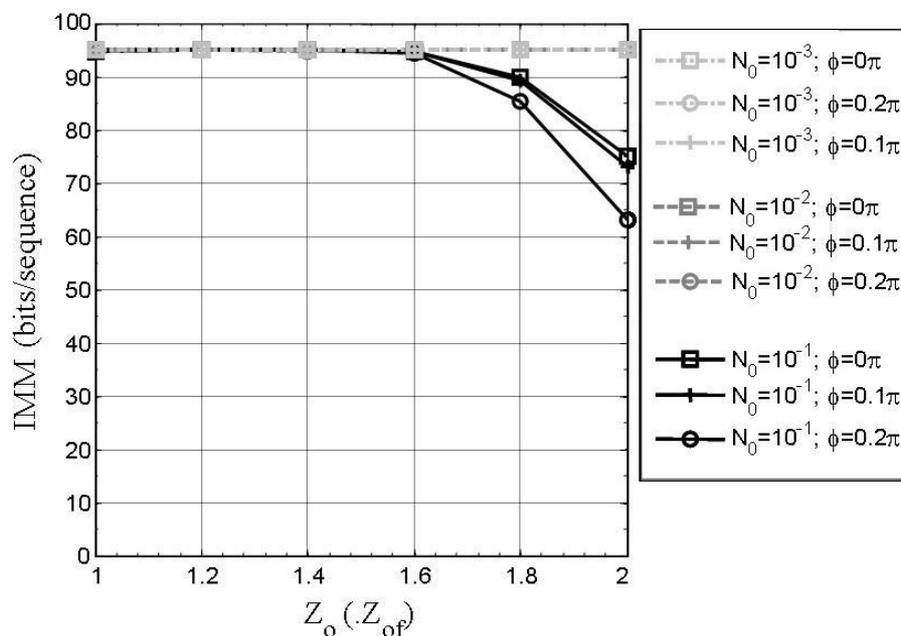


Figure 3.10 : Caméra CP - IMM par  $Z_o$

Ceci s'explique principalement par la valeur du grandissement de l'optique qui implique

l'intégration de chaque barre de  $C^{1D}$  par un grand nombre de pixels (près de 40 en moyenne pour un positionnement du plan objet en  $z_{of}$  contre 20 pour la LP et à peine 1 pour la cLP). Le calcul de la moyenne des niveaux de gris des pixels pour la sortie du système diminue la puissance du bruit par barre. De plus, le phénomène de désajustement entre les frontières des symboles projetés et l'agencement des pixels est négligeable dans cette configuration.

### 3.3.2.4 Caméras et seuils d'IMM, tableau récapitulatif

Nous avons tracé en figures 3.11 à 3.13 les variations de la PdCT des 3 caméras en fonction du RSB. Cette représentation illustre de manière claire les liens apportés par nos travaux au triplet {positionnement, IMM, RSB}. En fait, ce triplet reflète les trois pans de la problématique initialement posée, à savoir {monde réel, théorie, système}, et les courbes ci délivrées en constituent la réponse prototype. Elles rendent directement lisible les positions limites du plan du code correspondant à un seuil d'IMM donné tout en indiquant le RSB obtenu.

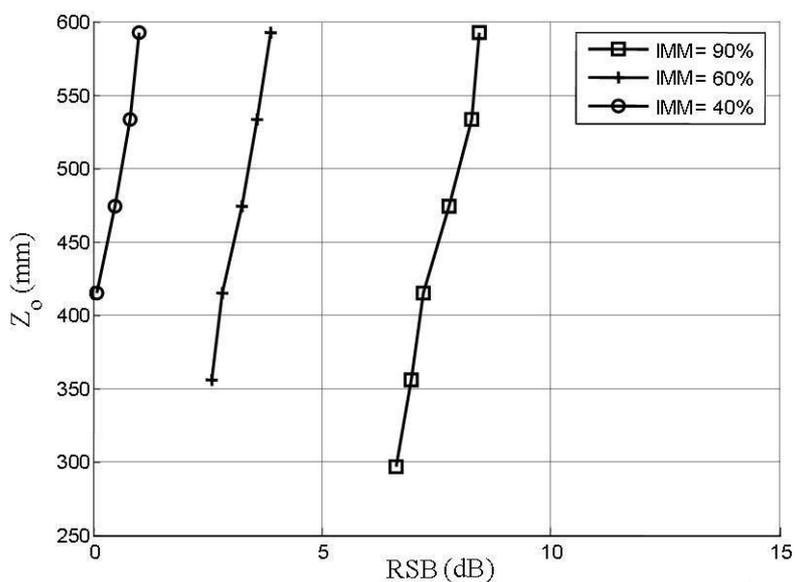


Figure 3.11 : Caméra LP - IMM par RSB

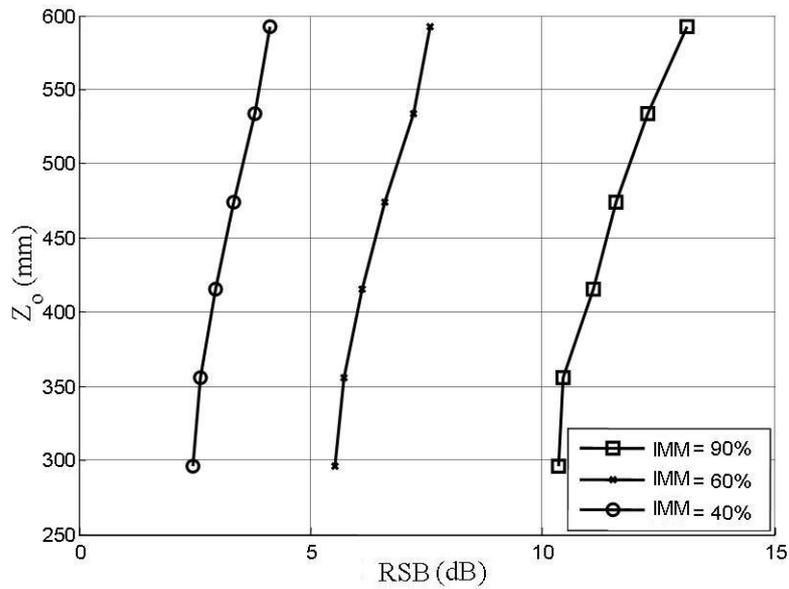


Figure 3.12: Caméra cLP - IMM par RSB

Dans le cas des caméras LP et cLP, les figures 3.11 et 3.12 illustrent bien le plus grand besoin en énergie de la version compacte par rapport à la version standard pour des performances similaires. Les deux caméras ayant même  $z_{of}$ , la comparaison des courbes est directe et ce type de représentation semble plus pertinente que d'autres lors des phases de conception du système de transmission.

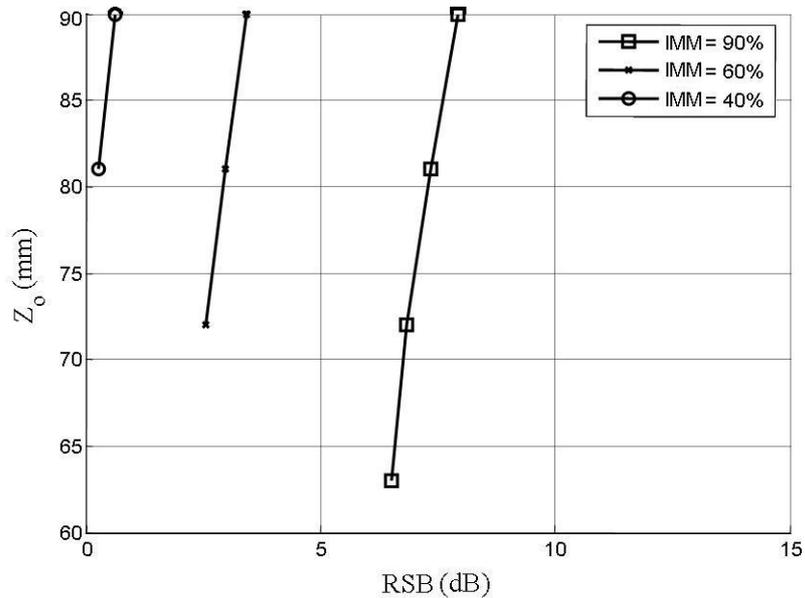


Figure 3.13: Caméra CP - IMM par RSB

Nous présentons enfin au tableau 3.2 un récapitulatif des performances des 3 caméras simulées. Les valeurs retenues sont dans le cas le plus bruité ( $N_0=10^{-1}$ ) et lorsque le code-barres est présenté de face. Quatre niveaux d'IMM ont été retenus et les positions correspondent à l'éloignement maximum du code par rapport au seuil du flux fixé. A titre de comparaison, la position du plan éloigné de la profondeur de champ géométrique (paragraphe 2.2.2.4) est également fournie pour chaque caméra. La PdCT s'avère beaucoup plus riche en terme d'interprétation que la PdCG, ceci principalement par le fait que la quantité d'information exigée est variable ce qui implique une plus grande flexibilité et finesse lors de la conception du système. Surtout, et cela est bien mis en évidence dans le cas cLP, la PdCG peut donner des valeurs extrêmement optimistes (voire trop) car elle ne prend pas en compte toutes les distorsions du canal. Comme cela a été présenté, un éloignement du code diminue le RSB ce qui ne peut qu'amoindrir la quantité d'information maximale récupérable. La PdCT, adossée à l'IMM et au modèle du canal, s'avère donc non seulement plus riche mais plus juste.

IMM (bits par code)	LP	cLP	CP
	$f = 7.6\text{mm} - f^{nb}_8 -$ $z_{of} = 296.4\text{mm}$	$f = 1.492\text{mm} - f^{nb}_{1.4} -$ $z_{of} = 296.4\text{mm}$	$f = 7.5\text{mm} - f^{nb}_8 -$ $z_{of} = 45\text{mm}$
	$z_{pe} = 1.25 z_{of}$	$z_{pe} = 51 z_{of}$	$z_{pe} = 1.1 z_{of}$
	PdCT : $z_o/z_{of}$		
90	$\approx 1.28$ (380mm)	$\approx 0.95$ (280mm)	$\approx 1.78$ (80mm)
80	$\approx 1.43$ (425mm)	$\approx 0.98$ (290mm)	$\approx 1.89$ (85mm)
60	$\approx 1.62$ (480mm)	$\approx 1.15$ (340mm)	$\approx 2.16$ (97mm)
40	$\approx 1.82$ (540mm)	$\approx 1.28$ (380mm)	$\approx 2.31$ (104mm)

Tableau 3.2 Position du code pour différents seuil d'IMM avec  $N_0=10^{-1}$

### 3.3.3 Effets d'un autofocus

Pour rappel, les résultats présentés pour illustrer la profondeur de champ théorique sont tirés de simulations pour lesquelles les blocs optiques sont à focale fixe. Le déplacement

du code-barres hors de sa position focalisée déterminée par l'équation (2.17) implique un flou dit de défocalisation. Ce flou, qui aggrave l'étalement induit par la diffraction, a été modélisé en section 2.2.3 par un système linéaire en puissance. Nous étudions dans ce paragraphe l'impact de la suppression du flou de défocalisation sur les performances du système. Nous simulons à ce titre un système autofocus associé aux blocs optiques. Il consiste à adapter la distance focale ainsi que la position du capteur  $z_c$  par rapport à l'éloignement  $z_o$  du plan objet.

Commençons par le bloc LP, nous avons tracé les variations de l'IMM par rapport à  $z_o$  pour différentes orientations et  $N_0=10^{-2}$ , puis par rapport au RSB. Les résultats sont tracés dans le cas avec autofocus et à focale fixe. Les courbes sont en figures 3.14 a/ et b/. Le tracé de la figure 3.14 a/ montre une légère amélioration de l'IMM par l'adaptation optique sans que cela ne dépasse toutefois quelques bits par séquence. Cette différence négligeable entre les deux cas se confirment également pour une large gamme du RSB (figure 3.14 b/). Ce faible gain signifie que la défocalisation n'altère pas en réalité la transmission. Cependant, un tel résultat n'est possible que par la parfaite connaissance de la matrice de distorsion du canal, ce qui est le cas dans le protocole expérimental que nous avons employé. L'avantage de cette parfaite connaissance est que les résultats obtenus correspondent aux maxima atteignables, tant du point de vue performances que robustesse aux distorsions.

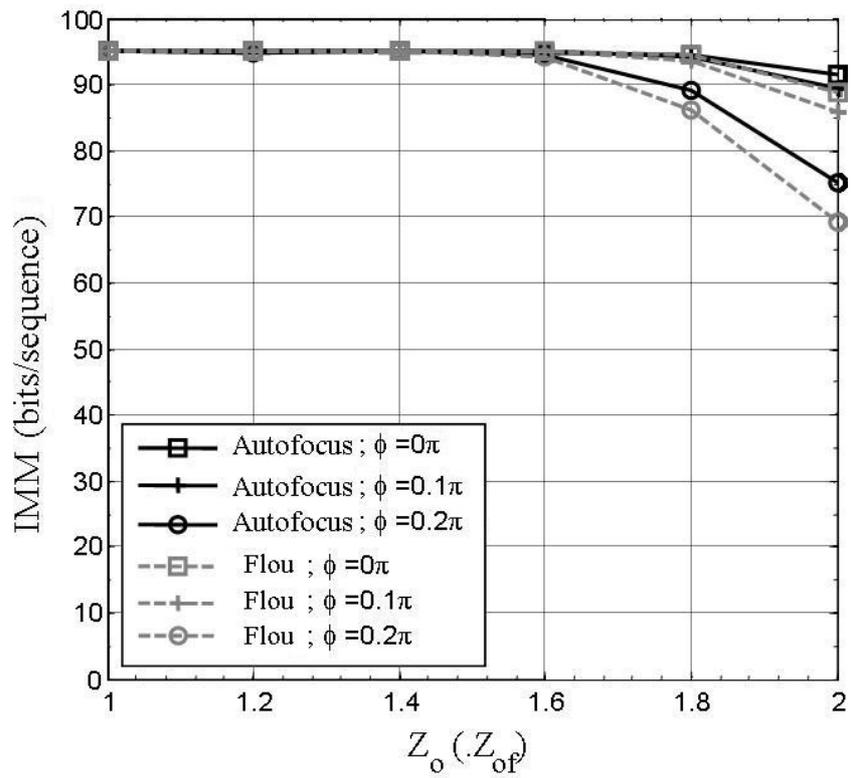


Figure 3.14 a/ : Caméra LP - IMM par  $Z_o$

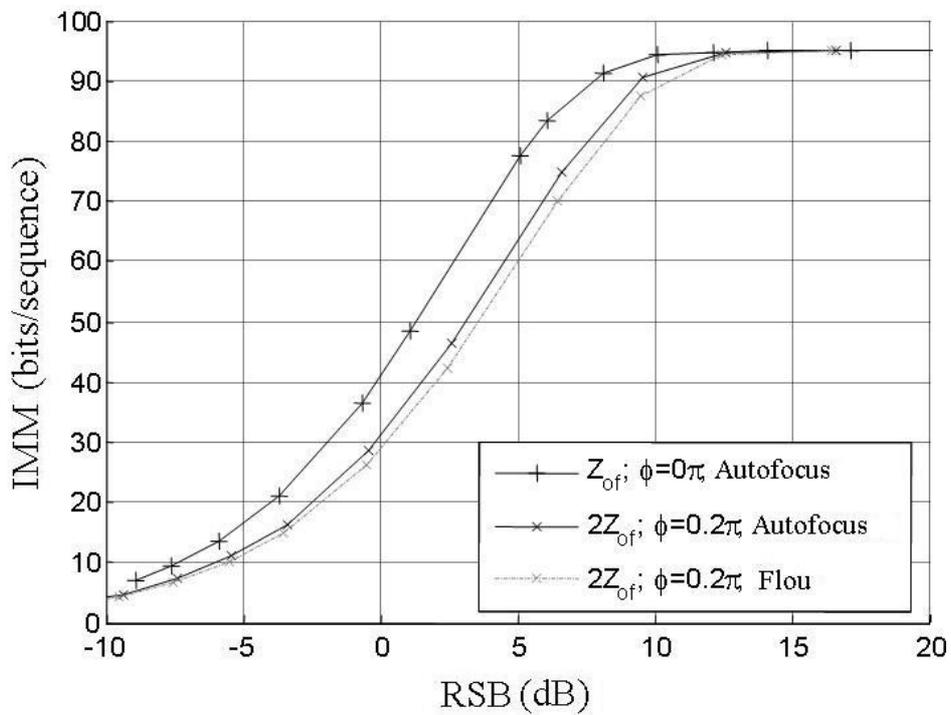


Figure 3.14 b/ : Caméra LP - IMM par RSB

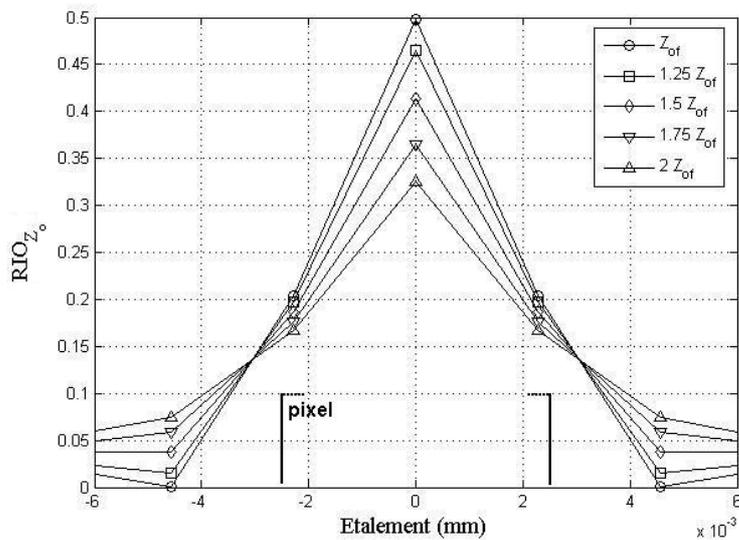


Figure 3.15 : Caméra LP - Dynamique de la RIO

Sachant les résultats du bloc LP, on peut prédire que le cas cLP sera peu différent quant au gain d'un autofocus, voire plus négligeable de par une dynamique de RIO plus lente due à une focale très courte (cf 2.2.3). Cette prévision est confirmée par les courbes d'IMM données en figures 3.16, ainsi que par la comparaison des dynamiques de réponses impulsionnelles (figures 3.15 et 3.17). Cette optique compacte est surtout sujette au désajustement spatial entre les barres projetées du code et les pixels du détecteur. Les performances du système ainsi configuré sont limitées par son détecteur, seule une modification de la largeur  $\delta$  des barres est susceptible d'améliorer le niveau de récupération de l'information.

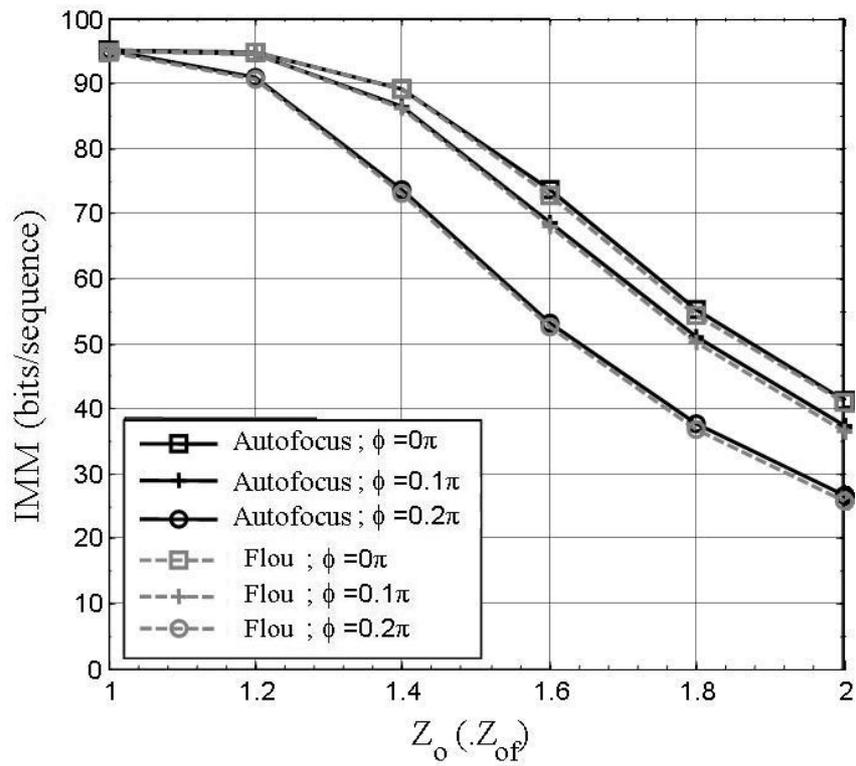


Figure 3.16 a/ : Caméra cLP - IMM par  $Z_o$

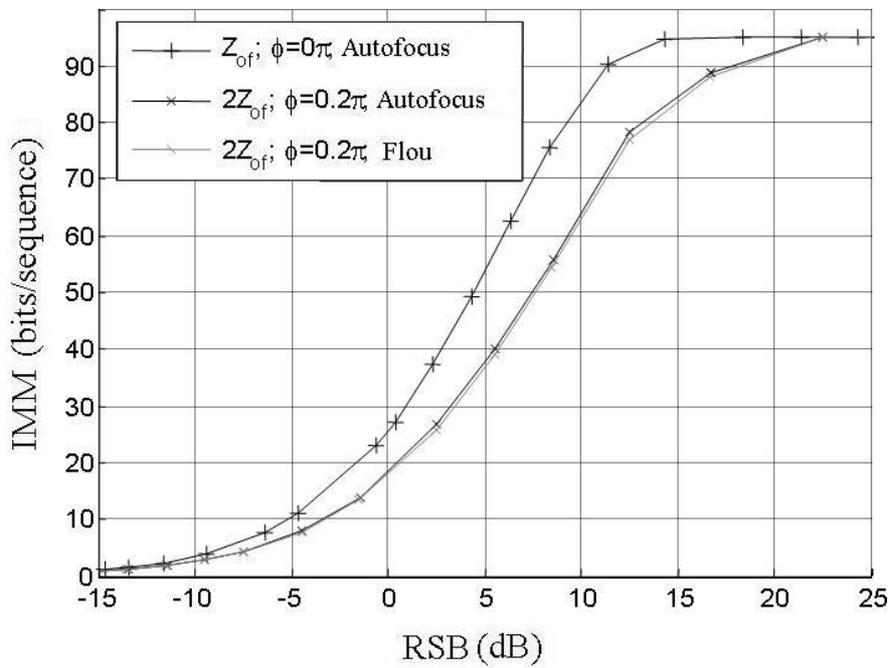


Figure 3.16 b/ : Caméra cLP - IMM par RSB

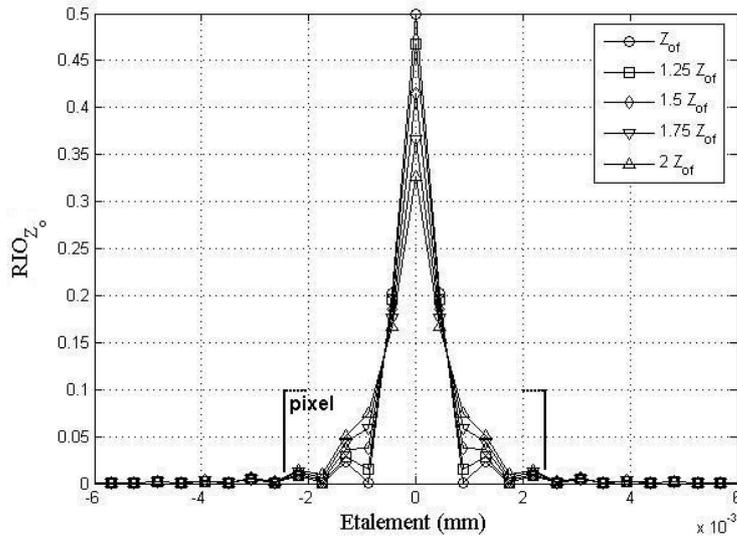


Figure 3.17 : Caméra cLP - Dynamique de la RIO

A contrario, l'adjonction d'un autofocus à la caméra CP amène un gain certain au système. En effet, on note sur la figure 3.18 a/ un gain de jusque 20 bits par séquence. Cette amélioration se produit également pour une large gamme du RSB.

