

THÈSE

pour obtenir le grade de

DOCTEUR DE L'UNIVERSITE DE MONTPELLIER II

Spécialité : "Génie Informatique, Automatique et Traitement du signal"

préparée au Laboratoire d'Informatique de Robotique et de Microélectronique de Montpellier
dans le cadre de l'École Doctorale
" Information, Structures et Systèmes "

présentée et soutenue publiquement

par

Florence Jacquey

le 21 12 2007

Traitements d'Images Omnidirectionnelles

Directeur de Thèse : William PUECH

Co-Encadrant : Frédéric COMBY et Olivier STRAUSS

JURY

M. Hichem MAAREF	Professeur à l'Université Evry Val d'Essonne	Président du jury
Mme. Isabelle BLOCH	Professeur à l'ENST Paris	Rapporteur
M. El Mustapha MOUADDIB	Professeur à l'Université d'Amiens	Rapporteur
M. William PUECH	Maître de conférence à l'Université de Nîmes - HDR	Directeur de thèse
M. Frédéric COMBY	Maître de conférence à l'Université de Montpellier II	Encadrant de thèse
M. Olivier STRAUSS	Maître de conférence à l'Université de Montpellier II - HDR	Encadrant de thèse

dedicaces ...

Remerciements

Table des matières

Introduction Générale	1
1 Etat de l'art	7
Introduction	7
1.1 Acquisition et géométrie des images omnidirectionnelles	8
1.1.1 Objectifs spéciaux	8
1.1.2 Systèmes rotatifs	10
1.1.3 Systèmes de caméras multiples	11
1.1.4 Systèmes catadioptriques	13
1.1.4.1 Capteurs catadioptriques centraux	16
1.1.4.1.1 Miroir plan	16
1.1.4.1.2 Miroir elliptique	18
1.1.4.1.3 Miroir hyperbolique	18
1.1.4.1.4 Miroir parabolique	18
1.1.4.1.5 Modèle de projection unifié	19
1.1.4.2 Capteurs catadioptriques centraux par approximation de la contrainte du point de vue unique	20
1.1.4.2.1 Miroir conique	20
1.1.4.2.2 Miroir sphérique	21
1.1.4.3 Système catadioptrique axial	22
1.1.4.4 Système catadioptrique caustiques	23
1.1.5 Combinaison des différentes approches	23
1.2 Traitement des images omnidirectionnelles	25
1.2.1 Utilisation des traitements classiques	25
1.2.2 Adaptation des traitements	26
1.2.2.1 Projection de l'image sur la sphère	27
1.2.2.2 Utilisation de la géométrie du miroir	29
Conclusion	34

2	Modélisation floue	35
	Introduction	35
2.1	Imprécisions sur les images numériques	36
2.1.1	Echantillonnage	37
2.1.2	Modélisation de l'image	41
2.1.2.1	Discrétisation du plan de l'image	41
2.1.2.2	Modélisation de l'information	42
2.2	Théorie des possibilités et traitement d'images	43
2.2.1	Noyaux sommatifs	44
2.2.2	Noyaux non-sommatifs	45
2.2.3	Lien entre les noyaux sommatifs et non-sommatifs	46
2.3	Intégrale de Choquet	49
2.3.1	Intégrale de Choquet continue	49
2.3.2	Intégrale de Choquet discrète	49
2.3.2.1	Vision verticale	51
2.3.2.2	Vision horizontale	54
	Conclusion	56
3	Adaptation des masques de convolution aux images omnidirectionnelles	57
	Introduction	57
3.1	Masques de convolution	57
3.2	Algorithme de projection	59
3.3	Approche proportionnelle	64
3.4	Approche floue	65
	Conclusion	68
4	Une nouvelle approche d'estimation de gradient non-additive	71
	Introduction	71
4.1	Détection de contours dans les images	72
4.1.1	Expression continue d'un contour	72
4.1.2	Estimation du gradient de l'image	72
4.1.3	Maximisation de gradient	74
4.1.4	Seuillage de l'image	75
4.1.5	Variations entre les contours obtenus avec différentes approches	79
4.2	Méthode non-additive d'estimation de gradient dans les images	81
4.2.1	Dérivation d'un signal avec des noyaux non-sommatifs	81
4.2.2	Détection de contour dans une image perspective	82
4.2.3	Détection de contour dans une image omnidirectionnelle	84

4.2.3.1	Modélisation du capteur	85
4.2.3.2	Projection sur le cylindre	86
4.2.3.3	Définition et projection du noyau	90
4.2.3.4	Estimation du noyau dérivé dans chaque direction	92
Conclusion	92
5	Expériences	93
Introduction	93
5.1	Quelques critères quantitatifs pour la détection de contours	93
5.1.1	Paramètres de Fram et Deutsch	93
5.1.2	Critère de bonne localisation	95
5.1.2.1	Mesure des écarts	95
5.1.2.2	Régression linéaire	96
5.1.3	Critère de bonne détection	96
5.2	Présentation des approches comparées	97
5.3	Validation de l'approche Gradient maxitif avec des images perspectives	98
5.3.1	Evaluation quantitative de l'approche du Gradient maxitif	98
5.3.1.1	Application de l'approche du Gradient maxitif sur des images de synthèse	98
5.3.1.2	Evolution du paramètre P_1 en présence de bruit	101
5.3.1.3	Estimation des paramètres P_1 en fonction de l'intensité du contour	102
5.3.1.4	Critère de bonne détection	105
5.3.1.5	Mesure des écarts	106
5.3.1.6	Régression linéaire	108
5.3.2	Robustesse de l'approche du Gradient maxitif sur des images réelles	109
5.4	Comparaison des différentes approches sur des images omnidirectionnelles	112
5.4.1	Influence de l'étalonnage de la caméra sur les traitements	112
5.4.2	Influence de la taille du noyau sur les traitements	116
5.4.3	Comparaison quantitative sur des images artificielles	119
5.4.3.1	Evolution du paramètre P_1 en présence de bruit	119
5.4.3.2	Estimation des paramètres P_1 et P_2 en fonction de l'intensité du contour	120
5.4.3.3	Mesure des écarts	124
5.4.3.4	Régression linéaire	125
5.4.3.5	Critère de bonne détection	126
5.4.4	Expériences sur des images réelles	128
Conclusion	140

Conclusion Générale	143
A Ensembles flous	147
B Dérivation par différences finies	149
C Critères de Canny	151
C.1 Critère de bonne détection	151
C.2 Critère de bonne localisation	151
C.3 Critère d'unicité de la réponse	152
C.4 Résolution du système	152
D Approche de Deriche	153
E Approche de Shen-Castan	155
Publications de l'auteur	157
Bibliography	163
Index	165

Table des figures

1	Schéma en coupe de l'œil.	1
2	(a) Dessin de la "caméra obscura". (b) Appareil photo fixant les images sur une plaque de verre.	2
3	Mouvement des pixels de l'image pour une translation de la caméra parallèlement à la scène (a) dans une image perspective et (b) dans une image omnidirectionnelle. Mouvement des pixels de l'image pour une rotation de la caméra (c) dans une image perspective et (d) dans une image omnidirectionnelle.	4
4	(a) Position de la grille par rapport à une caméra classique, (b) Une grille régulière projetée sur une image perspective.	4
5	(a) Position de la grille par rapport à une caméra omnidirectionnelle, (b) et (c) Images omnidirectionnelle de la grille.	5
6	(a) Image omnidirectionnelle d'une grille, (b) Zoom sur la zone centrale de l'image, (c) Zoom sur la zone périphérique de l'image.	5
1.1	(a) Schéma d'une rotonde. (b) Paysage urbain d'Edimbourg.	8
1.2	Appareil photo à soufflet grand angle.	8
1.3	(a) Caméra fisheye (b) Acquisition d'une mire avec une lentille fisheye afin de mettre en évidence (c) les distorsions des lignes de l'image et (d) la non-déformation des lignes radiales.	9
1.4	(a) Caméra PAL (b) Trajet des rayons lumineux dans la lentille PAL.	9
1.5	(a) Cyclographe de Jules Damoizeau. (b) Photorama de Louis Lumière.	10
1.6	(a) Caméra en rotation autour d'un axe vertical fixe, (b) Reconstitution d'une image panoramique à partir d'une mosaïque d'images perspective, (c) Capteur stéréoscopique dédié à l'acquisition d'images sphériques couleur.	11
1.7	(a) Le capteur omnidirectionnel RingCam, constitué de caméras perspectives. (b) Poly-camera à quatre objectifs grand-angle, (c) Système multi-caméras, (d) Images prélevées sur différentes parties de la sphère visuelle.	12
1.8	Dôme permettant de visualiser une exécution artistique ou sportive à l'université de Carnegie Mellon.	12

1.9	(a) Cuvette en bronze que les chinoises remplissaient d'eau pour se parer. (b) Gravure représentant les vaisseaux de l'armée romaine enflammés par les miroirs ardents d'Archimède.	13
1.10	(a) "Le mariage de Giovanni Arnolfini" de Jan van Eyck en 1434 (National Gallery de Londres), (d) Zoom sur le miroir accroché au mur, reflétant la pièce, les époux de dos et le peintre. (b) "Le prêteur et sa femme" de Quentin Metsis en 1514 (musée du Louvre à Paris), (e) Zoom sur le miroir posé sur la table, représentant une personne qui attends devant la fenêtre. (c) "Saint Éloi orfèvre" de Petrus Christus en 1449 (Metropolitan Museum de New-York.) (f) Zoom sur le miroir posé sur la table qui donne vue sur la rue.	14
1.11	"Hand with Reflecting Sphere" de Escher	15
1.12	Reconstruction d'une image perspective à partir d'une image omnidirectionnelle acquise avec un capteur catadioptrique possédant un miroir hyperbolique.	15
1.13	Famille de capteurs catadioptriques.	16
1.14	Formation de l'image P' d'un point P par réflexion sur un miroir plan.	17
1.15	(a) Miroir pyramidal carré, (b) Miroir pyramidal hexagonal avec deux paires de caméras (AA' , BB'), chacune associé à une face, dont la focale est en (+).	17
1.16	Réflexion d'un point P dans un miroir elliptique.	18
1.17	Réflexion d'un point P dans un miroir hyperbolique.	19
1.18	Reflection d'un point P dans un miroir parabolique.	19
1.19	Modèle de projection unifié pour les capteurs catadioptriques centraux.	20
1.20	Reflection d'un point P dans un miroir conique (a) lorsque la focale coïncide avec le sommet du cône, (b) lorsque le point de vue n'est pas unique.	21
1.21	Réflexion d'un point P dans un miroir sphérique.	22
1.22	Caméra axiale munie d'un miroir hyperbolique.	22
1.23	Surfaces caustiques 3D pour un système catadioptrique (a) symétrique constitué d'un miroir parabolique, (b) asymétrique constitué d'un miroir parabolique, (c) symétrique constitué d'un miroir hyperbolique.	23
1.24	Combinaison d'un miroir parabolique avec un miroir sphérique.	24
1.25	(a) Une caméra est mise en rotation pour acquérir des bandes d'images de 360° , (b) chaque bande correspond à un cercle sur la sphère, (c) la moitié d'une bande représente une ligne dans l'image panoramique.	25
1.26	Projection de l'image sur l'hémisphère.	27
1.27	(a) Définition des patchs sur le miroir puis projection sur l'image omnidirectionnelle. (b) Les pixels sont appareillés lorsque leurs patchs de différentes tailles ont le plus grand coefficient de confiance. (c) Définition de l'angle solide.	30
1.28	Image anamorphosée.	30

1.29	Reconstruction d'une image perspective.	31
1.30	Projection de l'image sur le cylindre englobant.	31
1.31	Projection de l'image sur le cylindre englobant.	32
1.32	Schématisation du traitement d'images	32
2.1	Fonctionnement d'un système d'acquisition d'images numériques. Un boîtier numérique constitué d'un corps, d'un objectif, d'un diaphragme, d'un obturateur, permet de créer une image sur la surface sensible. Lors de l'acquisition, un filtre "infrarouge" limite la sensibilité des capteurs aux radiations et un filtre "passe-bas" supprime les aberrations chromatiques et produit un léger flou pour éviter des erreurs d'interpolation de la couleur par le capteur. Le capteur réagit à l'intensité lumineuse et produit des signaux électriques analogiques permettant de restituer les couleurs de la scène observée. A la sortie du capteur, un amplificateur permet d'augmenter le niveau du signal. Le signal analogique est ensuite converti en signal numérique, les pixels "éteints" issus des défauts de fabrication du capteur sont traités et finalement l'image est enregistrée.	36
2.2	Echantillonnage du signal avec (a) une faible fréquence et (b) une fréquence deux fois plus importante.	37
2.3	Image de Léna échantillonnée (a) avec une fréquence élevée, (b) avec une fréquence moyenne, (c) avec une faible fréquence. Zoom sur l'oeil de Léna (d) avec une fréquence élevée, (e) avec une fréquence moyenne, (f) avec une faible fréquence.	38
2.4	Numérisation de l'espace.	40
2.5	Portion du signal intégré par les pixels d'une même ligne.	40
2.6	Cône de visibilité d'un photosite dans le cas de caméras C.M.O.S. ou de caméras à balayage.	41
2.7	Partitionnement régulier de l'image (a) couvrant vérifiant les propriétés 2.3 et 2.2, (b) non-couvrant vérifiant uniquement la propriété 2.2, (c) couvrant vérifiant uniquement la propriété 2.3	42
2.8	Représentation d'un pixel flou	43
2.9	Schéma général de traitement d'images.	44
2.10	(a), (b) et (c) Distributions de probabilité. (d) Distribution de possibilité. (e) Distribution de possibilité triangulaire.	48
2.11	Noyau non-sommatif W	50
2.12	Interactions entre un pixel et ses voisins.	51
2.13	(a) Chacune des deux cellules interagit avec l'autre, l'estimation des valeurs par l'intégrale de Choquet fournit pour chacun (b) une borne supérieure obtenue en privilégiant les cellules ayant les plus grandes valeurs et (c) une borne inférieure privilégiant les cellules ayant les plus petites valeurs.	52

2.14	(a) Application d'un noyau non-sommatif W à un ensemble de pixels. (b) Schématisation des résultats obtenus l'intégrale de Choquet inférieure. (c) Schématisation des résultats obtenus l'intégrale de Choquet supérieure.	54
2.15	Les interactions entre deux cellules dépendent des hauteurs d'intersection entre ces cellules.	55
2.16	Hauteur d'intersection des cellules avec le noyau W	56
3.1	Les filtres (a) Moyenneur, (b) Gaussien, (c) Prewitt, (d) Sobel, (e) Laplacien	58
3.2	(a) Projection du centre du pixel sur le cylindre. (b) Reconstruction d'un masque régulier autour du centre du pixel projeté. (b) Projection du masque sur l'image omnidirectionnelle.	61
3.3	Quasi-équivalence entre la projection d'un pixel sur le cylindre et sur l'image perspective	62
3.4	Projection des masques 3×3 sur l'image omnidirectionnelle.	63
3.5	Le masque projeté couvre plusieurs pixels de l'image omnidirectionnelle.	63
3.6	Interactions entre un élément du masque projeté et les pixels de l'image.	64
3.7	Zoom sur l'élément central du masque : le niveau de gris du noyau projeté est proportionnelle à la surface des pixels intersectés.	64
3.8	Simplification du masque approximé un quadrilatère.	65
3.9	Représentation floue de l'élément central du masque sur le cylindre.	66
3.10	Projection de l'élément central du masque sur l'image omnidirectionnelle.	66
3.11	Calcul de $\Pi(P_i, M_{u,v})$ comme la hauteur d'intersection entre P_i et $M_{u,v}$	67
3.12	Hauteur d'intersection entre les nombres flous.	68
3.13	Interactions entre les éléments du masque sur l'image omnidirectionnelle.	69
4.1	Etapas de détection de contours.	73
4.2	Vue 3D des contours présents dans l'image de Léna.	73
4.3	Maximisation de gradient : le pixel $P_{i,j}$ est un point de contour dans une image perspective si la norme de son gradient est maximum dans la direction du gradient.	74
4.4	Vue 3D des contours présents dans l'image de Léna à laquelle on a ajouté un bruit gaussien.	75
4.5	Détection de contours sur l'image de Léna pour différentes valeurs de seuil (a) $s_h = 29$ et $s_b = 26$, (b) $s_h = 0$ et $s_b = 0$, (c) $s_h = 60$ et $s_b = 55$, (d) $s_h = 32$ et $s_b = 30$	76
4.6	(a) Image originale des poivrons et détection de contours sur cette image avec (b) le même seuil que pour Léna $s_h = 29$ et $s_b = 26$ et (c) avec un seuil plus faible $s_h = 12$ et $s_b = 10$	77
4.7	(a) Image originale bruitée avec un bruit de Poisson. (b) $s_h = 10$ et $s_b = 8$, (c) $s_h = 20$ et $s_b = 18$, (d) $s_h = 25$ et $s_b = 23$, (e) $s_h = 108$ et $s_b = 105$	78

4.8	(a) Zoom sur l'oeil de Léna. (b) Contours obtenus avec le filtre de Canny-Deriche. (c) Contours obtenus avec le filtre de Shen-Castan. (d) Contours obtenus avec le filtre de Sobel. (e) Contours obtenus avec le filtre de Prewitt.	79
4.9	Présentation des noyaux de Canny-Deriche, Shen-Castan, gaussien, de Prewitt et de Sobel.	80
4.10	Noyau dérivé construit par soustraction de deux noyaux usuels.	82
4.11	Extension en 2D du noyau dérivé.	83
4.12	Maximisation de gradient dans une image omnidirectionnelle	84
4.13	Projection des noyaux sur l'image omnidirectionnelle.	85
4.14	Projection d'un point de l'image sur le cylindre.	86
4.15	Deux solutions possibles pour projeter un point sur le cylindre.	88
4.16	Définition des noyaux dérivatifs sur le cylindre dans chaque direction.	90
4.17	Calcul des coordonnées géométriques des noyaux sur le cylindre : (a) pour un noyau dérivatif horizontal, (b) pour un noyau dérivatif vertical.	91
4.18	Noyaux dérivatifs projetés dans chaque direction.	92
5.1	Modélisation d'un contour dans l'image.	94
5.2	Mesure des écarts-types : les pixels grisés appartiennent au contour. Les pixels gris foncé ont été marqués par le détecteur de contours, les pixels gris clair n'ont pas été détectés.	95
5.3	Droite de régression linéaire : les pixels grisés appartiennent à la zone de contour. Les pixels gris clair sont ceux détectés par l'opérateur et les pixels gris foncé ne le sont pas. Le contour théorique, représenté en pointillé est calculé à partir de l'ensemble des pixels appartenant au contour alors que la droite de régression linéaire représentée en noir est calculée uniquement à partir des pixels marqués par le détecteur de contour.	96
5.4	(a) Image de synthèse originale non bruitée. (b) Ajout d'un bruit gaussien de variance $\sigma = 10$. (c) Ajout d'un bruit gaussien de variance $\sigma = 20$. (d) Ajout d'un bruit de Poisson.	99
5.5	Image originale (a) et les contours obtenus avec un filtre de (b) Deriche sans seuillage, (c) Deriche avec un seuillage arbitraire adapté $s_h = 35$ et $s_b = 30$. (d) Shen sans seuillage, (e) Shen avec un seuillage arbitraire adapté $s_h = 24$ et $s_b = 20$. (f) Prewitt sans seuillage, (g) Prewitt avec un seuillage arbitraire adapté $s_h = 120$ et $s_b = 115$. (h) avec notre approche sans seuillage. (i) et (q) représentent l'image originale avec ajout d'un bruit gaussien tel que $\sigma = 10$ et $\sigma = 20$. Les mêmes estimations que (b), (c), (d), (e), (f), (g) et (h) sont fournies pour l'image (i) en (j), (k), (l), (m), (n), (o) et (p) et pour l'image (q) en (r), (s), (t), (u), (v), (w) et (x).	100

5.6 (a) représente l'image originale 5.4(a) avec ajout d'un bruit de poisson et les contours obtenus avec un filtre de Deriche sans seuillage, (b) Deriche avec un seuillage arbitraire adapté $s_h = 12$ et $s_b = 10$. (c) Shen sans seuillage, (d) Shen avec un seuillage arbitraire adapté $s_h = 8$ et $s_b = 6$. (e) Prewitt sans seuillage, (f) Prewitt avec un seuillage arbitraire adapté $s_h = 60$ et $s_b = 55$. (g) avec notre approche sans seuillage. 100

5.7 Valeurs comparatives de P_1 pour un bruit gaussien. (a) Sans seuillage, (b) Avec seuillage (pour toutes les approches excepté l'approche du Gradient maxitif qui fournit un seuillage naturel). 101

5.8 (a) Image originale bruitée avec un bruit de Poisson. (b) Résultat obtenu avec notre approche sans seuillage. Les résultats obtenus avec un filtre de Deriche sont présentés pour différentes valeurs de seuil : (c) $s_h = 10$ et $s_b = 8$, (d) $s_h = 20$ et $s_b = 18$, (e) $s_h = 25$ et $s_b = 23$, (f) $s_h = 108$ et $s_b = 105$. Les résultats obtenus avec un filtre de Shen sont présentés pour différentes valeurs de seuil : (g) $s_h = 1$ et $s_b = 0$, (h) $s_h = 3$ et $s_b = 2$, (i) $s_h = 10$ et $s_b = 8$, (j) $s_h = 14$ et $s_b = 13$. Les résultats obtenus avec un filtre de Prewitt sont présentés pour différentes valeurs de seuil : (k) $s_h = 34$ et $s_b = 33$, (l) $s_h = 48$ et $s_b = 46$, (m) $s_h = 60$ et $s_b = 55$, (n) $s_h = 83$ et $s_b = 79$. . 103

5.9 Image de la croix pour des sauts de (a) 3 niveaux de gris, (b) 6 niveaux de gris, (c) 12 niveaux de gris, (d) 20 niveaux de gris et (e) 100 niveaux de gris. 104

5.10 (a) Valeurs de P_1 obtenues sans seuillage des images. 104

5.10 (b) Valeurs de P_1 obtenues avec seuillage des images (sauf pour l'approche Gradient maxitif). 105

5.11 (a) Rapports signal sur bruit obtenus sans seuillage. 105

5.11 (b) Rapports signal sur bruit obtenus avec seuillage. 106

5.12 (a) Contour théorique. Les images obtenues pour des sauts de (b) 3 Niveaux de gris, (c) 6 Niveaux de gris, (d) 12 Niveaux de gris, (e) 20 Niveaux de gris. 106

5.13 (a) Moyennes obtenues pour des sauts d'amplitude 3, 6, 12 et 20 niveaux de gris pour chaque approches. 107

5.13 (b) Ecart-types entre les pixels du contour extrait et la transition théorique obtenus pour des sauts d'amplitude 3, 6, 12 et 20 niveaux de gris pour chaque approche. . . . 107

5.14 (a) Pourcentage d'erreur sur la pente de la droite détectée pour des sauts d'amplitude 20, 12, 6 et 3 niveaux de gris pour chaque approche. 108

5.14 (b) Pourcentage d'erreur sur l'ordonnée à l'origine pour des sauts d'amplitude 20, 12, 6 et 3 niveaux de gris pour chaque approche. 109

5.15 Image réelle originale. 110

5.16 (a), (b), (c), (d) et (e) représentent l'image originale après réhaussement de la luminosité et du contraste afin de mieux visualiser les contours. 111

5.16	(f), (g), (h), (i) et (j) montrent les contours obtenus avec notre approche. (k), (l), (m), (n) et (o) montrent les contours obtenus avec le filtre de Deriche pour les valeurs de seuil donnant le meilleur rapport entre contours désirés et contours non-désirés $s_h = 16$ et $s_b = 13$; (p), (q), (r), (s) et (t) les images obtenues avec des valeurs de seuil plus élevées permettant de détecter correctement la grille $s_h = 30$ et $s_b = 28$; (u), (v), (w), (x) et (y) les images obtenues avec des valeurs de seuil plus faibles permettant de détecter correctement le corps du poisson $s_h = 6$ et $s_b = 4$	112
5.17	Influence de φ sur le système de projection.	113
5.18	Influence de ξ sur l'image pour (a) $\xi = 1.54$, (b) $\xi = 1.88$, (c) $\xi = 2$, et (d) $\xi = 2.35$	114
5.19	Détection de contour obtenue avec (a) les paramètres de calibration correctes de la caméra (b) avec $k_u = 163.100$ et $k_v = 203.875$	115
5.20	Exemples de noyaux projetés (a) avec les bonnes coordonnées de projection de l'axe optique dans l'image et (b) avec des coordonnées erronées.	115
5.21	Exemples de noyaux projetés (a) avec $(u_0, v_0) = (300, 390)$ et (b) avec $(u_0, v_0) = (300, 350)$	116
5.22	Secteur angulaire minimal et maximal du noyau.	116
5.23	Projections de noyaux de différentes tailles sur l'image omnidirectionnelle : (a) Noyaux de grande taille (40 pixels) (b) Noyaux de taille moyenne (20 pixels) (c) Noyaux de petite taille (10 pixels).	117
5.24	(a) Image originale. (b) Détection de contours obtenue avec un masque quatre fois plus petit que celui que nous avons utilisé.	118
5.24	(c) Détection de contours obtenue avec un masque quatre fois plus grand que celui que nous avons utilisé, (d) la taille du masque que nous avons utilisé, résulte d'un bon compromis entre le lissage et la précision.	118
5.25	Image artificielle bruitée avec un saut de 20 niveau de gris.	119
5.26	(a) Valeurs comparatives de P_1 pour un bruit gaussien sans maximisation ni seuillage.	120
5.26	(b) Valeurs comparatives de P_1 pour un bruit gaussien avec maximisation et seuillage.	120
5.27	Adaptation de la zone de contour pour une transition oblique.	121
5.28	(a) Paramètres P_1 obtenus pour une transition verticale.	122
5.28	Paramètres P_2 obtenus pour une transition verticale.	122
5.29	(a) Paramètres P_1 obtenus pour une transition oblique.	123
5.29	(b) Paramètres P_2 obtenus pour une transition oblique.	123
5.30	(a) Moyennes obtenues pour des sauts d'amplitude 20, 12, 6 et 3 niveaux de gris pour chaque approche.	124
5.30	(b) Ecart-types entre les pixels du contour extrait et la transition théorique obtenus pour des sauts d'amplitude 20, 12, 6 et 3 niveaux de gris pour chaque approche.	125

5.31 (a) Pourcentage d'erreur sur la pente de la droite détectée pour des sauts d'amplitude 20, 12, 6 et 3 niveaux de gris pour chaque approche.	126
5.31 (b) Pourcentage d'erreur sur l'ordonnée à l'origine pour des sauts d'amplitude 20, 12, 6 et 3 niveaux de gris pour chaque approche.	126
5.32 (a) Rapport signal sur bruit obtenu sur des images sans maximisation ni seuillage. . .	127
5.32 (b) Rapport signal sur bruit obtenu sur des images après maximisation et seuillage (sauf l'approche maxitive qui fournit un seuillage naturel et n'a donc pas été seuillée). 127	
5.33 Notre capteur.	128
5.34 Deux images réelles	129
5.35 (a) Filtre de Prewitt classique appliqué à l'image omnidirectionnelle, (b) Filtre de Prewitt classique appliqué à l'image désanamorphosée, (c) Filtre de Prewitt utilisant l'approche proportionnelle présentée dans la Section 3.3, (d) Filtre de Prewitt utilisant l'approche floue présentée dans la Section 3.4 sans prendre en compte les interactions entre les éléments du masque.	132
5.35 (e) Filtre de Prewitt utilisant l'approche floue avec interactions présentée dans la section 3.4, (f) Estimation de gradient maxitif présentée dans la section 4.2, (g) Filtre de Canny-Deriche appliqué à l'image omnidirectionnelle, (h) Filtre de Shen-Castan appliqué à l'image omnidirectionnelle.	133
5.36 Zoom sur l'image précédente : (a) Filtre de Prewitt classique appliqué à l'image omnidirectionnelle, (b) Filtre de Prewitt classique appliqué à l'image désanamorphosée, (c) Filtre de Prewitt utilisant l'approche proportionnelle présentée dans la section 3.3, (d) Filtre de Prewitt utilisant l'approche floue présentée dans la section 3.4 sans prendre en compte les interactions entre les éléments du masque.	134
5.36 Zoom sur l'image précédente : (e) Filtre de Prewitt utilisant l'approche floue avec interactions présentée dans la section 3.4, (f) Estimation de gradient maxitif présentée dans la section 4.2, (g) Filtre de Canny-Deriche appliqué à l'image omnidirectionnelle, (h) Filtre de Shen-Castan appliqué à l'image omnidirectionnelle.	135
5.37 (a) Filtre de Prewitt classique appliqué à l'image omnidirectionnelle, (b) Filtre de Prewitt classique appliqué à l'image désanamorphosée, (c) Filtre de Prewitt utilisant l'approche proportionnelle présentée dans la section 3.3, (d) Filtre de Prewitt utilisant l'approche floue présentée dans la section 3.4 sans prendre en compte les interactions entre les éléments du masque.	136
5.37 (e) Filtre de Prewitt utilisant l'approche floue avec interactions présentée dans la section 3.4, (f) Estimation de gradient maxitif présentée dans la section 4.2, (g) Filtre de Canny-Deriche appliqué à l'image omnidirectionnelle, (h) Filtre de Shen-Castan appliqué à l'image omnidirectionnelle.	137

5.38	Zoom sur l'image précédente : (a) Filtre de Prewitt classique appliqué à l'image omnidirectionnelle, (b) Filtre de Prewitt classique appliqué à l'image désanamorphosée, (c) Filtre de Prewitt utilisant l'approche proportionnelle présentée dans la section 3.3, (d) Filtre de Prewitt utilisant l'approche floue présentée dans la section 3.4 sans prendre en compte les interactions entre les éléments du masque.	138
5.38	Zoom sur l'image précédente : (e) Filtre de Prewitt utilisant l'approche floue avec interactions présentée dans la section 3.4, (f) Estimation de gradient maxitif présentée dans la section 4.2, (g) Filtre de Canny-Derliche appliqué à l'image omnidirectionnelle, (h) Filtre de Shen-Castan appliqué à l'image omnidirectionnelle.	139
A.1	(a), (b) Intervalle flou. (c) Nombre flou (d) Nombre flou triangulaire.	147
B.1	Filtre de Prewitt (a) Filtre passe-bas effectuant le lissage, (b) Filtre passe haut effectuant la dérivation.	149
B.2	Filtre de Sobel (a) Filtre passe-bas effectuant le lissage, (b) Filtre passe haut effectuant la dérivation.	149
D.1	(a) Filtre passe-bas de Canny-Derliche, (b) Filtre passe haut de Canny-Derliche.	153
E.1	(a) Filtre passe-bas de Shen-Castan, (b) Filtre passe haut de Shen-Castan.	155

Liste des tableaux

5.1	Valeurs de P_1 avec un bruit de Poisson.	102
-----	--	-----

Liste des algorithmes

1	Algorithme de convolution d'une image avec un masque linéaire	59
2	Algorithme de projection des noyaux	60

Introduction Générale

L'œil ou globe oculaire est l'organe de la vision, c'est-à-dire l'organe du sens qui permet aux êtres vivants d'analyser la lumière pour interagir avec leur environnement. La lumière pénètre dans l'œil par la pupille, un orifice circulaire situé au centre de l'iris. Le cristallin, la cornée et l'humeur aqueuse constituent un objectif qui projette sur la rétine une image renversée de la scène observée. Les photons de la partie visible du spectre lumineux sont captés et convertis en impulsions électriques. Les informations sont ensuite transmises au cerveau par le nerf optique, engendrant la représentation et l'interprétation du monde extérieur. La vision est donc étroitement liée avec la notion d'images. Ceci explique en partie pourquoi, l'être humain a de tout temps cherché à représenter son environnement par le biais d'illustrations : peintures, sculptures, gravures puis photographie.