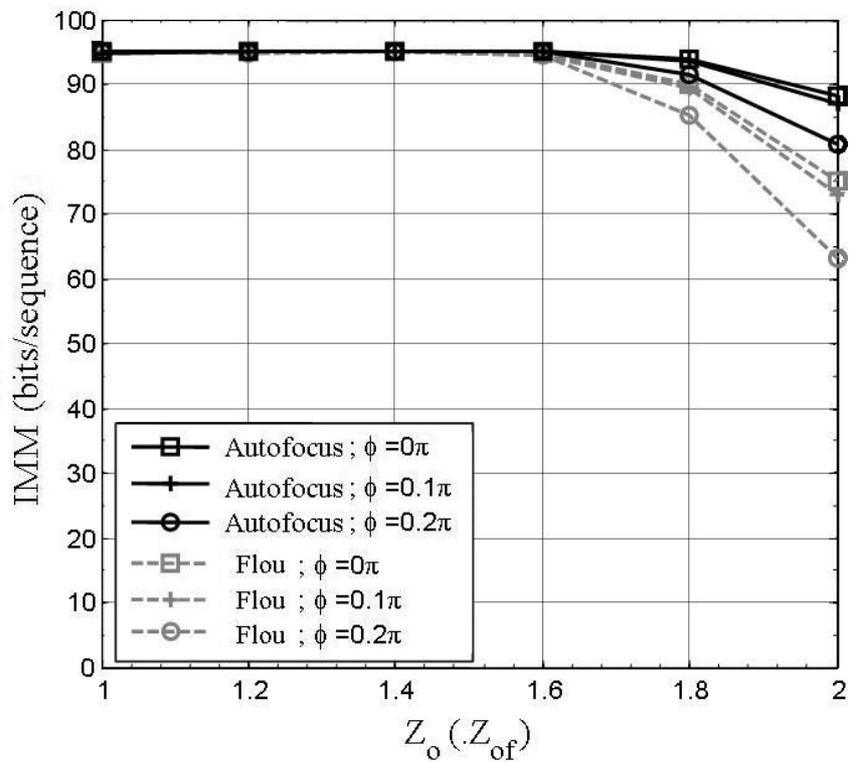


Figure 3.18 a/ : Caméra CP - IMM par  $Z_o$

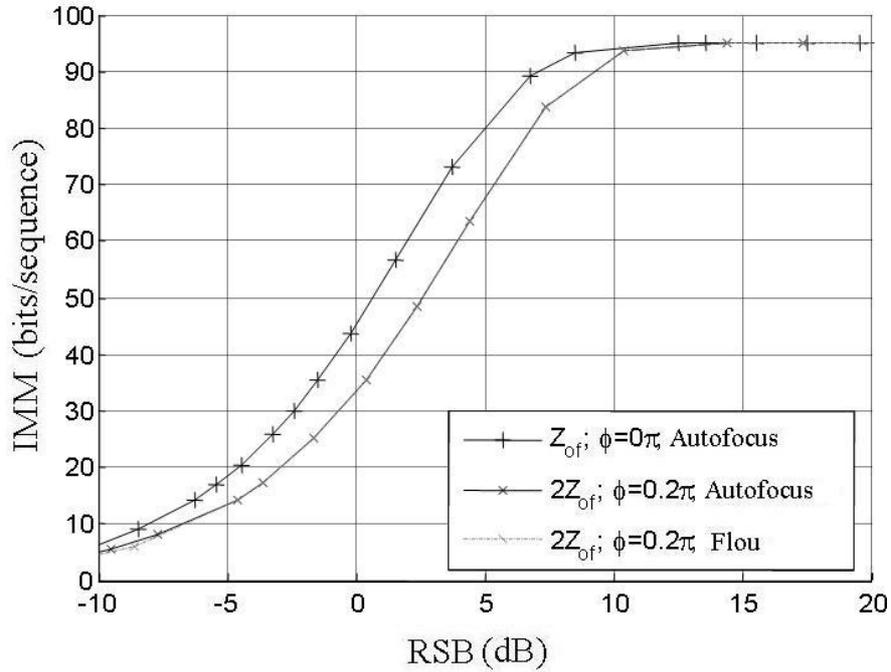


Figure 3.18 b/ : Caméra CP - IMM par RSB

Dans les faits, le bloc CP présente une diffraction faible avec la large ouverture de son diaphragme et n'est pas affecté par les problèmes de désajustement spatial (grandissement conséquent). La seule distorsion subsistante est donc le flou de défocalisation, dont l'impact est d'autant plus fort que l'étalement de la  $RIO$  croît rapidement avec l'éloignement du code pour la plage de variation de  $z_o$  (figure 3.19).

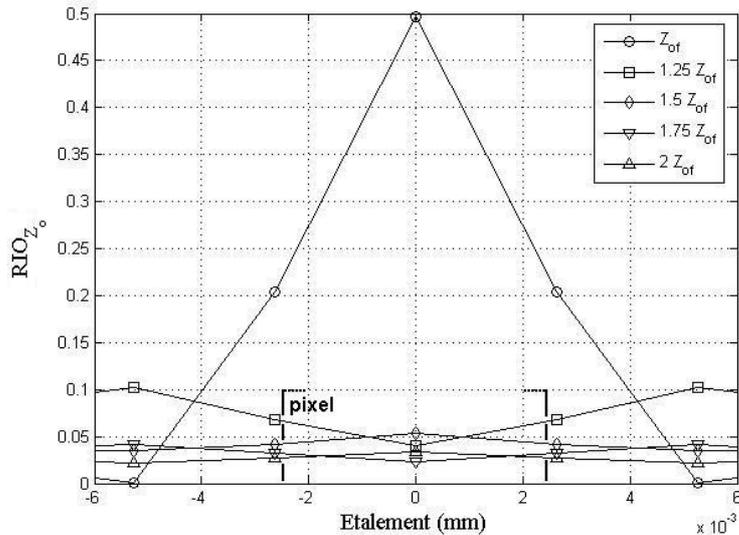


Figure 3.19 : Caméra CP - Dynamique de la RIO

### 3.3.4 Résolution spatiale théorique du système

En considérant la lecture de code-barres comme un système de télécommunications, nous avons proposé une modélisation du canal de transmission multi-paramétré. Ces paramètres (optique, position, largeur des barres, ...) composant la matrice caractéristique du canal sont liés entre eux. La modification de l'un agit sur un autre, cela pouvant affecter les performances du système. Par cette notion d'interaction, nous avons fait mention plus haut d'un cas vertueux où l'augmentation de la largeur des barres améliorerait la robustesse du système muni du bloc optique cLP. Nous allons ici plus loin dans cette idée, et plutôt que d'analyser le système en termes de bits transmis par rapport au RSB ou à la position, nous considérons le lien entre niveau d'IMM récupérable et taille de barre du code. En d'autres termes, à l'instar de la définition d'une profondeur de champ théorique, nous cherchons à déterminer une résolution spatiale théorique (RST).

L'énergie symbole, dont la définition a été introduite pour le calcul du RSB (équation (3.30)), est proportionnelle à l'éclairement produit par chacune des barres sur le plan du capteur et par conséquent à leur largeur  $\delta$ . Son augmentation pour une puissance de bruit capteur constante tend donc à améliorer le RSB en augmentant l'énergie symbole et diminuant la puissance du bruit par symbole. Lorsque le code est présenté face à la caméra et en se référant à (3.6), (3.8)

et (3.31), l'énergie symbole  $E_s$  pour les  $n$  symboles transmis est proportionnelle à la surface d'intégration des barres projetées et à la variation de l'éclairement sur le plan capteur par le vignettage :

$$E_s = \frac{1}{n} \sum_{k=1}^n \frac{\sum_{k_0=1}^{n+2} A_{k,k_0} x_{k_0}}{2^{n+2}} \Big]^2 \propto \frac{1}{n} \left( G \delta \rho \cdot \sum_{i=1}^n \cos^4 \alpha_i \right)^2 \quad (3.32)$$

où  $G = z_c / z_o$  est le grandissement transversal.

De plus, la puissance du bruit est moyennée par le nombre de pixels intégrant chaque barre et l'on peut écrire

$$N_0^s \approx N_0 \frac{\rho}{G \delta} \quad (3.33)$$

Dès lors le RSB associé à un point objet donné est proportionnel au rapport des largeurs de barre par la bruit pixelique :

$$RSB(z_o) = \left( \frac{E_s}{N_0^s} \right)_{z_o} \propto \frac{\delta^3}{N_0} \quad (3.34)$$

Cette proportionnalité nous permet d'écrire qu'une différence entre deux RSB est égale à :

$$\Delta_{RSB}(z_o) = RSB^{(2)}(z_o) - RSB^{(1)}(z_o) = 10 \log_{10} \left[ \frac{N_0^{s(1)} \left( \frac{\delta^{(2)}}{\delta^{(1)}} \right)^3 \frac{\sum_{i=1}^n \cos^8 \alpha_i^{(2)}}{n}}{N_0^{s(2)} \frac{\sum_{i=1}^n \cos^8 \alpha_i^{(1)}}{n}} \right] \quad (3.35)$$

En pratique cette différence de RSB se simplifie dans deux cas : lorsque le code est peu étendu, et lorsque le bloc optique est muni de lentilles compensant les pertes dues au vignettage [73] ce qui est de plus en plus répandu sur les systèmes actuels.

Ainsi, l'équation (3.35) se simplifie et on a :

$$\Delta_{RSB}(z_o) = 10 \log_{10} \left[ \frac{N_0^{(1)} \left( \frac{\delta^{(2)}}{\delta^{(1)}} \right)^3}{N_0^{(2)}} \right] \quad (3.36)$$

Les estimations de l'IMM par rapport au RSB lors de l'étude de la PdCT fournissent les variations de la capacité de transmission pour une largeur de barre donnée. L'équation (3.36), en établissant une proportionnalité entre RSB, puissance du bruit et largeur de barre, permet de déduire à partir de ces courbes quelle serait la largeur de barre nécessaire à une position donnée pour compenser une variation de la puissance du bruit et maintenir l'IMM. Ce mécanisme de compensation permis par (3.36) est l'astuce permettant de déduire la largeur de barre minimale pour atteindre un seuil de transmission donnée. En bref, on peut estimer la densité maximale du code-barres par rapport à sa position et au niveau de transmission théorique.

Signalons que le modèle proposé pour l'étude du système permet l'estimation de l'IMM pour n'importe quelle position du code, largeur de barre ou encore puissance du bruit. On peut alors se demander pourquoi passer par un gradient de RSB pour obtenir ce que l'on pourrait estimer directement. Notre motivation vient du fait que si le modèle n'impose effectivement aucune contrainte, le temps et les ressources nécessaires pour conduire les simulations varient suivant ces paramètres. Ainsi dans le cas d'un code dont les barres sont très fines et le flou important, le treillis décrit par la mémoire du canal comportera de nombreux états possibles rendant l'algorithme dérivé du BCJR concrètement long et gourmand en mémoire. Ceci est d'autant plus rédhibitoire que l'entropie est estimée par Monte Carlo. Tout le processus d'estimation directe de la résolution est à reproduire pour une large gamme de tailles de barre ce qui rend la démarche tout bonnement impossible. Par la méthode proposée pour l'évaluation de la RST, une seule configuration du code est nécessaire.

Nous appliquons la méthode d'estimation de la RST sur deux configurations du système de transmission. La première reprend les caractéristiques de la caméra longue portée LP, pour la seconde nous introduisons un nouveau bloc optique appelée moyenne portée (MP), qui garde un diaphragme carré, et dont la position focalisée est à 15cm du centre optique. En se référant aux notations de la figure 3.7 on a :

- un bloc Longue Portée (LP) avec  $f=7.6\text{mm}$  ,  $f^{nb}_8$  ,  $z_c=7.8\text{mm}$  et  $z_{of}=296.4\text{mm}$  .
- un bloc Moyenne Portée (MP) avec  $f=7.6\text{mm}$  ,  $f^{nb}_8$  ,  $z_c=8\text{mm}$  et  $z_{of}=152\text{mm}$  .

Dans le protocole expérimental, le code  $C^{1D}$  de 95 barres est déplacé le long de l'axe

optique, face caméra, et tel que  $z_o = \{0.5z_{of}, z_{of}, 1.5z_{of}, 2z_{of}\}$ . Le flux radiant est enfin considéré nominal en  $z_{of}$ .

### 3.3.4.1 IMM et RSB

Les résultats de l'estimation de l'IMM pour les caméras LP et MP en fonction du RSB sont présentés en figure 3.20 et 3.21.

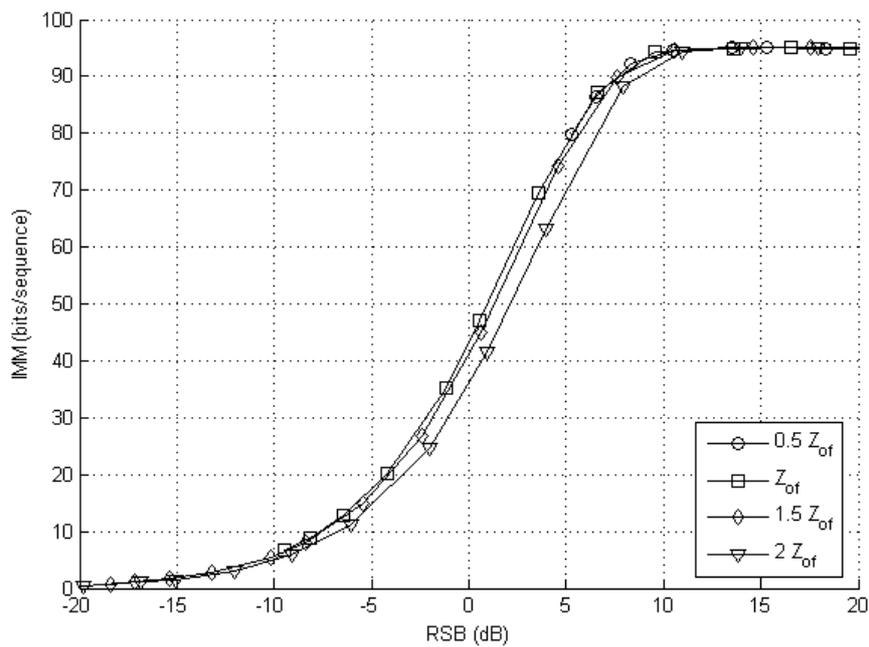


Figure 3.20 : Caméra LP - IMM par RSB

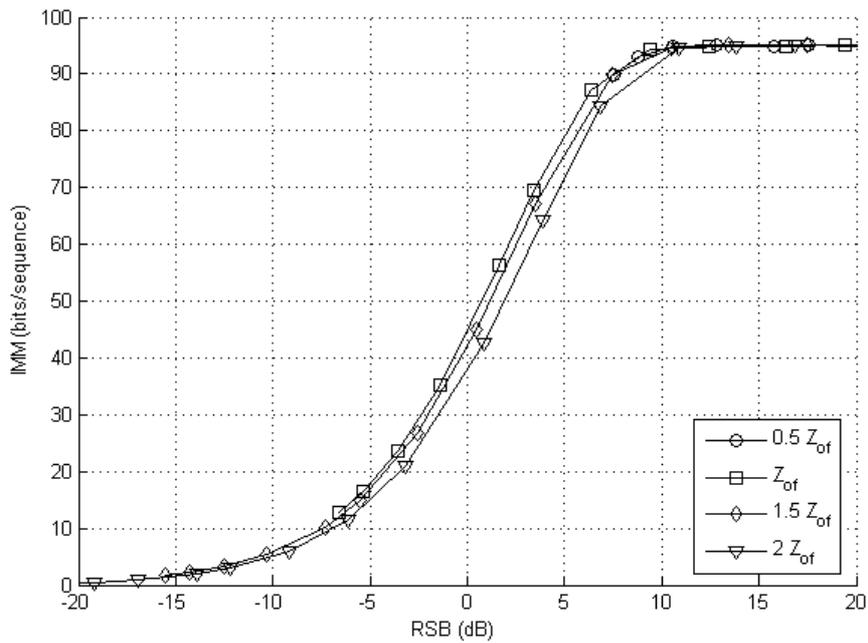


Figure 3.21 : Caméra MP - IMM par RSB

Les variations du flux des deux configurations ayant été obtenues, nous appliquons le procédé d'évaluation de la résolution spatiale théorique décrit dans le paragraphe précédent.

### 3.3.4.2 IMM et RST

Les variations de l'IMM des systèmes LP et MP par rapport à la largeur des barres du code sont représentées respectivement en figure 3.22 et 3.23.

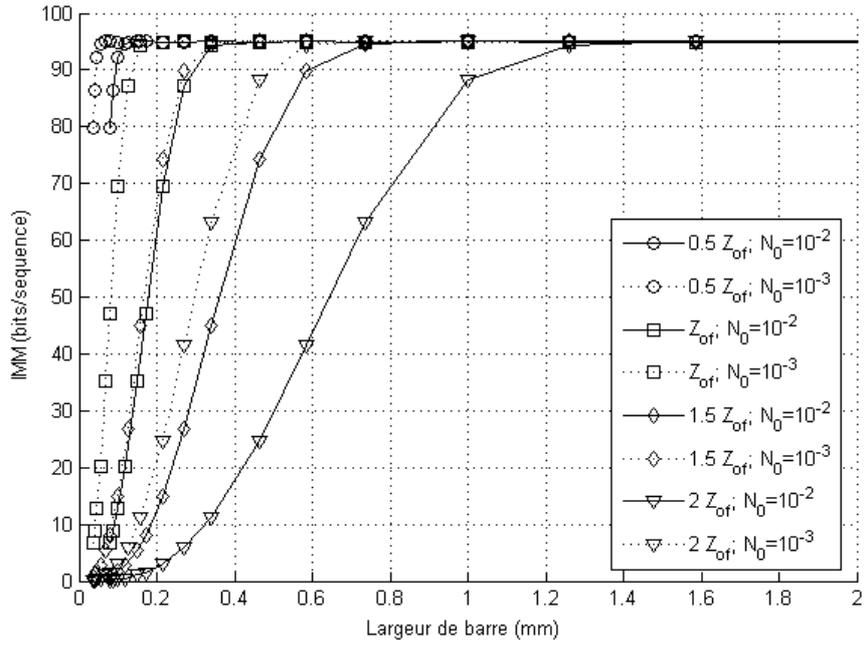


Figure 3.22 : Caméra LP - IMM par largeur de barres

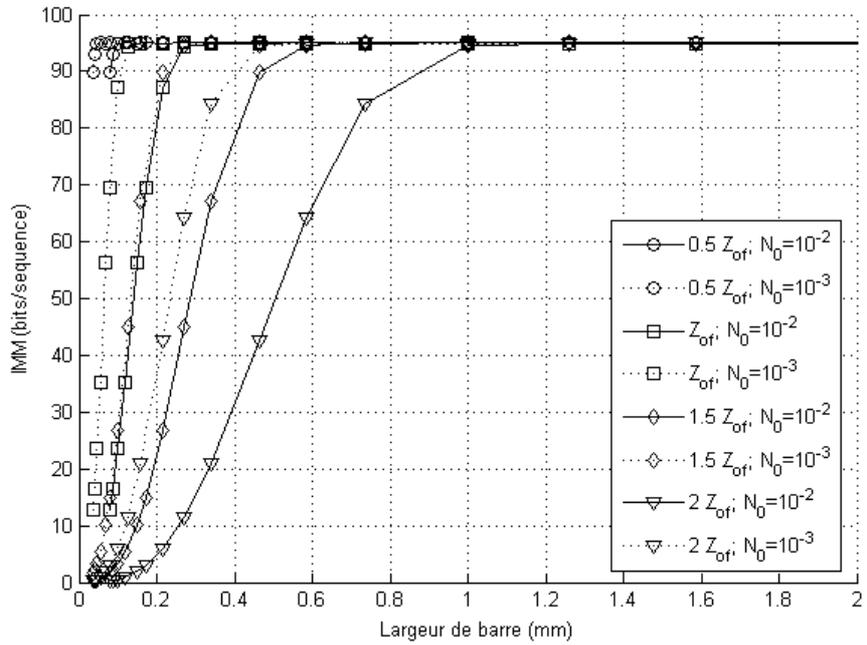


Figure 3.23 : Caméra MP - IMM par largeur de barres

Tant pour le bloc LP que MP, l'augmentation de la puissance du bruit tend à déplacer la résolution vers des barres plus larges. Cette tendance est renforcée par l'éloignement du code-

barres qui augmente le flou et par conséquent diminue l'énergie reçue. Cependant, pour une puissance et une position données, l'IMM du bloc LP diminue plus rapidement pour des barres plus fines. En effet, l'inconvénient de la configuration LP est un facteur de grandissement plus faible, diminuant le rapport signal/distorsions. Afin d'améliorer les performances de la caméra LP, une solution serait d'augmenter sa distance focale et déplacer le plan capteur afin d'obtenir une projection plus grande de la scène observée, ce qui correspond aux résultats classiques utilisés en photographie [18].

Dans les faits, le flou induit par la diffraction et les déplacements est relativement faible pour les deux blocs, de sorte que la taille du symbole projeté joue un rôle important dans les performances du système. Ceci est mis en évidence sur les figures 3.24 et 3.25. Nous y avons reporté la variation des largeurs de barres minimales par rapport à la position du code. Nous y avons ajouté pour comparaison une proposition de résolution spatiale géométrique (RSG) déterminant la taille minimale de barre permettant de respecter le critère d'échantillonnage de Nyquist (au moins 2 pixels effectuent l'intégration d'une barre). Le grandissement du bloc MP est le double de LP, le désignant donc comme plus robuste aux codes denses.

L'intérêt des figures 3.24 et 3.25 comparant résolution géométrique et théorique réside dans le lien qu'établit cette dernière entre l'IMM et la puissance du bruit. La RST ne consiste pas en une seule valeur valide à une position donnée, pour tous les types de codes et de capteurs. Le procédé proposé délivre plutôt une gamme de densités de code adaptée à l'imageur. Si le fait que différente qualité de capteur implique différente résolution a déjà été établi et étudié [18], la RST amène un élément nouveau. Cet apport est l'IMM qui devient un critère complémentaire aux précédents travaux de par le lien qu'il établit entre résolution et schéma de codage du code-barres.

La résolution spatiale théorique est par conséquent un outil adapté à l'évaluation des performances maximales d'un système de transmission graphique 1D. La considération des seuls principes géométriques mènerait à une transmission moins efficace dans le sens où ni la qualité du récepteur et ni le type de codage du code-barres ne sont pris en compte. Du point de vue de la théorie de l'information, la RST est la meilleure résolution atteignable par le système.

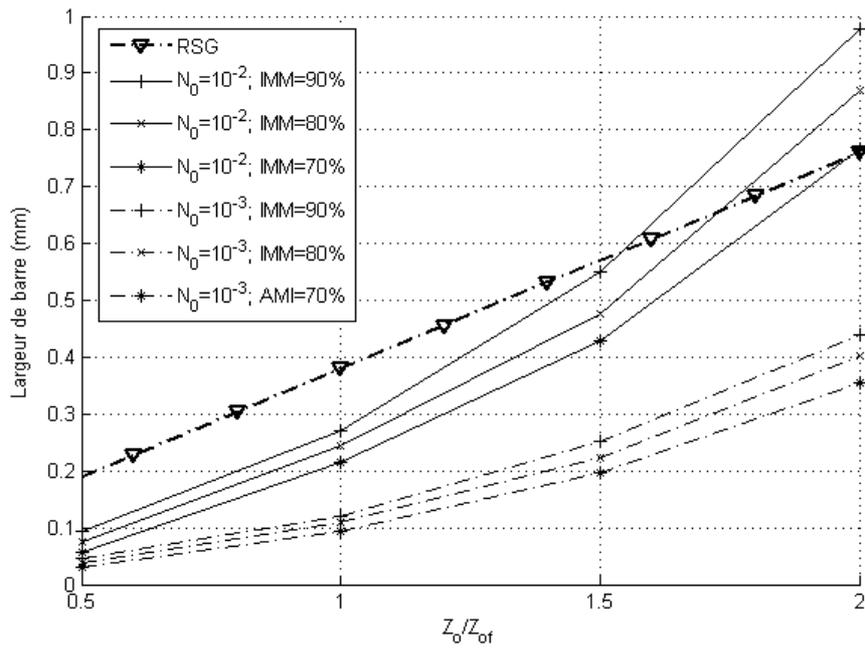


Figure 3.24 : Caméra LP - Largeur de barre par position

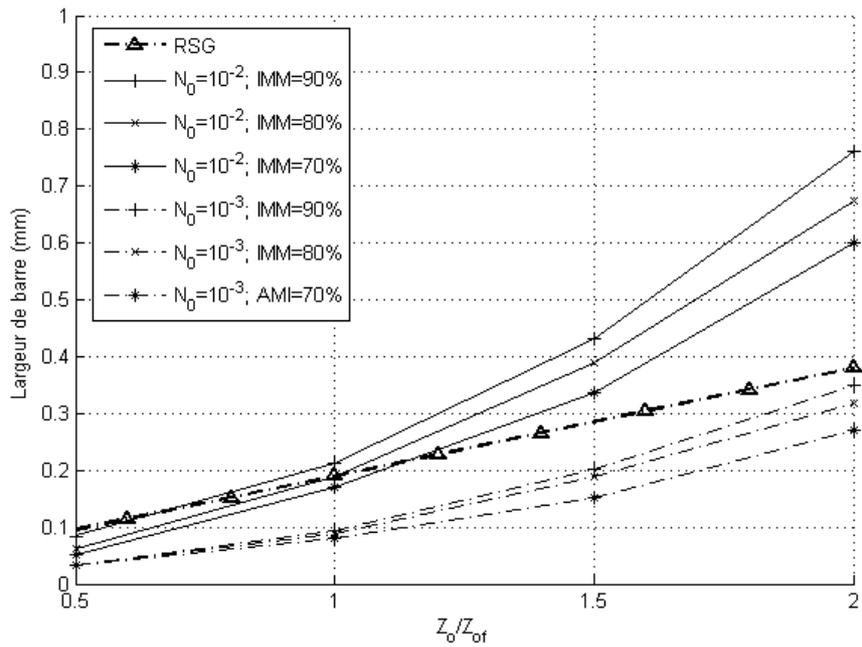


Figure 3.25 : Caméra MP - Largeur de barre par position

### 3.3.4.3 Diagrammes de résolution, codes-barres industriels

Comme nous avons essayé de l'illustrer dans le paragraphe précédent, bien que

l'approche adoptée pour l'étude des systèmes de transmission graphique soit théorique, l'estimation de l'IMM via un modèle multi-paramétré permet de dériver des grandeurs et critères intuitifs et proches du champ d'application. Nous continuons sur ce chemin en présentant ici deux types de résultats.

Le premier est un diagramme de résolution. Il s'agit de la représentation de la largeur de barre minimale par rapport à l'éloignement du code, ceci pour une puissance de bruit et un seuil d'IMM donné. Les diagrammes de résolution des caméras MP et LP sont donnés en figures 3.26 et 3.27, le seuil d'IMM est à 90% et  $N_0 = 10^{-3}$ .

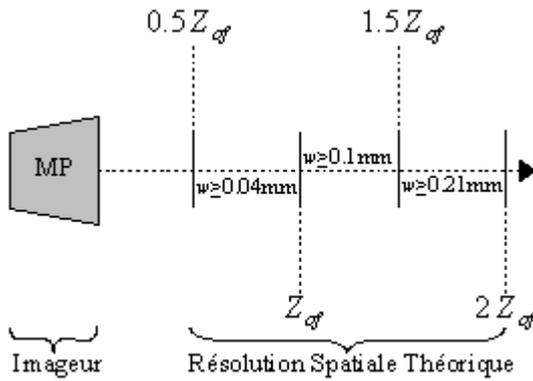


Figure 3.26 : Diagramme de résolution caméra MP

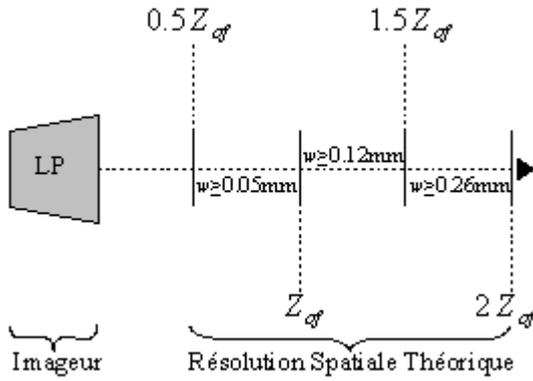


Figure 3.27 : Diagramme de résolution caméra LP

Le second type d'illustration utilise le lien entre le flux d'information, la résolution et le

schéma de codage. Pour rappel, l'IMM estimée ici est la quantité maximale d'information acheminable à travers le canal de communication. Ainsi, en se basant sur les résultats des travaux de Pavlidis et al. calculant la redondance intrinsèque des code-barres industriels provenant de leur schéma de codage [12], nous avons reporté au tableau 3.3 la largeur de barre correspondante pour trois types de codes usuels. Les résultats sont donnés pour une puissance de bruit égale à  $N_0=10^{-2}$  et  $N_0=10^{-3}$ .

		<b>UPC</b>	<b>Code128</b>	<b>Code93</b>
<b>LP</b>	$N_0=10^{-2}$	0.17 mm	0.2 mm	0.2 mm
	$N_0=10^{-3}$	0.08 mm	0.09 mm	0.09 mm
<b>MP</b>	$N_0=10^{-2}$	0.13 mm	0.15 mm	0.15 mm
	$N_0=10^{-3}$	0.06 mm	0.07 mm	0.07 mm

*Tableau 3.3 Résolution spatiale théorique de codes industriels*

### 3.4 Conclusion

L'analyse optique et radiométrique du processus affichage – lecture d'un code-barres par une caméra linéaire nous a amené à définir pour le canal de transmission un modèle à interférence entre symboles bruité. Cette équation du canal est paramétrée par les caractéristiques du système, tant optiques que spatiales, de sorte que son comportement lors d'une transmission est entièrement défini par une matrice de canal synthétisant l'ensemble des distorsions. L'écriture matricielle nous a permis de mettre en évidence la markovianité de la mémoire du système, propriété qui s'est révélée être l'articulation entre l'étape de modélisation du système et l'évaluation de ses performances. En effet, l'étude a été inscrite dans le cadre de la théorie de l'information de sorte que l'information mutuelle moyenne soit la mesure objective de l'efficacité de transmission. De par la markovianité, nous avons proposé un algorithme d'estimation de l'IMM basé sur l'étape "forward" du BCJR couplé à une méthode de Monte Carlo. A partir de cette estimation du flux théorique d'information, nous avons pu définir de nouveaux critères de performances : la profondeur de champ théorique et la résolution spatiale théorique. S'inscrivant comme complémentaires aux critères classiques

basés principalement sur la géométrie, ils illustrent concrètement l'apport de ce travail qui s'est voulu être la création d'un pont entre différentes disciplines: l'optique, la théorie de l'information et l'ingénierie des systèmes. Dans le chapitre suivant, nous abordons le cas bidimensionnel en conservant la même démarche. En nous appuyant sur les éléments introduits au chapitre 2, nous construisons un modèle du canal de communication lorsque l'émetteur est un code 2D capturé par une caméra numérique. Ce modèle, qui s'avérera similaire au cas 1D, nécessitera pour autant une approche différente pour la transposition du calcul de l'IMM. En ajoutant une dimension à la transmission, la distribution de probabilité du champ du code émis conditionnellement aux observations est une distribution de Gibbs, et ce sera par la physique statistique que nous approcherons l'évaluation du flux d'information.





## **Chapitre 4**

### **Modélisation et analyse des transmissions par code 2D**

L'utilisation de séquences de barres noires et blanches a été la première implémentation des travaux de Silver et Woodland [1] pour la transmission graphique d'un identifiant. Historiquement, ce choix de coder l'information dans le sens de la longueur correspondait à un contexte technologique peu pourvu en systèmes d'acquisition d'image. Les premiers lecteurs utilisés industriellement furent des lecteurs laser [66] permettant l'acquisition à des distances suffisamment longues (quelques dizaines de centimètres) pour un coût relativement faible. Les progrès dans les technologies à semi-conducteurs furent toutefois rapides, notamment concernant l'imagerie avec les capteurs CCD et CMOS [8, 9, 10]. C'est ainsi que portées à la fois par l'enthousiasme engrangé par le code-barres et les améliorations technologiques, de nouvelles méthodes d'agencement graphiques furent proposées, permettant de surcroît de transmettre plus d'information.

L'une des premières solutions a été trivialement d'empiler plusieurs codes-barres [14]. Plus précisément, une surface rectangulaire est partagée en plusieurs bandes horizontales où des codes-barres de hauteur réduite sont placés par impression ou affichage (figure 4.1). Un tel système de lecture peut être lu par un lecteur 1D identique à ceux utilisés pour les codes-barres. L'acquisition complète du code empilé s'effectue en faisant "glisser" le lecteur sur sa hauteur, ce qui correspond à une lecture séquentielle de chacun des codes-barres le composant. Par conséquent, la méthodologie et les outils d'analyse proposés au chapitre 3 de ce manuscrit s'appliquent directement.

La seconde évolution de la symbologie, cette fois majeure, est le codage matriciel. Le code est une matrice de points noirs et blancs où l'information à transmettre est répartie sur la totalité de la surface allouée : c'est le code 2D. Ces codes bidimensionnels sont destinés à être lus par une caméra effectuant l'acquisition sur une surface et non plus sur une seule ligne photosensible. De nombreuses symbologies ont été proposées depuis la fin des années 80, citons par exemple le DotCode de Philips [4], le Datamatrix [5], le QR Code [6] ou encore l'Aztec [7]. La plupart des codes 2D n'utilisent pas la totalité de leur surface pour l'information, une partie plus ou moins conséquente décrivant une cible est dédiée à la reconnaissance du code par le lecteur. Pour autant, le principe de codage matriciel subsiste et c'est à l'étude de ces "codes-points", par extension aux "codes-barres", que nous dédions ce chapitre.

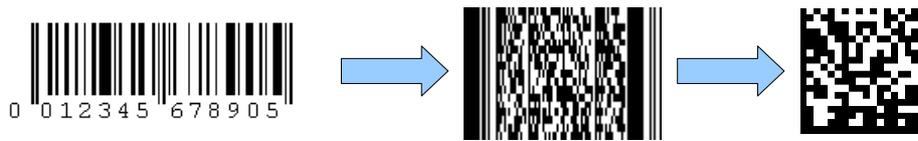


Figure 4.1 Evolution de l'occupation spatiale des codes graphiques : 1D (code EAN13), 1D empilé (PDF417), 2D (Datamatrix).

A l'instar du cas monodimensionnel nous adoptons la nomenclature suivante : nous notons  $C^{2D}$  le code bidimensionnel composé des éléments (ou modules)  $C_{i,j}^{2D}$ ,  $(i, j) \in \{1, 2, \dots, n\}$ . Ces modules sont de forme quelconque (carré, rectangulaire, hexagonale, etc.) et créent un partitionnement de  $C^{2D}$  en mosaïque. Le code est présenté sur un plan de couleur blanche définissant l'équivalent d'une marge autour de  $C^{2D}$ . Pour simplifier les futures notations, cette marge est morcelée en modules neutres de même couleur entourant le code et l'on a alors  $(i, j) \in \{0, 1, \dots, n+1\}$  (figure 4.2).

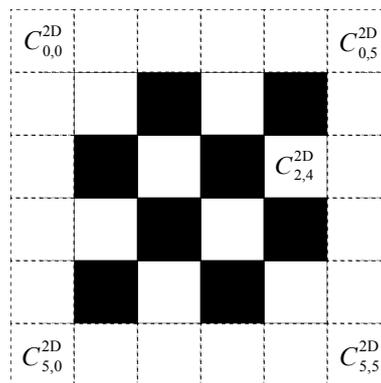


Figure 4.2 Nomenclature d'un code 2D dans le cas de modules carrés

La matrice  $C_{i,j}^{2D}$  est l'équivalent graphique de la matrice de symboles binaires  $X_{i,j}$ , matrice d'entrée du système et l'on pose  $C^{2D}(X_{i,j}) = C_{i,j}^{2D}$ . La marge du code quant à elle ne porte pas d'information. Pour faciliter l'écriture nous assignons toutefois à ses modules un symbole binaire constant égal à 1.

En suivant la même démarche que pour les systèmes de transmission graphiques monodimensionnels (chapitre 3), nous nous basons sur les principes optiques et radiométriques (chapitre 2) pour établir la relation entrée – sortie du processus de

transmission. Le cheminement de la modélisation du canal est présenté aux paragraphes suivants.

#### 4.1 Modélisation du canal de transmission

Le système de transmission est constitué d'un code bidimensionnel  $C^{2D}$ , dont le support est plan et noté  $\Pi$  (figure 4.3). Le centre de  $C^{2D}$  est le centre  $O$  du repère objet  $R(O, \vec{x}, \vec{y}, \vec{z})$ .

Le code est présenté devant une caméra munie d'un diaphragme de fonction de forme  $D(x', y')$  et de longueur focale  $f$ . Le bloc optique de la caméra est modélisé par une lentille mince (c.f. paragraphe 2.2.1) dont le centre optique est le centre  $O'$  du repère image  $R'(O', \vec{x}', \vec{y}', \vec{z}')$  et l'axe optique est l'axe  $[O' \vec{z}']$ . Le capteur d'image bidimensionnel est situé sur un plan  $\Pi'$  perpendiculaire à l'axe optique et coupant celui-ci en  $[0, 0, -z_c]^T$  dans le repère image. Par convention, les repères  $R$  et  $R'$  vérifient les conditions :

$$O=O' \text{ et } \{\vec{x}=-\vec{x}'; \vec{y}=\vec{y}'; \vec{z}=-\vec{z}'\}$$

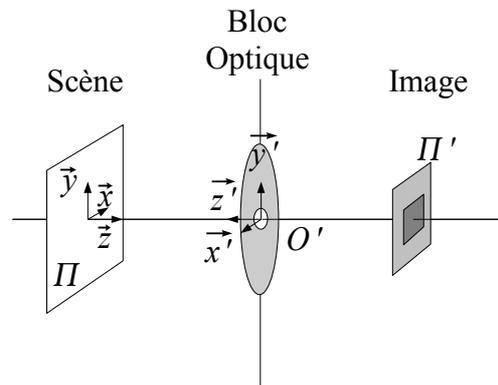


Figure 4.3 Schéma du système 2D

##### 4.1.1 Formation d'image et acquisition

L'équation du canal est déterminée à partir de l'analyse du processus de formation

d'image du bloc optique et de l'acquisition de l'éclairement par les pixels du capteur. Ces principes sont similaires à ceux présentés au chapitre 3 pour les systèmes 1D, c'est pourquoi nous invitons le lecteur à consulter le paragraphe 3.1 pour le détail des développements. L'essentiel est que la caméra effectue une projection perspective de la scène observée, et que le flou est modélisé par la convolution de l'éclairement avec la réponse impulsionnelle optique de la caméra. Les pixels quant à eux intègrent l'éclairement frappant leur surface, opération modélisée par un produit de convolution.

Ainsi le capteur composé de  $N_l$  lignes et  $N_c$  colonnes délivre une matrice de signaux  $s_{i,j}$ , associés aux pixels  $p_{i,j}$  avec  $1 \leq i \leq N_l$  et  $1 \leq j \leq N_c$ . Cette matrice est considérée comme l'information utile récupérée à l'issue de la transmission. On a :

$$s_{i,j} = RIC_{p_{i,j}} ** [RIO(\mathbf{M}') ** Ec'(\mathbf{M}')]_{\mathbf{M}' \in p_{i,j}} \quad (4.1)$$

avec  $Ec'(\mathbf{M}')$  l'éclairement de la projection sur le capteur,  $RIO(\mathbf{M}')$  la réponse impulsionnelle optique, et  $RIC_{p_{i,j}}$  la réponse impulsionnelle du capteur au pixel  $p_{i,j}$ .

#### 4.1.2 Equation du canal : construction

Les modules du code  $C^{2D}$  sont considérés comme des sources lambertiennes. La luminance énergétique est donc constante à leur surface et l'on pose :

$$[L(\mathbf{M})]_{\mathbf{M} \in C_{k,l}^{2D}} = X_{k,l} \quad (4.2)$$

avec  $X_{k,l} = \{0; 1\}$

Dès lors, en se basant sur (2.48) et (4.1), l'expression du niveau de gris  $s_{i,j}$  des pixels peut être développée en segmentant l'opération de convolution. Cette astuce a pour but de mettre en avant la participation de chacun des modules  $C_{k,l}^{2D}$  et leur image  $C_{k,l}'^{2D}$ . Nous obtenons la somme discrète suivante :

$$s_{i,j} = \sum_{k,l=0}^{n+1} \left( \kappa \cdot \frac{S_D}{Z_c^2} \cdot \left[ \cos^4 \alpha_M \cdot 1_{C_{k,l}^{2D}}(\mathbf{M}') ** RIC_{p_{i,j}} ** RIO(\mathbf{M}') \right]_{\mathbf{M}' \in p_{k,l}} \cdot X_{k,l} \right) \quad (4.3)$$

où  $\kappa$  est un facteur d'échelle lié aux caractéristiques du détecteur (sensibilité, temps

d'intégration) et  $\mathbf{M}'$  est le projeté de  $\mathbf{M}$  et  $S_D$  est la surface d'ouverture du diaphragme.

On considère que les symboles reçus sont les niveaux de gris moyens de chacun des modules de  $C'^{2D}$ . Le moyennage est effectué à partir des  $s_{i,j}$  des pixels intégrant ces modules. Les entrée et sortie du système sont ainsi deux matrices de taille respective  $(n+2)^2$  et  $(n)^2$ , et le niveau de gris moyen  $\bar{s}_{k,l}$  du module  $C'^{2D}_{k,l}$  est défini par :

$$\bar{s}_{k,l} = \sum_{i,j=1}^{N_i, N_c} g_{k,l,i,j} \cdot s_{i,j} \quad (4.4)$$

où  $g_{k,l,i,j}$  est l'aire normalisée de recouvrement du pixel  $p_{i,j}$  et du module  $C'^{2D}_{k,l}$ .

On a :

$$\begin{aligned} \bar{s}_{k,l} &= \sum_{i,j=1}^{N_i, N_c} g_{k,l,i,j} \cdot s_{i,j} \\ &= \sum_{i,j=1}^{N_i, N_c} \left[ g_{k,l,i,j} \cdot \left[ \sum_{k_0, l_0=0}^{n+1} \left( \kappa \cdot \frac{S_D}{z_c^2} \cdot \left[ \cos^4 \alpha_M \cdot 1_{C'^{2D}_{k_0, l_0}}(\mathbf{M}') ** RIC_{p_{i,j}} ** RIO(\mathbf{M}') \right]_{\mathbf{M}' \in p_{i,j}} \cdot X_{k_0, l_0} \right) \right] \right] \\ &= \sum_{k_0, l_0=0}^{n+1} \left[ \left[ \sum_{i,j=1}^{N_i, N_c} \left( \kappa \cdot \frac{S_D}{z_c^2} \cdot \left[ \cos^4 \alpha_M \cdot 1_{C'^{2D}_{k_0, l_0}}(\mathbf{M}') ** RIC_{p_{i,j}} ** RIO(\mathbf{M}') \right]_{\mathbf{M}' \in p_{i,j}} \cdot g_{k,l,i,j} \right) \right] \cdot X_{k_0, l_0} \right] \\ \bar{s}_{k,l} &= \sum_{k_0, l_0=0}^{n+1} h_{k,l,k_0, l_0} \cdot X_{k_0, l_0} \quad (4.5) \end{aligned}$$

A l'issu du processus d'acquisition, les perturbations inhérentes aux imperfections des éléments du système sont modélisées par un bruit. Ce bruit, associé à chaque pixel, est considéré comme additif, blanc et gaussien (ABG) de variance  $\sigma_p^2$ . Les symboles reçus étant une moyenne sur les pixels intégrant, le bruit associé à chaque symbole est également ABG mais dont la variance  $\sigma^2$  est égal au rapport entre  $\sigma_p^2$  et le nombre de pixels intégrant.