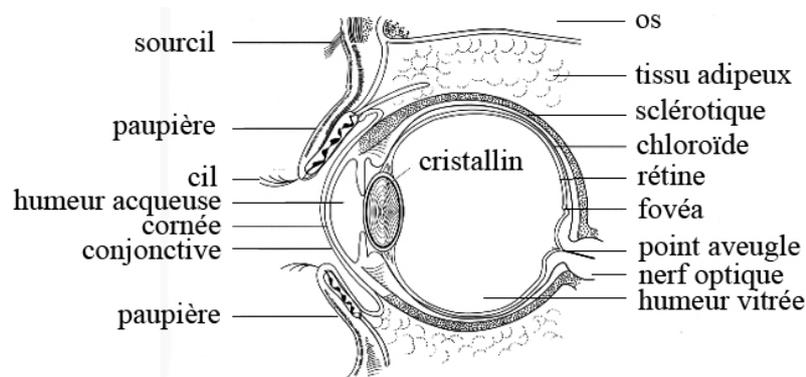


FIG. 1 – Schéma en coupe de l'œil.

Au IV^e siècle avant Jésus-Christ, Aristote découvre le principe de la chambre noire. Au cours d'une éclipse solaire, un rayon issu d'un orifice circulaire projetait, sur le mur opposé d'une pièce plongée dans l'obscurité, son image inversé. C'est le principe du sténopé, du grec "stenos qui signifie *étroit* et "ôps", *œil*. Ce phénomène de "caméra obscura" (Chambre Noire) est décrit en 1515 par Léonard de Vinci (Figure 2(a)). En 1540, Jérôme Cardan remplace le petit trou par une lentille afin de dessiner avec exactitude les perspectives. Ce n'est qu'au XIX^e siècle que Nicéphore Niepce inventa la photographie en fixant les images, en négatif à l'acide nitrique sur chlorure d'argent en 1816, puis en positif à l'aide du bitume de Judée étendu sur une plaque de verre en 1822 (Figure 2(b)). William Talbot met au point le procédé négatif-positif actuel ou calotype qui permet, à par-

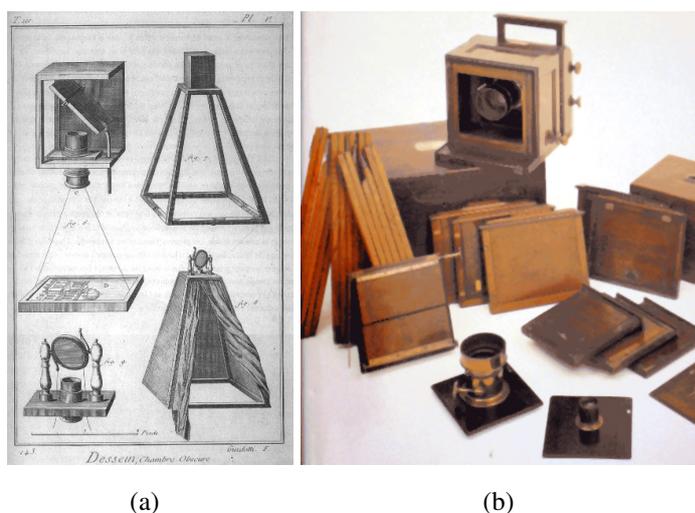


FIG. 2 – (a) Dessin de la "caméra obscura". (b) Appareil photo fixant les images sur une plaque de verre.

tir d'une image négative, d'obtenir par contact un nombre illimité de reproductions sur papier salé (au chlorure d'argent). En 1884, George Eastman met au point le film en celluloïd (surface sensible souple) permettant de stocker plusieurs images dans l'appareil photographique. Les frères Lumière inventent en 1903 le premier procédé de photographie en couleur, l'autochrome. 10 ans plus tard, le film 35 mm et les appareils de petit format font leur apparition avec la réalisation du Leica, premier prototype d'Oskar Barnack. Le premier appareil "Polaroid" à développement instantané est mis au point en 1948, par Edwin Land. Apparaissent ensuite le premier reflex avec contrôle automatique de l'exposition par mesure de la lumière à travers l'objectif en 1968 et le premier compact autofocus au monde en 1977. A la fin du XXe siècle, le premier appareil photo numérique de l'histoire est présenté par Sony en 1981. Le développement des techniques informatiques transforme l'image en une série de points, les pixels, de l'anglais "picture cells". Un format de codage des couleurs permet de reconstituer les couleurs visibles par l'œil humain. A l'écran, le principe de synthèse additive des couleurs est utilisé pour représenter une couleur visible : la lumière provenant de chacune des trois sources mono-chromatiques primaires Rouge Vert Bleu est additionnée, selon des proportions bien choisies. Cette technologie marque une rupture complète avec les procédés physico-chimiques initiaux. Sa difficulté se situe dans la conception des capteurs électroniques de l'image qui remplacent le film. Comme pour les anciennes techniques, cette qualité dépend aussi de l'optique de l'objectif et de la mécanique du boîtier. La photographie numérique présente des possibilités infinies de traitements et de retouches des images par ordinateur.

Héritée de la peinture et des spectacles populaires, la vision panoramique suscite l'intérêt des photographes. Elle a pris une place importante dans de nombreuses applications telles que la sur-

veillance, la téléconférence, l'acquisition de modèles pour la réalité virtuelle, l'immersion virtuelle ou bien tout simplement la navigation des robots.

Pour étendre le champ de vision, l'homme s'est inspiré du système de vision animale. En effet, dans leur milieu naturel, la connaissance de l'environnement est indispensable à la survie des espèces. Afin de détecter efficacement le mouvement ou de repérer les prédateurs, la géométrie oculaire s'est adaptée à l'animal considéré et à son environnement. Ainsi, le système de vision des proies est conçu pour examiner le plus large secteur possible. Leurs yeux sont placés de chaque côté du museau, de manière à étendre leur champ visuel jusqu'à 360°. Le zone de recouvrement du champ visuel des deux yeux est restreinte, voire nulle. Les prédateurs, en revanche, possèdent une meilleure acuité visuelle afin de localiser exactement leurs proies. Leurs yeux sont proches et offrent un champ de vision plus restreint, mais une excellente vision stéréoscopique¹.

Des mécanismes d'acquisition d'images omnidirectionnelles tels que les objectifs spéciaux à distance focale très courte, les systèmes rotatifs associés à une caméra, de multiples caméras disposées de manière à visualiser la totalité de l'espace 3D, etc. ont été créés. Ces solutions impliquent un raccordement spatial ou temporel des différentes images. Ce problème non trivial a conduit à l'utilisation de miroirs de révolutions couplés avec une caméra. Ces capteurs sont apparus sur le marché depuis une vingtaine d'année. Ils ont rendu possible l'acquisition, en une prise de vue, de scènes panoramiques complètes 360° à la fréquence vidéo de 25 images/seconde. Cependant, ces caméras ont une rétine de forme rectangulaire avec une distribution de pixels régulière qui n'est pas adaptée aux images omnidirectionnelles. Dans ce manuscrit, nous avons choisi de nous intéresser aux images acquises avec ces capteurs.

La représentation à 360° de la scène 3D est particulièrement intéressante pour la navigation de robots mobiles [19] [15]. Les caméras panoramiques permettent également de lever quelques ambiguïtés pour l'estimation de mouvement dans les séquences vidéo. En effet, sur les images perspectives, il est parfois difficile de distinguer un mouvement de translation (Figure 3(a)), d'un mouvement de rotation (Figure 3(c)), pour de petits déplacements de la caméra. Les caméras omnidirectionnelles en revanche contiennent toujours un point d'expansion et un point de contraction, ce qui ne laisse aucun doute possible sur la nature du mouvement, comme le montrent Svoboda, Padjla et Hlavac dans [65].

Qu'il s'agisse de détection de mouvement, d'extraction de contours, de lissage ou autres traitements, l'hypothèse utilisée pour définir les opérateurs est l'invariance par translation. Cette hypothèse traduit le fait que quelle que soit la position où l'on applique le traitement dans l'image, son effet reste le même. Dans le cas d'une image perspective, si la scène, parallèle au plan focal, présente peu de différences de profondeur (quasi-plane), alors on peut considérer que cette hypothèse est vérifiée. Ce phénomène est illustré Figure 4(a) et (b), où l'on peut observer l'image d'une grille dans

¹**Vision stéréoscopique** : Perception binoculaire de la profondeur et de la distance d'un objet résultant du chevauchement des champs de vision des deux yeux.

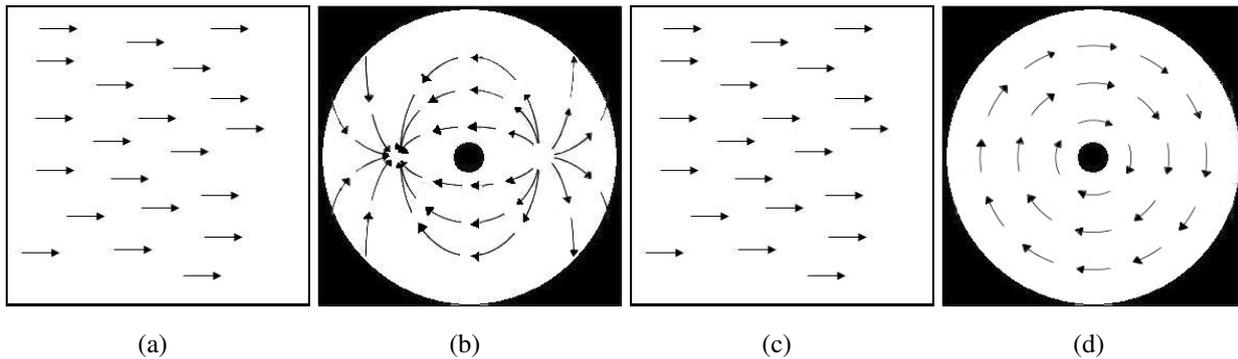


FIG. 3 – Mouvement des pixels de l’image pour une translation de la caméra parallèlement à la scène (a) dans une image perspective et (b) dans une image omnidirectionnelle. Mouvement des pixels de l’image pour une rotation de la caméra (c) dans une image perspective et (d) dans une image omnidirectionnelle.

un plan parallèle au plan focal. Sur cette image, la résolution est constante, les contours conservent

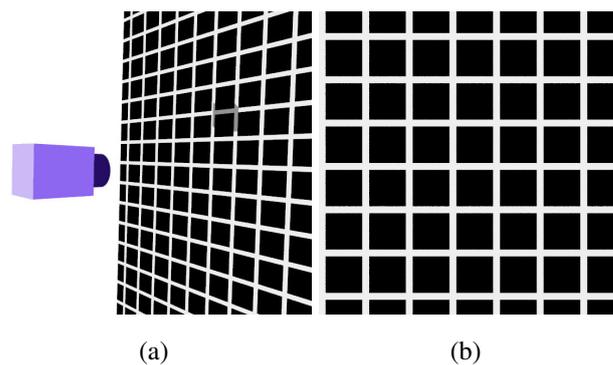


FIG. 4 – (a) Position de la grille par rapport à une caméra classique, (b) Une grille régulière projetée sur une image perspective.

la même épaisseur, etc. On peut donc penser que les opérateurs classiques qui ont le même effet en tout point de l’image sont appropriés à la géométrie de cette image. Dans le cas d’une image omnidirectionnelle, il est impossible de trouver une position de la caméra permettant de satisfaire la propriété d’invariance par translation, comme on peut le voir sur les Figures 5(a) et (b). L’image présente une résolution non-homogène. Un traitement classique, quel qu’il soit, n’aura pas le même effet au centre et à la périphérie de l’image. En effet, la géométrie du miroir de révolution du système d’acquisition introduit des distorsions importantes dans l’image et crée une concentration de l’information au centre de l’image. Ce phénomène est mis en évidence Figure 6(a), (b) et (c). Ce constat soulève un problème : Si les caméras omnidirectionnelles ont réussi à augmenter le champ de vision des caméras classiques, cela s’est fait au détriment de la simplicité des traitements. Si les conditions d’utilisation des traitements classiques ne sont pas respectées pour les images omnidirectionnelles,

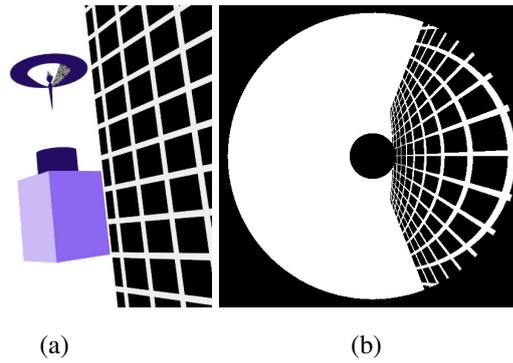


FIG. 5 – (a) Position de la grille par rapport à une caméra omnidirectionnelle, (b) et (c) Images omnidirectionnelle de la grille.

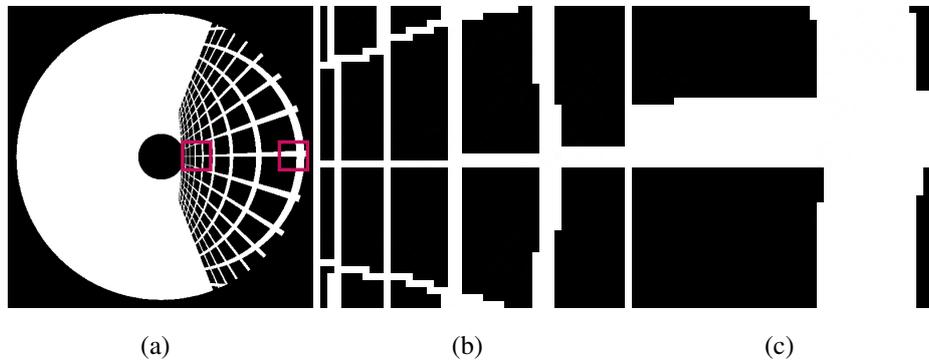


FIG. 6 – (a) Image omnidirectionnelle d’une grille, (b) Zoom sur la zone centrale de l’image, (c) Zoom sur la zone périphérique de l’image.

il va falloir définir de nouveaux opérateurs adaptés à la géométrie de ces images. C’est ce que nous proposons dans ce manuscrit.

Le premier Chapitre présente un état de l’art de la vision omnidirectionnelle. Un historique des méthodes mises en œuvre par l’homme pour élargir le champ visuel est décrit, à savoir le remplacement de l’optique de la caméra par une lentille à distance focale très courte, la conception d’un système rotatif fournissant différentes vues de l’environnement afin de reconstituer une image panoramique, l’assemblage de caméras permettant d’acquérir en une seule prise la totalité de la scène, l’association de miroirs de révolution aux caméras classiques. Cette dernière approche ayant retenu notre attention, nous nous sommes intéressés aux algorithmes développés pour traiter ces images omnidirectionnelles. La géométrie du miroir introduit des déformations dans ces images et modifie la notion de voisinage et la résolution de l’image. Les opérateurs de traitement d’images classiques, développés pour des images perspectives, ne sont pas satisfaisants pour les images omnidirectionnelles. Dans la littérature, les auteurs peuvent être regroupés en deux catégories, ceux qui ont traité les images omnidirectionnelles de la même manière que les images perspectives dans la mesure du

possible et ceux qui ont cherché des méthodes adaptées à la géométrie des images. De notre état de l'art, il est ressorti qu'il existait peu de méthodes permettant de traiter les images omnidirectionnelles et que ces méthodes étaient souvent complexes à réaliser. C'est pourquoi nous proposons, dans la suite de ce manuscrit, de développer des traitements adaptés à la géométrie des images omnidirectionnelles.

Notre objectif étant de traiter des images numériques, nous avons repris la chaîne d'acquisition de l'information de luminance depuis sa projection continue jusqu'aux données numériques représentées par les pixels. Nous avons pu constater les imprécisions sur la mesure de luminance introduite dans l'image lors de l'échantillonnage spatial. De plus dans les images omnidirectionnelles, la résolution n'est pas uniforme en tout point de l'image. L'information est plus concentrée au centre de l'image qu'à la périphérie. Ce phénomène ajoute des imprécisions supplémentaires sur les mesures de luminance. Pour prendre en compte l'ensemble de ces phénomènes, nous proposons d'utiliser une représentation floue des données.

Pour traiter nos images omnidirectionnelles, nous nous sommes orientés vers un modèle projectif utilisant le cylindre englobant le miroir de révolution. Dans un premier temps, nous avons proposé une adaptation des masques de convolutions aux images omnidirectionnelles prenant en compte les imprécisions sur la localisation du niveau de gris des pixels. Dans un second temps, nous nous sommes intéressés plus précisément aux détecteurs de contours. Nous avons développé une nouvelle approche d'estimation de gradient utilisant des noyaux possibilistes. L'estimation du gradient est fournie sous la forme d'un intervalle de valeurs. Cet intervalle permet de seuiller naturellement les valeurs de gradient obtenues.

Nous proposons ensuite une série d'expérimentations sur des images de synthèses visant à mettre en avant la robustesse de nos méthodes. Nous présentons entre autre les résultats obtenus avec les critères de Canny de bonne détection, bonne localisation et réponse unique du contour, ainsi que les paramètres de Fram et Deutsch visant à prouver que nous obtenons de bons résultats quelle que soit l'image utilisée et la nature du bruit qu'elle contient. Nous illustrons également notre approche en présentant les résultats obtenus sur des images réelles par comparaison avec les approches existantes.

Finalement, une conclusion générale dresse un bilan des apports de notre approche et fait ressortir les différentes perspectives de ce travail.

Chapitre 1

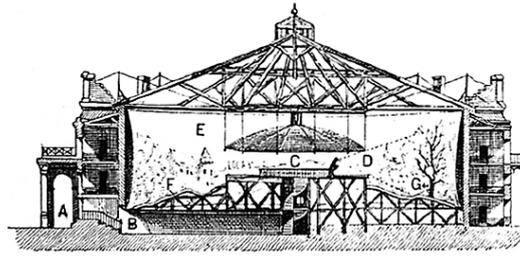
Etat de l'art

Introduction

Le mot panorama apparaît à la fin du siècle XVIIIe siècle, des mots grecs "pan" qui signifie *toute* et "horama", *la vue*. Il fait référence à l'invention de l'écossais Robert Barker breveté en 1767 : incarcéré pour dettes à la prison d'Edimbourg, il remarque l'effet singulier de la lumière sur les murs de son cachot et imagine alors une rotonde (Figure 1.1(a)) où des toiles 360°, tendues le long des murs, seraient éclairées par une verrière. Au centre, une plate-forme accueillerait les visiteurs qui auraient ainsi l'illusion de se trouver au cœur d'un vaste paysage. Un parajour cache la source lumineuse afin d'éviter la projection des ombres des spectateurs sur la peinture et de rendre les couleurs de la toile plus éclatantes. Le tableau, illusion de la réalité, conserve la perspective : pour juger l'effet de chaque trait, le peintre doit "grimper sur le point de vue" au centre de la plate-forme, à mi-hauteur de la toile (6,50 mètres). Malheureusement, la plupart des peintures panoramiques ont été perdues. La Figure 1.1(b) représente le paysage urbain d'Edimbourg peint par Robert Barker en 1787.

Le besoin d'élargir le champ de vision s'est également manifesté dans d'autres domaines tels que la photographie et la cinématographie puis pour des applications telles que la surveillance, la vidéoconférence, l'acquisition de modèles pour la réalité virtuelle, l'immersion virtuelle ou la navigation de robots mobiles. Cela vient du fait que les caméras traditionnelles, aussi appelées caméras perspectives, ont un champ de vue très limité, généralement compris entre 30° et 60° seulement. L'espace perçu par la caméra est restreint par la taille de la rétine et l'optique de la caméra. Plusieurs solutions optiques ou mécaniques ont alors été imaginées pour élargir le champ de vision.

L'objectif de ce chapitre est de présenter différentes méthodes utilisées pour étendre le champ de vision des caméras classiques. Dans un premier temps nous présenterons les solutions technologiques et nous préciserons leurs avantages et leurs inconvénients. Nous aborderons dans un deuxième temps le problème des traitements adaptés aux images omnidirectionnelles.



(a)



(b)

FIG. 1.1 – (a) Schéma d'une rotonde. (b) Paysage urbain d'Edimbourg.

1.1 Acquisition et géométrie des images omnidirectionnelles

1.1.1 Objectifs spéciaux

Les premiers appareils grand-angle remontent au début du XXe siècle (Figure 1.2). L'approche

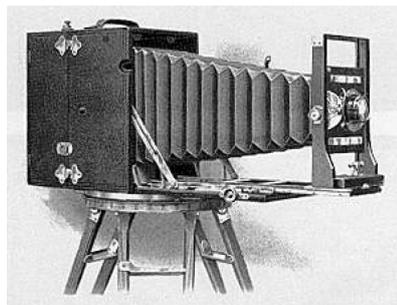


FIG. 1.2 – Appareil photo à soufflet grand angle.

consiste à remplacer l'optique de la caméra par une lentille de distance focale très courte appelée "Fish-eye" (œil de poisson) illustrée Figure 1.3(a). Cette lentille dévie les rayons lumineux et offre un angle d'ouverture dépassant les 180° . Dans [21], les lentilles Fish-eye sont utilisées par Bakstein et Pajdla pour acquérir des images $360^\circ \times 360^\circ$ à l'aide d'une technique de mosaïque. Cependant les images sont déformées : contrairement aux objectifs classiques, la distorsion n'est pas corrigée. Toutes les lignes sont fortement courbées (Figure 1.3(c)) à l'exception de celles passant par le centre

de l'image (Figure 1.3(d)). Pour retrouver les angles de ces images, elles doivent être projetées sur un écran sphérique. L'image perspective ne peut être reconstituée, ce qui rend l'analyse de ces images relativement difficile, principalement sur la périphérie, où la résolution est faible. De plus, ces lentilles sont complexes à réaliser et par conséquent très onéreuses ($\simeq 800$ €).

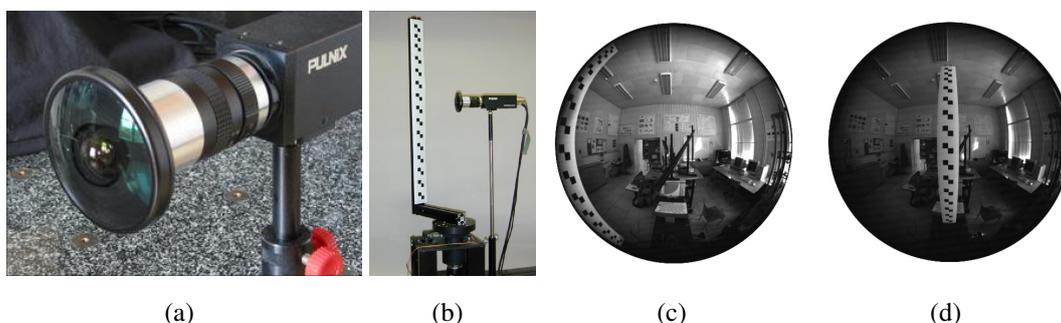


FIG. 1.3 – (a) Caméra fisheye (b) Acquisition d'une mire avec une lentille fisheye afin de mettre en évidence (c) les distorsions des lignes de l'image et (d) la non-déformation des lignes radiales.

Dans [20], Greguss a proposé une lentille appelée "Panoramic Annular Lens" (PAL) qui acquiert des images panoramiques sans distorsion. Le système est composé d'un bloc optique d'une seule pièce de verre sur laquelle deux plans de réflexion sont créés : les rayons lumineux après deux réflexions et deux réfractions sont projetés sur la lentille de la caméra (Figure 1.4). Une version avancée de cette lentille inclut une optique auxiliaire supplémentaire pour compenser l'angle mort dans la direction de l'axe optique. En raison de l'analogie avec l'oeil humain possédant une vision fovéale et une vision périphérique, cette image est appelée Humanoid Machine Vision System (HMVS). Le système est facilement miniaturisable, cependant, sa modélisation est complexe.

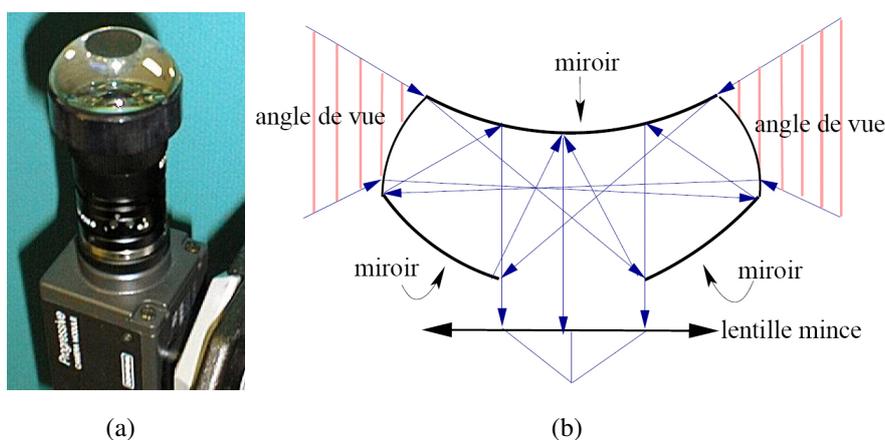


FIG. 1.4 – (a) Caméra PAL (b) Trajet des rayons lumineux dans la lentille PAL.

1.1.2 Systèmes rotatifs

La deuxième solution consiste à faire tourner une caméra perspective autour d'un axe vertical fixe à vitesse angulaire constante.

Au XIXe siècle, les procédés panoramiques permettaient d'obtenir des images couvrant jusqu'à 180°. Le premier panorama complet par rotation d'un appareil fut réalisé par Damoizeau, en 1891, avec son cyclographe (Figure 1.5(a)). Cependant, les essais de projection de la scène, lors de l'Exposition de Chicago, ne donnaient que des résultats imparfaits. Le 29 décembre 1900, Louis Lumière dépose le brevet du Photorama (Figure 1.5(b)). Ce procédé de photographie panoramique permet de reproduire une vue complète 360° de l'horizon sur un seul cliché et de la projeter sur un écran cylindrique de plus de 6 mètres de hauteur. La projection uniforme de l'ensemble de la photographie sur un cylindre est obtenue par la rotation continue (3 tours par seconde) d'un plateau supportant 12 objectifs. La persistance rétinienne confond l'ensemble des portions d'images projetées sur l'écran circulaire en une seule image continue, éliminant ainsi les aberrations dues à l'emploi d'objectifs grossissants.

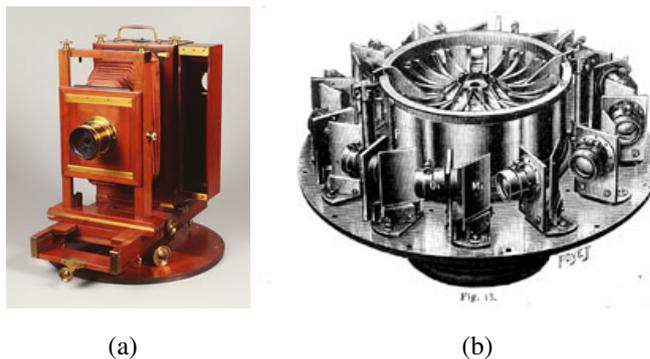


FIG. 1.5 – (a) Cyclographe de Jules Damoizeau. (b) Photorama de Louis Lumière.

Dans [55], Peleg et Ban-Ezra utilisent une caméra perspective rotative (Figure 1.6(a)) pour créer des images panoramiques. Une technique de projection circulaire des images est développée à partir de mosaïques d'images perspectives (Figure 1.6(b)). L'utilisation d'une caméra perspective permet de spécifier la résolution horizontale des images ainsi que la vitesse de rotation, il est donc possible d'acquérir des images de très haute résolution. Cependant, une grande précision au niveau du dispositif d'animation de la caméra est nécessaire : pour obtenir un centre de projection unique, le centre optique de la caméra doit coïncider avec l'axe de rotation du système. De plus, le temps d'acquisition d'une image panoramique est assez long, la scène doit donc rester fixe durant toute la phase d'acquisition. Ces capteurs sont, par conséquent, plutôt indiqués pour l'acquisition de scènes statiques, même si dans la littérature, Barth et Barrows [35] ont développé une méthode d'acquisition rapide d'images panoramiques pour des robots en mouvement en utilisant une caméra linéaire. Dans [40],

Gastaud, Gardan, Romain et Ea proposent d'utiliser deux caméras perspectives couplées à un projecteur de lumière structurée infrarouge. Ils réalisent, avec ce capteur, l'acquisition et la restitution 3D d'un couple d'images stéréoscopiques panoramiques avec des couleurs représentatives de la scène observée. Dans [61], un capteur stéréoscopique original dédié à l'acquisition d'images sphériques couleur (Figure 1.6 c)) est conçu à l'aide de deux caméras perspectives mises en rotation. Leur centre optique est placé sur un axe de rotation horizontal. Un axe de rotation vertical passant par le milieu des deux centres optiques, permet l'acquisition de l'ensemble de la scène sur $360^\circ \times 360^\circ$ en 3D pour des applications multi-médias et cinématographiques.

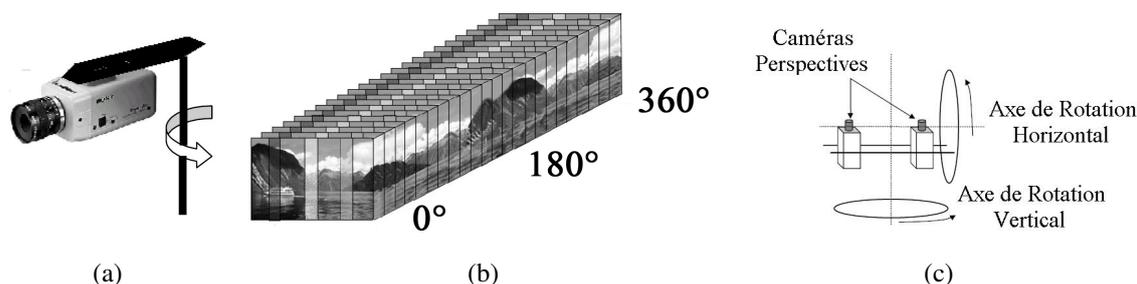


FIG. 1.6 – (a) Caméra en rotation autour d'un axe vertical fixe, (b) Reconstitution d'une image panoramique à partir d'une mosaïque d'images perspective, (c) Capteur stéréoscopique dédié à l'acquisition d'images sphériques couleur.

1.1.3 Systèmes de caméras multiples

Une solution alternative aux systèmes de caméras rotatives consiste à fusionner une multitude d'images acquises avec plusieurs caméras montées de manière à avoir une zone de recouvrement entre les différents champs visuels. L'acquisition des images est synchronisée pour toutes les caméras, ce qui permet de réduire le temps d'acquisition. L'inconvénient de ces techniques réside dans le fait que certaines parties de l'environnement, les angles morts, ne sont pas acquises par le système, alors que d'autres sont dupliquées à cause de l'intersection du champ de vue entre les caméras adjacentes. Les images doivent être traitées après enregistrement pour obtenir une image panoramique ; de plus il est difficile de s'assurer que toutes les caméras ont le même centre optique.

Nanda et Cutler présentent, dans [22], un système de caméras multiples de haute résolution, simple à réaliser et de coût très raisonnable permettant de créer des images panoramiques en temps réel (Figure 1.7(a)). L'utilisation d'une unité centrale de traitement d'images (pour modifier le contraste, la luminosité, la balance des couleurs, etc.) permet d'améliorer la qualité de l'image.

Dans [46], Swaminathan et Nayar utilisent une poly-camera à quatre objectifs grand-angle (Figure 1.7(b)) permettant d'accroître la résolution de l'image et proposent une méthode de correction des paramètres de déformation sans utiliser d'objets étalons.

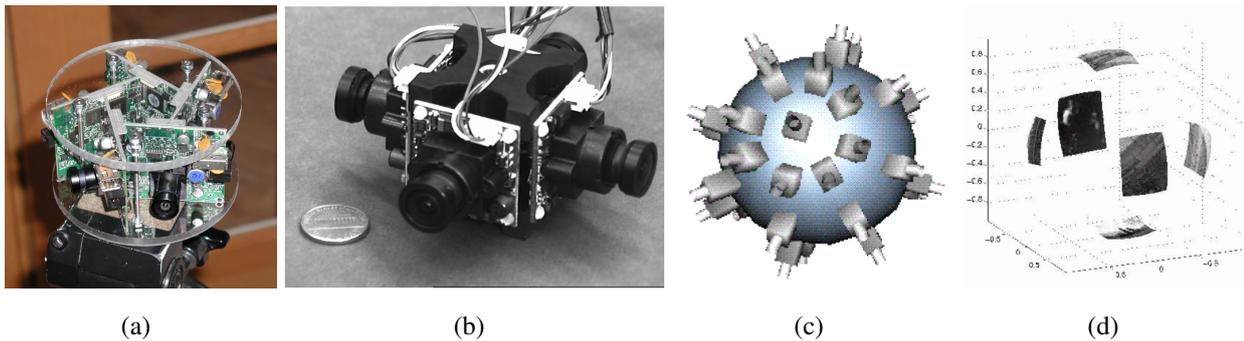


FIG. 1.7 – (a) Le capteur omnidirectionnel RingCam, constitué de caméras perspectives. (b) Polycamera à quatre objectifs grand-angle, (c) Système multi-caméras, (d) Images prélevées sur différentes parties de la sphère visuelle.

Dans [43], c'est un système de six caméras haute résolution qui prélèvent différentes parties de la sphère visuelle comme modélisé Figures 1.7(c) et (d). Baker, Fermüller, Aloimonos et Pless ont développé des algorithmes pour l'étalonnage du système basés sur la relation entre les projections sur l'image de lignes parallèles dans l'espace et la rotation entre les caméras. Cela leur permet d'estimer le mouvement des objets de l'image. Avec ce système, l'ambiguïté liée à la détection des translations et des rotations disparaît (cf. Introduction Générale).

La plupart des capteurs de ce type permettent d'acquérir les images tout autour de la caméra. Vedula, Saito, Kanade et Rander, ont réalisé dans [62] un capteur qui regarde vers l'intérieur (Figure 1.8). De nombreuses images stéréo sont acquises afin de numériser des exécutions de mouvement de danse par exemple.



(a)

FIG. 1.8 – Dôme permettant de visualiser une exécution artistique ou sportive à l'université de Carnegie Mellon.

1.1.4 Systèmes catadioptriques

Bien avant l'avènement de la science, certains surent que le reflet dans une eau calme était le leur. Dans la mythologie grecque, Narcisse, s'abreuvant à la source, voit son reflet dans l'eau et en tombe amoureux. En Chine, durant trois millénaires, ce phénomène fut mis à profit par des chinoises de toutes classes sociales qui remplissaient d'eau des cuvettes en terre cuite puis en bronze pour se parer (Figure 1.9(a)). Dans l'Antiquité (aux environs de 4000 ou 3000 ans avant Jésus-Christ), des

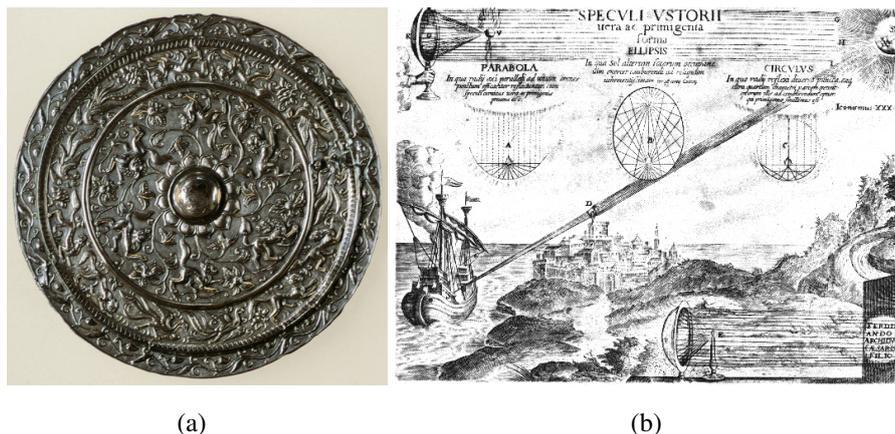


FIG. 1.9 – (a) Cuvette en bronze que les chinoises remplissaient d'eau pour se parer. (b) Gravure représentant les vaisseaux de l'armée romaine enflammés par les miroirs ardents d'Archimède.

artisans du Moyen-Orient polirent des surfaces métalliques en or, en bronze (airain) ou en argent, au point de les rendre réfléchissantes. Ces premiers *miroirs à main* renvoient une image légèrement déformée et nécessitent un entretien contre l'oxydation. La perception de soi choquant autant les humains que les animaux, ces miroirs furent utilisés pour chasser les animaux féroces. La bête, à la vue de son image, arrêtait sa course et s'en approchait pour se regarder avec étonnement. Le chasseur à l'affût en profitait pour la tuer. Les militaires s'en servaient également pour transmettre rapidement des messages à grandes distances, en réfléchissant les rayons du soleil. Des miroirs sphériques en or étaient utilisés par les Athéniens, afin de concentrer les rayons du soleil et rallumer le feu sacré de Hestia, déesse du Foyer. Selon la légende, entre 215 et 212 avant Jésus-Christ, Archimède réussit à enflammer à distance les vaisseaux de l'armée romaine venus envahir la Sicile grâce aux célèbres miroirs ardents (Figure 1.9(b)). Ces miroirs simulent un miroir concave à longue focale, assimilable à un paraboloïde de révolution, par juxtaposition de miroirs plans hexagonaux. Bien que découvert 4000 ans avant Jésus-Christ, ce n'est qu'à la Renaissance que le verre fut, pour la première fois, coulé à plat puis étamé par les artisans de Venise. Avec l'expérience, les défauts et déformations s'atténuèrent et la taille des miroirs augmenta. Les miroirs convexes furent appelés *miroirs de sorcières* car on leur attribuait des pouvoirs magiques. Placés en face d'une fenêtre, ils permettaient de diffuser la lumière dans les maisons européennes nordiques au XVI^e siècle. Un autre nom usuel

est *miroirs des banquiers* car ils permettaient aux banquiers, orfèvres et usuriers de surveiller l'ensemble d'une pièce (Figure 1.10). Les miroirs plans réfléchissent les rayons de manière symétrique

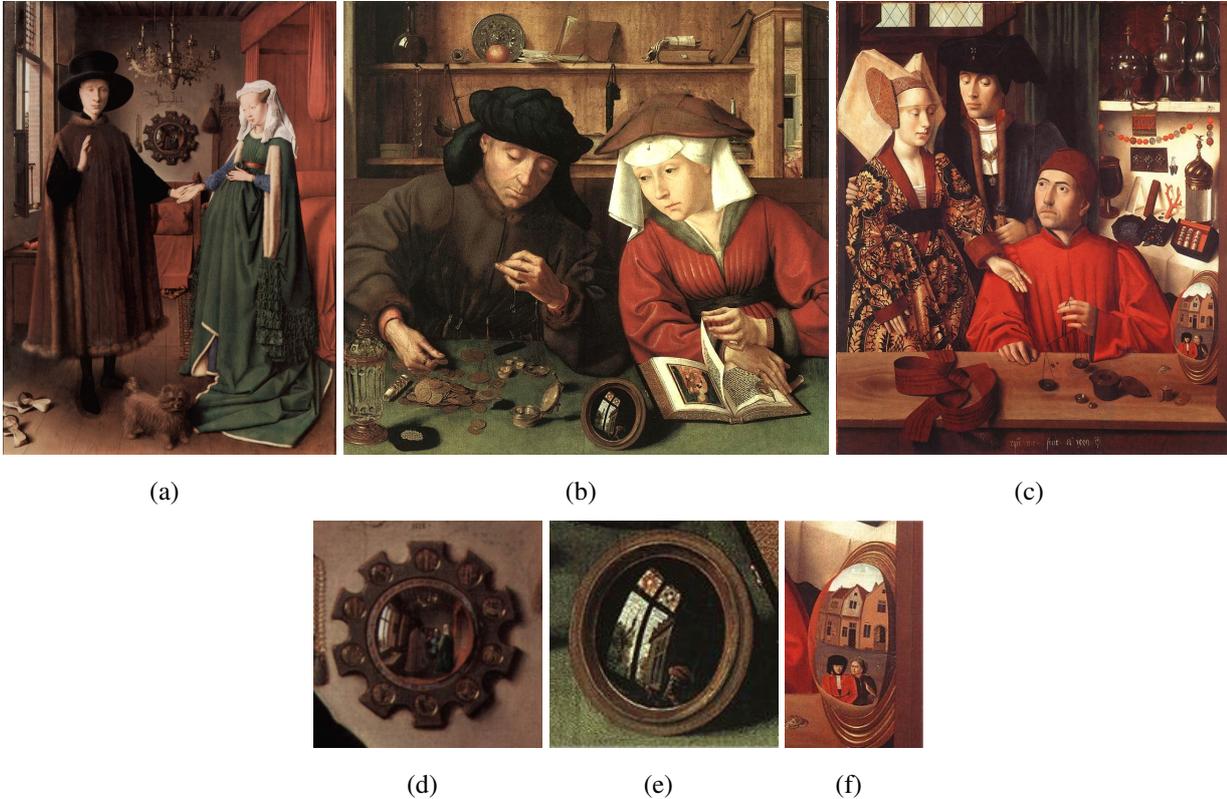


FIG. 1.10 – (a) "Le mariage de Giovanni Arnolfini" de Jan van Eyck en 1434 (National Gallery de Londres), (d) Zoom sur le miroir accroché au mur, reflétant la pièce, les époux de dos et le peintre. (b) "Le prêteur et sa femme" de Quentin Metsys en 1514 (musée du Louvre à Paris), (e) Zoom sur le miroir posé sur la table, représentant une personne qui attends devant la fenêtre. (c) "Saint Éloi orfèvre" de Petrus Christus en 1449 (Metropolitan Museum de New-York.) (f) Zoom sur le miroir posé sur la table qui donne vue sur la rue.

et inversent l'image, ces lois sont décrites par Descartes au XVIIe siècle. Ce principe est utilisé pour déterminer la hauteur (angle par rapport à l'horizon) et la distance (distance Terre-Lune) d'un astre. Les premiers miroirs, fabriqués en bronze blanc doivent être inlassablement polis, car leur surface ternit. Vers 1850, un Français, Foucault, réussit à déposer sur du verre une couche métallique uniforme. Pour que le verre soit homogène, chaque pièce est refroidie très lentement. En 1935, Escher représente *Hand with Reflecting Sphere* (Figure 1.11), un globe sphérique reflétant l'environnement de l'artiste. En astronomie, le miroir concave est utilisé dans la fabrication du télescope vers 1950. Les systèmes omnidirectionnels *catadioptriques* résultent de la combinaison des lentilles et des miroirs. "dioptrique" signifie *science des éléments réfractants* (lentilles) et "catoptrique", *science des surfaces réfléchissantes* (miroirs). En alignant l'axe optique de la caméra et l'axe de révolution d'un



FIG. 1.11 – "Hand with Reflecting Sphere" de Escher

miroir convexe, il est alors possible d'élargir le champ de vision à 360° et de percevoir ainsi l'environnement dans sa totalité.

Baker et Nayar [51] ont montré que les systèmes catadioptriques pouvaient être scindés en deux catégories : les systèmes catadioptriques centraux et les autres. Les systèmes catadioptriques cen-

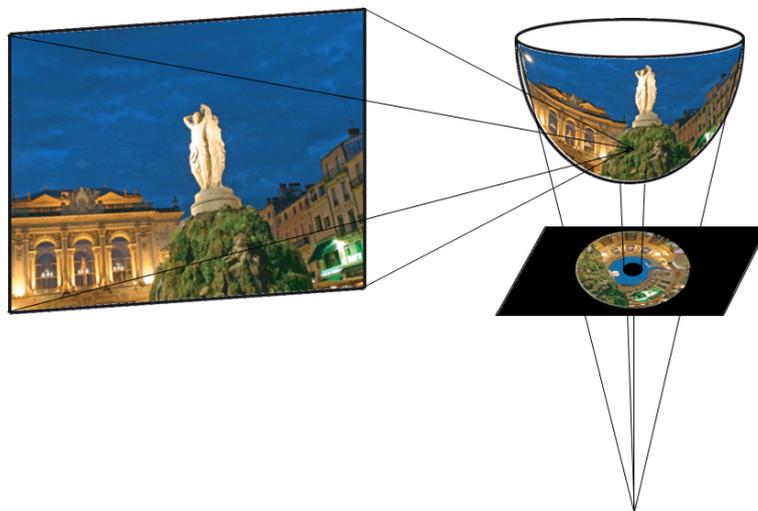


FIG. 1.12 – Reconstruction d'une image perspective à partir d'une image omnidirectionnelle acquise avec un capteur catadioptrique possédant un miroir hyperbolique.

traux vérifient la contrainte du point de vue unique "Single View Point" (SVP) : à une et une seule direction dans l'espace, on associe un point unique dans l'image. Avec un système vérifiant cette contrainte, il est donc possible de reconstruire n'importe quelle image perspective à partir de l'image

omnidirectionnelle (sous réserve qu'elle soit dans le champ de vue de la caméra omnidirectionnelle). Cela permet de visualiser des images géométriquement correctes (Figure 1.12). Chaque pixel de l'image donne alors une estimation de l'illumination de la scène 3D dans une direction spécifique. Si la géométrie du système catadioptrique est connue, cette valeur peut être calculée pour chaque pixel. Il est donc possible de projeter les valeurs de niveaux de gris de chaque pixel sur une image perspective. Cette image ne présente pas de déformation, ce qui la rend plus compréhensible pour l'être humain.

Les dispositifs catadioptriques centraux sont ceux dont les miroirs ont les sections coniques comme sections transversales. Ils regroupent les miroirs paraboliques, hyperboliques, elliptiques et plans. Certains capteurs utilisant des miroirs coniques ou sphériques peuvent également, par approximation, être considérés comme ayant un point de vue unique. Dans leur article [52], Nayar et Baker ont étudié trois critères essentiels pour les capteurs catadioptriques (Figure 1.13) vérifiant la contrainte du point de vue unique : la forme des miroirs, la résolution des caméras et la position du point de vue. Ils ont explicité la classe complète des miroirs à point de vue unique, trouvé une expression pour la résolution des capteurs catadioptriques et présenté une analyse détaillée du "defocus blur" provoqué par l'utilisation d'un miroir incurvé. Un grand nombre de miroir a été décrit parmi lesquels les miroirs coniques, sphériques, plans, hyperboloïdes, ellipsoïdes et paraboloides. Une question se pose : quel miroir utiliser ?

La réponse est complexe.



FIG. 1.13 – Famille de capteurs catadioptriques.

1.1.4.1 Capteurs catadioptriques centraux

1.1.4.1.1 Miroir plan D'après les lois de Descartes, si un rayon lumineux rencontre un miroir plan, il est réfléchi et reste dans le plan d'incidence, de plus, l'angle d'incidence est égal à l'angle de réflexion (Figure 1.14). Le point de vue unique existe. Il est sur la bissectrice perpendiculaire à la droite reliant le point focal et son image virtuelle.

Dans [38], Nalwa a réalisé un capteur omnidirectionnel à l'aide de quatre miroirs et de quatre caméras, disposés en pyramide (Figure 2.8(a)). Le capteur acquiert quatre images qui sont juxtapo-

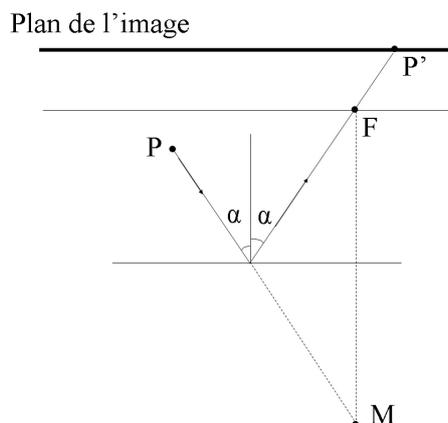


FIG. 1.14 – Formation de l'image P' d'un point P par réflexion sur un miroir plan.

sées pour calculer une image panoramique de haute résolution de $360^\circ \times 50^\circ$. Ce capteur vérifie la contrainte du point de vue unique, il est donc possible de reconstruire une image géométriquement correcte sur une surface projective donnée. Un autre exemple de système est présenté dans [32] : le capteur génère une image omnidirectionnelle grâce à un système composé de plusieurs caméras positionnées de manière stratégique et d'un miroir hexagonal (Figure 2.8(b)). On retrouve les problèmes évoqués Section 1.1.3 concernant les angles morts et les zones redondantes de l'environnement, ainsi que le traitement des intersections qui augmente le temps de calcul.

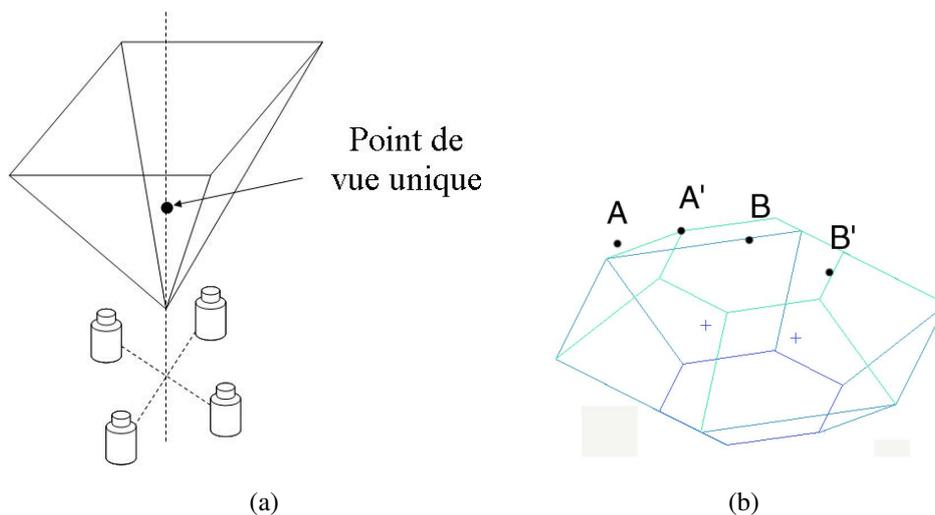


FIG. 1.15 – (a) Miroir pyramidal carré, (b) Miroir pyramidal hexagonal avec deux paires de caméras (AA' , BB'), chacune associé à une face, dont la focale est en (+).

1.1.4.1.2 Miroir elliptique Le point de vue unique est à l'intérieur du miroir (Figure 1.16). La limitation principale de l'ellipsoïde est qu'elle ne reflète qu'un seul hémisphère. De plus la forme concave de ce miroir le rend relativement peu pratique à utiliser. Nous n'avons pas trouvé d'application ni de travaux se servant d'un miroir elliptique.

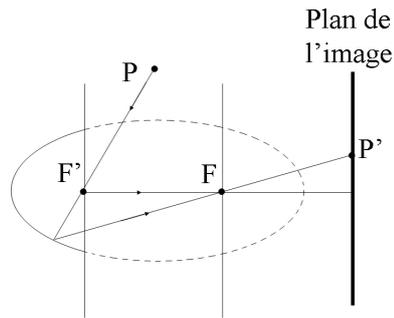


FIG. 1.16 – Réflexion d'un point P dans un miroir elliptique.

1.1.4.1.3 Miroir hyperbolique Le miroir hyperbolique a deux points focaux (Figure 1.17). Tout rayon incident dont le prolongement passe par le foyer F' d'une branche de l'hyperbole sera réfléchi en direction de l'autre foyer F . Le point de vue unique se trouve en F' . La réalisation des capteurs hyperboliques nécessite une coïncidence entre le centre optique F de la caméra et le deuxième foyer de l'hyperbole. Si la courbure du miroir est faible, alors l'astigmatisme¹ est faible. Ce miroir a été réalisé pour la première fois par Rees [50]. Les premières utilisations d'un capteur de ce type ont été effectuées par Yamazawa et Yagi [31] avec HyperOmni Vision. Dans [26] Ishiguro discute des dispositifs omnidirectionnels et de leurs problèmes de conception. Il conçoit une famille de capteurs omnidirectionnels, à faible coût et relativement compacts. Les réflexions internes causées par le support en verre cylindrique (Figure 1.13) sont supprimées par l'adjonction d'un "dard" central noir mat.

1.1.4.1.4 Miroir parabolique Le miroir parabolique (Figure 1.18) permet de contourner la contrainte de l'alignement des foyers lors de la fabrication. Une projection orthographique est réalisée : les points sont projetés suivant une direction déterminée \vec{k} sur un plan situé à une distance f de l'origine (la projection orthographique est parfois interprétée comme étant une projection perspective dans laquelle le point de perspective est rejeté à l'infini). Le miroir parabolique est l'un des derniers

¹L'astigmatisme apparaît lorsque les lignes verticales et horizontales d'un objet sont focalisées de manière différente. L'image d'un objet apparaît courbée : si le centre de l'image est net, les bords seront flous et inversement.

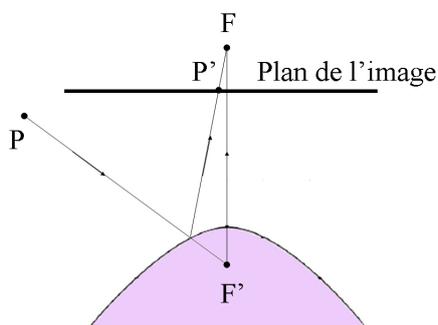


FIG. 1.17 – Réflexion d'un point P dans un miroir hyperbolique.

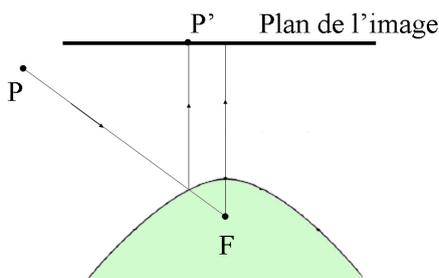


FIG. 1.18 – Reflection d'un point P dans un miroir parabolique.

à être utilisé. Il est relativement simple à réaliser et s'il présente une faible courbure, l'astigmatisme est faible. Nayar [39] a réalisé une étude sur un capteur omnidirectionnel composé d'un miroir parabolique et d'une caméra perspective munie d'une lentille télécentrique. Les lentilles télécentriques sont des lentilles spéciales permettant de sélectionner le faisceau lumineux des rayons parallèles à l'axe optique de la lentille. Si cette lentille est alignée avec l'axe du miroir, seuls les rayons incidents qui passent par le foyer du miroir sont réfléchis parallèlement à l'axe optique. Ce capteur, nommé OmniVideo, a été proposé pour des applications de vidéo-conférence. Ses principaux inconvénients sont la taille de la lentille télécentrique et son prix.

1.1.4.1.5 Modèle de projection unifié Tous miroirs catadioptriques centraux satisfont la contrainte du point de vue unique. Geyer et Daniilidis ont introduit, dans [5], un modèle de projection unifié

pour l'ensemble des capteurs catadioptriques centraux. Ils ont démontré que l'anamorphose obtenue avec un capteur catadioptrique central est isomorphe à une double projection via la sphère comme illustré sur la Figure 1.19. L'image S d'un point $M(X, Y, Z)$ de la scène 3D, sur la sphère, est obtenue par une projection centrale issue du centre de la sphère O . La projection centrale de S , issue d'un point B appartenant à l'axe de la sphère, donne l'image $m(x, y)$ de $M(X, Y, Z)$ sur l'image omnidirectionnelle.