Finalement, la matrice des symboles reçus  $\mathbf{Y}$  est la somme d'une combinaison linéaire des éléments de la matrices des symboles émis  $\mathbf{X}$  et d'une matrice de bruit  $\mathbf{B}$ . En utilisant (4.5), l'équation du canal de transmission par module est :

$$Y_{k,l} = \sum_{k_0, l_0=0}^{n+1} h_{k,l,k_0, l_0} \cdot X_{k_0, l_0} + B_{k,l} \quad (4.6)$$

### 4.1.3 Equation du canal : écriture matricielle, caractéristiques statistiques

Le cadre de travail dans lequel l'étude est placée est celui de la théorie de l'information. L'analyse du système est donc basée sur les caractéristiques statistiques du canal de communication. Afin de déterminer les distributions et densités de probabilité des processus du système, nous adoptons une écriture matricielle de l'équation du canal. La disposition naturelle des données  $\{X_{k,l}\}$ , des observations  $\{Y_{k,l}\}$  et du bruit  $\{B_{k,l}\}$  est ainsi modifiée : d'une structure en matrice à deux dimensions on génère les vecteurs  $\mathbf{X}$ ,  $\mathbf{Y}$  et  $\mathbf{B}$ . L'algorithme de réagencement :

- est appliqué à l'identique sur les trois matrices
- n'introduit pas de répétition
- n'introduit pas d'omission

Deux variantes simples d'un tel algorithme sont l'empilement des colonnes des matrices ou l'empilement des transposées des lignes. C'est cette dernière qui est retenue pour la suite de ce chapitre.

Dès lors, une matrice d'interférence  $\mathbf{A}$  est construite à partir des coefficients  $\{h_{k,l,k_0,l_0}\}$ . Cette matrice est rectangulaire, de  $(n)^2$  lignes et  $(n+2)^2$  colonnes. Elle caractérise le canal de transmission en synthétisant l'ensemble de ses distortions. Ainsi, l'équation (4.6) se réécrit :

$$\mathbf{Y} = \mathbf{A} \mathbf{X} + \mathbf{B} \quad (4.7)$$

avec

- $\mathbf{Y} = [Y_{1,1} \quad \dots \quad Y_{1,n} \quad Y_{2,1} \quad \dots \quad Y_{2,n} \quad \dots \quad Y_{n,1} \quad \dots \quad Y_{n,n}]^T$

- $$\mathbf{A} = \begin{bmatrix} h_{1,1,0,0} & h_{1,1,0,1} & \cdots & h_{1,1,0,n+1} & \cdots & h_{1,1,n+1,0} & \cdots & h_{1,1,n+1,n+1} \\ h_{1,2,0,0} & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & h_{1,2,n+1,n+1} \\ \vdots & & & & & & & \vdots \\ h_{1,n,0,0} & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & h_{1,n,n+1,n+1} \\ h_{2,1,0,0} & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & h_{2,1,n+1,n+1} \\ \vdots & & & & & & & \vdots \\ h_{2,n,0,0} & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & h_{2,n,n+1,n+1} \\ \vdots & & & & & & & \vdots \\ h_{n,1,0,0} & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & h_{n,1,n+1,n+1} \\ \vdots & & & & & & & \vdots \\ h_{n,n,0,0} & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & h_{n,n,n+1,n+1} \end{bmatrix}$$
- $$\mathbf{X} = [X_{0,0} \quad X_{0,1} \quad \cdots \quad X_{0,n+1} \quad X_{1,0} \quad \cdots \quad X_{1,n+1} \quad \cdots \quad X_{n+1,0} \quad \cdots \quad X_{n+1,n+1}]^T$$
- $$\mathbf{B} = [B_{1,1} \quad \cdots \quad B_{1,n} \quad B_{2,1} \quad \cdots \quad B_{2,n} \quad \cdots \quad B_{n,1} \quad \cdots \quad B_{n,n}]^T$$

Nous considérons à ce stade que la source du système de transmission est binaire équiprobable, en d'autres termes  $P(\mathbf{X}=\mathbf{x})=P_{\mathbf{X}}(\mathbf{x})=(1/2)^{n^2}$ . Le bruit étant gaussien et blanc, la densité conditionnelle  $p_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x})$  est donnée par :

$$p_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x}) = \left( \frac{1}{\sqrt{2\pi\sigma^2}} \right)^{n^2} e^{-\frac{1}{2\sigma^2} \sum_{k=1}^{n^2} (y_k - [A\mathbf{x}]_k)^2} \quad (4.8)$$

soit

$$p_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x}) = \left( \frac{1}{\sqrt{2\pi\sigma^2}} \right)^{n^2} e^{-\frac{1}{2\sigma^2} \|\mathbf{y} - A\mathbf{x}\|^2} \quad (4.9)$$

où  $\|\mathbf{u}\|$  est la norme du vecteur  $\mathbf{u}$ .

On en déduit que

$$p_{\mathbf{Y}}(\mathbf{y}) = \left( \frac{1}{2\sqrt{2\pi\sigma^2}} \right)^{n^2} \sum_{\{\mathbf{x}\}} e^{-\frac{1}{2\sigma^2} \|\mathbf{y} - A\mathbf{x}\|^2} \quad (4.10)$$

Sachant que classiquement :

$$P_{\mathbf{X}|\mathbf{Y}}(\mathbf{x}|\mathbf{y}) = \frac{p_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x})P_{\mathbf{X}}(\mathbf{x})}{p_{\mathbf{Y}}(\mathbf{y})} \quad (4.11)$$

la densité *a posteriori*  $p_{\mathbf{X}|\mathbf{Y}}(\mathbf{x}|\mathbf{y})$  est définie par :

$$P_{\mathbf{x}|\mathbf{y}}(\mathbf{x}|\mathbf{y}) = \frac{e^{-\frac{1}{2\sigma^2}\|\mathbf{y}-\mathbf{A}\mathbf{x}\|^2}}{\sum_{\{\mathbf{x}\}} e^{-\frac{1}{2\sigma^2}\|\mathbf{y}-\mathbf{A}\mathbf{x}\|^2}} = \frac{1}{Z} e^{-\frac{1}{2\sigma^2}\|\mathbf{y}-\mathbf{A}\mathbf{x}\|^2} \quad (4.12)$$

où  $Z = \sum_{\{\mathbf{x}\}} e^{-\frac{1}{2\sigma^2}\|\mathbf{y}-\mathbf{A}\mathbf{x}\|^2}$  est appelée *fonction de partition*.

La distribution (4.12) est celle d'un champ de Gibbs [74, 75] dont la fonction d'énergie est  $U_{\mathbf{x}|\mathbf{y}} = \frac{1}{2\sigma^2}\|\mathbf{y}-\mathbf{A}\mathbf{x}\|^2$ . L'équation (4.12) se réécrit:

$$P_{\mathbf{x}|\mathbf{y}}(\mathbf{x}|\mathbf{y}) = \frac{1}{Z} e^{-U_{\mathbf{x}|\mathbf{y}}} \quad (4.13)$$

La probabilité *a posteriori* du système de transmission par code 2D est donc régi par une loi de Gibbs munie de la fonction d'énergie  $U_{\mathbf{x}|\mathbf{y}}$ . Or la réécriture des matrices de données et d'observations en vecteurs a été purement formelle, et ceci dans le seul but de faciliter l'écriture de l'équation de canal de communication et la mise en évidence de ses propriétés. Les deux agencements étant équivalents, le processus conditionnel  $\mathbf{X}|\mathbf{Y}$  est un champ de Gibbs sur l'espace des configurations possibles  $\Omega^{n^2}$ . Le théorème d'Hammersley – Clifford [74, 76, 77] établit de plus qu'une distribution de Gibbs est équivalente à un champ aléatoire markovien. Rappelons à ce titre que la variable  $\mathbf{C} = \{c_i\}$ ,  $i \in [1; n^2]$ , est un champ aléatoire markovien sur l'ensemble des configurations possibles  $\Omega^{n^2}$  si :

- $P_{\mathbf{C}}(\mathbf{c}) > 0$
- $P_{\mathbf{C}}(c_i | \mathbf{c}_{\mathcal{V}(i)}) = P_{\mathbf{C}}(c_i | \mathbf{c}_{\mathcal{V}(i)})$  où  $\mathbf{c}_{\mathcal{V}(i)}$  est l'ensemble des sites de  $\mathbf{C}$  excepté le site  $c_i$ , et  $\mathcal{V}(i)$  est l'ensemble des sites appartenant à un voisinage  $\mathcal{V}(i)$  de  $c_i$

Il vient donc que le processus conditionnel  $\mathbf{X}|\mathbf{Y}$ , *a posteriori* du système de transmission, est un champ aléatoire markovien associé à un système de voisinage  $\mathcal{V}(i)$ , et dont la probabilité jointe est la distribution de Gibbs (4.13).

Le développement de la fonction d'énergie du champ permet de mettre en évidence les liens locaux entre les différents éléments de la matrice  $\mathbf{X}|\mathbf{Y}$  (ou liens intersites).

Soit la matrice  $\mathbf{R} = \mathbf{A}^T \mathbf{A}$ .  $\mathbf{R}$  peut être décomposée en une matrice  $\mathbf{D}$  formée de la première diagonale et une matrice  $\mathbf{R}_D$  égale à  $\mathbf{R}$  dont la première diagonale est remplacée par des zéros.

$$\mathbf{R} = \mathbf{R}_D + \mathbf{D} \quad (4.14)$$

L'énergie  $U_{\mathbf{x}|\mathbf{y}}$  s'écrit alors :

$$U_{\mathbf{x}|\mathbf{y}} = \|\mathbf{y}\|^2 + \mathbf{x}^T \mathbf{D} \mathbf{x} + \mathbf{x}^T \mathbf{R}_D \mathbf{x} - 2 \mathbf{y}^T \mathbf{A} \mathbf{x} \quad (4.15)$$

$U_{\mathbf{x}|\mathbf{y}}$  est donc la somme d'une énergie associée aux modules, ou sites dans la terminologie markovienne :

$$U'_{\mathbf{x}|\mathbf{y}} = \|\mathbf{y}\|^2 + \mathbf{x}^T \mathbf{D} \mathbf{x} - 2 \mathbf{y}^T \mathbf{A} \mathbf{x} \quad (4.16a)$$

et d'une énergie liant les modules entre eux, ou intersites :

$$U''_{\mathbf{x}|\mathbf{y}} = \mathbf{x}^T \mathbf{R}_D \mathbf{x} \quad (4.16b)$$

Ainsi, la distribution de  $\mathbf{X}|\mathbf{Y}$  (4.13) se met sous la forme :

$$P_{\mathbf{X}|\mathbf{Y}}(\mathbf{x}|\mathbf{y}) = \frac{1}{Z} e^{-U'_{\mathbf{x}|\mathbf{y}} - U''_{\mathbf{x}|\mathbf{y}}} \quad (4.17)$$

Celle-ci peut aisément être développée en produit d'exponentielles associées aux sites et aux liens intersites:

$$P_{\mathbf{X}|\mathbf{Y}}(\mathbf{x}|\mathbf{y}) = \frac{1}{Z} \prod_{\{k,l\}} e^{-U'_{\mathbf{x}|\mathbf{y}}(\{k,l\})} \prod_{\{k,l\} \Leftrightarrow \{k_0,l_0\}} e^{-U''_{\mathbf{x}|\mathbf{y}}(\{k,l\} \Leftrightarrow \{k_0,l_0\})}$$

Il s'agit donc d'un produit de facteurs, forme naturelle de la distribution  $P(\mathbf{X}|\mathbf{Y})$  que nous utiliserons plus tard dans ce mémoire (c.f. § 4.3.3.1).

Dans les paragraphes suivants nous présentons l'approche adoptée pour le calcul de l'IMM. Celle-ci se base classiquement sur les distributions et densités du système, notamment  $p_Y(\mathbf{y})$ . La particularité réside dans le fait que la densité de probabilité des observations est proportionnelle à la fonction de partition du champ  $\mathbf{X}|\mathbf{Y}$ . On montre ainsi que le calcul de l'IMM passe par l'évaluation d'une énergie, l'énergie de Helmholtz, grandeur fondamentale de la physique statistique.

## 4.2 Information Mutuelle Moyenne

L'analyse du processus de transmission de l'information portée par le pictogramme  $C^{2D}$  nous a mené à une modélisation linéaire du canal. Ce modèle est similaire au cas monodimensionnel, notamment par l'introduction d'une matrice de canal synthétisant l'ensemble des distorsions. Dans le troisième chapitre de ce mémoire, nous avons présenté l'information mutuelle moyenne comme un outil de mesure des performances du système (paragraphe 3.2.1). En s'inscrivant dans le cadre mathématique de la théorie de l'information, l'IMM s'est révélée être une mesure objective pour l'analyse du système, complétant les critères optiques classiques tels que la profondeur de champ et surtout dressant naturellement un pont avec le codage de l'information. C'est ainsi logiquement que nous adoptons de nouveau l'IMM pour le cas 2D.

Malgré la dimension supplémentaire, l'IMM reste une différence d'entropie. Ainsi :

$$IMM(X_{1:n,1:n}; Y_{1:m,1:m}) = H(Y_{1:m,1:m}) - H(Y_{1:m,1:m} | X_{1:n,1:n}) \quad (4.18)$$

où, dans le cas de variables aléatoires discrètes, l'entropie de la séquence  $X_{1:n,1:n}$  est par définition égale à :

$$H(X_{1:n,1:n}) = - \sum_{\{x_{1:n,1:n}\}} P(x_{1:n,1:n}) \log_2 P(x_{1:n,1:n}) \quad (4.19)$$

et l'entropie conditionnelle de  $Y_{1:m,1:m}$  par rapport à  $X_{1:n,1:n}$  est

$$H(Y_{1:m,1:m} | X_{1:n,1:n}) = - \sum_{\{y_{1:m,1:m}\}} \sum_{\{x_{1:n,1:n}\}} P(x_{1:n,1:n}, y_{1:m,1:m}) \log_2 P(y_{1:m,1:m} | x_{1:n,1:n}) \quad (4.20)$$

Dans le cas continu, l'entropie différentielle de la séquence aléatoire  $Z_{1:l,1:l}$  est une intégration multiple :

$$H(Z_{1:l,1:l}) = - \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p(z_{1:l,1:l}) \log_2 p(z_{1:l,1:l}) dz_1 \cdots dz_l \quad (4.21)$$

où  $p(z_{1:l,1:l})$  est la densité de probabilité jointe de  $Z_{1:l,1:l}$ .

### 4.2.1 Entropie conditionnelle $H(\mathbf{Y}|\mathbf{X})$

Le canal de transmission 2D est régi par l'équation (4.7) que nous rappelons ici :

$$\mathbf{Y} = \mathbf{A} \mathbf{X} + \mathbf{B} \quad (4.22)$$

L'entropie conditionnelle  $H(\mathbf{Y}|\mathbf{X})$  est totalement définie par les caractéristiques du bruit.

Les vecteurs d'observation et du bruit étant de taille  $n^2$ , on a le résultat ([13] p225, [3] p107):

$$H(\mathbf{Y}|\mathbf{X}) = H(\mathbf{B}) = \frac{n^2}{2} \log_2(2\pi e\sigma^2) \quad (4.23)$$

où  $\sigma^2$  est pour rappel la variance du bruit.

### 4.2.2 Entropie $H(\mathbf{Y})$

En se référant à l'équation (4.21) l'entropie du vecteur d'observation  $\mathbf{Y}$  est égale à :

$$H(\mathbf{Y}) = - \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} p_Y(\mathbf{y}) \log_2 p_Y(\mathbf{y}) d\mathbf{y} \quad (4.24)$$

Or nous avons montré au paragraphe 4.1.2 que  $p_Y(\mathbf{y})$  est définie par (équation (4.10)):

$$p_Y(\mathbf{y}) = \left( \frac{1}{2\sqrt{2\pi\sigma^2}} \right)^{n^2} \cdot Z$$

avec  $Z = \sum_{\{\mathbf{x}\}} e^{\frac{-1}{2\sigma^2} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|^2}$ . Le calcul de l'entropie passe donc par l'estimation de  $Z$ , fonction de partition du champ markovien  $\mathbf{X}|\mathbf{Y}$  (paragraphe 4.1.2) dont la distribution est une loi de Gibbs donnée par (équation (4.17)):

$$P_{\mathbf{X}|\mathbf{Y}}(\mathbf{x}|\mathbf{y}) = \frac{1}{Z} e^{-U'_{\mathbf{x}|\mathbf{y}} - U''_{\mathbf{x}|\mathbf{y}}}$$

avec  $U'_{\mathbf{x}|\mathbf{y}}$  la fonction d'énergie associée aux sites et  $U''_{\mathbf{x}|\mathbf{y}}$  la fonction d'énergie intersites (équations (4.16a) et (4.16b)).

Cette estimation de la fonction de partition est un fondamental de la physique statistique. Rappelons à ce titre que cette discipline consiste en l'étude par l'outil mathématique et la théorie de la probabilité d'ensembles de particules et de leurs interactions.

Les propriétés locales des éléments (échelle microscopique) sont ainsi mises en relation avec les propriétés globales du système (échelle macroscopique) permettant de passer d'une résolution à l'autre pour mesurer les effets de l'environnement ou la recherche d'un état d'équilibre. Un cas d'école est l'étude des matériaux ferromagnétiques et la recherche des transitions de phase, c'est à dire du passage entre états paramagnétique et diamagnétique [74, 75]. On montre qu'à l'équilibre thermodynamique la configuration d'un système de particules  $\mathbf{X}$  suit la distribution de probabilité suivante :

$$P_{\mathbf{X}}(\mathbf{x}) = \frac{1}{\sum_{\{\mathbf{x}\}} e^{\frac{-1}{k_B T} E_{\mathbf{x}}}} e^{\frac{-1}{k_B T} E_{\mathbf{x}}} = \frac{1}{Z} e^{\frac{-1}{k_B T} E_{\mathbf{x}}} \quad (4.25)$$

où  $k_B = 1.38 \cdot 10^{-23} \text{ J K}^{-1}$  est la constante de Boltzmann,  $T$  est la température exprimée en kelvins, et  $E_{\mathbf{x}}$  est l'énergie du champ  $\mathbf{X} = \mathbf{x}$ .

La fonction de partition  $Z$  dénombre ainsi l'ensemble des configurations possibles du système et l'on montre que ses principales grandeurs telles l'énergie moyenne ou l'entropie se déduisent de  $Z$ . Pour un plus ample développement de ses notions, importantes pour nos travaux mais en marge de leur objectif, nous invitons le lecteur à consulter [78, 79].

L'une de ces grandeurs caractéristiques dérivées de  $Z$  est l'énergie libre de Helmholtz  $EL_H = -\log Z$  qui va se révéler primordiale pour l'estimation de l'IMM de la chaîne de transmission 2D.

### 4.3 Calcul de l'IMM

Nous introduisons dans ce paragraphe les outils pour l'estimation de la fonction de partition de  $\mathbf{X}|\mathbf{Y}$  qui nous permet dans un deuxième temps d'évaluer l'IMM. Ce champ est analogue à un modèle d'Ising pour lesquels différentes stratégies d'estimation de la fonction de partition ont été publiées dans la littérature. Nous adoptons une approche variationnelle, qui vise à approcher l'énergie libre de Helmholtz en utilisant un algorithme itératif de minimisation d'une énergie. Ce choix est motivé par les performances reconnues de la méthode ainsi que de son utilisation des graphes factoriels, outils de plus en plus présents en télécoms. Finalement, à l'issue de la méthode variationnelle, on obtient une approximation de

l'énergie libre de  $X|Y$  d'où nous déduisons  $Z$  et donc  $p_Y(\mathbf{y})$ . L'algorithme global d'estimation de l'IMM est détaillé en fin de ce paragraphe.

### 4.3.1 Distribution du champ $X|Y$ et modèle d'Ising

La distribution jointe du champ aléatoire  $X|Y$  est définie par l'équation (4.17) que nous rappelons ici :

$$P_{X|Y}(\mathbf{x}|\mathbf{y}) = \frac{1}{Z} e^{-U'_{x|y} - U''_{x|y}}$$

avec  $U'_{x|y} = \|\mathbf{y}\|^2 + \mathbf{x}^T \mathbf{D} \mathbf{x} - 2 \mathbf{y}^T \mathbf{A} \mathbf{x}$  l'énergie associée aux modules du champ, et  $U''_{x|y} = \mathbf{x}^T \mathbf{R}_D \mathbf{x}$  l'énergie liant les modules deux à deux.

Nous nous sommes placés dans le cas où les modules ne peuvent prendre que deux valeurs avec un pictogramme noir et blanc. Le champ  $X|Y$  dont nous voulons calculer la fonction de partition est donc équivalent au modèle d'Ising [74]. Or il n'existe pas pour ce modèle de formulation analytique de  $Z$  [80, 81, 82], et le calcul direct de la somme devient rapidement dispendieux avec la taille de  $C^{2D}$ . Nous avons ainsi recours à une approximation de  $Z$  pour, au final, estimer l'IMM.

### 4.3.2 Estimation de la fonction de partition

L'une des approches rencontrées dans la littérature pour l'approximation de la fonction de partition d'un champ de Gibbs est l'utilisation de méthodes stochastiques. Ainsi différents algorithmes basés sur les méthodes de Monte Carlo sont proposés dans [83-87, 88, 89]. L'idée principale de ces propositions repose sur le calcul d'espérances de fonctionnelles dérivées de la distribution du champ. La complexité algorithmique de ces méthodes, ainsi que leur convergence sont comparées par Potamianos et Goutsias dans [90].