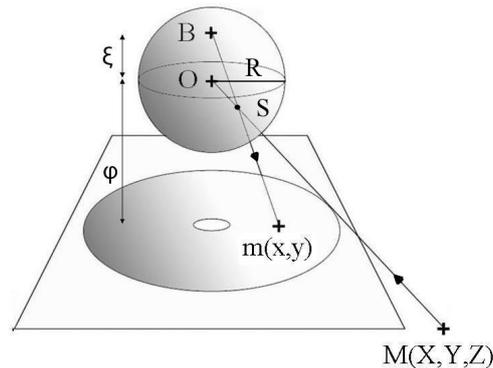


FIG. 1.19 – Modèle de projection unifié pour les capteurs catadioptriques centraux.

Pour un miroir parabolique, le point B se trouvera sur le pôle nord de la sphère. Pour un miroir elliptique ou hyperbolique, B se situera entre le pôle du nord et le centre de la sphère. Seulement deux paramètres (φ et ξ) sont nécessaires pour associer un point 3D $M(X, Y, Z)$ avec sa projection $m(x, y)$ sur l'image panoramique. φ représente la distance entre l'image omnidirectionnelle et le centre O de sphère, ξ la distance OB et $R^2 = X^2 + Y^2 + Z^2$. L'utilisation de ce modèle permet une transposition facile des algorithmes que nous avons développé à l'ensemble des capteurs catadioptriques centraux.

1.1.4.2 Capteurs catadioptriques centraux par approximation de la contrainte du point de vue unique

1.1.4.2.1 Miroir conique Un miroir conique associé à une caméra perspective satisfait la contrainte du point de vue unique si la focale de la caméra coïncide avec le sommet du cône. Cette solution est dégénérée (Figure 1.20 (a)) et ne peut être employée pour élargir le champ visuel. La solution la plus fréquemment utilisée consiste à éloigner la caméra du cône. Le point de vue n'est plus unique : il s'agit d'un cercle dont le rayon dépend de la distance entre le point de vue et le sommet du cône, ainsi que de l'angle au sommet du cône. Une étude menée par Lin et Bajcsy [60] a montré qu'il est néanmoins possible sous certaines conditions de considérer ce capteur comme ayant un point de vue unique (une distance focale courte peut être approchée avec un point de vue unique). Spa-

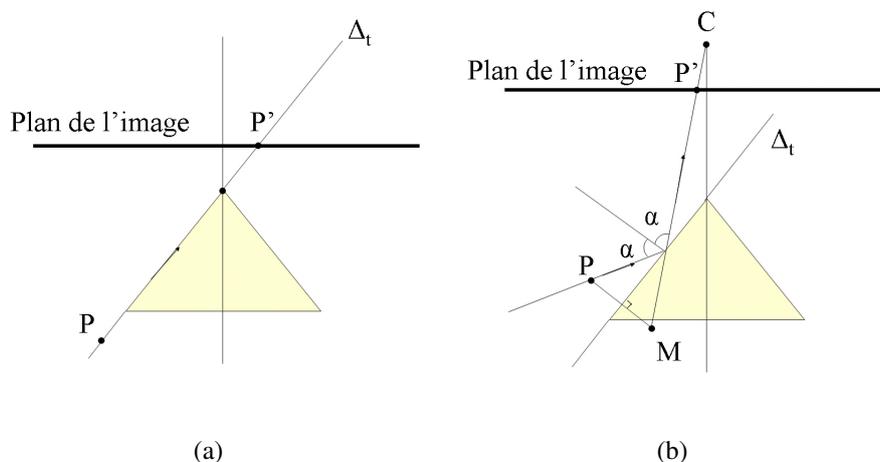


FIG. 1.20 – Reflection d'un point P dans un miroir conique (a) lorsque la focale coïncide avec le sommet du cône, (b) lorsque le point de vue n'est pas unique.

cek introduit dans [59] le concept de points de vue dimensionnellement séparables, grâce auquel une pseudo image perspective peut être reconstruite à partir d'une image conique. En effet, malgré son astigmatisme très important, le miroir conique est très facile à construire et présente des propriétés intéressantes. Les rayons lumineux acquis par la caméra couvrent la totalité de la scène observée (Figure 1.20 (b)). Le miroir conique n'introduit pas de distorsion longitudinale et se comporte, en coupe verticale, comme un miroir plan : la normale de réflexion suit la direction verticale. Cela permet d'obtenir des images de très bonne résolution à la périphérie de l'image. Yagi, dans [67], développe un système de perception à l'aide d'un miroir conique pour la navigation d'un robot mobile, appelé COPIS (COnic Projection Image Sensor). Dans [2], Cauchois, Brassart et Drocourt s'intéressent à l'étalonnage d'un capteur omnidirectionnel conique SYCLOP (Conical SYstem for Localization and Perception) utilisé en robotique mobile. Avec le miroir conique, si l'axe optique correspond avec l'axe de révolution du cône, les droites parallèles à cet axe se projettent de façon radiale dans l'image. Dans notre environnement, si la base du cône est parallèle au sol, les droites verticales se projettent en droites radiales dans l'image. Le miroir conique, par la simplicité de sa géométrie peut également être couplé facilement avec d'autres miroirs, comme le miroir plan utilisé par Zhu, Yang, Xu, Lin et Shi, dans [69].

1.1.4.2.2 Miroir sphérique Le miroir sphérique, comme le miroir conique, ne peut être employé pour étendre le champ de vision (Figure 1.21) : si le point de vue unique et le point focal de la caméra sont confondus, sur l'image acquise, seule la caméra serait visible (l'observateur ne peut voir que lui-même). Cependant, ces miroirs sont utilisées dans plusieurs applications de robotique et de vidéo-surveillance. Un miroir sphérique est utilisé par Winters et Santos-Victor dans [66] pour

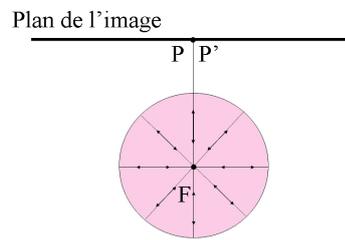


FIG. 1.21 – Réflexion d'un point P dans un miroir sphérique.

déterminer la position du robot. Matsui, Asoh et Thompson [63] ont monté ce type de capteur à l'avant d'un robot. L'image obtenue est dépliée et une mesure de corrélation est effectuée sur une base d'images de référence pour en déduire la localisation du robot. Une image acquise à partir d'un miroir sphérique présente une bonne résolution dans la région centrale et ce miroir est très facile à fabriquer : il suffit de déposer de l'aluminium sur une surface de verre. Cependant, le miroir sphérique est astigmatique. L'astigmatisme, faible près de l'axe optique de la caméra considéré comme une surface plane, est considérable sur la périphérie.

1.1.4.3 Système catadioptrique axial

Les caméras axiales sont celles dont les lignes de vue n'ont pas un foyer unique, mais où il existe une ligne qui les coupe toutes. Elles forment une classe intermédiaire entre les caméras centrales et les caméras non-centrales. Au niveau des systèmes catadioptriques, ces caméras sont moins restrictives que les modèles catadioptriques centraux. Ces caméras sont décrites par Sturm dans [49].

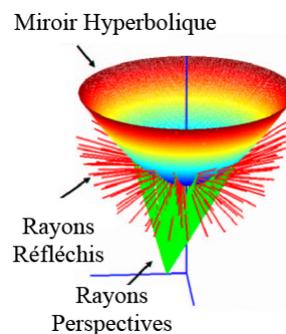


FIG. 1.22 – Caméra axiale munie d'un miroir hyperbolique.

1.1.4.4 Système catadioptrique caustiques

Comme nous l'avons constaté, la contrainte du point de vue unique est difficile à satisfaire et peut nécessiter des assemblages complexes. Certains auteurs se sont donc tournés vers d'autres types de capteurs plus simples ou permettant de répondre à certaines exigences. Le point de vue de ces capteurs n'est plus unique. Le centre peut prendre la forme complexe de plusieurs points de vue ou d'une surface appelée la surface caustique.

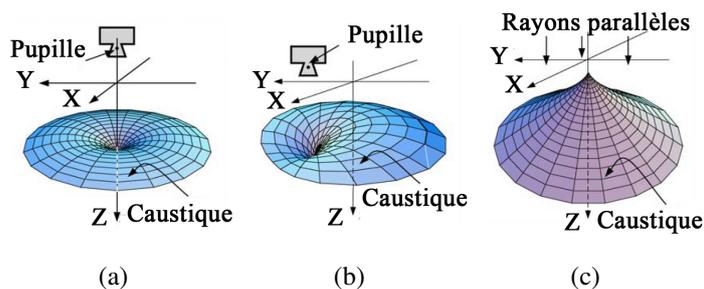


FIG. 1.23 – Surfaces caustiques 3D pour un système catadioptrique (a) symétrique constitué d'un miroir parabolique, (b) asymétrique constitué d'un miroir parabolique, (c) symétrique constitué d'un miroir hyperbolique.

La caustique est la courbe qui correspond aux lieux des foyers du miroir. C'est en théorie ce qu'il faut considérer lorsqu'il y a des défauts de conception du capteur catadioptrique. Des systèmes caustiques peuvent être conçus pour la formation d'images spécifiques ou pour les caractéristiques de ces images en termes de champ de vision et de résolution. Dans [47], Swaminathan, Grossberg et Nayar présentent une analyse détaillée des systèmes caustiques (Figure 1.23). Un modèle paramétrique simple décrit le système de formation d'images et fournit une carte des pixels de l'image associés à leurs points de vue et leurs directions dans l'espace 3D. Ce modèle est adaptable à l'ensemble des systèmes catadioptriques dont le réflecteur a une forme paramétrique.

1.1.5 Combinaison des différentes approches

Il est possible de concevoir un montage qui présente les mêmes caractéristiques que le miroir parabolique et la lentille télécentrique mais de coût et taille moindres, en combinant un miroir parabolique avec un miroir sphérique. Un tel système est introduit dans [45] par Orghidan, Mouaddib et Salvi. Les auteurs présentent les caractéristiques et la modélisation géométrique de ce capteur visuel panoramique capable d'estimer la profondeur de l'environnement et sa modélisation. Les deux composants omnidirectionnels qui forment le dispositif ont été calibrés et le modèle résultant est utilisé pour estimer la profondeur de la scène. Ce dispositif peut être utilisé dans la navigation des robots mobiles comme le propose Gonzalez-Barbosa dans [19]. Il s'intéresse au problème de

la localisation des robots mobiles par indexation des images dans des environnements naturels à l'aide d'une caméra panoramique. L'objectif est de donner au robot la capacité de modéliser et de reconnaître des éléments de son environnement afin de se localiser qualitativement.



FIG. 1.24 – Combinaison d'un miroir parabolique avec un miroir sphérique.

Dans ce capteur, le centre optique de la caméra doit être placé sur le foyer de la sphère (Figure 1.24). Seuls les rayons parallèles à l'axe optique du miroir sphérique traversent le foyer. Le miroir sphérique joue donc le rôle de la lentille télécentrique. Cependant, ce système est difficile à assembler : la focale de la caméra doit coïncider avec le foyer du miroir sphérique et l'alignement des miroirs parabolique et sphérique doit être parfait pour éviter de fortes aberrations de coma² et d'astigmatisme.

Nayar et Karmarkar [57] effectuent une combinaison des différentes méthodes afin de calculer une mosaïque sphérique complète avec des caméras perspectives. Cette opération pose de nombreux problèmes : le grand nombre d'images nécessaires pour capturer l'ensemble des vues de l'espace, les erreurs de recalage des images qui introduisent des artefacts au niveau des jonctions entre deux images, ou simplement l'assurance de couvrir la totalité de l'espace (angle solide³ de 4π). Deux approches sont présentées pour l'acquisition de mosaïques sphériques de haute résolution. La première consiste à effectuer une rotation de 360° autour d'un axe, d'une caméra grand-angle pour acquérir une séquence de bandes de 360° sur la sphère (Figure 1.25). Les rotations inconnues entre les bandes sont estimés pour construire une mosaïque sphérique. Dans la seconde, les auteurs cherchent à augmenter de manière significative la résolution de la mosaïque. La caméra de faible résolution proposée permet d'obtenir des bandes de 360° de haute résolution. Bien qu'un grand nombre d'images (bandes) soient nécessaire pour obtenir une mosaïque de haute résolution, le traitement de chaque image est minimal et est facilement réalisable en temps réel.

²**La coma** produit une sorte de traînée conique irisée dirigée sur les images : les rayons qui passent par la périphérie de la lentille ne sont pas focalisés au même endroit que ceux qui passent près du centre. La coma est un défaut récent, les films argentiques ne subissaient pas ce genre d'aberration.

³**Angle solide** : analogue tridimensionnel de l'angle plan ou bidimensionnel. Il mesure la surface sur laquelle un objet se projette radialement sur une sphère de rayon unité. Son unité est le stéradian.

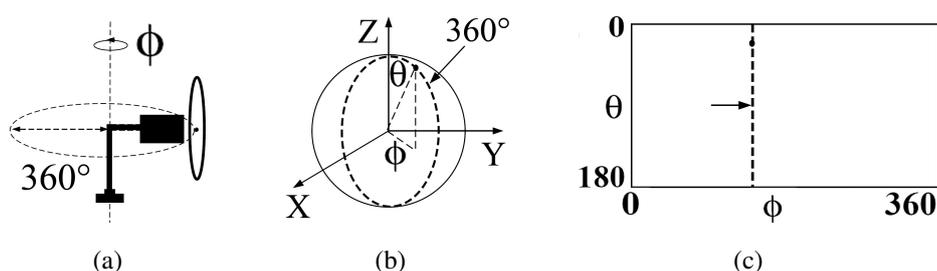


FIG. 1.25 – (a) Une caméra est mise en rotation pour acquérir des bandes d’images de 360° , (b) chaque bande correspond à un cercle sur la sphère, (c) la moitié d’une bande représente une ligne dans l’image panoramique.

1.2 Traitement des images omnidirectionnelles

1.2.1 Utilisation des traitements classiques

Si la caméra catadioptrique accroît le champ de vision, en revanche la géométrie des images formées est relativement complexe. Comme nous l’avons vu dans l’introduction générale, la résolution des images panoramiques varie en fonction de la position dans l’image. Certains auteurs pensent que les outils classiques donnent de bons résultats et qu’il est inutile d’augmenter le temps de calcul pour gagner en précision. Les images omnidirectionnelles sont alors traitées de la même manière que les images perspectives.

Dans l’article [37] par exemple, Vlasilis, Motomura, Hara et Asoh travaillent sur la détection de contours. Leur objectif est d’extraire l’information de bas-niveau d’une séquence d’images omnidirectionnelles pour la localisation et la navigation de robot mobile. La méthode proposée suit trois étapes : la détection de contours dans l’image, l’estimation de Parzen de la densité spatiale de ces contours⁴ et l’analyse des composantes principales⁵. Pour la détection de contours, deux filtres de Sobel décalés de 90° sont utilisés. La taille de ces filtres est toujours la même quelle que soit leur position dans l’image. L’image omnidirectionnelle est donc traitée de la même manière qu’une image perspective. Cette méthode fournit de bons résultats ; cependant, elle n’est pas robuste aux phénomènes d’occultation, ce qui engendre de sérieux problèmes en situation réelle.

Gonzalez-Barbosa et Lacroix [27] présentent un moyen de rectifier un banc stéréoscopique panoramique de sorte que les lignes épipolaires soient parallèles. Cette méthode permet d’exécuter un algorithme optimisé de stéréovision par corrélation de pixels. La méthode proposée est indépendante de la configuration du banc stéréo. Les équations qui permettent de faire de la reconstruction 3D en utilisant des caméras panoramiques ont été déterminées. Dans la plupart des modèles proposés dans

⁴Densité spatiale des contours : le nombre de pixels de contour par unité de surface.

⁵Analyse en composantes principales : méthode mathématique d’analyse des données qui consiste à rechercher les directions de l’espace qui représentent le mieux les corrélations entre n variables aléatoires.

la littérature, l'alignement entre la caméra et le miroir est supposé parfait. Dans cet article, une caractérisation complète d'une caméra catadioptrique avec miroir parabolique et lentille télécentrique a été réalisée. L'originalité de ce travail réside dans la prise en compte des paramètres intrinsèques du système (surface du miroir et paramètres intrinsèques de la caméra), des erreurs d'alignement entre le miroir et le repère caméra et de la distorsion optique radiale. Pour l'extraction des indices visuels, un extracteur de croix adapté aux images perspectives a été utilisé. Les auteurs soulignent l'importance d'améliorer la précision de ce détecteur en développant des extracteurs adaptés aux images panoramiques.

Dans [3], Demonceaux et Kachi-Akkouche s'intéressent au problème d'estimation de mouvement dans une séquence d'images acquise avec un capteur catadioptrique. Généralement, le lien entre les gradients d'intensités et les composantes du champ de vitesse est donné par contrainte du flot optique. Cette contrainte repose sur l'hypothèse que le flot est localement constant sur un voisinage. Cependant, les distorsions introduites par le capteur invalident cette hypothèse. Les auteurs ont donc développé une nouvelle méthode de décomposition de l'équation de contrainte du mouvement apparent dans une base d'ondelettes. La base d'ondelettes discrète permet de résoudre le problème d'ouverture généralisée (aperture problem⁶) et de diminuer le temps de calculs. Pour prendre en compte les déformations créées par le capteur, l'hypothèse de constance du flot optique pour un voisinage de points, utilisée pour les images perspectives, est remplacée par une hypothèse d'affinité du mouvement dans ce voisinage. Cette hypothèse permet d'obtenir de meilleurs résultats ; cependant, elle ne caractérise pas correctement les déformations non-linéaires des images omnidirectionnelles, ne prenant pas en compte de la géométrie du miroir.

1.2.2 Adaptation des traitements

Avant l'essor des images catadioptriques, d'autres auteurs se sont intéressés à l'adaptation des traitements à la géométrie d'images ayant subi des déformations. Bien qu'il ne s'agisse pas d'images omnidirectionnelles, les images échographiques présentent des déformations et une résolution non-uniforme. Ces images sont obtenues en coordonnées polaires ce qui crée une anisotropie importante dans la résolution spatiale. Habituellement, une transformation géométrique (ou reconstruction cartésienne) permet de passer des données en géométrie polaire, à une représentation cartésienne, par interpolation bilinéaire ou avec un algorithme de Bresenham. Dans [25], Herlin et Ayache ont montré les limites de cette approche et proposent de travailler dans l'espace de l'image ultra-sonore. Un processus de reconstruction utilisant un filtre passe-bas est appliqué directement sur l'ensemble des données polaires afin de prendre en compte la résolution variable des données. Cette approche est généralisée à l'extraction de contours et utilisée pour suivre des structures physiologiques dans une

⁶**Aperture problem** : dilemme de choisir une taille appropriée de la fenêtre d'analyse. Pour que l'estimation du mouvement présente une relative insensibilité au bruit, l'utilisation d'une grande fenêtre est préférable, cependant, plus la fenêtre est grande, moins sont réalistes les hypothèses de conservation de l'intensité et du mouvement.

séquence d'images.

Conscients des déformations présentes dans les images catadioptriques, de nombreux auteurs ont souligné l'importance d'adapter les traitements à leur géométrie. Plusieurs solutions ont été proposées afin d'obtenir des résultats significatifs. Cependant, les nouveaux algorithmes créés sont souvent partiellement adaptés aux images. On peut distinguer deux principales catégories d'adaptations :

- la projection et le traitement de l'image dans l'espace de la sphère,
- l'utilisation du miroir afin d'obtenir des informations sur les distorsions de la scène observée.

1.2.2.1 Projection de l'image sur la sphère

Dans [30], Daniilidis, Makadia et Bulow travaillent sur l'estimation de mouvement pour des images omnidirectionnelles. Ils proposent de projeter l'image omnidirectionnelle sur une sphère virtuelle (Figure 1.26) pour définir l'opérateur gradient. La fonction Gaussienne est introduite sur la sphère et permet de calculer les dérivées par convolution avec ses dérivées spatio-temporelles. Les dérivées sur la sphère sont utilisées pour estimer le flot optique en conservant l'échantillonnage de l'image omnidirectionnelle. En effet, la projection transforme l'échantillonnage régulier de l'image omnidirectionnelle en un échantillonnage non-régulier sur la sphère. Les convolutions sont effectuées sur la sphère mais leur réalisation préserve les échantillons discrets de l'image catadioptrique originale. Cependant, l'imprécision liée à cette irrégularité n'est pas prise en compte. Les résultats ont montré une amélioration sur le calcul de la direction et la qualité du champ de vecteur pour une même densité.



FIG. 1.26 – Projection de l'image sur l'hémisphère.

Dans le même esprit, Shakernia, Vidal et Sastry [41] généralisent la notion de rétro-projection pour l'estimation du flot optique avec des caméras catadioptriques centrales. La projection sur un capteur catadioptrique central peut être considérée comme une projection sur une rétine incurvée

virtuelle intrinsèque à la géométrie de la caméra. Le flot optique est obtenu en retro-projetant les points de l'image sur une rétine incurvée virtuelle. Cette méthode donne de meilleurs résultats que ceux obtenus avec l'hémisphère présentée dans [30] et les auteurs suggèrent qu'il existe un espace correct qui dépend du mouvement de la caméra. La rétine virtuelle proposée dans cet article est le support le plus approprié lorsque la caméra évolue dans le plan parallèle à la rétine. Pour appuyer cette conclusion, des tests ont été proposés pour différents mouvements de la caméra, différents paramètres de la caméra catadioptrique centrale et différents niveaux de bruit.

Bigot, Kachi, Durand et Mouaddib [53] proposent de créer les opérateurs sur cette sphère virtuelle. L'image omnidirectionnelle est présentée comme une fonction de la sphère. Une nouvelle méthode de lissage est décrite. Il s'agit d'une adaptation du filtre de Wiener⁷ suivi d'une régularisation de Tikhonov⁸. Cette approche est comparée avec les méthodes de convolution des noyaux usuels sur la sphère pour des images de synthèses et des images réelles. Le filtre est invariant par rotation et les résultats sont indépendants de la position sur la sphère. De plus, cette méthode donne un bon rapport signal sur bruit et peut être adaptée en temps réel.

Dans la continuité de leurs travaux [1], Daniilidis et Makadia considèrent que les images catadioptriques centrales peuvent être projetées sur la sphère. Ils proposent une méthode permettant de déterminer les invariances et les paramètres pour des mouvements de rotation 3D par généralisation de la transformée de Fourier (connue pour la sphère sous le nom de transformée en harmonique sphérique). Les angles d'Euler de ces rotations sont calculés par une généralisation du théorème de décalage. Le mouvement apparent est correctement détecté pour les images omnidirectionnelles et cet algorithme est robuste à la détection de faibles mouvements de translation. Cette méthode fonctionne pour tout espace homogène. Cependant, comme nous venons de le dire, l'échantillonnage des images omnidirectionnelles sur la sphère n'est pas régulier. Une interpolation des valeurs des niveaux de gris de ces images est donc nécessaire pour obtenir un échantillonnage uniforme en coordonnées sphériques traditionnelles.

Geyer, Sastry et Bajcsy [6] combinent les notions de géométrie, traitement du signal et analyse harmonique. Dans cet article, les auteurs démontrent que la fonction de vraisemblance peut être obtenue dans l'espace des matrices essentielles par convolution de deux signaux. Le premier signal exprime la géométrie épipolaire de deux points de vue, et le second encode les similitudes des intensités entre deux pixels appartenants à deux images différentes. La méthode proposée est analogue à une transformation de Radon ou une transformation de Hough sur l'espace des matrices essen-

⁷Le **filtrage de Wiener** est un problème d'estimation où on dispose d'une connaissance a priori sur le paramètre à estimer. Cette connaissance se présente sous la forme de données probabilistes. Typiquement on veut estimer un signal noyé dans un bruit et on sait que le signal est a priori centré, blanc, etc. En inférence statistique, lorsque l'on prend en compte une connaissance probabiliste sur le paramètre à estimer on parle d'estimation bayésienne.

⁸Le **principe de la régularisation de Tikhonov** est de choisir le meilleur estimateur possible grâce à un opérateur de régularisation par minimisation d'une fonctionnelle. L'objectif est de réduire de l'erreur empirique et d'augmenter la régularité de la sortie.

tielles. Elle permet de généraliser cette détection de mouvement à l'ensemble des mouvements. La nouveauté de ce travail est que nous pouvons calculer la convolution en utilisant les rotations harmoniques sphériques et les représentations des signaux. Un des principaux avantages de cette approche est qu'elle permet de déterminer les ambiguïtés lors de la détection et peut prendre en compte des bruits d'acquisition non gaussien des capteurs.

Tosic, Bogdanova, Frossard et Vandergheynst [24] ont également utilisé des images sphériques pour réaliser une estimation de mouvement par corrélation d'images successives. Une approche de multi-résolution a été proposée afin d'améliorer l'exactitude du mouvement détecté, et de limiter la complexité informatique. La détection locale du mouvement est assez efficace. L'erreur résiduelle est relativement faible et souvent à proximité des contours ou des composantes hautes fréquences dans l'image prédite.

1.2.2.2 Utilisation de la géométrie du miroir

Ieng, Benosman et Devars ont étudié la mise en correspondance de points dans des images acquises avec deux caméras catadioptriques ayant des points de vue différents. La mise en correspondance d'images consiste dans un premier temps à extraire des primitives (points, lignes, courbes) caractéristiques des images et à faire une association entre les deux images, le plus souvent à l'aide de fonctions de corrélation ou de similarité. La non-uniformité de l'espace des images omnidirectionnelles nécessite une adaptation de ce traitement. Dans leur article [54], l'algorithme proposé s'effectue suivant quatre étapes. Tout d'abord, les points d'intérêt sont extraits avec un détecteur de Harris. Les auteurs précisent que ce détecteur ne tient pas compte de la géométrie des images catadioptriques acquises avec un miroir hyperbolique et qu'une approche mieux adaptée serait souhaitable. Les points détectés sont ensuite projetés sur le miroir hyperbolique. Des quadrilatères de différentes tailles en forme de diamant, centrés autour de chacun des points projetés, sont définis. Chaque diamant représente un angle solide correspondant à une portion de l'espace projeté sur le miroir hyperbolique (Figure 1.27(c)). Les sommets des diamants sont projetés sur l'image omnidirectionnelle (Figure 1.27(a)). La projection des diamants permet d'éliminer les faux appariements entre les points des deux images. Toutes les combinaisons possibles entre les diamants des différentes tailles des deux images sont analysées (Figure 1.27(b)). Le vainqueur est celui auquel on accorde le plus grand coefficient de confiance. Cette méthode permet d'estimer la géométrie épipolaire et de faire de la reconstruction 3D.

Avec un capteur catadioptrique central, la contrainte du point de vue unique est vérifiée. À partir de l'image anamorphosée (Figure 1.28), il est donc possible de reconstituer n'importe quelle image perspective de la scène observée (Figure 1.29). Sur un plan choisi dans le champ de vision du capteur, on définit un maillage (échantillonnage de l'image), puis on calcule la valeur de niveau de gris de chaque point du maillage par interpolation. Pour traiter une image omnidirectionnelle, il est donc

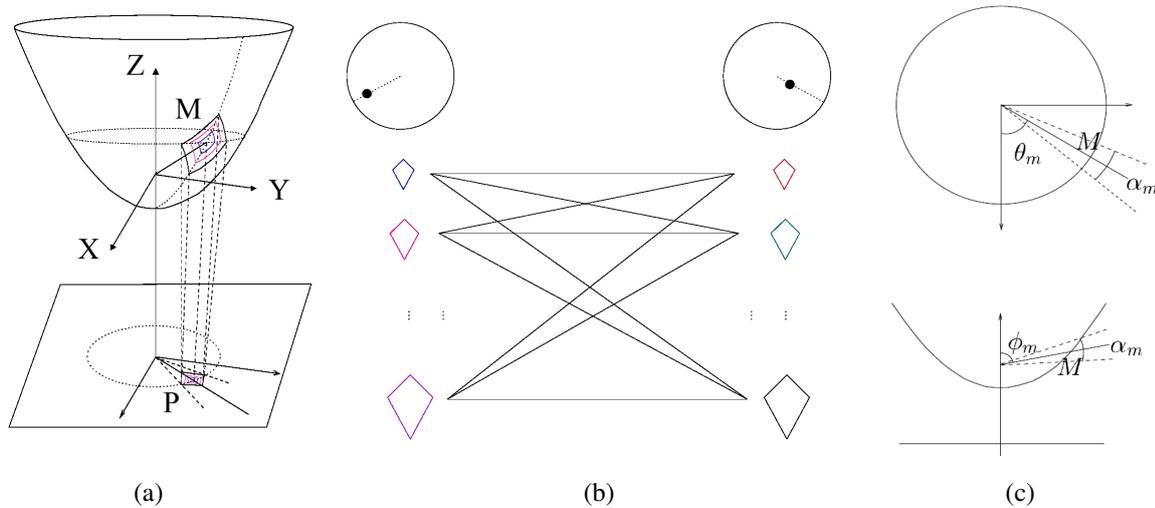


FIG. 1.27 – (a) Définition des patches sur le miroir puis projection sur l'image omnidirectionnelle. (b) Les pixels sont appareillés lorsque leurs patches de différentes tailles ont le plus grand coefficient de confiance. (c) Définition de l'angle solide.



FIG. 1.28 – Image anamorphosée.

possible de recréer toutes les images perspectives issues de cette image et de leur appliquer le traitement désiré. Dans certains domaines d'application où les utilisateurs des images ne sont pas habitués aux déformations des images omnidirectionnelles, par exemple l'inspection, la vidéo-surveillance ou la visio-conférence, la restitution d'images perspectives est souhaitable. Cependant, il existe une infinité d'images perspectives ; autant que de lignes radiales dans l'image omnidirectionnelle. Un très grand nombre d'images perspectives devra être reconstruit, ce qui rend le processus très lourd. De plus, le choix arbitraire du maillage de l'image perspective a évidemment une forte influence sur les valeurs interpolées des niveaux de gris des pixels. La partie extérieure de l'image aura donc une

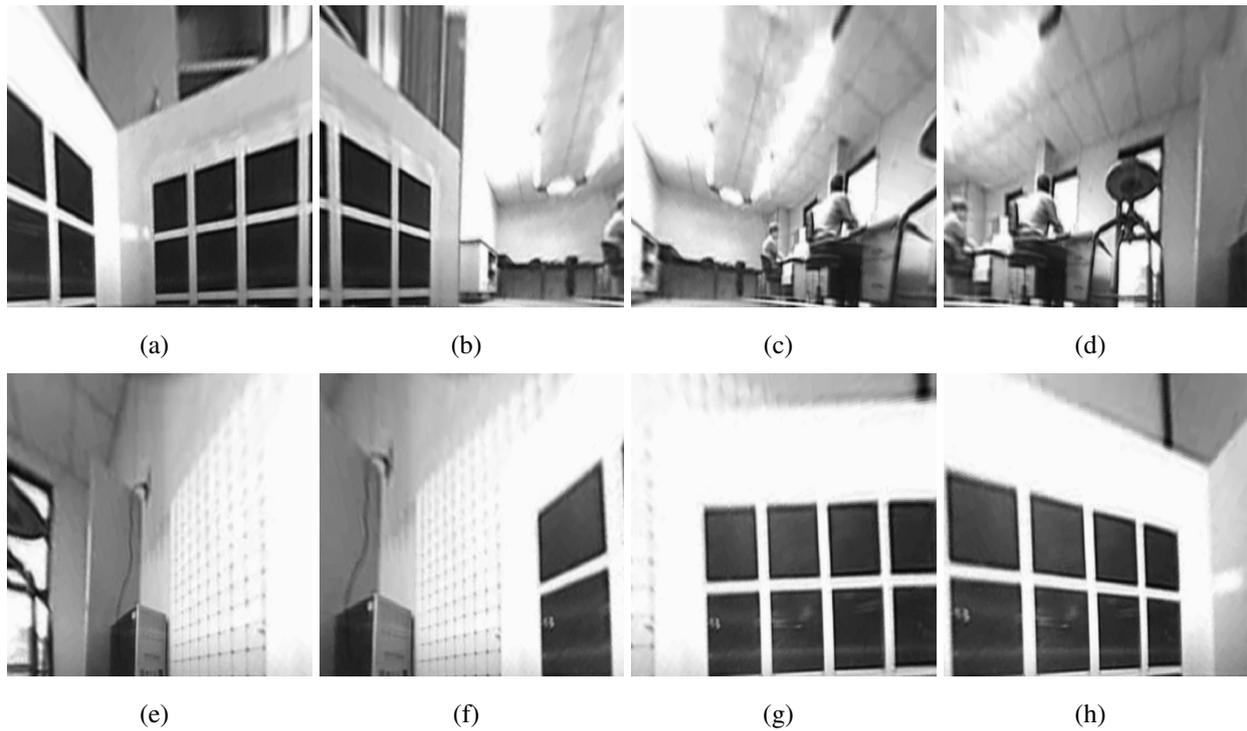


FIG. 1.29 – Reconstruction d’une image perspective.

bonne résolution, mais la perte d’information dans la partie centrale de l’image produira une forte pixelisation dans l’image perspective obtenue.

Cependant, la reconstruction des images perspectives ne permet qu’une vision partielle de la scène dans chaque image. Dans la littérature, la projection sur le cylindre englobant le miroir de révolution est fréquemment utilisée pour visualiser l’ensemble de l’image (Figure 1.30). La Figure

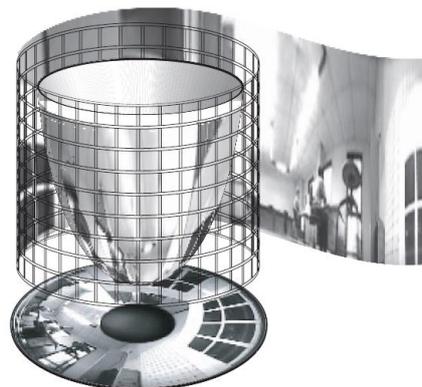


FIG. 1.30 – Projection de l’image sur le cylindre englobant.

1.31 représente l’image obtenue par projection sur le cylindre.



FIG. 1.31 – Projection de l'image sur le cylindre englobant.

Bunschoten et Krose [44] proposent une méthode de reconstruction 3D de l'environnement à partir d'une séquence d'images panoramiques. Cette méthode est basée sur des lignes de niveaux. Les images sont acquises avec un capteur catadioptrique monté sur un robot mobile puis projetée sur un cylindre virtuel. Les lignes de références stéréo utilisent la redondance contenue dans ces images.

Dans leur article [64], Pajdla et Hlavac s'intéressent à la localisation dans une image panoramique. Pour se faire, ils proposent de déplier l'image omnidirectionnelle sur le cylindre. Lorsque la caméra effectue une rotation, la scène observée sur l'image cylindrique est la même, à un décalage près. Lors d'une translation, la caméra change de position. Il est alors impossible de trouver une transformation (rotation et/ou translation) permettant de mettre en correspondance ces deux images. En effet, on constate une variation de la profondeur, ainsi que des phénomènes d'occultation de la scène. Ils proposent d'éviter cette étape de recherche de l'orientation en trouvant une représentation invariante par translation pour chaque classe d'image acquises à la même position.

L'espace du cylindre permet une vision panoramique de la scène, très proche de la vue perspective. C'est un support avantageux car il est possible de l'échantillonner avec une grille régulière dans un repère cylindrique. Cette projection est très lisible par l'oeil humain. Cependant, l'image reste anamorphosée au sens où des lignes droites perpendiculaires à l'axe de rotation du miroir ne se projettent pas comme des droites sur l'image projetée. Les lignes horizontales dans la scène 3D sont donc déformées dans l'image cylindrique. De plus, comme pour l'image perspective, la reconstruction génère une interpolation des valeurs des niveaux de gris des pixels, liée au lissage et à l'échantillonnage. L'information contenue dans l'image est donc modifiée et l'application d'un traitement ajoute des imprécisions supplémentaires.

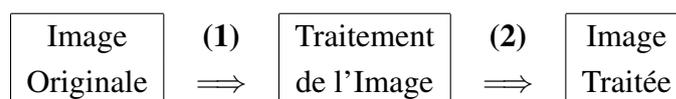


FIG. 1.32 – Schématisation du traitement d'images

Le schéma 1.32 résume très succinctement le processus du traitement des images. La recons-

truction de l'image sur un nouveau support engendre des phénomènes d'interpolation qui modifient l'information présente initialement dans l'image et introduit ainsi des imprécisions sur les valeurs des niveaux de gris de l'image. Cette reconstruction a lieu avant l'étape de traitement de l'image (étape **(1)** sur le schéma). Lors du traitement, ce seront donc les valeurs imprécises induites par l'interpolation qui seront utilisées. Les imprécisions seront donc propagées dans la chaîne de calcul de l'image filtrée.

Pour éviter l'ajout d'informations arbitraires à l'image et minimiser l'impact de la chaîne de calculs, une solution consiste à conserver les données originales de l'image le plus longtemps possible. Ainsi, les imprécisions sont introduites le plus tard possible dans la chaîne de calculs (étape **(2)** sur le schéma).

Plusieurs auteurs se sont intéressés à ce problème et ont proposé plusieurs solutions permettant de rester le plus fidèle possible à la géométrie du miroir et à l'information présente dans l'image.

Le fait que la géométrie du capteur joue un rôle central pour les traitements des images a servi de fil directeur à Bogdanova, Bresson, Thiran et Vanderghyest [23] dans leur méthode de segmentation. Ils restent fidèles à cette contrainte géométrique pour segmenter des images catadioptriques. Une nouvelle fonction d'énergie est introduite de manière à ce que les contours actifs obtenus prennent en compte les distorsions de l'image. Pour la classe des surfaces de Riemann étudiées, les calculs comportent seulement des différences finies euclidiennes standard sur le plan de l'image.

Une autre approche proposée par Demonceaux et Vasseur dans [4], utilise les champs aléatoires de Markov, sur des images catadioptriques. La topologie d'un voisinage de pixels étant modifiée par la projection sur le miroir, un nouveau voisinage est défini pour les champs aléatoires de Markov, à l'aide du modèle de projection unifié présenté Section 1.1.4.1.5. Avec ce système de voisinage, les interactions entre deux pixels différents s'ils sont situés à la périphérie ou au centre de l'image. Cependant, cette technique peut poser des problèmes liés à la minimisation markovienne d'énergie, en termes de temps de calcul. Des algorithmes déterministes, plus rapides, sont mis en place, mais ne permettent pas de trouver de minimum global.

Quant à Strauss et Comby [42], ils proposent de projeter, sur chaque pixel de l'image omnidirectionnelle, un noyau dont la géométrie est définie dans l'espace du cylindre en utilisant les paramètres d'étalonnage du système. Des opérateurs morphologiques flous à noyaux variables pour images omnidirectionnelles permettent de compenser les imprécisions liées à l'échantillonnage de l'image et la localisation du voisinage. Cette méthode est utilisée pour le traitement morphologique d'image omnidirectionnelle.

Conclusion

Depuis le début de la photographie, les hommes essaient de représenter le monde qu'ils voient. Pour élargir le champ de vision, de nombreuses techniques de visions panoramiques ont été imaginées, de la réalisation de lentilles grand-angle, à la conception de systèmes multi-caméras, en passant par les systèmes rotatifs permettant de reconstruire une image panoramique à partir d'une mosaïque d'images. Depuis une vingtaine d'années, l'utilisation de miroirs de révolution s'est beaucoup développée et l'approche catadioptrique tient une place privilégiée en vision omnidirectionnelle. Lorsqu'ils vérifient la contrainte du point de vue unique, ces capteurs permettent de reconstruire n'importe quelle image perspective issue d'une image omnidirectionnelle. Cependant, les images acquises présentent d'importantes distorsions radiales qui les rendent difficiles à interpréter et à traiter. Dans la littérature, de nombreux auteurs ont été confrontés à ces problèmes. Certains n'ont pas tenu compte de la géométrie de ces images et les ont traitées comme des images perspectives. D'autres ont cherché à reconstruire une image géométriquement correcte, perspective ou cylindrique, pour appliquer les traitements existants. Cependant, la projection de l'image omnidirectionnelle dans un autre espace modifie l'information et ajoute des imprécisions à l'image. Il est donc crucial de conserver l'image dans son espace d'origine. Si les données restent inchangées, c'est l'opérateur qu'il faut modifier. C'est ce que nous proposons comme alternative.

Chapitre 2

Modélisation floue

Introduction

Les signaux informatifs (sonores, visuels, etc.) de notre environnement sont des grandeurs continues dans le temps. On dit qu'elles sont analogiques. Pour garder une trace de ces signaux, il est possible de les enregistrer sur un support physique. Deux choix s'offrent en fonction de l'enregistrement : analogique (cassette vidéo, audio, disque vinyle, etc.) ou numérique (valeurs binaires). Les enregistrements numériques nécessitent une transformation du signal composée d'au moins deux étapes : l'échantillonnage et la quantification. On dira alors que le signal numérique, par opposition au signal continu, est discret. L'image est acquise par une caméra numérique. Le type de signaux qui nous intéresse dans ce manuscrit constitue alors un signal bidimensionnel discret.

Dans ce chapitre, nous rappellerons rapidement les principales étapes d'acquisition des images numériques. Nous présenterons quelques imprécisions qu'elles introduisent sur la mesure de luminance des images puis proposerons une modélisation des données permettant de prendre en compte ces imprécisions lors des traitements. Les traitements usuels, issus de la théorie du signal, sont généralement élaborés pour des signaux continus. Le fait de stocker ces signaux sous forme numérique (et donc discrète) nécessite quelques adaptations concernant ces algorithmes. Une étape d'interpolation est souvent utilisée pour passer du signal discret à une représentation s'approchant du signal continu. Cette étape fait habituellement appel à des voisinages pondérés, s'appuyant sur l'utilisation de noyaux probabilistes. Cependant, il existe une autre manière de définir un voisinage pondéré, utilisant des agrégations non-additives et permettant de prendre en compte les imprécisions présentes dans l'image. C'est cette solution que nous avons choisie.

2.1 Imprécisions sur les images numériques

Pour traiter une image par ordinateur, celle-ci doit être numérisée. Lors de l'acquisition d'une image numérique, le signal lumineux continu émis par la scène doit être converti en un signal discret échantillonné. Cette étape est appelée numérisation. La Figure 2.1 décrit l'ensemble du processus d'acquisition d'images numériques.

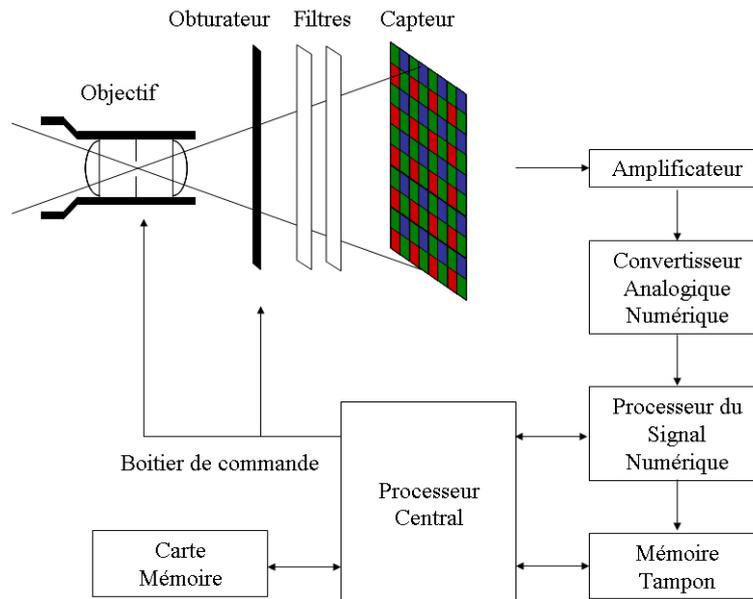


FIG. 2.1 – Fonctionnement d'un système d'acquisition d'images numériques. Un boîtier numérique constitué d'un corps, d'un objectif, d'un diaphragme, d'un obturateur, permet de créer une image sur la surface sensible. Lors de l'acquisition, un filtre "infra-rouge" limite la sensibilité des capteurs aux radiations et un filtre "passe-bas" supprime les aberrations chromatiques et produit un léger flou pour éviter des erreurs d'interpolation de la couleur par le capteur. Le capteur réagit à l'intensité lumineuse et produit des signaux électriques analogiques permettant de restituer les couleurs de la scène observée. A la sortie du capteur, un amplificateur permet d'augmenter le niveau du signal. Le signal analogique est ensuite converti en signal numérique, les pixels "éteints" issus des défauts de fabrication du capteur sont traités et finalement l'image est enregistrée.

La numérisation s'effectue au niveau du convertisseur Analogique/Numérique. Ce convertisseur mesure l'amplitude du courant entrant et le transforme en données binaires. La conversion du signal s'effectue suivant au moins deux étapes : l'échantillonnage et la quantification.

L'étape de quantification de l'image consiste à déterminer la plage de valeurs que peuvent prendre les échantillons : elle dépend du nombre de bits sur lequel on code les valeurs des pixels (plus on alloue de bits, meilleure est la résolution). Si plusieurs valeurs sont très proches dans le signal analogique, elles peuvent correspondre à la même valeur numérique, d'où les notions d'erreurs de quantification

et de bruit de quantification. L'erreur de quantification n'est pas corrigible : le signal est irrémédiablement dégradé. Nous ne pouvons donc pas prendre en compte cette erreur et ne développerons pas cette étape. En revanche, les imprécisions liées à l'échantillonnage spatial du signal peuvent être prises en compte. Dans la suite de ce chapitre, nous expliciterons ces imprécisions et présenterons la méthode permettant de les gérer lors des traitements.

2.1.1 Echantillonnage

Un signal analogique est, par définition, d'une précision infinie. L'échantillonnage permet de réduire ce signal à une suite de points discrets, afin de le stocker numériquement. Si la fréquence

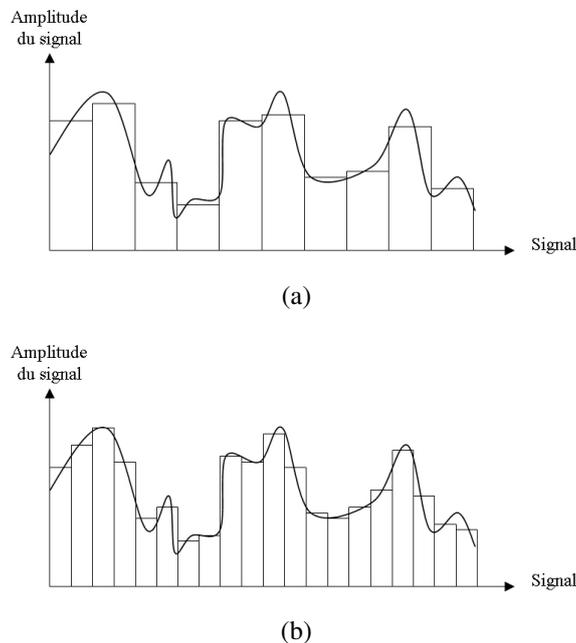


FIG. 2.2 – Echantillonnage du signal avec (a) une faible fréquence et (b) une fréquence deux fois plus importante.

d'échantillonnage est très faible (Figure 2.2(a)), les acquisitions seront très espacées. Tous les détails entre ces deux positions de capture seront perdus. Ainsi, la qualité du signal numérique est déterminée par la fréquence d'échantillonnage. Plus la fréquence d'échantillonnage sera élevée, plus la traduction numérique sera proche de l'original analogique (Figure 2.2(b)). Cependant, l'augmentation du nombre d'échantillons implique une augmentation de la quantité d'informations à conserver et par conséquent une augmentation de l'espace de stockage. Il est donc important de bien choisir la fréquence d'échantillonnage : suffisamment grande pour restituer correctement l'ensemble des informations du signal analogique mais sans être excessive afin de ne pas gaspiller l'espace de stockage.

D'après le théorème de Nyquist-Shannon, la fréquence d'échantillonnage d'un signal doit être égale ou supérieure au double de la fréquence maximale contenue dans ce signal.

L'image peut être vue comme un signal bi-dimensionnel. Dans le cas d'une image, le signal est découpé en petites surfaces carrées élémentaires d'une matrice rectangulaire. Chaque carré représente une cellule élémentaire de l'image appelée pixel. Un pixel est repéré par ses coordonnées entières dans l'image i et j et caractérisé par le niveau de gris quantifié qui lui est associé. Cette mosaïque de pixels donne de loin l'impression d'une image continue.

La Figure 2.3 représente l'image de Léna numérisée avec différentes fréquences d'échantillonnage. Pour chacune de ces images numérisées, notre oeil intègre les signaux lumineux. Lorsque la fré-

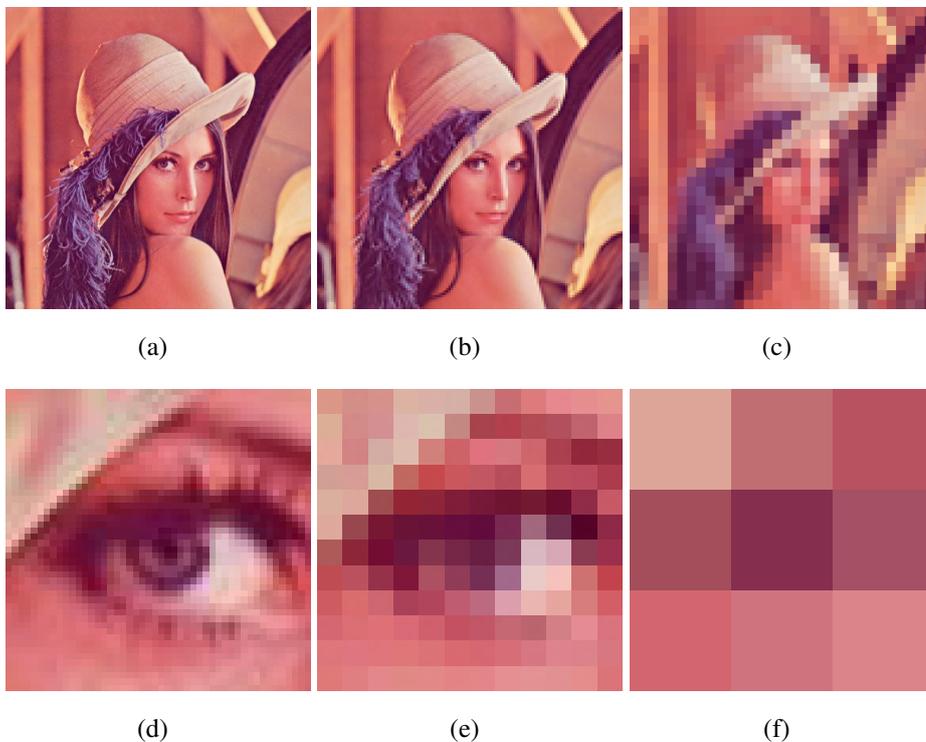


FIG. 2.3 – Image de Léna échantillonnée (a) avec une fréquence élevée, (b) avec une fréquence moyenne, (c) avec une faible fréquence. Zoom sur l'oeil de Léna (d) avec une fréquence élevée, (e) avec une fréquence moyenne, (f) avec une faible fréquence.

quence d'échantillonnage de l'image est suffisamment importante (a), notre système visuel interpole naturellement le signal discret de l'image, ce qui nous donne une impression de continuité en visualisant cette image. En zoomant sur une partie de cette image (d), les phénomènes de discrétisation sont à peine visibles. Le choix d'un pas d'échantillonnage plus grand, illustré sur les images (b) et (c), provoque une perte de détail et de précision importante dans l'image. Les phénomènes de discrétisation deviennent plus visibles et notre oeil discerne distinctement les valeurs des surfaces

de chaque pixel. Il devient alors impossible de zoomer sur une partie de ces images (e) et (f). On remarque que plus on grossit le pas d'échantillonnage, plus la quantité de données contenue par un pixel est importante. L'intégration du flux lumineux sur la surface du capteur fournit une valeur moyenne de la luminance du pixel. Les informations du signal continu initial ont été perdues.

Mais comment transfère-t-on la luminance de la scène observée aux pixels de l'image ? Pour répondre à cette question, il est nécessaire d'explicitier les mécanismes d'acquisition l'intensité lumineuse de la scène par la caméra.

C'est au niveau du capteur de la caméra que se crée une réaction à l'intensité lumineuse de la scène observée. Le capteur produit des signaux électriques analogiques. Dans sa thèse, Rabaud [48] distingue deux types de capteurs : les capteurs des caméras à balayage et les capteurs des caméras matricielles.

Avec une caméra à balayage, l'image est projetée sur une plaque photo-sensible balayée par un faisceau d'électrons. L'image est décomposée en une série d'impulsions électriques dont l'amplitude est représentative de l'illumination d'un point de l'image. Le signal électrique créé est déchargé en permanence par un système de lecture donnant naissance à un courant électrique : le signal vidéo. La cible est lue ligne par ligne, des impulsions venant signaler le changement de ligne et le changement d'image. Cependant, la largeur de la bande de balayage est inférieure à la taille d'un pixel. Le signal lumineux est donc intégré sur une surface beaucoup plus petite que la taille d'un échantillon. L'ensemble des portions de l'espace échantillonné n'est donc pas représentatif de la totalité de la scène observée. Certaines parties de la scène ne sont pas prises en compte lors de l'estimation de la valeur du niveau de gris des pixels associés.

Avec une caméra matricielle, la plaque photographique est recouverte de plusieurs millions de cellules, appelées "photosites", sensibles au rouge, vert et bleu, comme les cônes de notre rétine. Ces photosites sont regroupés en matrices contenant chacune deux photosites verts, un rouge et un bleu. Cette structure est inspirée de notre perception visuelle non-uniforme sur l'ensemble du spectre (elle est plus sensible aux verts 50% qu'aux rouges 39% représentatifs des hautes fréquences et aux bleus 11% représentatifs des basses fréquences et peut distinguer beaucoup de nuances dans les ombres du noir au blanc). Un photosite est associé à un pixel de l'image. Les photosites ont généralement une forme rectangulaire ou hexagonale. Avec des caméras de type C.C.D. (Coupled Charge Device), ils sont jointifs. Il y a coïncidence entre le photosite et le pixel. En revanche, pour des caméras de type C.M.O.S. (Complementary Metal Oxide Semi-conductor), le photosite est a priori inclus dans le pixel. Nous en arrivons donc à la même conclusion que pour les caméras à balayage : certaines parties de la scène n'interviennent pas dans l'estimation de la valeur du niveau de gris des pixels associés.

La Figure 2.4 illustre la relation existant entre le pixel et la portion de l'espace échantillonné par les différentes caméras présentées. Il apparaît clairement que tous les points de l'espace ne sont pas pris en compte par ces différents capteurs. Lors de l'échantillonnage, chaque pixel n'intègre donc

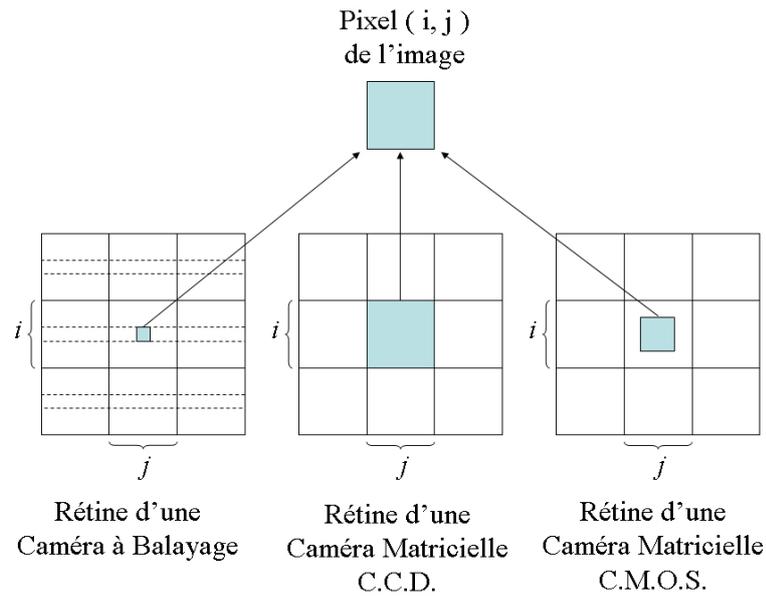


FIG. 2.4 – Numérisation de l'espace.

qu'une partie du signal lumineux. La Figure 2.5 montre la quantité d'information du signal intégrée par une ligne de pixels de l'image. En effet, le cône de visibilité du photosite des caméras C.M.O.S.

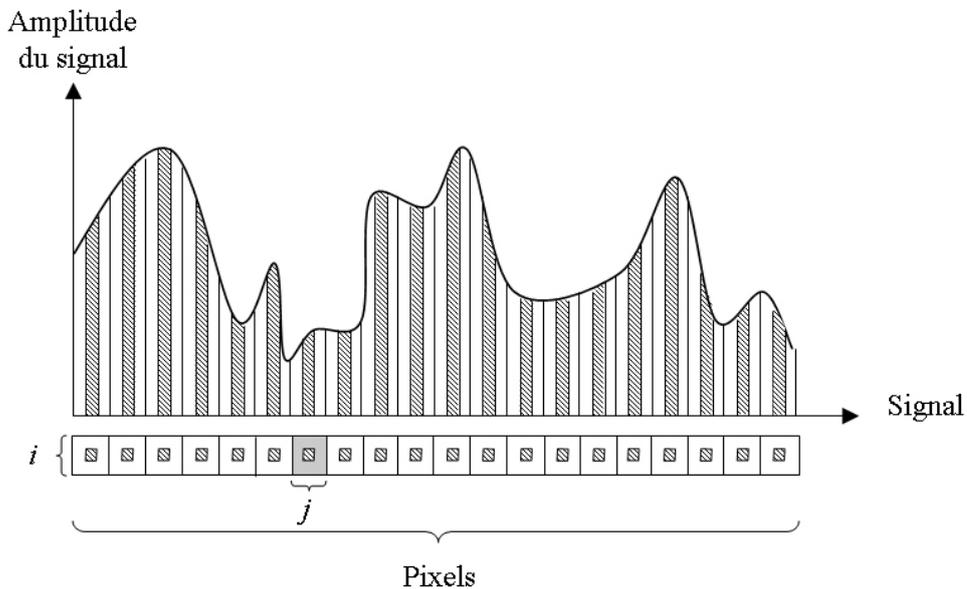


FIG. 2.5 – Portion du signal intégrée par les pixels d'une même ligne.

ou de caméras à balayage, représenté Figure 2.6, est beaucoup plus petit que le cône de visibilité d'un pixel. Rabaud explique que sous l'hypothèse de continuité de la matière ¹ et l'hypothèse de

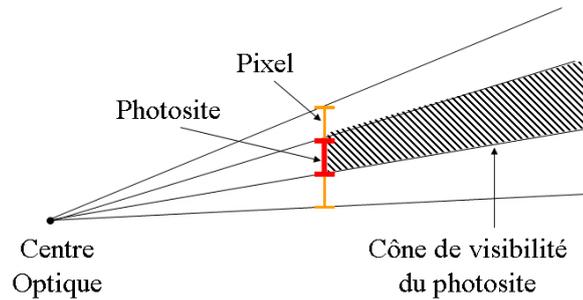


FIG. 2.6 – Cône de visibilité d'un photosite dans le cas de caméras C.M.O.S. ou de caméras à balayage.

surface ayant une granularité faible, il y a une forte corrélation entre les points de l'espace qui sont à l'extérieur du cône de visibilité du photosite et ceux qui sont à l'intérieur.