Une autre stratégie adoptée par la communauté pour atteindre la fonction de partition est l'application de méthodes variationnelles [91, 92] pour l'estimation de l'énergie libre  $EL_H$  d'un champ de Gibbs  $X$ . L'idée consiste à contourner la difficulté du calcul direct de la distribution  $P_X(\mathbf{x}) = e^{-E(\mathbf{x})}/Z$  en utilisant une distribution  $b(\mathbf{x})$  dite "d'essai", homogène à une probabilité, et pour laquelle on définit une *énergie libre variationnelle*  $EL_v(b)$  telle que :

$$EL_v(b) = En_v(b) - Ent_v(b) \quad (4.26)$$

où  $En_v(b)$  est l'énergie variationnelle moyenne :

$$En_v(b) = \sum_{\{\mathbf{x}\}} b(\mathbf{x}) E(\mathbf{x}) \quad (4.27)$$

et où  $Ent_v(b)$  est l'entropie variationnelle :

$$Ent_v(b) = - \sum_{\{\mathbf{x}\}} b(\mathbf{x}) \ln b(\mathbf{x}) \quad (4.28)$$

Dès lors il en découle directement le lien suivant entre l'énergie variationnelle et l'énergie de Helmholtz:

$$EL_v(b) = EL_H + D_{KL}(b||P) \quad (4.29)$$

où  $D_{KL}(b||P) = \sum_{\{\mathbf{x}\}} b(\mathbf{x}) \ln \frac{b(\mathbf{x})}{P_{\mathbf{x}}(\mathbf{x})}$  est la divergence de Kullback – Leibler [13]. Sachant que  $D_{KL}(b||P)$  est toujours positive ou nulle ([13] théorème 2.6.3), on a  $EL_v(b) \geq EL_H$ , avec égalité lorsque  $b(\mathbf{x}) = P_{\mathbf{x}}(\mathbf{x})$ . Minimiser l'énergie libre variationnelle par rapport à  $b(\mathbf{x})$  est donc un moyen sûr d'atteindre  $EL_H$  et obtenir de surcroît  $P_{\mathbf{x}}(\mathbf{x})$ . Notre intérêt pour cette approche est dû au fait que l'une des méthodes performantes pour la minimisation de  $EL_v(b)$  récemment publiée par Yedidia et al. [92] s'appuie sur les graphes factoriels, outil de plus en plus présent dans le monde des télécoms [93]. Ce formalisme permet de répondre élégamment au problème du choix de  $b(\mathbf{x})$  en introduisant le principe de *passage de messages*.

### 4.3.3 Energie libre variationnelle et graphes

#### 4.3.3.1 Graphes factoriels et Propagation de Croyance

Nous introduisons dans un premier temps les graphes factoriels. Soit le vecteur aléatoire  $\mathbf{X} = \{X_1, \dots, X_N\}$  dont les réalisations  $\{X_1 = x_1, \dots, X_n = x_N\}$  sont dans l'espace discret  $\Gamma^N$ . Nous considérons que la distribution de probabilité associée à  $\mathbf{X}$  peut se mettre sous la forme :

$$P_{\mathbf{x}}(\mathbf{x}) = \frac{1}{Z} \prod_{\{a\}} f_a(\mathbf{x}_a) \quad (4.30)$$

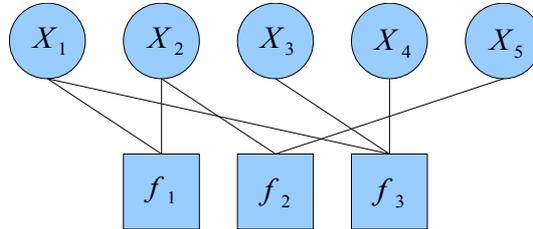
où  $a$  indexe  $M$  fonctions  $\{f_a\}$  dont l'argument est le vecteur  $\mathbf{x}_a$  qui constitue un sous-ensemble de  $\{x_1, \dots, x_N\}$ . Les  $\{f_a\}$  sont supposées positives et finies et  $Z$  est une constante de normalisation.

Notons que la forme développée de  $P(\mathbf{X}|\mathbf{Y})$  dans (4.17) est analogue à (4.30)

$$P_{\mathbf{x}|\mathbf{y}}(\mathbf{x}|\mathbf{y}) = \frac{1}{Z} \prod_{\{k,l\}} e^{-U'_{x|y}(\{k,l\})} \prod_{\{k,l\} \Leftrightarrow \{k_0,l_0\}} e^{-U''_{x|y}(\{k,l\} \Leftrightarrow \{k_0,l_0\})} = \frac{1}{Z} \prod_{\{a\}} f_a(\mathbf{x}_a|\mathbf{y})$$

où  $\{a\}$  regroupe l'ensemble des sites  $\{k, l\}$  et des liens inter-sites  $\{k, l\} \Leftrightarrow \{k_0, l_0\}$ .

Un graphe factoriel (GF) est un graphe bipartite représentant la distribution prototype (4.30). Il est constitué de noeuds de deux types, les uns associés aux variables  $\{x_i\}$  et les autres aux fonctions (ou facteurs)  $\{f_a\}$ , de plus la relation facteur – argument est symbolisé par un arc reliant les deux noeuds. Nous donnons en figure 4.4 une illustration de cette représentation.



$$P_{\mathbf{x}}(X_1, X_2, X_3, X_4, X_5) = Z^{-1} \cdot f_1(X_1, X_2) \cdot f_2(X_2, X_3, X_4, X_5) \cdot f_3(X_1, X_3, X_4)$$

Figure 4.4 Graphe factoriel

Le principe de marginalisation appliquée à (4.30) revient à une sommation sur un sous-ensemble de  $\{x_1, \dots, x_N\}$ . La probabilité marginale du vecteur  $\mathbf{x}_S$  est ainsi donnée par :

$$P_{\mathbf{x}_S}(\mathbf{x}_S) = \sum_{\{\mathbf{x} \setminus \mathbf{x}_S\}} P_{\mathbf{x}}(\mathbf{x}) \quad (4.31)$$

où  $\sum_{\{\mathbf{x} \setminus \mathbf{x}_S\}}$  est la sommation sur l'ensemble des noeuds qui ne sont pas dans le sous-ensemble

$S$  englobant le vecteur  $\mathbf{x}_S$ .

Notons qu'en définissant une énergie  $E_{GF}(\mathbf{x}) = -\sum_{a=1}^M \ln f_a(\mathbf{x}_a)$ , l'équation (4.30) est homogène à une distribution de Gibbs. La distribution  $P(\mathbf{X}|\mathbf{Y})$  a ainsi la forme de (4.30).

Conjointement à ce formalisme, l'algorithme de propagation de croyance (PC) est une méthode de calcul de marginales sur un graphe factoriel [92]. Il se traduit par l'introduction de messages, dont la circulation et la combinaison permet de déduire une croyance (*i.e.* probabilité) quant à l'état d'un noeud ou groupe de noeuds. Ces messages sont de deux types :

- le message  $m_{a \rightarrow i}(x_i)$  du noeud du facteur  $a$  au noeud de la variable  $i$ , qui est une assertion sur l'état du noeud  $i$  par rapport à tous ceux possibles et en se basant sur le facteur  $f_a$
- le message  $n_{i \rightarrow a}(x_i)$  du noeud de la variable  $i$  au noeud du facteur  $a$ , qui est une assertion sur l'état du noeud  $i$  par rapport à tous ceux possibles et en se basant sur toute l'information disponible si ce n'est le facteur  $f_a$ .

Les messages sont mis à jour selon les règles suivantes :

$$n_{i \rightarrow a}(x_i) = \prod_{c \in V(i) \setminus a} m_{c \rightarrow i}(x_i) \quad (4.32)$$

et

$$m_{a \rightarrow i}(x_i) = \sum_{\mathbf{x}_a \setminus x_i} f_a(\mathbf{x}_a) \prod_{j \in V(a) \setminus i} n_{j \rightarrow a}(x_j) \quad (4.33)$$

où  $V(i) \setminus a$  est l'ensemble des noeuds dans le voisinage de  $i$  excepté  $a$ .

Signalons que cette description de l'algorithme PC est également appelé dans la littérature "algorithme Somme – Produit" de par l'équation (4.33) [94].

Nous définissons la croyance  $b_i(x_i)$  sur le noeud  $i$  comme étant l'approximation de la marginale  $P_{x_i}(x_i)$  et l'on a :

$$b_i(x_i) \propto \prod_{a \in V(i)} m_{a \rightarrow i}(x_i) \quad (4.34)$$

où la proportionnalité implique une normalisation.

La croyance jointe  $b_a(\mathbf{x}_a)$  est quant à elle déduite de :

$$b_a(\mathbf{x}_a) \propto f_a(\mathbf{x}_a) \prod_{i \in V(a)} n_{i \rightarrow a}(x_i) \quad (4.35)$$

$$b_a(\mathbf{x}_a) \propto f_a(\mathbf{x}_a) \prod_{i \in V(a)} \prod_{c \in V(i) \setminus a} m_{c \rightarrow i}(x_i) \quad (4.36)$$

Le calcul des marginales du graphe est effectué itérativement, en initialisant les messages  $m$  et  $n$  (le plus souvent à 1), et en calculant les équations de mise à jour ((4.32) et (4.33)) et d'approximation ((4.34) et (4.36)) jusqu'à convergence ou suivant un critère d'arrêt.

Signalons enfin que l'équation (4.34) peut se déduire de (4.36) en appliquant une marginalisation :

$$b_i(x_i) = \sum_{\mathbf{x}_a \setminus x_i} b_a(\mathbf{x}_a) \quad (4.37)$$

Les mécanismes de la propagation de croyance, avec notamment le passage de message, va être reproduit et généralisé sur une super – structure se greffant sur le graphe factoriel : le graphe de régions. Cette évolution est la clef de voûte pour le calcul de l'énergie libre variationnelle, tenant plus finement compte des caractéristiques statistiques du champ.

#### 4.3.3.2 Graphe de régions et Propagation de Croyance Généralisée

La méthode proposée dans [92] pour l'évaluation de l'énergie libre relie les graphes factoriels et leurs algorithmes associés à la méthode variationnelle de Kikuchi [91] en adoptant une technique d'approximation de l'énergie libre par la définition de régions sur le champ aléatoire. L'objectif de l'approche par région est de morceler le graphe factoriel en zones suffisamment grandes pour que tous les noeuds du champ soient inclus. L'énergie libre du système est alors la somme des énergies libres de ces régions.

De manière générale, on définit une région  $R$  d'un graphe factoriel comme l'ensemble  $V_R$  de noeuds variables et l'ensemble  $A_R$  de noeuds facteurs de sorte que lorsqu'un noeud facteur  $a$  appartient à  $A_R$ , l'ensemble des noeuds variables qui lui sont liés sont dans  $V_R$ . Par extension, nous définissons l'énergie  $E_R(\mathbf{x}_R)$  de  $R$  par :

$$E_R(\mathbf{x}_R) = - \sum_{a \in A_R} \ln f_a(\mathbf{x}_a) \quad (4.38)$$

Ainsi pour la région  $R$ , on définit l'énergie moyenne  $U_R(\mathbf{b}_R)$  :

$$U_R(\mathbf{b}_R) = \sum_{\{\mathbf{x}_R\}} b_R(\mathbf{x}_R) E_R(\mathbf{x}_R) \quad (4.39)$$

l'entropie  $H_R(\mathbf{b}_R)$

$$H_R(\mathbf{b}_R) = - \sum_{\{\mathbf{x}_R\}} b_R(\mathbf{x}_R) \ln b_R(\mathbf{x}_R) \quad (4.40)$$

et l'énergie libre  $F_R(\mathbf{b}_R)$

$$F_R(\mathbf{b}_R) = U_R(\mathbf{b}_R) - H_R(\mathbf{b}_R) \quad (4.41)$$

La superposition éventuelle des régions implique que certains noeuds sont pris en compte plusieurs fois. On définit ainsi des compteurs  $c_R$  associés à chaque région qui vont servir de pondération pour le calcul de l'énergie libre totale. Cette pondération est établie de sorte que *chaque facteur du graphe ne soit compté qu'une seule fois*. Finalement l'énergie libre par région du champ est définie par :

$$F_{\{R\}}(\{\mathbf{b}_R\}) = U_{\{R\}}(\{\mathbf{b}_R\}) - H_{\{R\}}(\{\mathbf{b}_R\}) \quad (4.42)$$

avec

$$U_{\{R\}}(\{\mathbf{b}_R\}) = \sum_{\{R\}} c_R U_R(\mathbf{b}_R) \quad (4.43)$$

et

$$H_{\{R\}}(\{\mathbf{b}_R\}) = \sum_{\{R\}} c_R H_R(\mathbf{b}_R) \quad (4.44)$$

Le calcul de l'énergie libre du champ implique donc la connaissance des croyances associées aux noeuds des régions. Nous devons donc pouvoir calculer les marginales de ces régions. On définit ainsi un ensemble de régions formant une super-structure (le graphe de régions) sur laquelle vont transiter les messages d'une version généralisée de la propagation de croyance.

Soit  $I$  l'ensemble des indices des noeuds facteurs et variables dans un graphe factoriel. Un graphe de régions  $GR(S, T, L)$  est un graphe orienté et étiqueté dans lequel chaque sommet  $s \in S$  est étiqueté par un sous-ensemble de  $I$ . Cette étiquette est notée  $l(s) \in L$ . Un arc orienté  $t \in T$  existe entre les sommets  $s_p$  et  $s_e$  si on a  $e(s_e) \subset e(s_p)$ . Si tel

est le cas, on dira que  $s_p$  est le parent de  $s_e$ . Enfin s'il existe un chemin orienté du sommet  $s_a$  au sommet  $s_d$ , on dira que  $s_a$  est un ancêtre de  $s_d$  et réciproquement que  $s_d$  est un descendant de  $s_a$ .

Les compteurs  $c_s$  de chaque sommet du graphe sont définis par :

$$c_s = 1 - \sum_{u \in A(s)} c_u \quad (4.45)$$

où  $A(s)$  est l'ensemble des sommets qui sont ancêtres de  $s$ .

Enfin, un graphe de région  $GR$  est qualifié de valide pour l'estimation de l'énergie libre variationnelle si pour chaque indice  $i \in I$ , le sous-graphe  $GR(i)(S(i), T(i), L(i))$ , formé par les sommets dont les étiquettes incluent  $i$ , vérifie la relation suivante sur les compteurs :

$$\sum_{u \in S(i)} c_u = 1 \quad (4.46)$$

Il s'agit de nouveau de la nécessité de ne tenir qu'une seule fois compte de chaque sommet du graphe, et l'on associe directement les compteurs de régions  $c_R$  aux compteurs  $c_s$  de  $GR$ .

Ainsi l'énergie libre  $F_{GR}$  du graphe de région est la combinaison des énergies libres  $\{F_R\}$  de chaque région  $R$  :

$$F_{GR} = \sum_R c_R F_R \quad (4.47)$$

Les règles de création d'un graphe de régions ayant été posées, nous allons y introduire le principe de passage de messages, à l'instar de la propagation de croyance sur un graphe factoriel. Cette transposition du principe à la super-structure nouvellement définie est une généralisation et est liée à la structure du graphe. A ce propos, il apparaît de manière assez évidente qu'il existe pour un même graphe factoriel de nombreuses possibilités de choix de régions qui, dès lors que la condition (4.46) est satisfaite, minimisent la même énergie libre. Nous adaptons une approche *Parent vers Enfant* [92], qui se démarque des autres versions de l'algorithme par une formulation plus simple de la mise à jour des messages. Il convient également de signaler que son principe de passage des messages soutient l'analogie avec la propagation de croyance simple, ce qui renforce leur lien inclusif.

Dans l'algorithme de propagation de croyance généralisée *Parent vers Enfant*, noté

PCG-PvE, chaque région  $R$  d'un graphe  $GR$  a une croyance  $b_R(\mathbf{x}_R)$  définie par :

$$b_R(\mathbf{x}_R) \propto \prod_{a \in \{f_R\}} f_a(\mathbf{x}_a) \left( \prod_{P \in P(R)} m_{P \rightarrow R}(\mathbf{x}_R) \right) \cdot \left( \prod_{D \in D(R)} \prod_{P' \in P(D) \setminus E(R)} m_{P' \rightarrow D}(\mathbf{x}_D) \right) \quad (4.48)$$

où  $\{f_R\}$  est l'ensemble de noeuds facteurs appartenant à  $R$ ,  $P(R)$  est l'ensemble des régions parents de  $R$ ,  $D(R)$  est l'ensemble des régions descendants de  $R$ , l'ensemble  $E(R)$  est défini par  $E(R) = R \cup D(R)$ , et  $P(D) \setminus E(R)$  est l'ensemble des régions parents de  $D$  mis à part les descendants de  $R$  et  $R$  elle-même.

La mise à jour des messages de PCG-PvE est définie par :

$$m_{P \rightarrow R}(\mathbf{x}_R) = \frac{\sum_{\{x_{P \setminus R}\}} \prod_{a \in \{f_{P \setminus R}\}} f_a(x_a) \prod_{(I, J) \in N(P, R)} m_{I \rightarrow J}(x_J)}{\prod_{(I, J) \in D(P, R)} m_{I \rightarrow J}(x_J)} \quad (4.49)$$

où

- $\{f_{P \setminus R}\}$  est l'ensemble de noeuds facteurs appartenant au parent  $P$  mais pas à  $R$
- $N(P, R)$  est l'ensemble des paires de régions connectées  $(I, J)$  tel que  $J \in E(P)$ ,  $J \notin E(R)$  et  $I \notin E(P)$
- $D(P, R)$  est l'ensemble des paires de régions connectées  $(I, J)$  tel que  $J \in E(R)$ ,  $I \in E(P)$  et  $I \notin E(R)$

On montre qu'à la condition que les croyances dérivées soient homogènes à une mesure de probabilité, les points de convergence de l'algorithme généralisé correspondent aux minima de l'énergie libre variationnelle ([92]). En d'autres termes, l'algorithme PCG-PvE converge vers les marginales d'un champ aléatoire qui minimisent la divergence de Kullback – Leibler, délivrant de par la même une approximation de l'énergie libre d'Helmoltz et donc de la fonction de partition du système.

Nous présentons désormais un exemple d'application de PCG-PvE afin d'en clarifier les mécanismes et illustrer sa simplicité.

Exemple : application de l'algorithme PCG-PvE sur un graphe factoriel de taille 3x3

Soit le graphe factoriel  $GF_1$  de la figure 4.5.

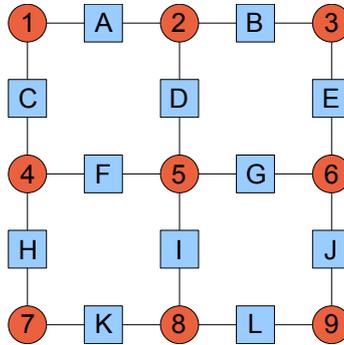


Figure 4.5 Graphe Factoriel  $GF_1$

$GF_1$  est associé à un processus aléatoire  $X|Y$  de taille 3x3 dont la distribution a la forme (4.17). Les sites  $\{1; 2; \dots; 9\}$  correspondent aux variables  $\{x_1; x_2; \dots; x_9\}_y$  et ont pour facteurs :  $f_{a_1} = e^{-U'_{x|y}(x_k)}$  avec  $a_1 = \{1; 2; \dots; 9\}$ . Les liens inter-sites  $\{A; B; \dots; L\}$  ont quant à eux pour facteurs :  $f_{a_2} = e^{-U''_{x|y}(x_k \leftrightarrow x_{k'})}$  avec  $a_2 = \{A; B; \dots; L\}$ .

Nous adoptons pour la construction du graphe de régions  $GR_1$  sur  $GF_1$  la méthode de variation de groupes proposée par Kikuchi [91, 92] (nous conserverons cette méthode pour le reste de l'étude). Un premier ensemble de régions  $\{R_0\}$  est défini de sorte que chaque noeud facteur ou variable soit dans une région. De plus aucune région n'est une sous-région d'une autre. Puis nous définissons l'ensemble des régions  $\{R_1\}$  qui sont les intersections des régions  $\{R_0\}$  mais en veillant à ce qu'aucune de ces régions-intersection ne soit une sous-région de  $\{R_1\}$ . Si possible, nous construisons l'ensemble  $\{R_2\}$  qui sont les intersections de  $\{R_0\} \cup \{R_1\}$  et en excluant les régions déjà contenues dans  $\{R_1\}$  et celles qui sont sous-régions de  $\{R_2\}$ . Nous poursuivons ainsi tant que des intersections non référencées subsistent.

Une fois les régions définies, celles-ci sont connectées en suivant les règles suivantes : les régions de  $\{R_1\}$  sont reliées à leurs super-régions dans  $\{R_0\}$ , les régions de  $\{R_2\}$  sont reliées à leurs super-régions de  $\{R_1\}$  puis à celles de  $\{R_0\}$  qui ne sont pas liées à une région de  $\{R_1\}$  identifiées pour  $\{R_2\}$ , puis il en va de même pour les autres niveaux de régions

jusqu'au dernier. Les compteurs associés aux régions  $GR_1$  sont calculées suivant (4.45). L'application de la méthode est illustrée figure 4.6. Le graphe  $GR_1$  est dessiné figure 4.7.

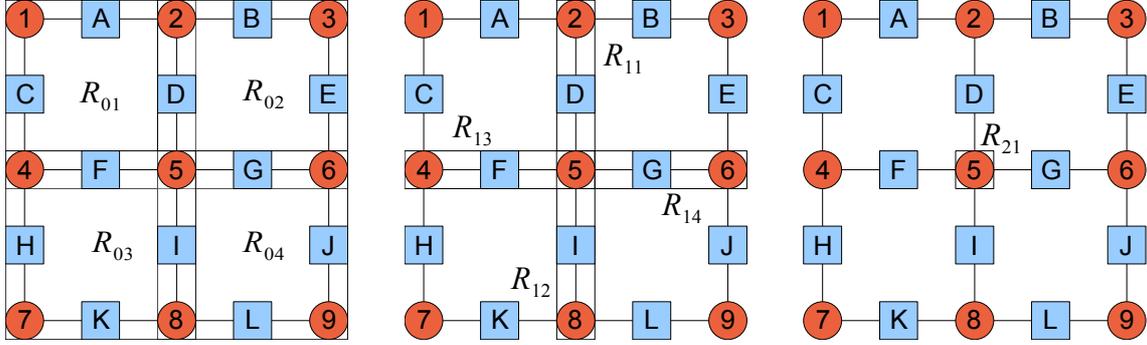


Figure 4.6 Construction du graphe de région  $GR_1$  par la méthode de variation de groupes

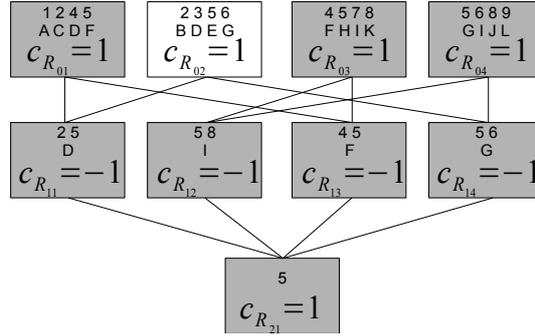


Figure 4.7 Graphe de région  $GR_1$

En se basant sur l'équation (4.48) , les croyances des régions sont données par :

- Niveau 0

$$b_{R_{01}}(\mathbf{x}_{R_{01}}) \propto f_1 f_2 f_4 f_5 f_A f_C f_D f_F m_{R_{02} \rightarrow R_{11}}(\mathbf{x}_{R_{11}}) m_{R_{03} \rightarrow R_{13}}(\mathbf{x}_{R_{13}}) m_{R_{12} \rightarrow R_{21}}(\mathbf{x}_{R_{21}}) m_{R_{14} \rightarrow R_{21}}(\mathbf{x}_{R_{21}}) \quad (4.50)$$

$$b_{R_{02}}(\mathbf{x}_{R_{02}}) \propto f_2 f_3 f_5 f_6 f_B f_D f_E f_G m_{R_{01} \rightarrow R_{11}}(\mathbf{x}_{R_{11}}) m_{R_{04} \rightarrow R_{14}}(\mathbf{x}_{R_{14}}) m_{R_{12} \rightarrow R_{21}}(\mathbf{x}_{R_{21}}) m_{R_{13} \rightarrow R_{21}}(\mathbf{x}_{R_{21}}) \quad (4.51)$$

$$b_{R_{03}}(\mathbf{x}_{R_{03}}) \propto f_4 f_5 f_7 f_8 f_F f_H f_I f_K m_{R_{04} \rightarrow R_{12}}(\mathbf{x}_{R_{12}}) m_{R_{01} \rightarrow R_{13}}(\mathbf{x}_{R_{13}}) m_{R_{11} \rightarrow R_{21}}(\mathbf{x}_{R_{21}}) m_{R_{14} \rightarrow R_{21}}(\mathbf{x}_{R_{21}}) \quad (4.52)$$

$$b_{R_{04}}(\mathbf{x}_{R_{04}}) \propto f_5 f_6 f_8 f_9 f_G f_I f_J f_L m_{R_{03} \rightarrow R_{12}}(\mathbf{x}_{R_{12}}) m_{R_{02} \rightarrow R_{14}}(\mathbf{x}_{R_{14}}) m_{R_{11} \rightarrow R_{21}}(\mathbf{x}_{R_{21}}) m_{R_{13} \rightarrow R_{21}}(\mathbf{x}_{R_{21}}) \quad (4.53)$$

- Niveau 1

$$b_{R_{11}}(\mathbf{x}_{R_{11}}) \propto f_2 f_5 f_D m_{R_{01} \rightarrow R_{11}}(\mathbf{x}_{R_{11}}) m_{R_{02} \rightarrow R_{11}}(\mathbf{x}_{R_{11}}) m_{R_{12} \rightarrow R_{21}}(\mathbf{x}_{R_{21}}) m_{R_{13} \rightarrow R_{21}}(\mathbf{x}_{R_{21}}) m_{R_{14} \rightarrow R_{21}}(\mathbf{x}_{R_{21}}) \quad (4.54)$$

$$b_{R_{12}}(\mathbf{x}_{R_{12}}) \propto f_5 f_8 f_I m_{R_{03} \rightarrow R_{12}}(\mathbf{x}_{R_{12}}) m_{R_{04} \rightarrow R_{12}}(\mathbf{x}_{R_{12}}) m_{R_{11} \rightarrow R_{21}}(\mathbf{x}_{R_{21}}) m_{R_{13} \rightarrow R_{21}}(\mathbf{x}_{R_{21}}) m_{R_{14} \rightarrow R_{21}}(\mathbf{x}_{R_{21}}) \quad (4.55)$$

$$b_{R_{13}}(\mathbf{x}_{R_{13}}) \propto f_4 f_5 f_F m_{R_{01} \rightarrow R_{13}}(\mathbf{x}_{R_{13}}) m_{R_{03} \rightarrow R_{13}}(\mathbf{x}_{R_{13}}) m_{R_{11} \rightarrow R_{21}}(\mathbf{x}_{R_{21}}) m_{R_{12} \rightarrow R_{21}}(\mathbf{x}_{R_{21}}) m_{R_{14} \rightarrow R_{21}}(\mathbf{x}_{R_{21}}) \quad (4.56)$$

$$b_{R_{14}}(\mathbf{x}_{R_{14}}) \propto f_5 f_6 f_D m_{R_{02} \rightarrow R_{14}}(\mathbf{x}_{R_{14}}) m_{R_{04} \rightarrow R_{14}}(\mathbf{x}_{R_{14}}) m_{R_{11} \rightarrow R_{21}}(\mathbf{x}_{R_{21}}) m_{R_{12} \rightarrow R_{21}}(\mathbf{x}_{R_{21}}) m_{R_{13} \rightarrow R_{21}}(\mathbf{x}_{R_{21}}) \quad (4.57)$$

- Niveau 2

$$b_{R_{21}}(\mathbf{x}_{R_{21}}) \propto f_5 m_{R_{11} \rightarrow R_{21}}(\mathbf{x}_{R_{21}}) m_{R_{12} \rightarrow R_{21}}(\mathbf{x}_{R_{21}}) m_{R_{13} \rightarrow R_{21}}(\mathbf{x}_{R_{21}}) m_{R_{14} \rightarrow R_{21}}(\mathbf{x}_{R_{21}}) \quad (4.58)$$

En se basant sur l'équation (4.48) et (4.49), la mise à jour des messages suit les équations suivantes :

- Niveau 1 vers 2

$$\circ m_{R_{11} \rightarrow R_{21}}(\mathbf{x}_{R_{21}}) = \sum_{\mathbf{x}_{R_{11} \setminus R_{21}}} f_2 f_D m_{R_{01} \rightarrow R_{11}}(\mathbf{x}_{R_{11}}) m_{R_{02} \rightarrow R_{11}}(\mathbf{x}_{R_{11}}) \quad (4.59)$$

$$\circ m_{R_{12} \rightarrow R_{21}}(\mathbf{x}_{R_{21}}) = \sum_{\mathbf{x}_{R_{12} \setminus R_{21}}} f_8 f_I m_{R_{03} \rightarrow R_{12}}(\mathbf{x}_{R_{12}}) m_{R_{04} \rightarrow R_{12}}(\mathbf{x}_{R_{12}}) \quad (4.60)$$

$$\circ m_{R_{13} \rightarrow R_{21}}(\mathbf{x}_{R_{21}}) = \sum_{\mathbf{x}_{R_{13} \setminus R_{21}}} f_4 f_F m_{R_{01} \rightarrow R_{13}}(\mathbf{x}_{R_{13}}) m_{R_{03} \rightarrow R_{13}}(\mathbf{x}_{R_{13}}) \quad (4.61)$$

$$\circ m_{R_{14} \rightarrow R_{21}}(\mathbf{x}_{R_{21}}) = \sum_{\mathbf{x}_{R_{14} \setminus R_{21}}} f_6 f_G m_{R_{02} \rightarrow R_{14}}(\mathbf{x}_{R_{14}}) m_{R_{04} \rightarrow R_{14}}(\mathbf{x}_{R_{14}}) \quad (4.62)$$

- Niveau 0 vers 1

$$\circ m_{R_{01} \rightarrow R_{11}}(\mathbf{x}_{R_{11}}) = \frac{\sum_{\mathbf{x}_{R_{01} \setminus R_{11}}} f_1 f_4 f_A f_C f_F m_{R_{03} \rightarrow R_{13}}(\mathbf{x}_{R_{13}})}{m_{R_{13} \rightarrow R_{21}}(\mathbf{x}_{R_{21}})} \quad (4.63)$$

$$\circ m_{R_{02} \rightarrow R_{11}}(\mathbf{x}_{R_{11}}) = \frac{\sum_{\mathbf{x}_{R_{02} \setminus R_{11}}} f_3 f_6 f_B f_E f_G m_{R_{04} \rightarrow R_{14}}(\mathbf{x}_{R_{14}})}{m_{R_{14} \rightarrow R_{21}}(\mathbf{x}_{R_{21}})} \quad (4.64)$$

$$\circ m_{R_{03} \rightarrow R_{12}}(\mathbf{x}_{R_{12}}) = \frac{\sum_{\mathbf{x}_{R_{03} \setminus R_{12}}} f_4 f_7 f_F f_H f_K m_{R_{01} \rightarrow R_{13}}(\mathbf{x}_{R_{13}})}{m_{R_{13} \rightarrow R_{21}}(\mathbf{x}_{R_{21}})} \quad (4.65)$$

$$\circ m_{R_{04} \rightarrow R_{12}}(\mathbf{x}_{R_{12}}) = \frac{\sum_{\mathbf{x}_{R_{04} \setminus R_{12}}} f_6 f_9 f_G f_J f_L m_{R_{02} \rightarrow R_{14}}(\mathbf{x}_{R_{14}})}{m_{R_{14} \rightarrow R_{21}}(\mathbf{x}_{R_{21}})} \quad (4.66)$$

$$\circ m_{R_{01} \rightarrow R_{13}}(\mathbf{x}_{R_{13}}) = \frac{\sum_{\mathbf{x}_{R_{01} \setminus R_{13}}} f_1 f_2 f_A f_C f_D m_{R_{02} \rightarrow R_{11}}(\mathbf{x}_{R_{11}})}{m_{R_{11} \rightarrow R_{21}}(\mathbf{x}_{R_{21}})} \quad (4.67)$$

$$\circ \quad m_{R_{03} \rightarrow R_{13}}(\mathbf{x}_{R_{13}}) = \frac{\sum_{\mathbf{x}_{R_{01} \setminus R_{03}}} f_7 f_8 f_H f_I f_K m_{R_{04} \rightarrow R_{12}}(\mathbf{x}_{R_{12}})}{m_{R_{12} \rightarrow R_{21}}(\mathbf{x}_{R_{21}})} \quad (4.68)$$

$$\circ \quad m_{R_{02} \rightarrow R_{14}}(\mathbf{x}_{R_{14}}) = \frac{\sum_{\mathbf{x}_{R_{02} \setminus R_{04}}} f_2 f_3 f_B f_D f_E m_{R_{01} \rightarrow R_{11}}(\mathbf{x}_{R_{11}})}{m_{R_{11} \rightarrow R_{21}}(\mathbf{x}_{R_{21}})} \quad (4.69)$$

$$\circ \quad m_{R_{04} \rightarrow R_{14}}(\mathbf{x}_{R_{14}}) = \frac{\sum_{\mathbf{x}_{R_{04} \setminus R_{14}}} f_8 f_9 f_I f_J f_L m_{R_{03} \rightarrow R_{12}}(\mathbf{x}_{R_{12}})}{m_{R_{12} \rightarrow R_{21}}(\mathbf{x}_{R_{21}})} \quad (4.70)$$

La mise à jour des messages de PCG-PvE est opérée itérativement jusqu'à convergence. En pratique, nous employons une heuristique afin d'éviter autant que possible les problèmes d'oscillation ou de valeurs très grandes durant le processus. Cette heuristique est la suivante :

- deux versions de messages sont stockées, les valeurs à l'itération précédente  $m^{-1}$  et les valeurs mises à jour  $m^0$ . Les nouveaux messages sont tirés de ces deux versions :

$$m^1 = (m^{-1} + m^0) / 2 \quad (4.71)$$

- le calcul des messages  $m^0$  de l'itération courante ne s'effectue qu'avec les valeurs de l'itération précédente  $m^{-1}$ .

#### 4.3.4 Algorithme d'estimation de l'IMM

En nous basant sur le modèle de canal proposé (équation (4.7)), nous avons pu mettre en évidence le lien de proportionnalité entre la distribution de probabilité du vecteur d'observation  $p_Y(\mathbf{y})$  et la fonction de partition  $Z$  du champ aléatoire *a posteriori*  $\mathbf{X}|\mathbf{Y}$ . Cette fonction de partition est le plus souvent inaccessible analytiquement et en rechercher une approximation s'avère un choix judicieux bien que difficile. Cependant, le formalisme des graphes factoriels et le lien entre le modèle du système et la physique statistique nous a permis d'adopter l'algorithme PCG-PvE et ses modalités d'application sur  $\mathbf{X}|\mathbf{Y}$ . Les caractéristiques intrinsèques de l'algorithme nous permettent d'estimer  $Z$  et conséquemment  $p_Y(\mathbf{y})$ . Dès lors, et à l'instar du cas monodimensionnel (cf paragraphe 3.2.2.2), nous adoptons une méthode de Monte – Carlo pour approcher l'entropie  $H(\mathbf{Y})$ , calculant via  $N_s$

séquences de sortie simulées  $\{\mathbf{Y}=\mathbf{y}^{(s)}, s \in \{1, \dots, N_s\}\}$  l'approximation  $\hat{H}(\mathbf{Y})$ . L'estimée de l'entropie est donnée par l'égalité :

$$\hat{H}(\mathbf{Y}) = \frac{-1}{N_s} \sum_{s=1}^{N_s} \log_2 p_{\mathbf{Y}}(\mathbf{Y}=\mathbf{y}^{(s)}) \quad (4.72)$$

Les densités  $p_{\mathbf{Y}}(\mathbf{Y}=\mathbf{y}^{(s)})$  quant à elles sont estimées par l'approximation des fonctions de partition  $Z^{(s)}$  des  $N_s$  champs  $\{\mathbf{X}|\mathbf{Y}\}^{(s)}$  par l'application de PCG-PvE.

Finalement l'IMM du système 2D est donnée par l'équation :

$$IMM(\mathbf{X}; \mathbf{Y}) = \hat{H}(\mathbf{Y}) - \frac{n^2}{2} \log_2(2\pi e \sigma^2) \quad (4.73)$$

Le récapitulatif synthétique de l'algorithme d'estimation de l'IMM est présenté au tableau 4.1.

<p><b>Positionner</b> <math>C^{2D}</math></p> <p><b>Calculer</b> la matrice de canal : <math>\mathbf{A}</math></p> <p><b>Pour</b> numéroSimulation = <math>\{1 \text{ à } N_s\}</math> <b>faire</b></p> <p style="padding-left: 40px;"><b>Simuler</b> <math>\mathbf{Y}^{(\text{numéroSimulation})}</math></p> <p style="padding-left: 40px;"><b>Construire</b> <math>GF^{(\text{numéroSimulation})}</math></p> <p style="padding-left: 40px;"><b>Construire</b> <math>GR^{(\text{numéroSimulation})}</math></p> <p style="padding-left: 40px;"><b>Estimer</b> <math>Z^{(\text{numéroSimulation})}</math> <b>par</b> PCG-PvE</p> <p><b>En Dédire</b> <math>IMM(n, N_s, \sigma^2)</math></p> <p><i>Tableau 4.1 : Algorithme de calcul du flux d'information</i></p>
---

#### 4.4 Simulations et exploitation de résultats

Nous présentons dans ce paragraphe l'application des outils et concepts introduits aux précédents paragraphes pour l'analyse du système de transmission bidimensionnel. A l'instar du cas 1D, nous évaluons de nouveau le rapport signal sur bruit moyen (RSB) des transmissions dont la définition est pour rappel :

$$RSB = \frac{E_s}{N_0^s} \quad (4.74)$$

où cette fois

- $E_s$  est l'énergie moyenne du symbole binaire reçu avec, en se reportant à l'équation du canal (4.7),

$$E_s = \frac{1}{n^2} \sum_{k=1}^{n^2} \frac{\sum_{k_0=1}^{(n+2)^2} A_{k, k_0} x_{k_0}}{2^{(n+2)^2}} \quad (4.75)$$

- $N_0^s$  est la puissance par symbole binaire reçu du bruit ABG du canal.

#### 4.4.1 Configuration du système

La caméra est dotée d'un diaphragme carré de côté égal à  $2l_D$ , et a pour distance focale  $f$ . Le capteur est doté de pixels contigus carrés de  $\rho = 5\mu\text{m}$  de côté, et est placé à la distance  $z_c$  du centre optique.

Nous avons retenu la configuration optique de la caméra LP (longue portée) définie par:

- une focale  $f = 7.6\text{mm}$ , un nombre d'ouverture  $f_8^{nb}$  et  $z_c = 7.8\text{mm}$ . La position sur l'axe optique du plan objet correctement focalisé est donc à  $z_{of} = 296.4\text{mm}$ .

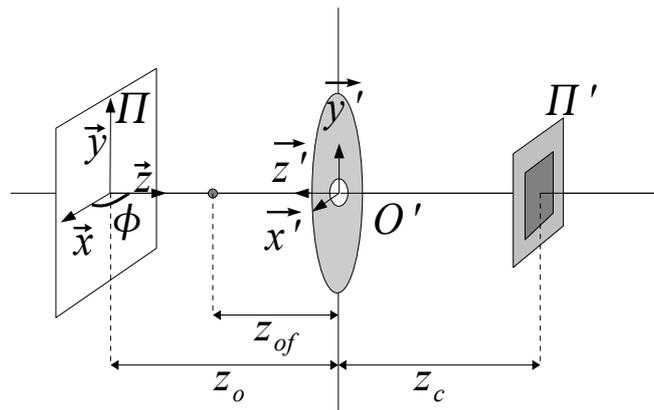


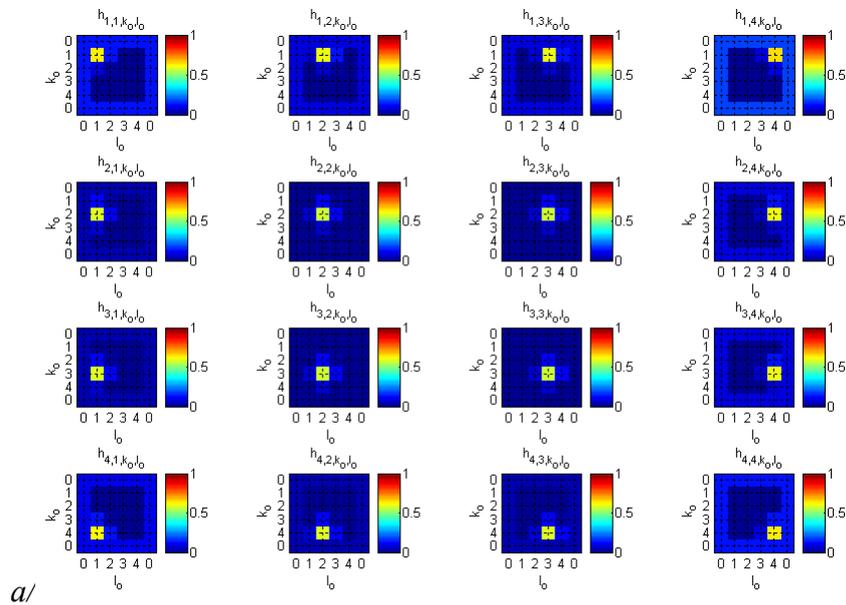
Figure 4.8 : Configuration et notations

Le code  $C^{2D}$  simulé est une matrice carrée de 4 par 4 modules dont les côtés sont de

longueur  $\delta=1\text{mm}$ . Le calcul de l'IMM est effectué en se basant sur la modélisation par module du canal (équation (4.6)). La position focalisée de la caméra est prise comme point de référence et le flux radiant intégré sur le plan du capteur est normalisé par rapport au flux reçu lorsque le code est en  $z_{of}$ . La source des symboles  $X_i$  associés au modules  $C_{i,j}^{2D}$  est une source binaire et équiprobable. Le bruit est enfin simulé à différentes puissances afin d'avoir une plage de RSB suffisamment large.

#### 4.4.2 Comportement du canal

En nous référant au schéma de la figure 4.8,  $C^{2D}$  est présenté face à la caméra et déplacé le long de l'axe optique tel que  $z_o \in [z_{of}; 2z_{of}]$ . L'ensemble des distorsions affectant le code lors de la transmission est synthétisé lors du calcul de la matrice  $A$  du canal. Nous avons représenté aux figures 4.9 a/ à 4.9 c/ l'évolution de ses éléments  $h_{k,l,k_0,l_0}$ , facteurs de l'interférence entre symboles (IES) associés aux modules  $C_{k,l}^{2D}$  et définis en (4.5).