Les niveaux de gris des pixels peuvent donc varier en fonction du temps d'intégration (ou temps de pose), du biais (ou offset) de chaque pixel², du courant d'obscurité (ou dark), de la sensibilité et du flux de lumière atteignant réellement le pixel. Tous ces facteurs ajoutent des imprécisions sur les valeurs des niveaux de gris des pixels. Dans les cas des images omnidirectionnelles viennent s'ajouter les imprécisions liées à la non-uniformité de l'information contenue dans chaque pixel en fonction de leur position dans l'image. De plus, la présence d'un bruit, inhérent à toute mesure, doit être prise en compte. Ainsi, deux images d'une même scène, acquises avec un système d'acquisition donné, ne seront pas strictement identiques. Avec un autre système d'acquisition, on obtiendrait une troisième image encore différente des deux premières.

2.1.2 Modélisation de l'image

2.1.2.1 Discrétisation du plan de l'image

Chaque pixel est représenté par un élément de la partition du plan de l'image. La partition du plan Ω de l'image I est donc constituée d'un ensemble de N éléments non-vides $(H_n)_{n \in \{1, \dots, N\}}$. Elle est définie suivant deux critères :

- le critère de continuité

$$\bigcup_{n=1}^N H_n = \Omega \quad (2.1)$$

¹**Hypothèse de continuité de la matière** : pas de discontinuité forte le long d'une surface.

²**Offset** : Sans exposition à la lumière et sans temps de pose, nous devrions obtenir une matrice de pixels nulle. Ce n'est pas le cas du fait du convertisseur analogique-numérique qui transforme un courant électrique (les électrons piégés dans les puits de potentiel que sont chacun des pixels) en nombre. Le courant électrique résiduel de chaque pixel étant différent, il est impossible de trouver un courant électrique qui correspondrait à zéro.

– le critère d’additivité

$$m \neq n \Rightarrow H_n \cap H_m = \emptyset \quad (2.2)$$

$$\bigcup_{n=1}^N H_n = \Omega \quad (2.3)$$

Si les H_n sont de taille identique, la partition est dite régulière. Considérons une partition régulière satisfaisant les propriétés de continuité 2.3 et d’additivité 2.2. Chaque pixel de cette partition est caractérisé par une valeur. Cependant, nous n’avons aucune information sur la localisation précise de cette valeur au sein du pixel, par conséquent elle est supposée répartie uniformément en tout point du pixel. Les pixels peuvent être vus comme des nombres binaires auxquels on a attribué une valeur en niveau de gris ou en couleur. L’ensemble de ces pixels constitue alors une partition qui satisfait les propriétés de continuité 2.3 et d’additivité 2.2.

Cependant, il existe d’autres approches qui ne vérifient pas ces propriétés. En réalité, on peut distinguer trois sortes de discrétisation binaire régulière qui peuvent être utilisées pour une image. Ces partitions sont décrites Figure 2.7. Dans la partition (a), les $(H_n)_{n \in \{1, \dots, N\}}$ de Ω vérifient les propriétés 2.3 et 2.2. Dans la partition (b), les $(H'_n)_{n \in \{1, \dots, N\}}$ ne couvrent pas la totalité de l’espace, ils ne vérifient que la propriété 2.2. Enfin les $(H''_n)_{n \in \{1, \dots, N\}}$ de la partition (c) se chevauchent ; seule la propriété 2.3 est satisfaite.

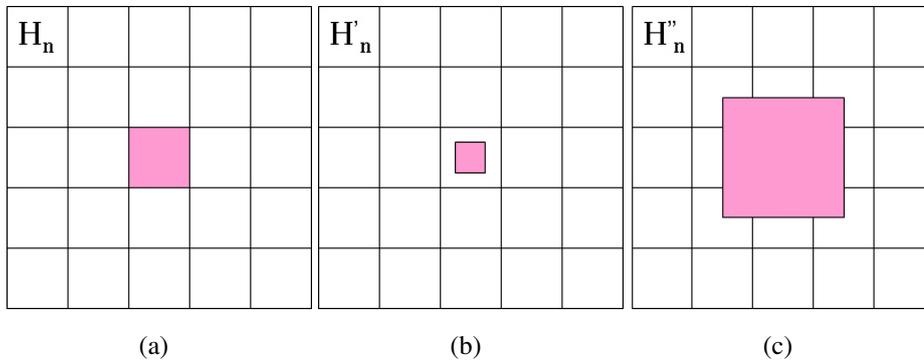


FIG. 2.7 – Partitionnement régulier de l’image (a) couvrant vérifiant les propriétés 2.3 et 2.2, (b) non-couvrant vérifiant uniquement la propriété 2.2, (c) couvrant vérifiant uniquement la propriété 2.3

2.1.2.2 Modélisation de l’information

Le choix d’une partition comme celle décrite Figure 2.7(c) permet de prendre en compte l’imprécision sur les valeurs des niveaux de gris des pixels présentée Section 2.1.1. Pour chaque pixel $H_{i,j}$ à la position (i, j) dans l’image, la localisation du niveau de gris est inconnue dans l’intervalle $2D [i - \Delta_i, i + \Delta_i] \times [j - \Delta_j, j + \Delta_j]$, où $2 \times \Delta_i$ représente la taille d’un pixel horizontal et $2 \times \Delta_j$

la taille d'un pixel vertical. Cette partition ne vérifie que la propriété 2.3. Cependant, elle prend en compte la totalité de l'image. On parle de partition couvrante de l'image.

Cette partition peut être modélisée avec des sous-ensembles flous. La théorie des ensembles flous est introduite par Zadeh [70] en 1965. Associée à la théorie des possibilités, elle permet de représenter mathématiquement l'imprécision relative de certaines informations. Dans cette théorie interviennent les notions de mesures de possibilité et de nécessité présentées par Dubois et Prade [11]. Lorsque l'on raisonne avec cette théorie, les statistiques classiques ne fonctionnent plus. Il faut utiliser des outils provenant de la théorie des possibilités. Nous avons rappelé en Annexe A les notions de base de cette théorie concernant les ensembles flous.

Dans [9] les auteurs ont montré que les nombres flous triangulaires dominent tous les nombres flous mono-modaux symétriques de même mode et même support. Nous avons représenté chaque pixel de l'image par des nombres triangulaires flous. Les nombres 1D triangulaires flous ont été étendus en nombres 2D flous pyramidaux (Figure 2.8) en considérant la t-norme min (produit cartésien de deux nombres 1D flous triangulaires). Tous les pixels constituent alors une partition floue de l'image. Prendre une modélisation floue de l'image avec des cellules qui se chevauchent permet d'éviter l'aspect arbitraire du partitionnement.

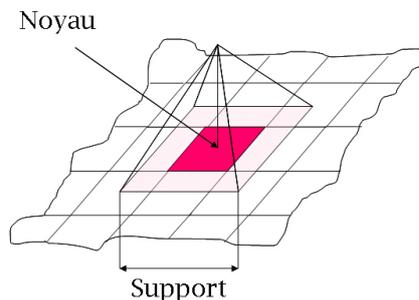


FIG. 2.8 – Représentation d'un pixel flou

2.2 Théorie des possibilités et traitement d'images

En traitement du signal, la plupart des algorithmes sont formulés dans le domaine continu. Or bien souvent le signal à notre disposition est discret. L'approche classique consiste à reconstruire le signal continu de l'image à partir des valeurs discrètes échantillonnées. Cette reconstruction est basée sur l'utilisation de noyaux.

Les noyaux sont des fonctions qui utilisent un voisinage pondéré du point considéré. Les algorithmes de traitement d'images sont souvent des généralisations en 2D de ceux issus du traitement du signal. A ce titre, ils héritent du même problème : appliquer un traitement continu sur une image discrète. Un schéma général du traitement d'images est présenté par la Figure 2.9. Une opération continue

est utilisée pour filtrer les images dans l'espace continu. Un noyau d'interpolation η est utilisé pour assurer la transition entre le signal discret de l'image originale et le signal continu interpolé ; et un noyau d'échantillonnage ν assure le lien entre l'image traitée dans le domaine continu et son homologue dans l'espace discret.

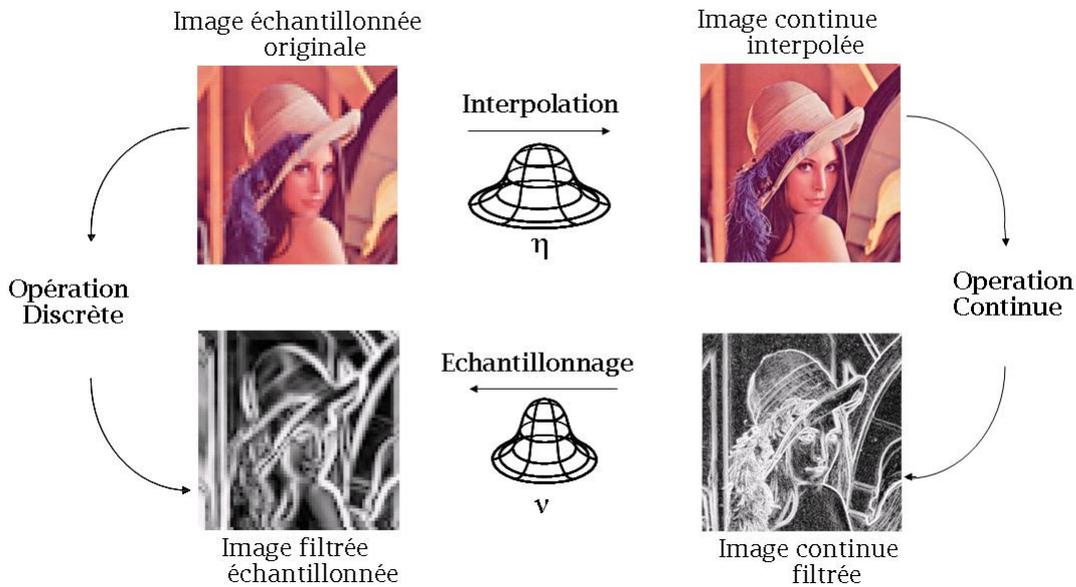


FIG. 2.9 – Schéma général de traitement d'images.

2.2.1 Noyaux sommatifs

Généralement, l'interpolation du signal est réalisée à l'aide de fonctions appelées noyaux. Ces noyaux permettent d'associer, à chaque mesure d'illumination échantillonnée, un domaine probabilisé dans le plan de l'image. Ce domaine probabilisé représente indifféremment : les phénomènes d'intégration donnant lieu à la mesure de la luminance, le domaine spatial sur lequel la mesure de luminance peut être considérée comme ergodique ou le domaine dans lequel cette valeur de luminance peut être prise en compte pour en obtenir une interpolation. On parle de noyaux probabilistes ou noyaux sommatifs.

Un noyau sommatif ou probabiliste est une fonction κ , à valeurs dans \mathbb{R} , définie sur un domaine Ω , vérifiant la propriété de sommativité normalisée :

$$\int_{\Omega} \kappa(\omega) d\omega = 1. \quad (2.4)$$

En reprenant notre schéma général de traitement d'images, on a :

$$\kappa = f(\eta) * \nu. \quad (2.5)$$

où f est l'opération continue fonction de η .

En traitement d'images, ce noyau est positif, centré, borné, mono-modal, symétrique. Le critère de positivité permet de le voir comme une fonction de distribution probabiliste de mesure $P_\kappa(A)$:

$$\forall A \subseteq \Omega, P_\kappa(A) = \int_A \kappa(\omega) d\omega. \quad (2.6)$$

vérifiant :

- (i) $P_\kappa : \wp(\Omega) \longrightarrow [0, 1]$ où $\wp(\Omega)$ est l'ensemble des parties de Ω ,
- (ii) $P_\kappa(\Omega) = 1$,
- (iii) Si $a \cap b = \emptyset$, $P_\kappa(a \cup b) = P_\kappa(a) + P_\kappa(b)$.

La modélisation classique par noyau sommatif consiste à considérer que l'espérance d'un signal S via un noyau sommatif κ est obtenue par convolution de ce signal avec le noyau sommatif κ . L'estimation d'un signal S convolué au noyau sommatif κ est donnée par :

$$\mathbb{E}_\kappa(S) = \int_\Omega S dP_\kappa = \int_\Omega S(\omega) \kappa(\omega) d\omega. \quad (2.7)$$

De nombreuses approches utilisent des noyaux sommatifs (ou éléments structurants en morphologie mathématique) pour traiter les images. Cependant, il existe une autre représentation du voisinage pondéré utilisée en morphologie floue . Cette représentation utilise des noyaux flous. Ces noyaux ne vérifient pas la propriété de sommativité normalisée (2.4). Nous les appellerons donc noyaux non-sommatifs ou maxitifs.

2.2.2 Noyaux non-sommatifs

Un noyau non-sommatif est une fonction π , à valeurs dans $[0, 1]$, définie sur un domaine Ω , vérifiant la propriété de maximisation normalisée :

$$\max_{\omega \in \Omega} \pi(\omega) = 1. \quad (2.8)$$

De même qu'un noyau sommatif peut être associé à une distribution de probabilité, un noyau non-sommatif peut être associé à une distribution de possibilité [11]. Celle-ci peut être vue comme la fonction d'appartenance à un sous-ensemble flou. Sa mesure est donnée par :

$$\forall A \subseteq \Omega, \Pi_\pi(A) = \sup_{\omega \in A} \pi(\omega). \quad (2.9)$$

et satisfait :

- (i) $\Pi_\pi : \wp(\Omega) \longrightarrow [0, 1]$ où $\wp(\Omega)$ est l'ensemble des parties de Ω ,
- (ii) $\Pi_\pi(\Omega) = 1$, $\Pi_\pi(\emptyset) = 0$,

$$(iii) \forall I \subset \mathbb{N}, \forall i \in I, \forall A_i \in \wp(\Omega), \Pi_\pi(\cup A_i) = \sup_i(\Pi_\pi(A_i)).$$

A toute mesure de possibilité Π_π on associe une fonction duale N_π , appelée mesure de nécessité, telle que :

$$(i) N_\pi : \wp(\Omega) \longrightarrow [0, 1] \text{ où } \wp(\Omega) \text{ est l'ensemble des parties de } \Omega,$$

$$(ii) N_\pi(\Omega) = 1, N_\pi(\emptyset) = 0,$$

$$(iii) \forall I \subset \mathbb{N}, \forall i \in I, \forall A_i \in \wp(\Omega), N_\pi(\cap A_i) = \inf_i(N_\pi(A_i)).$$

Ces deux mesures de confiance vérifient

$$\Pi_\pi(A) \geq N_\pi(A), \tag{2.10}$$

$$N_\pi(A) = 1 - \Pi_\pi(A^c). \tag{2.11}$$

2.2.3 Lien entre les noyaux sommatifs et non-sommatifs

Les noyaux sommatifs et non-sommatifs sont liés par le fait qu'une capacité concave ν peut représenter une famille spéciale des mesures de probabilité $core(\nu^c)$, telle que

$$core(\nu^c) = \{P_\kappa \mid \forall A \subseteq \Omega, \nu^c(A) \leq P_\kappa(A) \leq \nu(A)\}. \tag{2.12}$$

Une mesure de possibilité est un cas particulier de capacité concave de Choquet notée ν . Son conjugué ν^c est une capacité convexe qui est donc définie par :

$$\forall A \subseteq \Omega, \nu^c(A) = 1 - \nu(A^c). \tag{2.13}$$

où A^c est le complémentaire de A dans Ω .

Cela signifie qu'une distribution de possibilité englobe une famille de distributions de probabilité, ou encore, qu'un noyau non-sommatif représente une famille de noyaux sommatifs.

Dubois et Prade ont défini, dans [10], une transformation optimale permettant de passer d'une distribution de probabilité à une distribution de possibilité, selon 3 critères :

(i) La condition de cohérence ou principe de domination :

La représentation possibiliste est moins informative que la représentation probabiliste. Une transformation probabilité-possibilité perd une partie de l'information (plusieurs distributions de probabilités peuvent donc aboutir à la même distribution de possibilité), alors qu'une transformation possibilité-probabilité en ajoute arbitrairement de l'information. Cette condition permet d'associer, à tout noyau sommatif κ , un noyau non-sommatif π satisfaisant :

$$\forall A \subseteq \Omega, P_\kappa(A) \leq \Pi_\pi(A). \tag{2.14}$$

(ii) Le principe de préservation de la préférence :

Il s'agit de préserver l'ordre des éléments entre eux.

$$\forall (a, b) \in \Omega^2, \begin{cases} \kappa(a) > \kappa(b) \Rightarrow \pi(a) > \pi(b), \\ \kappa(a) = \kappa(b) \Rightarrow \pi(a) = \pi(b). \end{cases} \quad (2.15)$$

(iii) Le principe du maximum de spécificité :

La plus spécifique des deux distributions de possibilités est la plus petite au sens de la cardinalité de l'ensemble flou.

$$\forall A \subseteq \Omega, \begin{cases} N_{\pi_1} \leq N_{\pi_2}, \\ \Pi_{\pi_2} \leq \Pi_{\pi_1}. \end{cases} \quad (2.16)$$

Considérons κ une distribution probabilité normalisée, mono-modale dans \mathbb{R} , ω_m son mode et $f :]-\infty, \omega_m] \rightarrow [\omega_m, \infty[$, la fonction définie par $f(x) = \max\{\omega \mid \kappa(\omega) \geq \kappa(x)\}$. Alors, la fonction de distribution de possibilité π la plus spécifique, compatible avec κ est :

$$\pi_{optimal}(x) = \pi_{optimal}(f(x)) = \int_{-\infty}^x \kappa(\omega) d\omega + \int_{f(x)}^{\infty} \kappa(\omega) d\omega. \quad (2.17)$$

Si κ est symétrique, on a

$$\pi_{optimal}(x) = 2 \int_{-\infty}^x \kappa(\omega) d\omega \quad (2.18)$$

$$\frac{\delta}{\delta x} \pi_{optimal}(x) = 2(\kappa(x) - \lim_{\omega \rightarrow -\infty} \kappa(\omega)) \quad (2.19)$$

$$= 2(\kappa(x) - 0) \quad (2.20)$$

$$= 2\kappa(x) \quad (2.21)$$

$\kappa(x)$ étant croissante sur l'intervalle $]-\infty, \omega_m[$, la dérivée de $\pi_{optimal}(x)$ est également croissante cet intervalle $]-\infty, \omega_m[$. Donc, la distribution $\pi_{optimal}(x)$ n'a pas de point d'inflexion sur l'intervalle $]-\infty, \omega_m[$. De même $\kappa(x)$ étant décroissante sur l'intervalle $]\omega_m, \infty[$, la dérivée de $\pi_{optimal}(x)$ est également décroissante cet intervalle. Donc, la distribution $\pi_{optimal}(x)$ n'a pas de point d'inflexion sur l'intervalle $]\omega_m, \infty[$. La courbe prend donc une forme incurvée à gauche et à droite du mode ω_m . Quelle que soit la distribution de probabilité mono-modale, de mode ω_m , symétrique et bornée sur l'intervalle $[-\Delta, \Delta]$, sa distribution de possibilité optimale associée est majorée par la distribution de possibilité triangulaire, de même mode et même support.

Dubois, Foulloy, Mauris et Prade [9] ont montré ainsi que les nombres flous triangulaires généralisent tous les nombres flous mono-modaux symétriques. Ainsi, une distribution de possibilité triangulaire, symétrique, positive, de support donné, domine l'ensemble des distributions de probabilités symétrique, positive, de même mode et même support. Cela signifie qu'un noyau non-sommatif triangulaire, symétrique, positif, de support donné, domine l'ensemble des noyaux sommatifs symétrique, positifs, de même mode et même support. Cela explique que dans la suite de nos travaux nous ayons choisi d'utiliser des noyaux non-sommatifs triangulaires.

Les Figures 2.10(a), (b) et (c) illustrent des noyaux sommatifs κ bornés de même support $[-\Delta, \Delta]$, mono-modaux, symétriques, normalisés, positifs. Leur aire est donc égale à 1 ce qui explique la différence de hauteurs entre ces noyaux. Le premier a la forme d'une cloche, le second est plus fin et le troisième est un intermédiaire entre les deux précédents. La transformation probabilité-possibilité permet d'associer à chacun de ces noyaux un noyau non-sommatif de la forme de (d). La distribution de probabilité triangulaire (e) domine l'ensemble des distribution de probabilités.

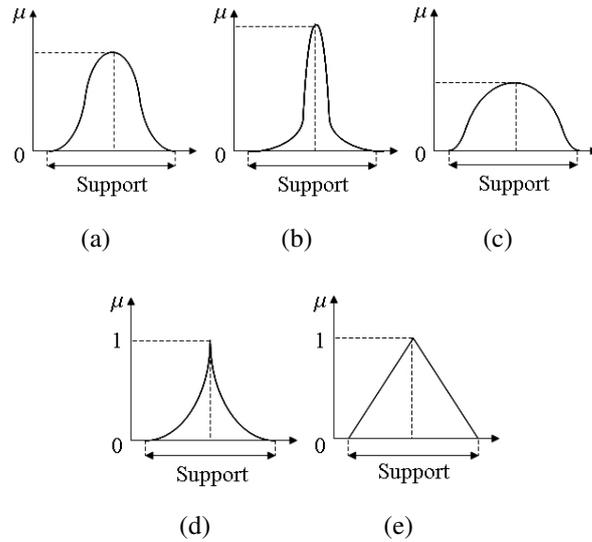


FIG. 2.10 – (a), (b) et (c) Distributions de probabilité. (d) Distribution de possibilité. (e) Distribution de possibilité triangulaire.

Le principe de domination des mesures de possibilité peut être étendu à plusieurs dimensions. En dimension 2, dans un espace $\Omega_1 \times \Omega_2$, un produit des mesures marginales de possibilité Π_1 et Π_2 est défini dans [11] par

$$\forall (A, B) \subseteq \Omega_1 \times \Omega_2, \Pi_{12}(A, B) = \min(\Pi_1(A), \Pi_2(B)). \quad (2.22)$$

Supposons que marginalement, $\Pi_1 \geq P_1$ et $\Pi_2 \geq P_2$, avec P_1 et P_2 des mesures de probabilité respectivement sur Ω_1 et Ω_2 . Alors Π_{12} domine P_{12} défini par

$$\forall (A, B) \subseteq \Omega_1 \times \Omega_2, P_{12}(A, B) = P_1(A)P_2(B). \quad (2.23)$$

$$P \leq 1 \Rightarrow \min(\Pi_1(A), \Pi_2(B)) \geq \min(P_1(A), P_2(B)) \geq P_1(A)P_2(B). \quad (2.24)$$

Avec les définitions de P_{12} et Π_{12} , on a

$$\Pi_{12}(A, B) \geq P_{12}(A, B). \quad (2.25)$$

Il est ainsi possible d'étendre nos noyaux non-sommatifs en 2D afin de pouvoir les appliquer à l'image.

2.3 Intégrale de Choquet

2.3.1 Intégrale de Choquet continue

En traitement d'image, la convolution de l'image par un noyau sommatif κ fournit une espérance $\mathbb{E}_\kappa(S)$ de l'image qui est une forme de moyenne pondérée. Avec des noyaux non-sommatifs, cette espérance ne peut être utilisée. Lorsqu'un noyau sommatif est remplacé par une capacité concave de Choquet, l'Equation (2.7) est remplacée par une espérance plus générale appelée une intégrale de Choquet [13, 36] :

$$\mathbb{C}_\nu(S) = (C) \int_{\Omega} S d\nu = \int_0^{+\infty} \nu(\{\omega : S(\omega) > \alpha\}) d\alpha. \quad (2.26)$$

$$\mathbb{C}_{\nu^c}(S) = (C) \int_{\Omega} S d\nu^c = \int_0^{+\infty} \nu^c(\{\omega : S(\omega) > \alpha\}) d\alpha, \quad (2.27)$$

L'intégrale de Choquet peut être vue comme une généralisation de l'intégrale de Lebesgue, définie pour les mesures non-additives. Elle se distingue par la représentation de phénomènes d'interaction mutuelle qui peuvent exister entre certains éléments.

L'utilisation de l'intégrale de Choquet permet de conserver la propriété de domination décrite formule (2.12). Ainsi, l'espérance d'une image via un noyau sommatif κ est inférieure à l'espérance plus générale, donnée par l'intégrale de Choquet via un noyau non-sommatif. De plus si κ décrit l'ensemble des noyaux de même support, mono-modaux, symétriques, positifs, normalisés, l'intégrale de Choquet est le plus petit des majorants de toutes les espérances obtenues avec ces noyaux :

$$\mathbb{C}_\nu(S) = \sup_{\kappa | P_\kappa \in \text{core}(\nu)} \mathbb{E}_\kappa(S), \quad (2.28)$$

$$\mathbb{C}_{\nu^c}(S) = \inf_{\kappa | P_\kappa \in \text{core}(\nu^c)} \mathbb{E}_\kappa(S). \quad (2.29)$$

D'où le théorème suivant :

Théorème :

$$\forall \kappa \subseteq \text{core}(\nu^c), \mathbb{C}_{\nu^c}(S) \leq \mathbb{E}_\kappa(S) \leq \mathbb{C}_\nu(S) \quad (2.30)$$

2.3.2 Intégrale de Choquet discrète

Pour traiter des données discrètes, comme par exemple le signal échantillonné de l'image, c'est l'intégrale de Choquet discrète qui est utilisée. Cette intégrale permet de répercuter l'imprécision des données sur la valeur estimée.

Le signal continu est supposé connu avec une localisation imprécise représentée par une partition floue $(H_i)_{i \in \{1, \dots, N\}}$. Le noyau assure le lien entre le signal continu et le signal discret de l'image.

L'objectif est de transférer les connaissances discrètes du signal à tout voisinage pondéré W , représenté Figure 2.11, à l'aide d'une intégrale de Choquet qui fournit une borne supérieure $\mathbb{C}_{\nu_W}(S)$ et une borne inférieure $\mathbb{C}_{\nu_W^c}(S)$ de l'espérance de l'image via tout noyau sommatif de même mode et même support que W . Si de plus W est un noyau centré, symétrique, positif, alors choisir pour le

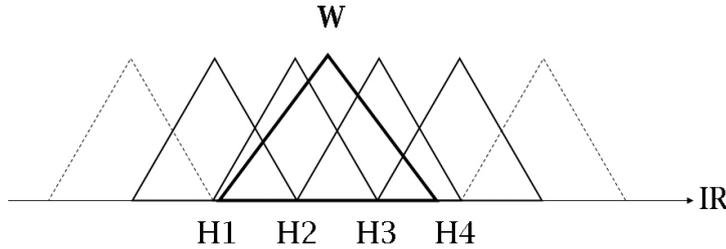


FIG. 2.11 – Noyau non-sommatif W .

représenter le noyau triangulaire comme distribution de possibilité permet de conserver le principe de domination à savoir :

$$\mathbb{C}_{\nu_W}(S) = \sum_{n=1}^N S_{(n)}[\nu_W(A_{(n)}) - \nu_W(A_{(n+1)})], \quad (2.31)$$

$$\mathbb{C}_{\nu_W^c}(S) = -\mathbb{C}_{\nu_W}(-S) = -\sum_{n=1}^N (-S_{(n)})[\nu_W(A_{(n)}) - \nu_W(A_{(n+1)})]. \quad (2.32)$$

où S_n est la valeur réelle positive associée à la $n^{\text{ème}}$ cellule et $(.)$ indique une permutation telle que $S_{(1)} \leq S_{(2)} \leq \dots \leq S_{(N)}$. Les $A_{(n)} = \{H_{(n)}, \dots, H_{(N)}\}$ sont les coalitions binaires des cellules de valeurs supérieures à $S_{(n)}$.