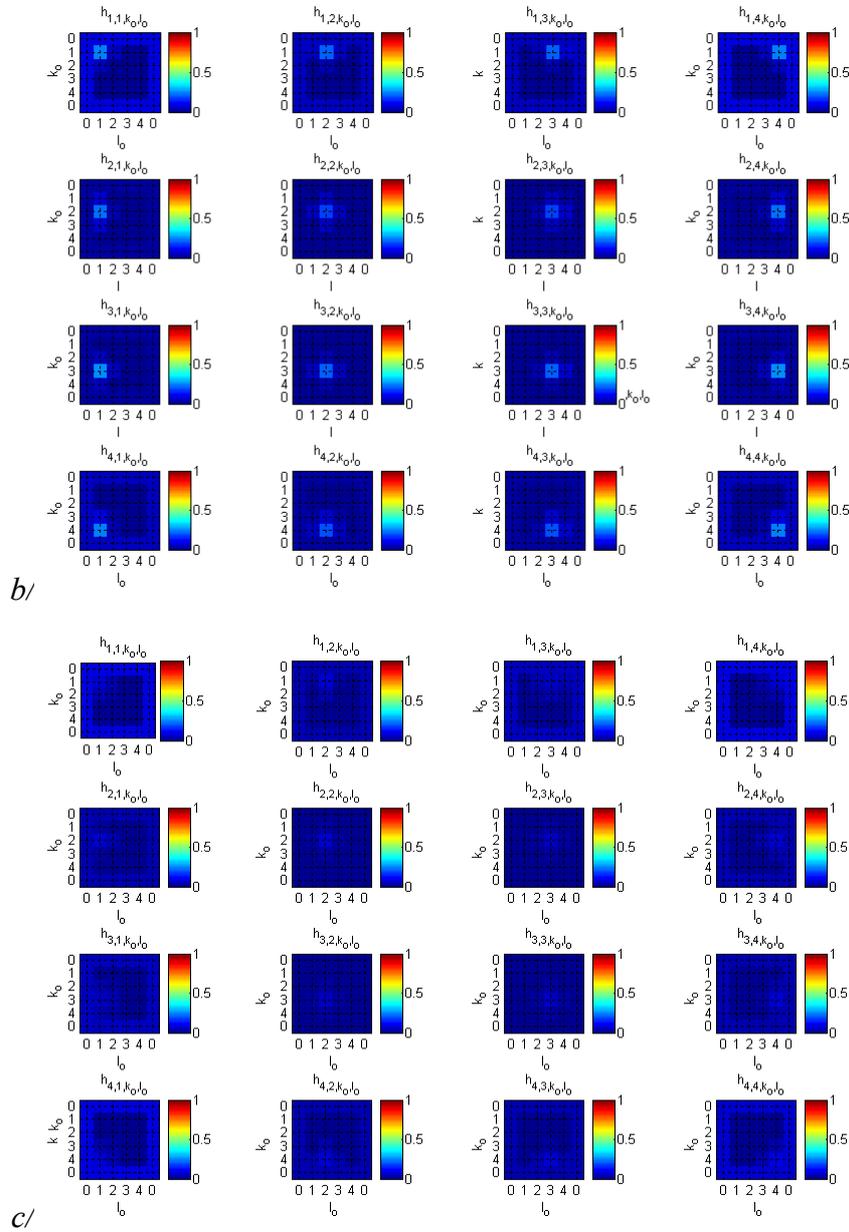
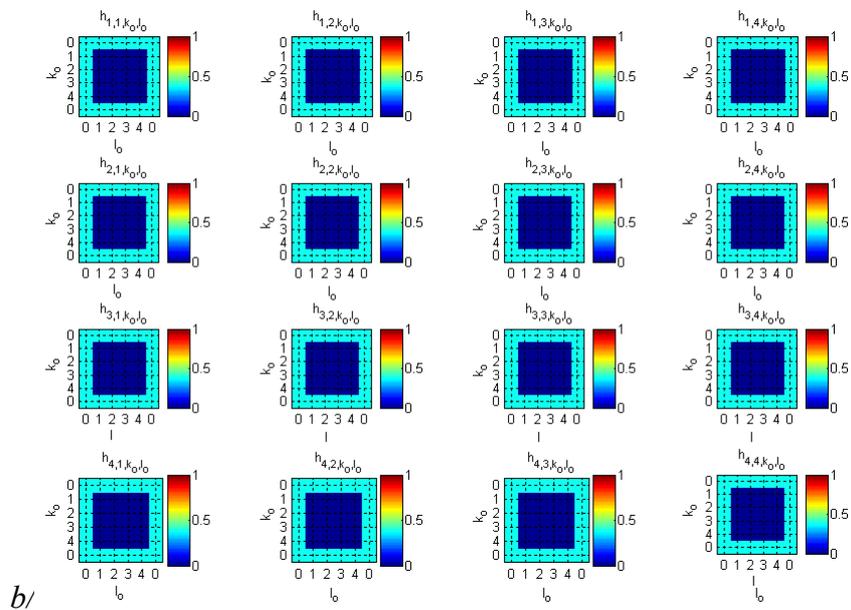
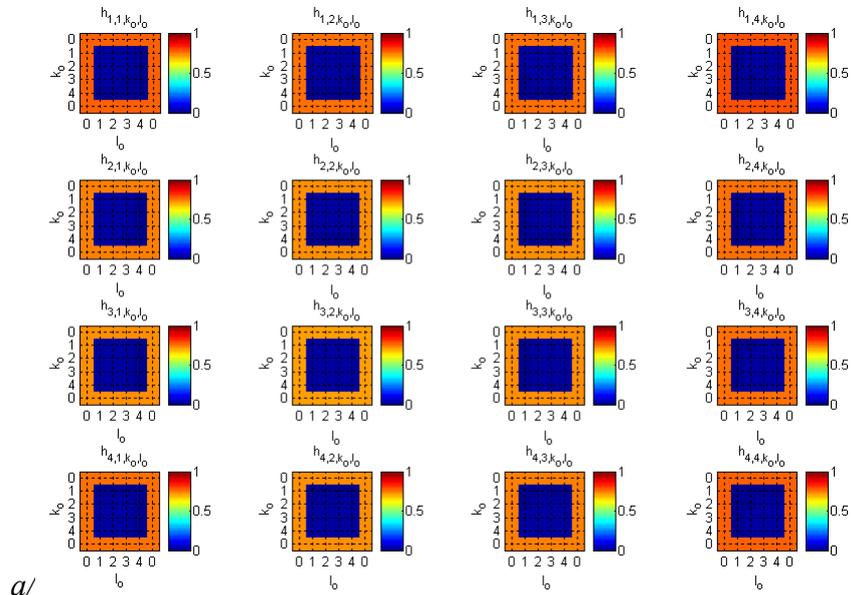


Figure 4.9 : Bloc LP - Coefficients  $h_{k,l,k_0,l_0}$  à : a/  $z_{of}$ , b/  $1.5 z_{of}$ , c/  $2 z_{of}$

Ces représentations établissent une cartographie des perturbations du canal induites par le système en soi et par les déplacements du code 2D. Ces cartes sont un premier indicateur des performances que l'on peut escompter du système ou, tout du moins, mesurent sa sensibilité aux paramètres pour chaque module du code. Pour preuve, comparons les cartes d'IES du bloc LP avec celles de sa version compacte cLP, défini au chapitre 3, et dont l'analyse pour la lecture de codes 1D a mis en évidence une plus forte sensibilité aux

déplacements. Les résultats sont en figures 4.10 a/ à 4.10 c/.



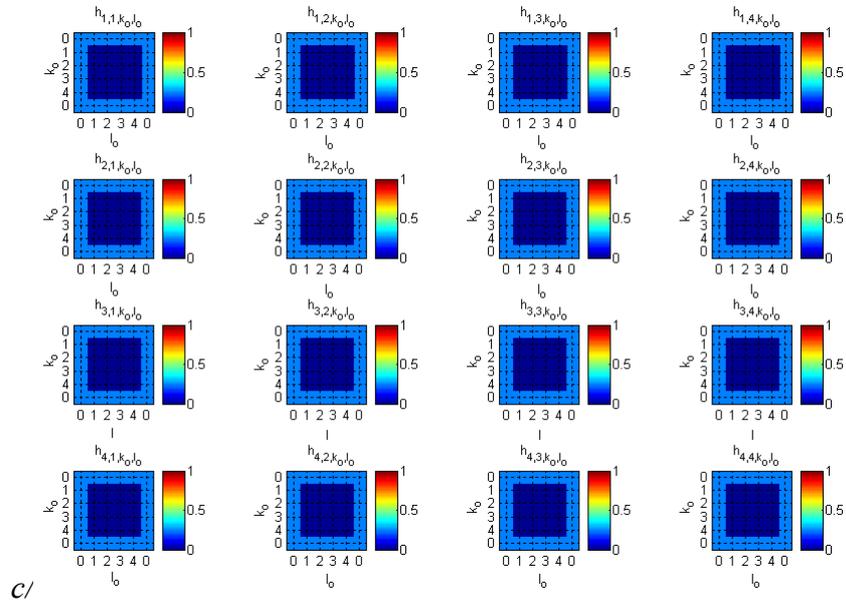


Figure 4.10 : Bloc LP - Coefficients  $h_{k,l,k_0,l_0}$  à :  $a/z_{of}$ ,  $b/1.5z_{of}$ ,  $c/2z_{of}$

Le faible facteur de grandissement du bloc cLP par rapport à LP se retrouve ici illustré par une IES qui s'étend systématiquement à tous les modules du code. Alors que sur les figures 4.9 a/ et b/, les symboles centraux n'étaient affectés que par leurs plus proches voisins, il ressort que le symbole neutre  $C_{0,0}^{2D}$  porté par la marge est l'élément le plus perturbant avec cLP. Le rapport signal / distorsion est donc plus faible laissant présager des performances en deçà du bloc LP concernant l'IMM. Ce résultat concorde avec ceux du chapitre précédents où l'IMM du système 1D chutait rapidement avec l'éloignement du code et l'augmentation de la puissance du bruit.

La construction des cartes d'IES, basée sur le calcul de la matrice du canal, est une représentation rapide et lisible de l'amplitude des distorsions du canal. Sans amener de réponse définitive quant aux performances du système, elles permettent néanmoins d'en établir des pronostics avant de se lancer dans l'algorithme d'évaluation de l'IMM qui est par essence long et gourmand en ressources. Elles constituent donc à notre sens une bonne étape de pré-analyse.

Dans le paragraphe suivant nous procédons au détail de cette analyse en évaluant le flux d'information du système LP par l'algorithme PCG-PvE.

### 4.4.3 Evaluation de l'IMM du système LP

L'IMM du système LP a été évalué pour la transmission des 16 bits d'information portés par  $C^{2D}$ .

Nous avons appliqué l'algorithme détaillé au paragraphe 4.2 en simulant 1000 transmissions par puissance de bruit, dans une plage de 16 dB, et pour un déplacement du code entre  $z_{of}$  et  $2z_{of}$ .

Les variations de l'IMM par rapport au positionnement et pour différentes puissance du bruit sont tracés en figure 4.11. La puissance du bruit est à ce titre représentée par un rapport signal sur bruit réduit défini par  $RSB_r = 1/N_0$  et exprimé en dB.

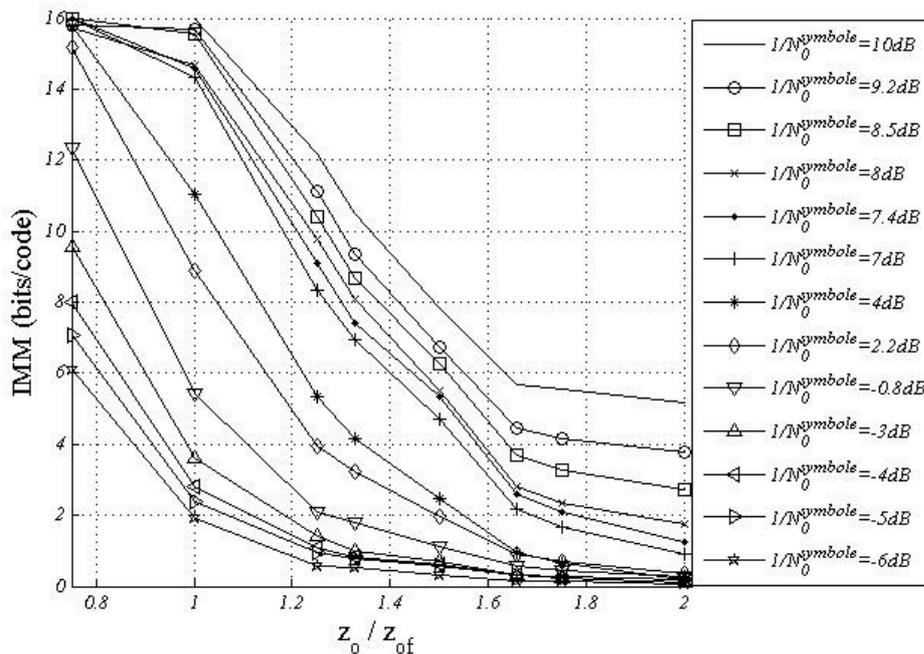


Figure 4.11 : Bloc LP – variations de l'IMM par rapport à l'éloignement du code

Les cartes d'IES du bloc optique (figures 4.9) indiquaient une diminution du rapport signal / distorsion avec l'éloignement dont la conséquence serait une chute des performances de transmission. Ce pronostic est validé par la figure 4.11 en ajoutant la mesure de l'impact de la puissance du bruit. Un  $RSB_r$  de 10 dB permet la récupération de presque l'ensemble des bits transmis à  $z_{of}$ . On note toutefois que les performances, même dans le cas le plus

favorable, chutent fortement dès l'éloignement du code. De plus, les résultats montrent que le bruit limite fortement la quantité récupérable d'information puisqu'à peine 4 bits sont récupérés, en position focalisée, pour un rapport de 0 dB.

La sensibilité du système au bruit est illustrée en figure 4.12, où les variations de l'IMM sont données par rapport à  $RSB_r$ .

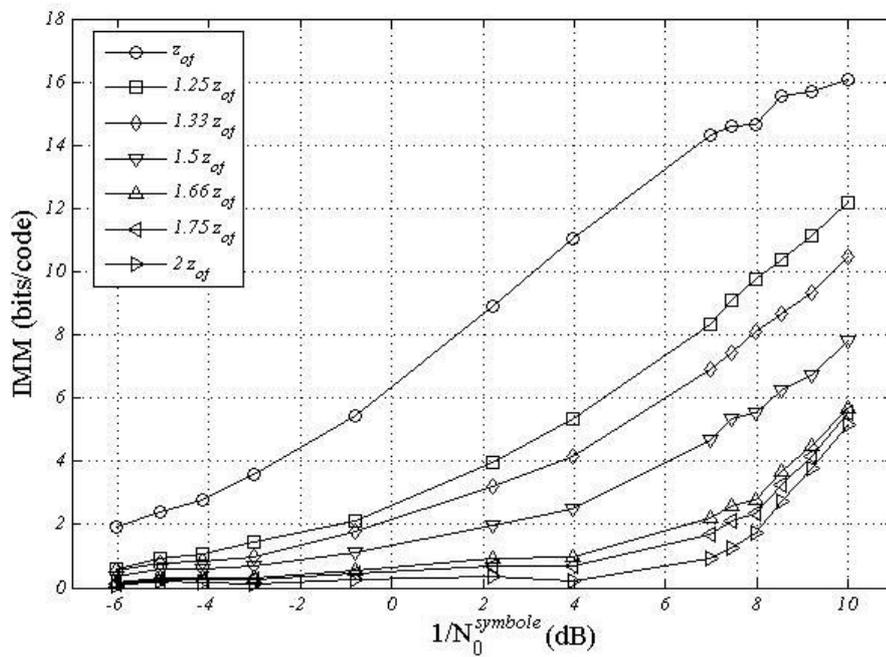


Figure 4.12 : Bloc LP – variations de l'IMM par rapport à la puissance du bruit

L'intérêt de cette représentation est de permettre une vue générale des performances du système. Bien que le modèle intègre de nombreux paramètres, tant optiques que liés au système, et dont les effets sur la transmission se cumulent et se compensent de manière non triviale, les résultats délivrés par la méthode établissent finement la dynamique des performances du système.

## **4.5 Conclusion**

Nous avons proposé dans ce chapitre un modèle d'analyse des systèmes de transmission numérique par code graphique 2D. Celui-ci est, à l'instar du cas 1D présenté au chapitre 3, formalisé par l'équation du canal de propagation qui est un système linéaire à bruit additif et caractérisé par une matrice d'interférence englobant l'ensemble des distorsions. Toutefois, l'analogie au cas monodimensionnel s'arrête là. En effet, la recherche d'une stratégie pour l'évaluation de performances théoriques du système indexées sur l'IMM a mis en évidence une plus grande complexité pour l'estimation du flux d'information. L'algorithme BCJR permettait pour le code-barres de simplifier le calcul de l'entropie du système, mais le BCJR est intrinsèquement en une dimension. La modélisation du canal a néanmoins permis d'établir la proximité du système avec un champ markovien ainsi que le lien entre l'entropie des observations et une fonction de partition. A partir de cette analogie avec la physique statistique, nous avons proposé un algorithme basé sur la propagation de croyance généralisée et Monte Carlo permettant l'estimation de cette fonction de partition et ainsi obtenir l'IMM. Signalons pour conclure que l'implémentation de la méthode a révélé que la méthode de calcul du flux est plus lente que dans le cas 1D, nécessitant d'importantes ressources mémoires pour une quantité d'information transmise moins importante (95 bits contre 16). Si ce fait est d'une importance moindre pour une étude théorique, elle peut se révéler rédhibitoire en pratique pour une phase de conception. Le calcul des cartes d'interférences permettent de répondre simplement à ce problème en donnant une indication qualitative du comportement du système par rapport au positionnement. Cette indication sera complétée dans un second temps par l'estimation exhaustive du flux.





## Conclusion Générale

Notre objectif dans cette thèse était l'étude de la transmission d'information par codes graphiques. L'utilisation de pictogrammes tels que les codes-barres ou les codes 2D s'est beaucoup intensifiée durant la dernière décennie. Comme identifiants en logistique ou URL à lire avec son photophone, ils sont une solution performante et à bas coût pour cet élan socio-industriel qu'est la dématérialisation. Les travaux menés à ce jour sur cette technologie ont, comme nous l'avons précédemment introduit, principalement portés sur le traitement d'image autour de ces codes ainsi que sur la définition de nouveaux usages. Or notre problématique est centrée sur les capacités et le comportement de ce mode de transmission. En d'autres termes, la question était d'évaluer la quantité d'information transmissible par un code graphique lu par une caméra numérique, questionnement pertinent lorsque l'on se place du point de vue d'une communication numérique. Les travaux se justifient également au regard de l'usage des codes graphiques. Le positionnement et l'orientation du code est libre devant la caméra, de plus l'imageur influe directement sur les performances du système. Il s'avère donc important de déterminer cette quantité transmissible d'information ainsi que les impacts du système sur celle-ci.

En se rapprochant des communications numériques, notre postulat de base a alors été de considérer la lecture de codes graphiques comme un système singulier de télécommunications. L'évaluation du flux d'information et de ses variations s'inscrit alors naturellement dans le cadre de la théorie de l'information de Shannon, théorie dont nous avons utilisé le concept d'information mutuelle moyenne (IMM) comme équivalent mathématique à ce flux et, de fait, mesure des performances du système. Notre réponse à la problématique posée est ainsi une mesure objective, car quantitative, du flux d'information et tenant compte des différentes perturbations. Ce calcul de l'IMM s'est révélé non trivial dans notre contexte car il s'effectue à partir des caractéristiques statistiques d'un canal de communication qui restait à définir. Nous avons sur ce point distingué le cas monodimensionnel, c'est-à-dire d'un code à barres lu par une caméra linéaire, du cas bidimensionnel où un code 2D est lu par un imageur matriciel. Cette différenciation a été

motivée par des propriétés statistiques différentes du canal dans ces deux cas, et dont ont découlées les stratégies proposées pour l'IMM.

Nous avons ainsi introduit au chapitre 2 de ce mémoire les notions nécessaires à la construction du modèle du canal. Il s'agit principalement des outils radiométriques et optiques, les premiers car un imageur est fondamentalement un quantificateur d'énergie rayonnée, les seconds de par les implications du processus de formation d'image. Nous avons ainsi identifié les principales distorsions du système que sont la déformation géométrique due à la projection perspective, l'intégration opérée par le capteur et surtout le flou, les lecteurs de codes étant pour la majorité des cas dépourvus d'autofocus. La projection a été modélisée directement par les lois de l'optique géométrique. En revanche, nous avons adapté les principes de l'optique de Fourier ainsi que la théorie des systèmes linéaires afin de modéliser le plus finement possible l'impact de la défocalisation, ceci pour aboutir à un filtre comme modèle du flou, paramétré par les spécifications du bloc optique. L'intégration des pixels a été également modélisée par un filtre, en l'occurrence moyenneur, pour garder la cohérence de la démarche et s'acheminer vers une écriture compacte de l'ensemble des distorsions.

Ces prérequis théoriques posés, nous avons développé au chapitre 3 la modélisation et l'étude de la transmission d'information par codes graphiques dans le cas 1D. La liberté de positionnement du code a été pris en compte avec, par conséquent les perturbations identifiées au chapitre 2, à savoir la déformation géométrique et le flou dus au bloc optique, et l'intégration de l'éclairement et l'ajout d'un bruit dus au capteur. L'ensemble des distorsions a pu être synthétisé pour aboutir à une écriture linéaire du canal, reliant directement les symboles binaires émis aux symboles reçus, et paramétrée par les caractéristiques du système (distance focale, position et orientation du code, puissance du bruit, ...). Une fois le modèle du canal défini, nous avons proposé un algorithme stochastique d'estimation de l'IMM utilisant une adaptation du BCJR de Bahl et al. [70]. Cette méthode est légitimée par le fait que le canal 1D est à interférences entre symboles, et possède une mémoire markovienne de taille variable mais finie. L'IMM est finalement paramétrée par la position et l'orientation du code ainsi que par les caractéristiques du système de sorte que la quantité d'information transmissible par le système est déterminée quelle que soit la configuration du système. Le réel apport d'un autofocus en terme d'information transmise a également été évalué, les

simulations ayant montré qu'un tel ajout n'est pertinent que pour les caméras dont la dynamique du flou est forte. De plus, nous avons déduit de l'IMM deux critères faisant partie des principales contributions des travaux: la profondeur de champ théorique (PdCT) ainsi que la résolution spatial théorique (RST). La première correspond à l'éloignement maximum du code pour lequel le récepteur récupère une quantité maximale d'information. Les résultats de nos simulations avec trois types de caméra ont montré que la PdCT complète la définition géométrique classique de la profondeur de champ en tenant compte des distorsions globales du système de manière plus exhaustive. La RST quant à elle établit le lien entre l'information transmissible et la densité du pictogramme. Cette grandeur est donc un critère extrêmement intuitif, liant simplement les capacités du système à la configuration de son émetteur.

Dans le cas 2D, l'évaluation de l'IMM revient au calcul de la fonction de partition d'un champ de Gibbs. Nous avons en effet démontré que les propriétés statistiques du canal impliquent que la distribution de probabilité du message émis conditionnellement aux observations est celle d'un champ de Gibbs dont la fonction de partition est, à une constante près, la probabilité d'apparition du message reçu. Nous nous sommes ainsi basés sur les outils de la physique statistique pour le calcul de cette fonction de partition à partir duquel nous déduisons l'IMM par une approche stochastique. Plus précisément, nous nous sommes concentrés sur les méthodes variationnelles d'approximation de l'énergie libre d'un modèle d'Ising en adaptant à notre système la propagation de croyance généralisée introduite par Yedidia [92]. Le champ de Gibbs du système, équivalent à un champ de Markov par le théorème d'Hammersley – Clifford, est ainsi transposé en un graphe factoriel régionalisé sur lequel transitent les messages de la propagation généralisée. Au terme de la "discussion", nous obtenons à l'instar du cas 1D une estimation de l'IMM paramétrée par les caractéristiques du système. De plus, le canal étant à interférences entre symboles, nous avons proposé le concept de cartes d'interférences, représentations illustrant l'étendue et l'amplitude de la mémoire du canal.

En adoptant une approche orientée télécoms, ces travaux ont apporté une réponse originale à la question de la capacité en transmission d'un système à code graphique. Le choix de l'IMM comme mesure des performances ainsi que les méthodes d'estimations proposées, adossées au modèle du canal, se sont donc révélés pertinents pour cette problématique. Les investigations menées ont également soulevé quelques questions et ouvert quelques portes

pour d'éventuels futurs travaux.

En premier lieu, l'algorithme d'estimation du flux d'information dans le cas 2D s'est révélé être beaucoup plus lent et gourmand en ressources que dans le cas 1D pour un message binaire de taille équivalente. La raison principale réside dans l'estimation de la fonction de partition par la propagation de croyance généralisée dont la complexité, contrairement au BCJR, croît exponentiellement avec la taille du code. Une solution serait d'appliquer une démarche équivalente à celle de Kumar et Torr [95] pour la suppression des calculs redondants et la gestion des configurations irréalistes (code tout blanc ou tout noir par exemple).