Remarque :

Pour déterminer la valeur d'un noyau sommatif chevauchant plusieurs cellules H_1, \dots, H_n , les modèles additifs constituent l'approche traditionnellement adoptée. L'utilisation de noyaux sommatifs suppose que la partition $(H_n)_{n \in \{1, \dots, N\}}$ satisfait les propriétés 2.3 et 2.2 et que la capacité est additive. On a donc $\nu(A_n) - \nu(A_{n+1}) = \nu(H_n)$. L'interpolation du signal continu entre les cellules est réalisée à l'aide d'un opérateur additif de type moyenne pondérée :

$$M_\nu(S) = \sum_{i=1}^n S_i \nu(H_n) \quad (2.33)$$

où S_i est la valeur de la cellule H_i et ν_i son poids.

Cette approche revient à estimer la contribution de chaque élément pour estimer la valeur finale. L'utilisation d'un modèle non-additif permet d'affaiblir les hypothèses nécessaires à la représentation additive.

Il existe deux visions différentes d'une distribution de possibilité détaillées dans la thèse de Lasserre [34] : une distribution de possibilité peut être vue comme une borne supérieure d'une famille de distributions de probabilité (c'est la vision verticale), ou par ses α -coupes comme un intervalle de confiance (c'est la vision horizontale). Une distribution possibiliste peut donc être construite à partir de la connaissance d'un ou plusieurs de ces intervalles. Dans la suite de ce document, nous expliciterons chacune des deux visions, illustrées d'exemples simples afin de faciliter leur compréhension.

2.3.2.1 Vision verticale

Une distribution de possibilité peut être vue comme une borne supérieure d'une famille de distributions de probabilité :

$$P(\Pi) = \{P | \forall A \subseteq \Omega, P(A) \leq \Pi(A)\} \quad (2.34)$$

Pour définir la capacité concave de l'équation (2.31), une solution consiste à utiliser un transfert pignistique. La capacité est alors donnée par :

$$\nu_W(A_{(n)}) = \frac{|W \cap A_{(n)}|}{|W \cap A_{(1)}|}. \quad (2.35)$$

avec $|X|$ est la cardinalité de l'ensemble X .

Considérons une partition floue. Ses cellules sont représentées par des nombres flous triangulaire, normalisés, symétriques, de hauteur $h = 1$ et de support $Supp = 2$, illustrés Figure 2.12. Chaque cellule interagit avec les cellules voisines.

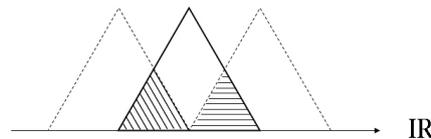


FIG. 2.12 – Interactions entre un pixel et ses voisins.

Si nous connaissons la valeur S_1 d'une cellule H_1 de cette partition (ie. nous n'avons aucune information à priori sur les valeurs des cellules voisines), alors la valeur de cette cellule vaut S_1 .

Supposons à présent que l'on connaisse la valeur S_2 de la cellule voisine H_2 de H_1 (Figure 2.13(a)). Les cellules H_1 et H_2 se chevauchant, on a donc $H_1 \cap H_2 \in [S_1, S_2]$. Pour estimer la valeur de chaque cellule, on utilise une intégrale de Choquet.

Posons $W = H_1$ et supposons que

$$\begin{cases} S_1 = 90 \\ S_2 = 180 \end{cases}$$

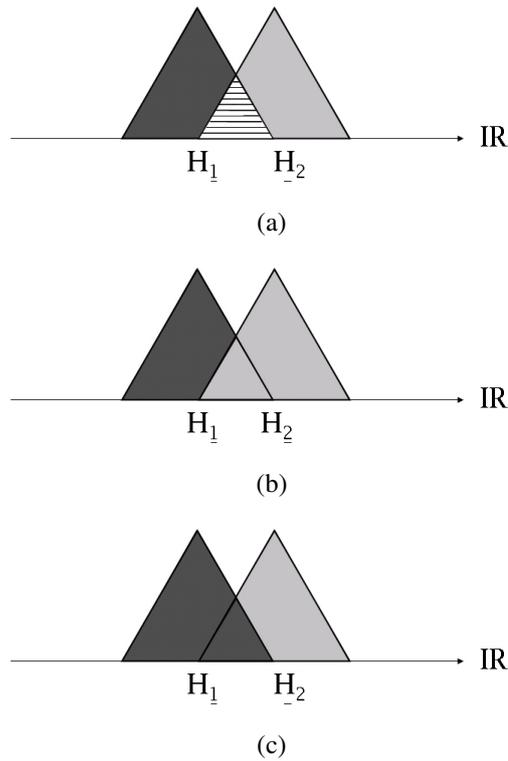


FIG. 2.13 – (a) Chacune des deux cellules interagit avec l'autre, l'estimation des valeurs par l'intégrale de Choquet fournit pour chacun (b) une borne supérieure obtenue en privilégiant les cellules ayant les plus grandes valeurs et (c) une borne inférieure privilégiant les cellules ayant les plus petites valeurs.

Les coalitions obtenues sont donc :

$$\begin{cases} A_{(1)} = \{H_1, H_2\} \\ A_{(2)} = \{H_2\} \end{cases}$$

Nos cellules sont définies telles que :

$$\begin{cases} |H_1| = |H_2| = 1 \\ |H_1 \cap H_2| = \frac{1}{4} \end{cases}$$

L'estimation de l'intégrale de Choquet supérieure (Figure 2.13(b)) s'écrit donc :

$$\begin{aligned}
 \mathbb{C}_{\nu_{H_1}}(S) &= \sum_{n=1}^2 S_{(n)}[\nu_{H_1}(A_{(n)}) - \nu_{H_1}(A_{(n+1)})] & (2.36) \\
 &= S_{(1)}[\nu_{H_1}(A_{(1)}) - \nu_{H_1}(A_{(2)})] + S_{(2)}[\nu_{H_1}(A_{(2)}) - 0] \\
 &= S_{(1)}\left[\frac{|H_1 \cap A_{(1)}| - |H_1 \cap A_{(2)}|}{|H_1 \cap A_{(1)}|}\right] + S_{(2)}\left[\frac{|H_1 \cap A_{(2)}|}{|H_1 \cap A_{(1)}|}\right] \\
 &= S_{(1)}\left[1 - \frac{|H_1 \cap H_2|}{|H_1|}\right] + S_{(2)}\frac{|H_1 \cap H_2|}{|H_1|} \\
 &= 90\left[1 - \frac{1}{4}\right] + \frac{180}{4} \\
 &= 112.5
 \end{aligned}$$

Pour l'estimation inférieure (Figure 2.13(c)), l'ordre des éléments est inversé : $-S_{(1)} \leq -S_{(2)}$, les coalitions sont donc modifiées :

$$\begin{cases} A_{(1)} = \{H_1, H_2\} \\ A_{(2)} = \{H_1\} \end{cases}$$

L'estimation inférieure vaut donc :

$$\begin{aligned}
 \mathbb{C}_{\nu_{H_1}^c}(S) &= -\sum_{n=1}^2 (-S_{(n)})[\nu_{H_1}(A_{(n)}) - \nu_{H_1}(A_{(n+1)})] & (2.37) \\
 &= -((-S_{(1)})[\nu_{H_1}(A_{(1)}) - \nu_{H_1}(A_{(2)})] + (-S_{(2)})[\nu_{H_1}(A_{(2)}) - 0]) \\
 &= -((-S_{(1)})\left[\frac{|H_1 \cap A_{(1)}| - |H_1 \cap A_{(2)}|}{|H_1 \cap A_{(1)}|}\right] - (-S_{(2)})\left[\frac{|H_1 \cap A_{(2)}|}{|H_1 \cap A_{(1)}|}\right]) \\
 &= -((-S_{(1)})\left[\frac{|H_1| - |H_1|}{|H_1|}\right] + (-S_{(2)})\left[\frac{|H_1|}{|H_1|}\right]) \\
 &= -(-S_{(2)}) \\
 &= 90
 \end{aligned}$$

La valeur de H_1 est donc connue de manière imprécise comme appartenant à l'intervalle $[90, 112.5]$. De même la valeur de H_2 sera comprise entre $[157.5, 180]$.

Considérons à présent le noyau non-sommatif W représenté Figure 2.11. Ce noyau recouvre de nombreuses cellules. Chaque cellule interagit avec les cellules voisines. Nous avons attribué des valeurs arbitraires à chacune des cellules comme illustré Figure 2.14(a) :

$$\begin{cases} S_1 = 150 \\ S_2 = 100 \\ S_3 = 200 \\ S_4 = 50 \end{cases}$$

Les Figures (b) et (c) illustrent les niveaux de gris et aires des pixels intervenant lors de l'estimation de l'intégrale de Choquet. On peut remarquer que notre noyau non-sommatif n'est pas entièrement recouvert par nos cellules floues. Il faudra donc normaliser les résultats pour estimer correctement l'intervalle de valeurs obtenu.

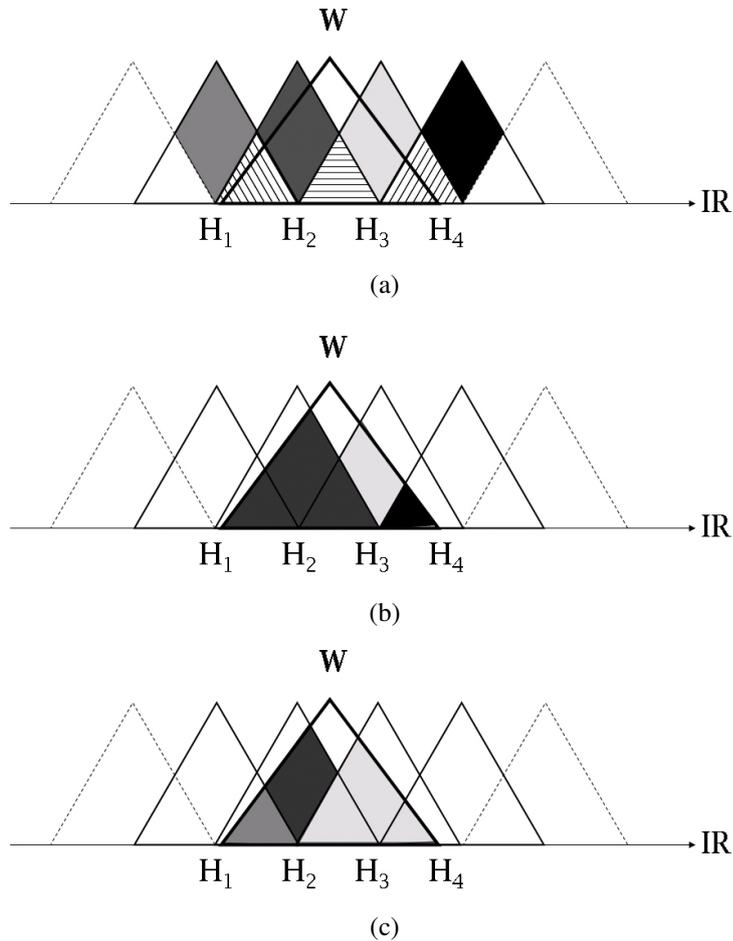


FIG. 2.14 – (a) Application d'un noyau non-sommatif W à un ensemble de pixels. (b) Schématisation des résultats obtenus l'intégrale de Choquet inférieure. (c) Schématisation des résultats obtenus l'intégrale de Choquet supérieure.

2.3.2.2 Vision horizontale

L'ensemble flou normalisé que constitue une distribution de possibilité peut être vu par ses α -coupes comme un ensemble d'intervalles de confiance. Une distribution de possibilité peut donc être construite à partir de la connaissance d'un ou plusieurs de ces intervalles.

$$\nu_W(A_{(n)}) = \text{Sup}_{i=n..N} \Pi(H_{(i)}, W) \quad (2.38)$$

2.3. Intégrale de Choquet

Les valeurs des $\Pi(H_{(i)}, W)$ sont les hauteurs d'intersection des $H_{(i)}$ avec W . La mesure d'interaction Π est la possibilité que l'ensemble flou des éléments de \mathbb{R} appartienne à la fois à $H_{(i)}$ et W . Elle peut être vue comme la plus grande valeur de α telle que les coupes de niveau de W et de $H_{(i)}$ aient une intersection non-nulle. On note β_i la matrice d'interaction $N \times N$ des $(\Pi(H_{(i)}, W))_{i=1..N}$. Elle peut être estimée simplement par dichotomie .

Reprenons l'exemple de nos deux cellules H_1 et H_2 (Figure 2.15) de valeurs $S_1 = 90$ et $S_2 = 180$.

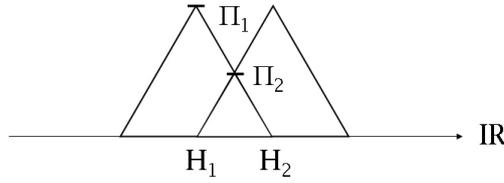


FIG. 2.15 – Les interactions entre deux cellules dépendent des hauteurs d'intersection entre ces cellules.

On a :

$$\begin{cases} \Pi(H_1) = \Pi(H_2) = 1 \\ \Pi(H_1, H_2) = \frac{1}{2} \end{cases}$$

L'estimation supérieure de l'intégrale de Choquet vaut :

$$\begin{aligned} \mathbb{C}_{\nu_W}(S) &= \sum_{n=1}^2 S_{(n)}[\nu_{H_1}(A_{(n)}) - \nu_{H_1}(A_{(n+1)})] & (2.39) \\ &= S_{(1)}[Sup_{i=\{1,2\}}\Pi(H_i, H_1) - Sup_{i=2}\Pi(H_i, H_1)] + S_{(2)}[Sup_{i=2}\Pi(H_i, H_1) - 0] \\ &= S_{(1)}[\Pi(H_1, H_1) - \Pi(H_1, H_2)] + S_{(2)}[\Pi(H_1, H_2) - 0] \\ &= 90[1 - \frac{1}{2}] + 180[\frac{1}{2} - 0] \\ &= 135 \end{aligned}$$

et l'estimation inférieure :

$$\begin{aligned} \mathbb{C}_{\nu_W^c}(S) &= - \sum_{n=1}^2 -S_{(n)}[\nu_{H_1}(A_{(n)}) - \nu_{H_1}(A_{(n+1)})] & (2.40) \\ &= S_{(1)}[Sup_{i=\{1,2\}}\Pi(H_i, H_1) - Sup_{i=2}\Pi(H_i, H_1)] + S_{(2)}[Sup_{i=2}\Pi(H_i, H_1) - 0] \\ &= S_{(1)}[\Pi(H_1, H_1) - \Pi(H_1, H_1)] + S_{(2)}[\Pi(H_1, H_1) - 0] \\ &= 180[1 - 1] + 90[1 - 0] \\ &= 90 \end{aligned}$$

On obtient donc $H_1 \in [90, 135]$. De la même manière, $H_2 \in [135, 180]$.

Reprenons à présent le noyau non-sommatif W représenté Figure 2.11. L'estimation de l'intégrale de Choquet sera effectuée à partir des hauteurs d'intersection du noyau avec les cellules comme l'illustre la Figure 2.16. De la même manière que précédemment, si le mode de notre noyau ne coïncide pas avec le mode d'une cellule, il devra être normalisé.

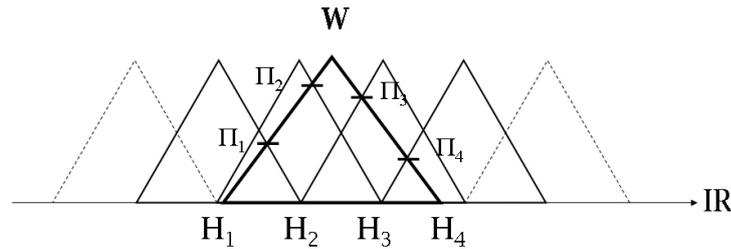


FIG. 2.16 – Hauteur d'intersection des cellules avec le noyau W .

Conclusion

Dans ce chapitre, nous avons présenté les principales étapes d'acquisition d'une image numérique. Nous avons mis en évidence la présence d'imprécisions géographiques sur les mesures de luminance des pixels de l'image. Pour prendre en compte ces imprécisions, une étude nous a conduit à une représentation possibiliste des données. Cette représentation est liée à la représentation probabiliste classiquement utilisée par le fait qu'un noyau possibiliste triangulaire domine une famille de noyaux probabilistes de même mode et même support. Cependant, les outils de traitement des données probabilistes ne peuvent être utilisés avec des données possibilistes. Pour estimer l'espérance de l'image, il faut utiliser une espérance plus générale, l'intégrale de Choquet. Cette intégrale fournit une estimation du noyau sous la forme d'un intervalle représentatif de l'ensemble des valeurs obtenues avec des noyaux sommatifs de même mode et même support que notre noyau non-sommatif.

Chapitre 3

Adaptation des masques de convolution aux images omnidirectionnelles

Introduction

Une image est un signal continu en deux dimensions. Comme en traitement du signal, les traitements d'images ont été définis dans le domaine continu. Cette représentation permet de décrire les phénomènes physiques. En revanche, pour le traitement de données numérisées, les traitements doivent être discrétisés. Ainsi, les noyaux de convolution, permettant de traiter les images, ont été échantillonnés régulièrement de manière coïncider avec les pixels de l'image numérique. Les masques ainsi obtenus vérifient la propriété d'invariance par translation. Ces opérateurs cartésiens sont convolués avec la partition cartésienne de l'image. Nous avons expliqué précédemment que la résolution des images catadioptriques, n'est pas la même en tout point de l'image. Ces traitements ne sont donc pas adaptés aux images catadioptriques. Dans ce chapitre nous proposons deux méthodes d'adaptation des traitements aux images catadioptriques. Dans un premier temps, nous expliciterons ce qu'est un masque de convolution et la manière de l'appliquer aux images perspectives. Ensuite nous présenterons chacune des deux approches d'adaptation des masques de convolution aux images catadioptriques. La première consiste à utiliser un modèle projectif pour définir les masques sur l'image catadioptriques. La seconde utilise les noyaux non-sommatifs. Un modèle de pixel flou permet prendre en compte les imprécisions liées à l'échantillonnage spatial de l'image. Finalement, une conclusion fait une synthèse sur les apports et les limites de ces approches.

3.1 Masques de convolution

Comme nous l'avons vu Section 2.2.1, le produit de convolution est fréquemment utilisé en traitement du signal pour effectuer des opérations de filtrage de l'image. Le signal de l'image étant un signal bidimensionnel, l'équation de convolution 2.7 doit être étendue en 2D pour être appliquée

à l'image. Pour un signal continu 2D $S(x, y)$ et un noyau sommatif 2D $\kappa(x, y)$, le signal de sortie $\mathbb{E}_\kappa(S)$ sera :

$$S(x, y) * \kappa(x, y) = \int_{\Omega_1} \int_{\Omega_2} S(x, y) \kappa(x, y) dx dy. \quad (3.1)$$

Dans le cas des signaux numériques tels que les images numériques, cette opération de convolution s'effectue de façon discrète. En pratique, la convolution numérique d'une image se fera simplement par une sommation de multiplications. La convolution discrète permet de construire des filtres linéaires, invariants par translation.

Si on considère un filtre linéaire $(2N + 1) \times (2N + 1)$, le niveau de gris $F_{i,j}$ du pixel filtré de coordonnées (i, j) est donné par la formule :

$$F_{i,j} = \sum_{u=-N}^N \sum_{v=-N}^N C_{u,v} I_{i+u,j+v} \quad (3.2)$$

où $C_{u,v}$ est la valeur du coefficient du masque à la coordonnée (u, v) et $I_{i+u,j+v}$ le niveau de gris du pixel de l'image à la coordonnée $(i + u, j + v)$. Le produit de convolution de l'image par le masque remplace chaque pixel par une somme pondérée de ses voisins.

Les masques de convolution les plus fréquemment utilisés sont les masques de lissage (filtre moyenneur, gaussien, etc.) et les masques de dérivation (gradient, laplacien, etc.).

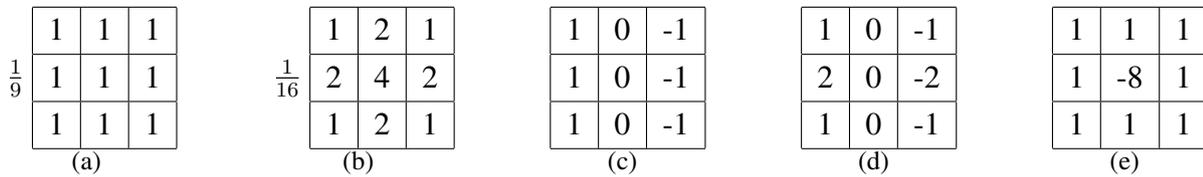


FIG. 3.1 – Les filtres (a) Moyenneur, (b) Gaussien, (c) Prewitt, (d) Sobel, (e) Laplacien

La taille de ces noyaux peut être paire ou impaire, petite ou grande. Plus la taille de la fenêtre est grande, moins le signal filtré contiendra de bruit et plus le signal sera flou, lissé et le niveau de détail réduit. Par conséquent, pour des images de taille raisonnable (on exclut les images très haute résolution), on choisit généralement des fenêtres de petites tailles, 3×3 ou 5×5 , même si elles sont plus sensibles au bruit. Dans la suite de ce document, nous présenterons nos méthodes avec des noyaux 3×3 .

Ces filtres sont souvent séparables, c'est à dire qu'un filtre de réponse impulsionnelle $C(x, y)$ séparable en x et y vérifie $C(x, y) = C_x(x)C_y(y)$. Lors du filtrage d'image, on a donc $I'(x, y) = C(x, y) * I(x, y) = C_y(y) * (C_x(x) * I(x, y))$. La séparabilité permet de ramener le problème de filtrage d'un signal bidimensionnel à celui du filtrage d'un signal monodimensionnel, de réduire le temps de calcul et de donner la possibilité d'une implémentation récursive. Par exemple, le filtre de

Sobel peut être décomposé suivant l'Equation 3.3 :

$$\begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 \end{bmatrix} * \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}. \quad (3.3)$$

Pour éviter de modifier l'entropie de l'image, la somme des coefficients doit être égale à

$$\begin{cases} 1 \text{ pour les filtres de lissage} \\ 0 \text{ pour les filtres de dérivation} \end{cases}$$

La convolution d'une image avec un masque linéaire est décrite par l'algorithme 1.

Algorithme 1 Algorithme de convolution d'une image avec un masque linéaire

Pour Tout pixels $P_{i,j}$ de l'image numérique **Faire**

Positionner le centre du masque sur le pixel à filtrer.

Calculer le produit de convolution de l'image avec ce masque.

Donner au pixel filtré la valeur obtenue.

Fin Pour

Normaliser l'image filtrée.

Comme nous l'avons montré dans le chapitre, les masques de convolution, invariants par translation, ne sont pas appropriés pour traiter les images omnidirectionnelles. Nous avons donc proposé une méthode d'adaptation de ces filtres aux images acquises avec un capteur catadioptrique central.

3.2 Algorithme de projection

Les systèmes catadioptriques sont constitués d'une caméra couplée avec un miroir catadioptrique. Comme nous l'avons évoqué dans l'introduction générale, les caméras perspectives fournissent un échantillonnage uniforme de la scène observée. Cependant, la projection de la scène sur le miroir introduit des déformations radiales importantes. La résolution de la scène observée n'est plus uniforme. Pour pallier cette non-uniformité, une solution possible consiste à utiliser une caméra C.C.D. munie d'un capteur conçu spécialement pour les caméras omnidirectionnelles. Cependant un tel capteur est très coûteux et par conséquent peu utilisé en vision omnidirectionnelle. La plupart des systèmes catadioptriques sont munis de caméras perspectives. Les images omnidirectionnelles acquises avec ces caméras ont donc une résolution non-uniforme. Lors de la mesure spatiale de la luminance, la quantité d'information intégrée n'est pas la même pour chaque pixel. Les pixels plus proches du centre de l'image intègrent une plus grande quantité d'informations que les pixels situés

à l'extérieur de l'image. Cette résolution non-uniforme ajoute donc des imprécisions sur la valeur d'illumination obtenue pour chaque pixel de l'image omnidirectionnelle.

Pour traiter l'image catadioptrique, on a intuitivement l'idée de projeter l'image sur un support permettant une vision perspective de la scène ou une vision s'en approchant comme le cylindre par exemple. Les traitements classiques deviennent alors applicables à cette image. Cependant, cette opération de projection nécessite d'échantillonner le support et de calculer l'intensité en chacun de ses points. Cet échantillonnage est arbitraire et risque de modifier le contenu ou la nature de l'image. Cette étape de désanamorphose contribuerait alors à ajouter des imprécisions dans l'image. Pour rester le plus fidèle possible au signal original, nous avons choisi de travailler dans l'espace de l'image catadioptrique afin de modifier le moins possible son contenu. Par analogie, on choisit donc de projeter l'opérateur sur l'image omnidirectionnelle. Cette contrainte permet de s'affranchir des phénomènes d'interpolation engendrés par la reconstruction d'une image sur un autre support. Cependant, elle nécessite une adaptation des opérateurs de filtrage à la géométrie de l'image. La forme des noyaux de convolution doit être modifiée pour les appliquer à l'image omnidirectionnelle. Pour construire ces noyaux dans l'espace de l'image, nous avons choisi d'associer le modèle de projection unifié et l'espace du cylindre englobant. Le centre de chaque pixel de l'image omnidirectionnelle est projeté sur le cylindre via la sphère virtuelle (Figure 3.2(a)). Un noyau de convolution régulier est défini géométriquement autour du point projeté. La Figure 3.2(b) présente l'exemple d'un masque de convolution 3×3 . Ce sont ces noyaux qui ont été utilisés dans la suite de nos travaux. Cependant, il aurait tout à fait été possible de choisir un noyau de taille différente, tant que sa taille reste relativement petite par rapport à la taille de l'image. Le noyau est ensuite rétro-projeté sur l'image panoramique (Figure 3.2(b)). A chaque élément du noyau projeté est associée une valeur de niveau de gris. Cette valeur est estimée en fonction des valeurs des niveaux de gris des pixels intersectés par l'élément projeté. Le calcul du pixel filtré est effectué en fonction des coefficients du noyau et des niveaux de gris associés. Cette approche est décrite par l'Algorithme 2.

Algorithme 2 Algorithme de projection des noyaux

Pour Tout pixels $P_{i,j}$ de l'image omnidirectionnelle **Faire**

Projeter le centre P du pixel sur le cylindre.

Calculer les coordonnées des noyaux sur le cylindre autour de le centre projeté Q du pixel.

Rétro-projeter sur l'image omnidirectionnelle les noyaux obtenus.

Estimer la valeur du gradient du pixel $P_{i,j}$.

Fin Pour

Un avantage de cet algorithme est que notre espace projectif est utilisé uniquement pour définir les coordonnées géométriques des noyaux projetés dans l'image omnidirectionnelle. L'étape de rétro-projection est donc pré-calculable, ce qui permet de réduire d'autant le temps de calcul.

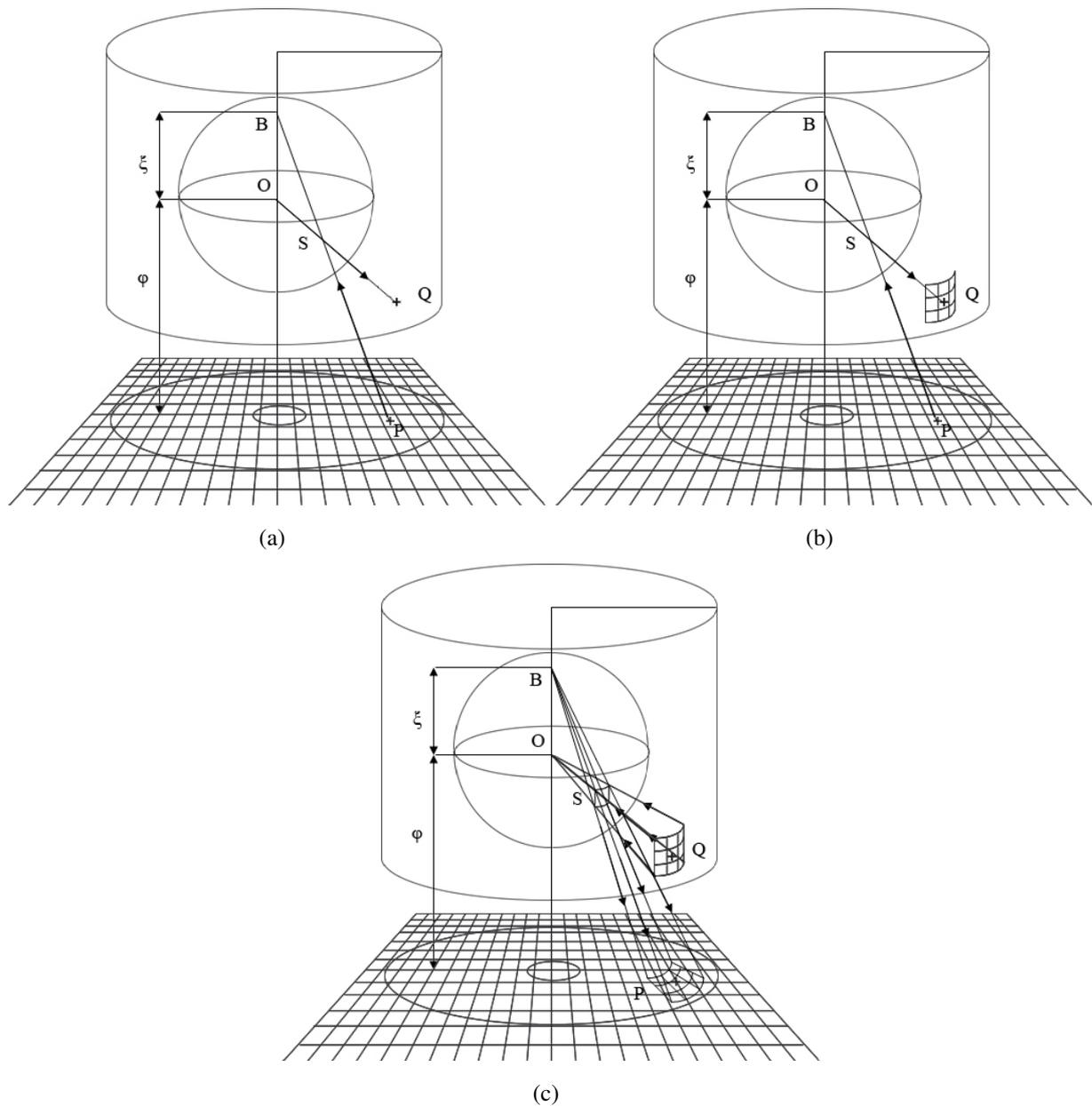


FIG. 3.2 – (a) Projection du centre du pixel sur le cylindre. (b) Reconstruction d'un masque régulier autour du centre du pixel projeté. (b) Projection du masque sur l'image omnidirectionnelle.