Deuxièmement, l'extension logique de ces travaux serait l'utilisation de la couleur afin d'augmenter la capacité des codes graphiques. Cet ajout nécessiterait une adaptation du modèle proposé, bien que mineure dans le sens où dans la majorité des cas, l'acquisition des couleurs repose sur une combinaison linéaire prédéterminée du résultat de l'intégration des pixels (filtres de Bayer et caméra tri-capteurs [96]). En revanche, la complexité des algorithmes d'estimation du flux d'information pourrait devenir rédhibitoire.

Enfin, nous avons considérés dans cette étude que l'information contenue dans les codes graphiques est exempte de tout codage canal. Cette simplification a permis de définir les bornes des performances du système en considérant la source émettrice comme binaire et équiprobable. L'introduction d'une redondance et de sa prise en compte dans l'évaluation du flux permettrait néanmoins d'améliorer la robustesse du système notamment par rapport au bruit. En plus de l'intérêt théorique de tels travaux avec la définition de nouveaux schémas de codage, l'industrie pourrait ainsi bénéficier de lecteurs de codes performants malgré l'utilisation de capteurs peu sensibles et donc bon marché.

Enfinement le fil conducteur de ces travaux s'est voulu être la construction de ponts autour d'une même problématique provenant de l'industrie. Liens scientifiques reliant les domaines pourtant vastes que sont l'optique, l'ingénierie des systèmes et la théorie de l'information, ils ont permis d'apporter un éclairage nouveau sur une technologie que nous côtoyons depuis maintenant quelques décennies. Cela met ainsi clairement en évidence, si tant est que cela fut nécessaire, l'intérêt de la collaboration entre théorie et pratique, où tantôt l'un questionne et l'autre répond, et tous avancent.





## **Références bibliographiques**

- [1] Woodland N.J. & Silver B., "Classifying apparatus and Method", USPTO 2612994, 7 Octobre 1952
- [2] Savir D. and Laurer G.J. "The characteristics and decodability of the Universal Product Code". *IBM Systems Journal*. Vol 14, pp. 16-33. 1975
- [3] Proakis J.G. *Digital Communications 4<sup>th</sup> Edition*. MacGraw – Hill. 2001
- [4] van Gils W.J. "Two-Dimensional Dot Codes for Product Identification". *IEEE Transactions on Information Theory*. Vol. 33, No 5, Sept 1986
- [5] Automatic Identification Manufacturers. *International Symbology Specification : Datamatrix*. Pittsburgh. AIM 1997
- [6] Automatic Identification Manufacturers. *International Symbology Specification : QR Code*. Pittsburgh. AIM 1997
- [7] Automatic Identification Manufacturers. *International Symbology Specification : Aztec*. Pittsburgh. AIM 1997
- [8] Ackland B. & Dickinson A. "Camera on a chip". 43<sup>rd</sup> International Solid-State Circuits Conferences. IEEE. 1996
- [9] Tu X. & Dubuisson B. "CCD Camera Model and Its Physical Characteristics Consideration in Calibration Task for Robotics". *Proc. IEEE International Workshop on Robot and Human Communication*, Tsukuba, Japan. pp. 75-77. 1992
- [10] Fossum E.R. "CMOS image sensors: Electronic camera-on-a-chip," *IEEE Trans. On Electronic Devices*. Vol.44, No.10, pp.1689–1698, 1997.
- [11] Pavlidis J.E. et al. "Bar code waveform recognition using peak locations". *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Vol. 16, Issue 6, pp. 630-640, 1994.
- [12] Pavlidis T., Swartz J., Wang Y.P. "Fundamentals of bar code information theory" *Computer*. Vol. 23, Issue 4, pp. 74 – 86, 1990.
- [13] Cover T.M. & Thomas J.A. *Elements of Information Theory*, John Wiley & Sons, 1991.
- [14] Pavlidis T., Swartz J., Wang Y.P. "Information Encoding with Two-Dimensional Bar Codes". *Computer*, Vol. 25, Issue 6, pp. 18-28, 1992.
- [15] Chan J.P. "Machine readable binary codes" (Snowflake Code). USPTO 5825015. 18 Octobre 1996

- [16] Hirokazu Y. "Method of decoding a two-dimensional code symbol mark" (CP Code). USPTO 5343031. 30 Août 1994
- [17] Rivailier J. "Procédé de codage de l'information de façon duale sur supports physiques et sous forme informatique (DoteM)". Brevet FR2882173. 18 Août 2006.
- [18] Tsi D., Marom E., Katz J., and Swartz J. "System Analysis of CCD-based Bar Code Readers". *Applied Optics*. Vol 32. pp. 3504-3512. 1993.
- [19] Riyanto T., Le Cong T., Ayache A., Noailles J. "Optimal Incoherent Lighting for CCD Bar Code Scanner". 4<sup>th</sup> International Conference On Control, Automation, Robotics And Vision, ICARCV'96, Singapour. Decembre 1996
- [20] Vicenzi D., Oliva G. "Tailoring freeform lenses for barcode lighting". 2007 IEEE Workshop on Automatic Identification Advanced Technologies. pp.134-139. Juin 2007
- [21] Il Hahn H. Joung Koo J. "Implementation of algorithm to decode two-dimensional barcode PDF-417". 2002 6th International Conference on Signal Processing. Vol.2. pp. 1791-1794. August 2002.
- [22] Muniz R., Junco L., Otero A. "A robust software barcode reader using the Hough transform". International Conference on Information Intelligence and Systems. Proceedings. pp. 313-319. 31 Octobre - 3 Novembre 1999.
- [23] Xianyong F., Fuli W., Bin L. et al. "Automatic Recognition of Noisy Code-39 Barcode". 16th International Conference on Artificial Reality and Telexistence - Workshops, ICAT '06. pp. 79-82. Novembre 2006
- [24] Shellhammer S.J., Goren D.P., Pavlidis T. "Novel signal-processing techniques in barcode scanning". *IEEE Robotics & Automation Magazine*. Vol. 6, Issue 1, pp. 57-65. 1999.
- [25] Kun Z., Baifeng W., Qiang Y., "Real-time recognition of two-dimensional barcode based on ARM and DMA". 7th International Conference on Computer-Aided Industrial Design and Conceptual Design, CAIDCD '06. pp. 1-5. Novembre 2006
- [26] Safran M. I., Oktem R. "A Fast Hough Transform Approximation and Its Application for Barcode Localization". IEEE 15th Signal Processing and Communications Applications, SIU 2007. pp. 1-4, 11-13 Juin 2007
- [27] Lu X., Fan G., Wang Y. "A Robust Barcode Reading Method Based on Image Analysis of a Hierarchical Feature Classification". IEEE/RSJ International Conference

- on Intelligent Robots and Systems. pp. 3358-3362. Octobre 2006
- [28] Ando S., Hontani H. "Automatic visual searching and reading of barcodes in 3-D scene". Proceedings of the IEEE International Vehicle Electronics Conference, IVEC 2001. pp. 49-54. 25-28 Septembre 2001
- [29] Ohbuchi E., Hanaizumi H., Hock L.A. "Barcode readers using the camera device in mobile phones". International Conference on Cyberworlds. pp. 260-265. 18-20 Novembre 2004
- [30] Chu C-H., Yang D-N., Chen M-S. "Extracting Barcodes from a Camera-Shaken Image on Camera Phones". IEEE International Conference on Multimedia and Expo. pp. 2062-2065. 2-5 Juillet 2007
- [31] Peizhuo Lv., Shengli L., Leian L., Yongming Z., Kai L. "Research and implementation of automatic recognition of two dimensional barcode mobile computing". 2nd International Conference on Mobile Technology, Applications and Systems. 15-17 Novembre 2005
- [32] Kato H. & Tan K.T. "2D barcodes for mobile phones". 2nd International Conference on Mobile Technology, Applications and Systems. 15-17 Novembre 2005
- [33] Kato H. & Tan K.T. "Pervasive 2D Barcodes for Camera Phone Applications". *IEEE Pervasive Computing*. Vol. 6, Issue 4, pp. 76-85. 2007
- [34] Sriram T., Vishwanatha Rao K., Biswas S., Ahmed B. "Applications of barcode technology in automated storage and retrieval systems". Proceedings of the IEEE 22nd International Conference on Industrial Electronics, Control, and Instrumentation IECON. Vol 1, pp. 641-646. 1996
- [35] Garm Y., ZhiYan W., Yi L., Ling H., "An application and implementation of two-dimensional symbols for circuit board quality control system". 2nd IEEE International Conference on Industrial Informatics. INDIN '04. pp. 397-401. 24-26 Juin 2004
- [36] Bing-Nan L., Ming-Chui D., Mang I.V. "From Codabar to ISBT 128: Implementing Barcode Technology in Blood BankAutomation System". 27th Annual International Conference of the Engineering in Medicine and Biology Society IEEE-EMBS 2005. pp. 542- 545. 1<sup>er</sup>-4 Septembre 2005
- [37] Gao J. Z., Prakash L., Jagatesan R., "Understanding 2D-BarCode Technology and

Applications in M-Commerce - Design and Implementation of A 2D Barcode Processing Solution". Computer Software and Applications Conference, COMPSAC 2007. Vol. 2, pp. 49-56. 24-27 Juillet 2007

- [38] Adelman R. "Mobile Phone Based Interaction with Everyday Products - On the Go". The 2007 International Conference on Next Generation Mobile Applications, Services and Technologies, NGMAST '07. pp. 63-69. 12-14 Septembre 2007
- [39] Chaisatien P., Akahori K. "A Pilot Study on 3G Mobile Phone and Two Dimension Barcode in Classroom Communication and Support System". Seventh IEEE International Conference on Advanced Learning Technologies, ICALT 2007. pp. 111-113. 18-20 Juillet 2007
- [40] Liu T-Y., Tan T-H., Chu Y-L. "2D Barcode and Augmented Reality Supported English Learning System". 6th IEEE/ACIS International Conference on Computer and Information Science, ICIS 2007. pp. 5-10. 11-13 Juillet 2007
- [41] Mooser J., You S., Neumann U. "Tricodes: A Barcode-Like Fiducial Design for Augmented Reality Media". IEEE International Conference on Multimedia and Expo. pp. 1301-1304. 9-12 Juillet 2006
- [42] Rekimoto J. "Matrix: a realtime object identification and registration method for augmented reality". 3rd Asia Pacific Computer Human Interaction, Proceedings. pp. 63-68. 15-17 Julliet 1998
- [43] Ebrahim Y., Abdelsalam W., Ahmed M., Siu-Cheung C. "Proposing a hybrid tag-camera-based identification and navigation aid for the visually impaired". Second IEEE Consumer Communications and Networking Conference, CCNC 2005. pp. 172-177. 3-6 Janvier 2005
- [44] Iannizzotto G., Costanzo C., Lanzafame P., La Rosa F., "Badge3D for Visually Impaired". IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Vol 3. pp. 29-29. 20-26 Juin 2005.
- [45] Premaratne P., Safaei F., "2D Barcodes as Watermarks in Image Authentication". 6th IEEE/ACIS International Conference on Computer and Information Science, ICIS

2007. pp. 432-437. 11-13 Juillet 2007.

- [46] Huck F. O. et al., "Information-theoretic assessment of sampled imaging systems," *Opt. Eng.* Vol 38, pp. 742-762. 1999.
- [47] Huck F. O., Fales C. L., Jobson D. J., and Rahman Z., "Electro-optical design for efficient visual communication," *Opt. Eng.* Vol 34, pp. 795-813. 1995
- [48] Huck F. O., C Fales. L. and Rahman Z., "An information theory of visual communication," *Philos. Trans. Roy. Soc. London A.* Vol 354, pp. 2193-2248. 1996
- [49] Fales C. L., Huck F. O., Alter-Gartenberg R. and Rahman Z., "Image gathering and digital restoration," *Philos. Trans. Roy. Soc. London A.* Vol 354, pp. 2249-2287. 1996
- [50] O'Sullivan J.A., Blahut R.E., Snyder D.L. "Information-theoretic image formation". *IEEE Transactions on Information Theory.* Vol 44, Issue 6, pp. 2094-2123. 1998
- [51] Snyder D.L., Hero A.O., Moulin P., Moura J.M.F., O'Sullivan J.A. "Introduction to the special issue on information-theoretic imaging". *IEEE Transactions on Information Theory.* Vol 46, Issue 5, pp. 1709-1713. 2000
- [52] Chen J., Siegel P.H. "On the symmetric information rate of two-dimensional finite-state ISI channels". *IEEE Transactions on Information Theory.* Vol 52, Issue 1, pp. 227-236. 2006
- [53] Immink A.H.J., Coene W.M.J., van der Lee A.M. et al. "Signal processing and coding for two-dimensional optical storage". *IEEE Global Telecommunications Conference, 2003. GLOBECOM '03.* Vol 7, pp. 3904-3908. 1<sup>er</sup>-5 Décembre 2003.
- [54] Pearl J. "Reverend Bayes on inference engines: A distributed hierarchical approach". *American Association of Artificial Intelligence Proceedings National Conference on AI, Pittsburgh, PA,* pp. 133-136. 1982.
- [55] Forsyth D. A. & Ponce J. *Computer Vision: A Modern Approach.* Prentice Hall. 2003
- [56] Marion A. *Acquisition et visualisation des images.* Eyrolles, Paris.1997
- [57] Pérez J.P. *Optique fondamentaux et applications, 6<sup>ème</sup> Edition.* Dunod. 2000
- [58] Meyzonnette J.L., *Radiométrie et détection optique dans Systèmes Optiques* , EDP Sciences, 2002

- [59] Csele M. *Fundamentals of Light Sources and lasers*. Wiley-IEEE. 2004
- [60] Born M. & Wolf E. *Principles of Optics, 7<sup>ème</sup> Edition*. Cambridge University Press. 2001
- [61] Bloomenthal J. & Rokne J. *Homogeneous Coordinates*. [en ligne]. 1993
- [62] Maxwell J.C. *A Treatise on Electricity and Magnetism Vol 1 & 2*. Dover Publications. 1954
- [63] Maxwell J.C. "A Dynamical Theory of the Electromagnetic Field". *Philosophical Transactions of the Royal Society of London*. Vol 155, pp. 459-512. 1865
- [64] Goodman J. *Introduction to Fourier Optics, 3<sup>rd</sup> Edition*. Roberts & Co Publishers. 2005
- [65] Hertz H. *Annalen der Physik*. Vol 33, p983. 1887
- [66] Palmer R.C., *The Bar Code Book, 4<sup>ème</sup> Edition*. Helmers Publishing. 2001
- [67] Lingrand D. "Approximations de la projection perspective : modèle combiné et étude des singularités homographiques". INRIA Rapport de Recherche 3682. 1999
- [68] Shannon C. E. "A Mathematical Theory of Communication". *Bell System Technical Journal*. Vol. 27, pp. 379–423, 623–656. 1948
- [69] Fishman G.S. *Monte Carlo: Concepts, Algorithms, and Applications*. Springer. 1996
- [70] Bahl L.R., Cocke J., Jelinek F., Raviv J., "Optimal Decoding of Linear Codes, for Minimizing Symbol Error Rate". *IEEE Transactions on Information Theory*. Vol. 20 No 2, pp. 284-287. 1974.
- [71] Baum L. E., Petrie T., Soules G., Weiss N. "A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains". *Ann. Math. Statistic*. Vol. 41, pp. 164-171. 1970
- [72] Devijver P. A. "Baum's forward-backward algorithm revisited" *Pattern Recogn. Lett.* Vol. 3, pp.363-373. 1985
- [73] Asada N., Amano A. and Baba M. "Photometric Calibration of Zoom Lens Systems". *Proceedings of the 13th International Conference on Pattern Recognition*. Vol 1, pp. 186-190. 1996
- [74] Winkler G. *Image Analysis, Random Fields and Markov Chain Monte Carlo Methods 2<sup>ème</sup> Edition*. Springer. 2003
- [75] Li S.Z. *Markov Random Field : Modeling in Computer Vision*. Springer Verlag. 1995

- [76] Hammersley J. M. & Clifford P. "Markov field on finite graphs and lattices". non publié, [en ligne]. 1971
- [77] Besag J. "Spatial interaction and the statistical analysis of lattice systems" (with discussions). *Journal of the Royal Statistical Society, Series B*. Vol 36. pp192-236. 1974
- [78] Gibbs J.W. *Principes élémentaires de mécanique statistique*. Hermann. 1998
- [79] McQuarrie D. *Statistical Mechanics (2<sup>ème</sup> Edition)*. University Science Books. 2000
- [80] Bhattacharjee S.M., Khare A. "Fifty Years of the Exact Solution of the Two-Dimensional Ising Model by Onsager". [en ligne] arXiv. 1995
- [81] Shental O., Shental N., Shamai S. "On the achievable information rates of finite-state input two-dimensional channels with memory". International Symposium on Information Theory, ISIT 2005, Proceedings. pp. 2354-2358. 4-9 Septembre 2005.
- [82] Baxter R. J. *Exactly Solved Models in Statistical Mechanics*. London, England: Academic, [en ligne]. 1982
- [83] Ogata Y. "A Monte Carlo method for an objective Bayesian procedure". *Ann. Inst. Statist. Math.* Vol. 42, pp. 403-433. 1990
- [84] Potamianos G., *Stochastic simulation algorithms for partition function estimation of Markov random field images*. Ph.D. dissertation. Baltimore, MD: The Johns Hopkins Univ., Dept. Elec. Comput. Eng., 1994.
- [85] Ogata Y. & Tanemura M. "Estimation of interaction potentials of spatial point patterns through the maximum likelihood procedure". *Ann. Inst. Statist. Math.* Vol. 33, pp. 315-338. 1981.
- [86] Ogata Y. & Tanemura M. "Likelihood analysis of spatial point patterns". *J. Roy. Statist. Soc. Serie B*, Vol. 46, pp. 496-518. 1984.
- [87] Ogata Y. & Tanemura M. "Likelihood estimation of soft-core interaction potentials for Gibbsian point patterns". *Ann. Inst. Statist. Math.* Vol. 41, pp. 583-600. 1989.
- [88] Potamianos G. & Goutsias J. K. "Partition function estimation of Gibbs random field images using Monte Carlo simulation". *IEEE Trans. Inform. Theory*. Vol. 39, pp. 1322-1332. 1993.
- [89] Jerrum M. & Sinclair A. "Polynomial-time approximation algorithms for the Ising model" *SIAM J. Comput.* Vol. 22, pp. 1087-1116. 1993.

- [90] Potamianos G., Goutsias J. "Stochastic approximation algorithms for partition function estimation of Gibbs random fields". *IEEE Transactions on Information Theory*. Vol 43, Issue 6, pp. 1948-1965. 1997
- [91] Kikuchi R., Masuda-Jindo K. "Cluster variation method in the computational materials science ". *Calphad*. Vol. 26, No 1, pp. 33-54. 2002
- [92] Yedidia J.S., Freeman W.T., Weiss Y. "Constructing free-energy approximations and generalized belief propagation algorithms". *IEEE Transactions on Information Theory*. Vol 51, Issue 7, pp. 2282-2312. 2005
- [93] Loeliger H.-A. "An introduction to factor graphs". *Signal Processing Magazine, IEEE*. Vol 21, Issue 1, pp. 28-41. 2004
- [94] Kschischang F.R., Frey B.J., Loeliger H.-A. "Factor graphs and the sum-product algorithm". *IEEE Transactions on Information Theory*. Vol 47, Issue 2, pp. 498-519. 2001
- [95] Kumar M. P. & Torr P."Fast memory-efficient generalized belief propagation". 9th European Conference on Computer Vision ECCV 2006. Graz, Austria. 7-13 Mai 2006
- [96] Holst G.C. *CCD Arrays Cameras and Displays, 2<sup>ème</sup> Edition*. SPIE-International Society for Optical Engineering. 1998

---

Publications issus des présents travaux:

- "Spatial Resolution of 1D Image-based Barcode Reading: An Information Theoretical Approach"; ISCCSP 2008,
- "One-Dimensional Barcode Reading : an Information Theoretic Approach"; *Applied Optics*, Vol. 47, No. 8, Mars 2008,
- "Information Mutuelle Moyenne pour les systèmes d'acquisition par lecteur 1D de code-barres 1D"; TAIMA 2007.





C'est depuis les années 70 que les principaux acteurs de l'industrie ont utilisé les codes-barres pour la logistique et l'identification automatique. De par sa simplicité, cette technologie s'est rapidement imposée, prenant de surcroît une nouvelle dimension avec les codes 2D utilisés de nos jours dans, entre autres, le m-Commerce. L'affichage et la lecture de ces pictogrammes bicolores revient à transmettre une information binaire et c'est précisément sur cette notion de transmission qu'ont porté les travaux. Le couple pictogramme – caméra constitue un système singulier de communication numérique, ce qui nous a amené à considérer son analyse en nous basant sur les principes de la théorie de l'information.

Ainsi, nous avons proposé une modélisation du canal composite de transmission par un filtre linéaire bruité par un processus additif, blanc et gaussien. Les paramètres du modèle sont fonction des caractéristiques du code, de son positionnement, de sa puissance de rayonnement et des caractéristiques de la caméra. L'algorithme proposé de calcul des paramètres du canal est basé sur l'optique géométrique et ondulatoire, notamment pour le flou, ainsi que sur le modèle sténopé de la caméra pour les distorsions géométriques induites par la projection perspective.

L'analyse à proprement parler des performances du système est effectuée en calculant l'information mutuelle moyenne (IMM). Pour le cas 1D (codes à barres), nous estimons l'IMM en utilisant la markovianité du canal de transmission. En effet, son comportement peut être décrit en un treillis sur lequel nous appliquons une version normalisée de l'étape "Forward" de l'algorithme BCJR. L'IMM est alors la mesure objective des performances du système par rapport au positionnement relatif du code ou du rapport signal à bruit. Sur la base de ces estimations du flux, nous définissons une Profondeur de Champ Théorique ainsi qu'une Résolution Spatiale Théorique. Ces deux grandeurs complètent et enrichissent les notions classiques de profondeur de champ et de résolution en tenant compte des distorsions et du bruit du canal.

Dans le cas 2D, l'équation du canal, basé sur une construction analogue au cas 1D, est un filtre linéaire bidimensionnel bruité. Nous avons montré que ce modèle implique que la distribution de probabilité des données sachant les observation est celle d'un champ markovien. Nous avons ainsi mis en évidence le lien entre l'IMM du système et l'énergie libre de ce champ et proposé un algorithme stochastique d'estimation basé sur la propagation de croyance généralisée.

Parmi les perspectives de ces travaux citons : la prise en compte de la couleur, l'étude du codage canal et l'application des principes turbo, ou encore l'application des concepts de Markov couple et Markov triplet pour un décodage non supervisé.