Trois questions se posent :

Pourquoi utiliser ce modèle projectif ? Avec le modèle de projection unifié présenté Section 1.1.4.1.5, notre modèle est généralisable à l'ensemble des capteurs catadioptriques centraux. L'utilisation du cylindre est plus discutable. En effet, pourquoi ne pas travailler directement dans l'espace de l'image perspective ? Si nous conservons le même algorithme, à chaque centre de pixel de l'image omnidirectionnelle, on associe une image perspective sur laquelle le noyau est défini. Cette approche nécessite donc de définir autant d'images perspectives qu'il y a de pixels dans l'image omnidirectionnelle. Elle

est donc relativement lourde à implémenter. L'espace du cylindre permet, à lui seul, de représenter l'ensemble des noyaux de l'image. De plus, avec une caméra perspective, la scène acquise est supposée invariante par translation. Cela implique que la distance entre la scène et la caméra est constante. Avec une caméra omnidirectionnelle, la scène observée est à une distance constante d'une caméra ayant tourné sur elle-même. Par analogie, on retrouve l'espace du cylindre. L'espace projectif du cylindre englobant peut être considéré comme la collection de bandes centrales des images projectives reconstruites localement en tournant autour de l'axe de miroir (Figure 3.3). Comme mentionné Section 1.2, cet espace est relativement simple à échantillonner avec une grille cartésienne dans un repère cylindrique. Localement, dans le voisinage du filtre qui est petit (de l'ordre de 3×3 pixels),

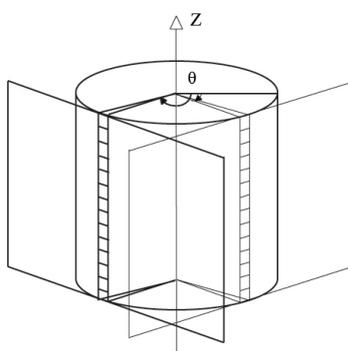


FIG. 3.3 – Quasi-équivalence entre la projection d'un pixel sur le cylindre et sur l'image perspective

il se comporte de la même manière qu'une image perspective. De plus, l'invariance par translation est conservée lors du traitement de l'image catadioptrique. Pour finir, nous avons estimé l'erreur maximale sur le calcul des coordonnées des éléments du masque de convolution afin de nous en convaincre. Cette erreur est de l'ordre de 10^{-5} pixels.

Comment définir le masque sur le cylindre ? La solution présentée consiste à définir le masque centré sur la projection du centre du pixel considéré. Cela permet de garantir qu'à chaque pixel de l'image, on associe un masque. Une étude a été réalisée dans la Section 5.4.2 pour montrer l'influence de la taille du noyau de convolution sur les contours détectés.

A l'aide de l'Algorithme 2, nous avons déterminé la géométrie des masques de convolution 3×3 dans l'image omnidirectionnelle et les avons représentés Figure 3.4. Sur cette figure, les masques ont été volontairement grossis 50 fois afin de mettre en évidence leurs formes et les déformations introduites par le miroir.

Considérons un des masques projetés sur lequel nous avons zoomé Figure 3.5. Chaque élément du masque recouvre plusieurs pixels de l'image omnidirectionnelle. Le but est de transférer l'information des pixels intersectés de l'image omnidirectionnelle vers un élément du masque projeté.

Comment calculer le niveau de gris associé au pixel projeté ? La valeur du niveau de gris $I_{i+u,j+v}$

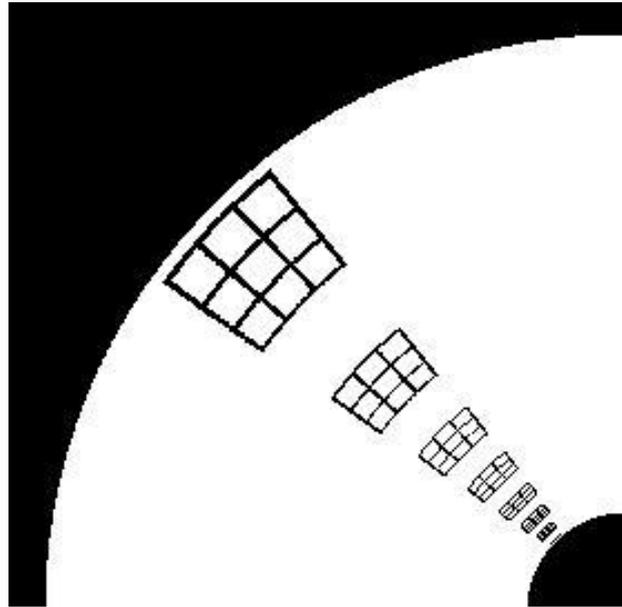


FIG. 3.4 – Projection des masques 3×3 sur l'image omnidirectionnelle.

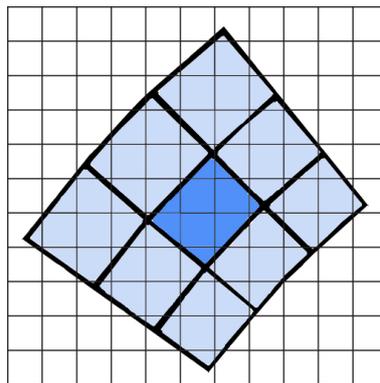


FIG. 3.5 – Le masque projeté couvre plusieurs pixels de l'image omnidirectionnelle.

de l'équation (3.2) doit être modifiée.

La solution la moins spécifique consiste à utiliser l'enveloppe englobante de chaque élément du masque projeté comme représentation discrète du noyau et la solution la plus spécifique, à ne considérer que les pixels inclus strictement dans chaque élément. La Figure 3.6 illustre le degré d'appartenance des pixels à l'élément projeté en fonction de la surface intersectée. L'ensemble des pixels grisés représente l'enveloppe englobante de l'élément projeté du masque. Les pixels gris foncés représentent la solution la plus spécifique. Si on souhaite disposer d'un modèle plus fin, il faut adopter une mesure de compatibilité entre les éléments du masque projeté et chaque pixel de l'image. Deux solutions sont proposées dans les sections 3.3 et 3.4 pour comparer les niveaux de gris obtenus en projetant les éléments du masque.

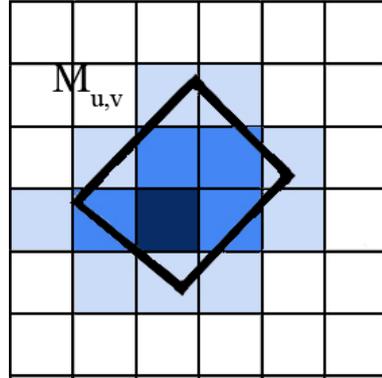


FIG. 3.6 – Interactions entre un élément du masque projeté et les pixels de l’image.

3.3 Approche proportionnelle

La première solution s’inspire du transfert de croyance : le transfert pignistique [58]. Chaque élément du pixel et du masque est considéré comme un intervalle 2D défini par son support et avec une fonction d’appartenance uniforme. La valeur de niveau gris d’un élément du masque est une somme pondérée de tous les niveaux gris des pixels intersectés. Le poids est proportionnel à l’aire d’intersection du pixel de l’image omnidirectionnelle avec le masque cylindrique projeté. La Figure 3.7 illustre un zoom sur un élément du masque. Le poids de la valeur du pixel foncé est proportionnel à sa surface d’intersection avec l’élément du masque considéré.

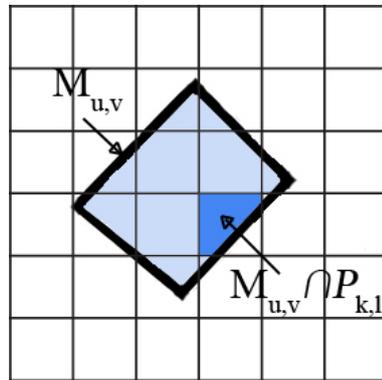


FIG. 3.7 – Zoom sur l’élément central du masque : le niveau de gris du noyau projeté est proportionnelle à la surface des pixels intersectés.

L’estimation du niveau de gris du pixel $I_{i+u,j+v}$ de l’équation (3.2) doit être remplacé par :

$$I_{i+u,j+v} \Leftrightarrow \sum_{\Omega(k,l)} \frac{|M_{u,v} \cap P_{k,l}|}{|M_{u,v}|} I_{k,l} \quad (3.4)$$

où (i, j) sont les coordonnées du pixel filtré dans l’image omnidirectionnelle, $M_{u,v}$ est l’intervalle 2D associé à la projection d’un élément (u, v) du masque sur l’image omnidirectionnelle, $\Omega_{k,l}$ (région

claire sur la Figure 3.5) est l'ensemble de tous les pixels $P_{k,l}$ intersectés par $M_{u,v}$ et $I_{k,l}$ est le niveau de gris du pixel $P_{k,l}$.

Pour simplifier le calcul de la région d'intersection, une approximation de la projection du masque est utilisée. La surface est approximée par un trapèze. En effet, l'accroissement du temps de calcul apporté par la prise en compte de la forme exacte du masque projeté est non négligeable. Dans le cas de l'approche proportionnelle, on doit calculer les aires d'intersection entre des parallélépipèdes et des formes géométriques correspondant à l'intersection de cercles et de polygones. Les calculs sur les arcs de cercles sont longs, difficiles à mettre en oeuvre et augmentent d'autant la complexité des algorithmes. De plus, l'approximation effectuée sur l'arc de cercle supérieur provoque une erreur de 0.605327% sur la valeur du niveau de gris et l'arc de cercle inférieure, une erreur de 0.604925%. Cela fait une erreur maximale de inférieure à 1,22% sur le niveau de gris du pixel soit 3 niveaux de gris si cette valeur est codée entre 0 et 255.

Pour visualiser cette approximation qui est vraiment infime dans la réalité, nous avons schématisé le masque de convolution obtenu Figure 3.8. Dans ce schéma, nous avons volontairement accentué les incurvations des éléments projetés de manière à mettre en évidence l'approximation effectuée. La Figure 3.8 illustre cette approximation.

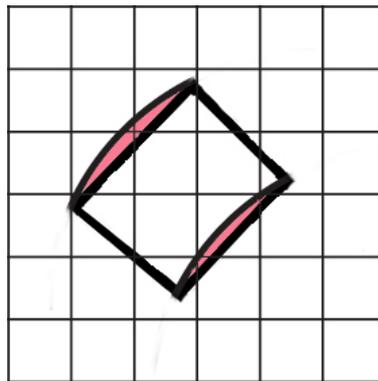


FIG. 3.8 – Simplification du masque approximé un quadrilatère.

3.4 Approche floue

L'approche probabiliste fournit une mesure de la compatibilité entre un pixel et un domaine de l'espace en intégrant le noyau sommatif sur l'élément du masque projeté. Cependant, le choix de la distribution de probabilité influe considérablement sur les résultats. Les travaux réalisés par Coquerez, Philipp, Bolon, Chassery, Demigny, Graffigne, Montanvert, Zéboudj et Zérubia [28] sur l'estimation des gradients d'illumination en attestent.

Dans cette approche, nous proposons une autre manière de définir la compatibilité entre un pixel et l'élément du masque projeté utilisant une modélisation floue des pixels présentée Section 2.1.2.2.

En appliquant l'Algorithme 2 de projection des opérateurs, chaque centre de pixel de l'image omnidirectionnelle (noyau du pixel flou 2D) est projeté sur le cylindre. Un masque de convolution est défini autour de ce point projeté puis est rétro-projeté sur l'image omnidirectionnelle. Ce masque est également une quantité échantillonnée. Par conséquent, il y a une imprécision sur la localisation des coefficients du masque. Chacun de ces éléments peut être modélisé par un nombre flou 2D. La Figure 3.9 illustre le masque de convolution défini sur le cylindre. Pour simplifier les illustrations, seul l'élément central flou du masque est représenté.

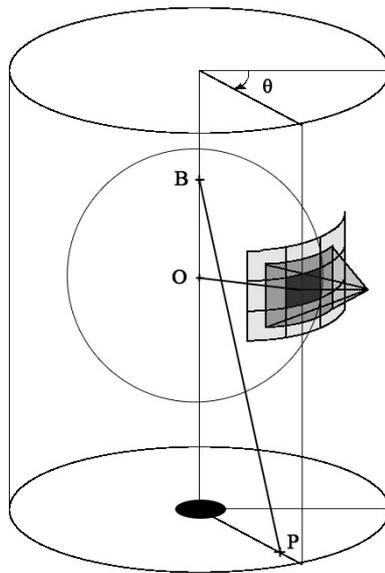


FIG. 3.9 – Représentation floue de l'élément central du masque sur le cylindre.

Les nombres flous associés au masque sont ensuite projetés sur l'image omnidirectionnelle comme le montre la Figure 3.10.

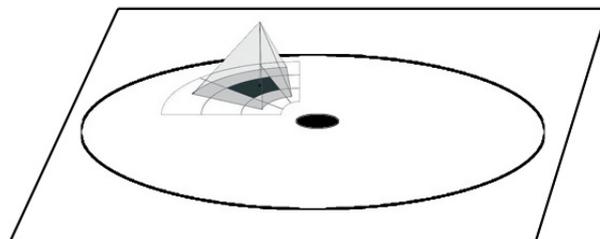


FIG. 3.10 – Projection de l'élément central du masque sur l'image omnidirectionnelle.

Comme dans la section 3.3 la projection de chaque élément flou du masque couvre plusieurs pixels flous de l'image omnidirectionnelle. La question est maintenant comment calculer le niveau de gris du pixel filtré. La solution proposée consiste à employer l'intégrale de Choquet présentée Section

2.3 pour en évaluer une borne supérieure et inférieure :

$$\mathbb{C}_{\nu_W}(S) = \sum_{n=1}^N S_{(n)}[\nu_W(A_{(n)}) - \nu_W(A_{(n+1)})], \quad (3.5)$$

$$\mathbb{C}_{\nu_W^c}(S) = -\mathbb{C}_{\nu_W}(-S) = -\sum_{n=1}^N (-S_{(n)})[\nu_W(A_{(n)}) - \nu_W(A_{(n+1)})]. \quad (3.6)$$

L'équation (2.38) de la Section 2.3.2.2 ($\nu_W(A_{(n)}) = \text{Sup}_{i=n..N} \Pi(H_{(i)}, W)$) est étendue en 2D :

$$v(A_n) = \text{Sup}_{i=n..N} \Pi(P_i, M_{u,v}) \quad (3.7)$$

où $M_{u,v}$ est la projection des éléments du masque flou, les $(P_i)_{i=1..N}$ représentent les pixels flous intersectés par $M_{u,v}$. Les $A_n = P_n, \dots, P_N$ sont les coalitions 2D binaires des pixels dont les niveaux gris sont supérieurs ou égaux à I_n . La plus grande valeur de confiance qui peut être donnée à la coalition A_n est $v(A_n)$ définie dans (3.7). Dans les équations (2.31) et (2.32), I_i indique le niveau gris du pixel flou P_i , $(.)$ indique une permutation telle que $I_1 \leq I_2 \leq \dots \leq I_N$ pour (2.31) et $-I_1 \leq -I_2 \leq \dots \leq -I_N$ pour (2.32). Les valeurs des $\Pi(P_i, M_{u,v})$ sont les hauteurs d'intersections des P_i avec $M_{u,v}$ comme illustré Figure 3.11. Le calcul de $\Pi(P_i, M_{u,v})$ est effectué par dichotomie.

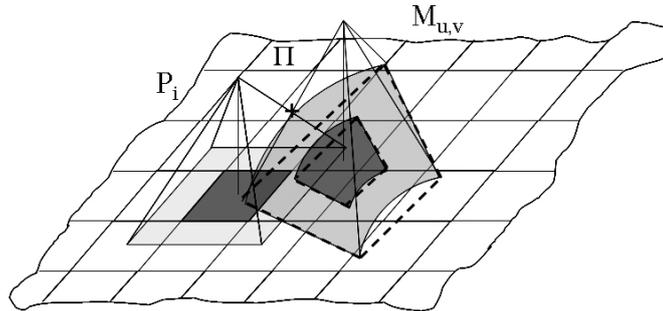


FIG. 3.11 – Calcul de $\Pi(P_i, M_{u,v})$ comme la hauteur d'intersection entre P_i et $M_{u,v}$.

Une fois que le niveau de gris de chaque élément du masque évalué, il faut calculer le niveau gris du pixel filtré. Nous avons envisagé deux solutions.

La première consiste à appliquer directement l'équation (3.2) d'estimation de la compatibilité entre le pixel et l'élément projeté du masque avec les valeurs des niveaux gris associées aux éléments du masque. Les images obtenues correspondent à une estimation supérieure et inférieure de l'image filtrée. Ces images présentent des différences au niveau de l'intensité des niveaux de gris des pixels. Pour représenter l'image filtrée, nous avons réalisé une moyenne des deux images.

La seconde revient à considérer que les éléments flous du masque se chevauchent et par conséquent interagissent entre eux. Le niveau de gris associé au masque $M_{u,v}$ n'est donc pas connu précisément. L'intégrale de Choquet sur chaque élément flou du masque permet alors de définir une borne supérieure et inférieure de ce niveau gris. Ainsi, le niveau de gris de chaque pixel filtré sera également

caractérisé par un intervalle. Le calcul du niveau gris du masque est relativement simple étant donné que la possibilité représentant l'interaction de deux cellules contiguës du masque est toujours égale à 0.5.

En effet, nos nombres triangulaires flous sont tels que les segments de droite de deux éléments contiguës ont une pente inverse et s'intersectent en une valeur égale à 0.5 comme illustré Figure 3.12.

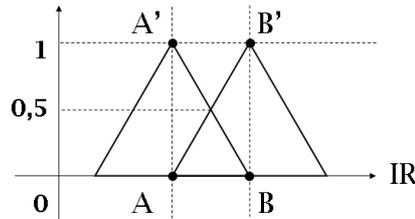


FIG. 3.12 – Hauteur d'intersection entre les nombres flous.

Lorsque ces nombres flous, définis sur le cylindre, sont projetés sur l'image omnidirectionnelle, les origines et extrémités de ces nombres seront les mêmes dans le plan de l'image. Nous conservons donc la même hauteur d'intersection entre les éléments du masques :

$$\Pi(M_{u,v}; M_{u',v'}) = 0.5 \quad (3.8)$$

avec

$$\begin{cases} u' = u \pm k \\ v' = v \pm k' \\ k, k' \in \{-1, 0, 1\}. \end{cases}$$

La Figure 3.13 représente les interactions entre 3 éléments du masque projeté.

Conclusion

Ce chapitre présente une solution d'adaptation des masques de convolution aux images omnidirectionnelles, respectant la géométrie de l'image. L'opérateur de filtrage est défini géométriquement sur un espace régulier, l'espace du cylindre englobant le miroir du révolution et retro-projeté sur l'image omnidirectionnelle. Cette méthode évite de projeter les pixels et leurs niveaux gris dans un autre espace et de ce fait d'approximer les données. La première solution s'inspire du transfert de croyance pignistique. Il s'agit d'une approche probabiliste qui fournit une mesure de la compatibilité entre un pixel et les éléments projetés du masque de convolution. La valeur de niveau gris d'un élément du masque est obtenue par une somme pondérée de tous les niveaux gris des pixels intersectés. Cependant, le choix de la distribution de probabilité influe considérablement sur les résultats. Nous

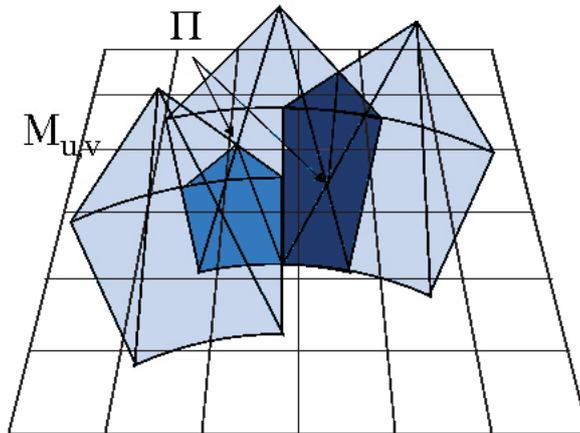


FIG. 3.13 – Interactions entre les éléments du masque sur l’image omnidirectionnelle.

avons donc proposé une seconde approche, une approche possibiliste. Des nombres flous triangulaires symétriques bidimensionnels ont été étendus en 3 dimensions pour modéliser l’imprécision sur la localisation des niveaux de gris des pixels, induite par l’échantillonnage spatial. Le masque de convolution est également une quantité échantillonnée. Ce masque est défini de manière à coïncider exactement avec les pixels de l’image. L’imprécision sur les valeurs des niveaux de gris des pixels induit donc une imprécision sur la localisation des coefficients du masque. Nous avons modélisé ces éléments de la même manière que les pixels par des nombres flous triangulaires symétriques bidimensionnels. Le pixel filtré est alors estimé à l’aide d’une intégrale de Choquet et fournit ainsi un intervalle de valeurs possibles du gradient estimé en chaque pixel.

La perception d'un objet comme désirable ou indésirable ne réside pas dans l'objet lui-même, mais dans la façon dont on le perçoit.

Aldous Huxley

Chapitre 4

Une nouvelle approche d'estimation de gradient non-additive

Introduction

La détection de contour est une étape préliminaire à de nombreuses applications de l'analyse d'images. Les contours sont caractérisés par des discontinuités de la fonction d'intensité de l'image. Un des principes de détection de contours repose sur l'étude des dérivées de cette fonction. Dans ce chapitre, nous nous sommes intéressés à la méthode de détection de contour par estimation du gradient de l'image. De nombreux opérateurs utilisant des noyaux sommatifs ont été proposés pour estimer ce gradient. Une question se pose : lequel choisir pour traiter au mieux nos images omnidirectionnelles ? En regardant de plus près les contours obtenus avec ces détecteurs, nous avons pu constater leurs faiblesses. Nous avons donc choisi de définir une nouvelle approche dérivative utilisant des noyaux non-sommatifs. Notre opérateur permet d'estimer le gradient sous forme d'intervalle qui fournit un seuillage naturel des valeurs estimées du gradient de l'image. Avant d'appliquer cette nouvelle approche aux images omnidirectionnelles, nous l'avons validée sur des images perspectives.

Ce chapitre expose, dans une première partie, ce qu'est la détection de contours et présente brièvement les principales approches mises en oeuvre pour cette opération. Ensuite, nous citerons quelques unes des difficultés rencontrées avec l'ensemble de ces approches pour éliminer le bruit présent dans ces images. Enfin, nous présenterons notre approche dérivative et expliquerons les avantages de cette méthode. Nous montrerons que cette approche peut être adaptée aux images omnidirectionnelles et simplifier les étapes d'adaptation de la détection de contours.

4.1 Détection de contours dans les images

4.1.1 Expression continue d'un contour

Comme évoqué dans l'introduction, la détection de contours consiste à rechercher les discontinuités de la fonction d'intensité de l'image. Le signal lumineux est un signal continu. Nous allons donc dans un premier temps expliciter ce qu'est un contour dans le domaine continu.

Considérons la fonction d'intensité continue de l'image $I(x, y)$.

Soit $\overrightarrow{\nabla I(x, y)}$ le gradient de $I(x, y)$.

$$\overrightarrow{\nabla I(x, y)} = \left[\frac{\delta I(x, y)}{\delta x}, \frac{\delta I(x, y)}{\delta y} \right]^t \quad (4.1)$$

Il est caractérisé par son module m et sa direction \vec{d} dans l'image :

$$m = |\overrightarrow{\nabla I(x, y)}| = \left[\left(\frac{\delta I(x, y)}{\delta x} \right)^2 + \left(\frac{\delta I(x, y)}{\delta y} \right)^2 \right]^{\frac{1}{2}} \quad (4.2)$$

$$\vec{d} = \frac{\overrightarrow{\nabla I(x, y)}}{|\overrightarrow{\nabla I(x, y)}|} \quad (4.3)$$

La direction peut aussi s'exprimer sous forme de l'angle formé avec le repère de l'image ϕ :

$$\phi = \text{Arctan} \left[\frac{\delta I(x, y)}{\delta y} / \frac{\delta I(x, y)}{\delta x} \right] \quad (4.4)$$

L'extraction de contours dans les images s'effectue généralement suivant trois étapes décrites Figure 4.1 : l'estimation de la dérivée en chaque pixel de l'image, l'extraction des maxima locaux qui permet d'affiner les contours obtenus et le seuillage afin d'éliminer les transitions les plus faibles assimilées au bruit.

4.1.2 Estimation du gradient de l'image

Comme on peut le voir sur la formule 4.1, le calcul du gradient fait intervenir la dérivée du signal image. Il s'agit de la dérivée d'un signal continu. Cependant, dans le cas numérique, l'image est de nature discrète. Il s'agit donc de trouver une approximation de la dérivée pour un signal discret. Plusieurs approches ont été envisagées ; les plus célèbres sont la dérivation par différences finies présentée Annexe B et la dérivation par optimisation de certains critères. On peut citer par exemple les critères de Canny de bonne détection, bonne localisation et réponse unique des contours, décrits dans l'Annexe C. En minimisant ces critères, on obtient une équation différentielle dont la solution fournit un filtre à réponse impulsionnelle finie. Suivant les critères de Canny, d'autres filtres ont été

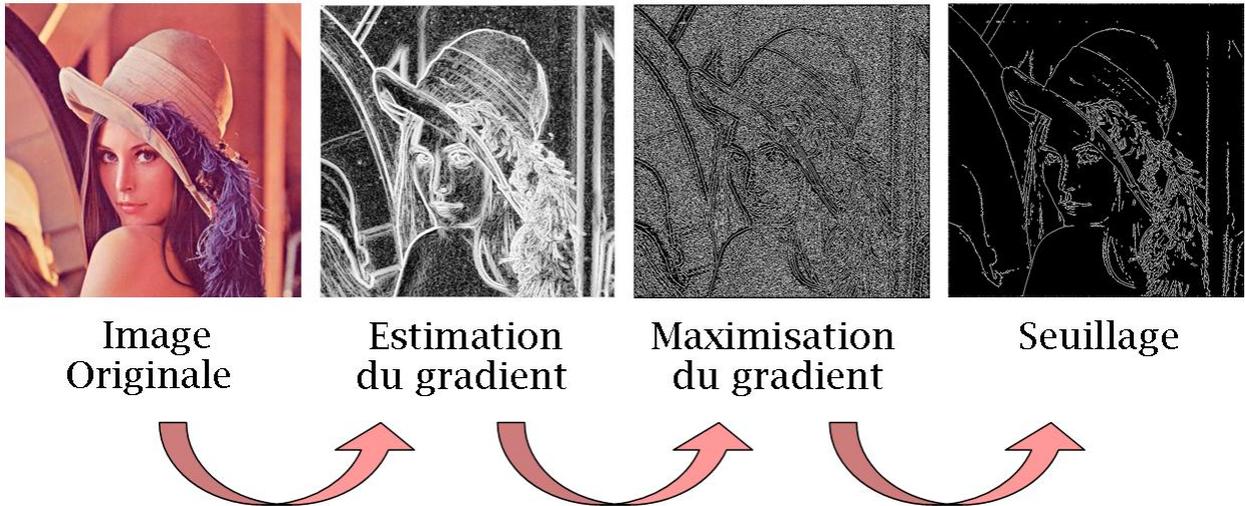


FIG. 4.1 – Etapes de détection de contours.

proposés comme celui de Deriche présenté en Annexe D ou de Shen-Castan présenté en Annexe E qui possèdent une réponse impulsionnelle infinie et sont implémentables récursivement.

Toutes ces approches fournissent une estimation de la valeur du gradient en chaque pixel de l'image. Cependant, comment déterminer si le pixel considéré est un point de contour ou non ?

La Figure 4.2(a) représente l'estimation du gradient sur l'image de Léna. Les contours correspondent aux valeurs des niveaux de gris les plus élevés de l'image du gradient. Si on passe cette image en 3D, on remarque que les contours correspondent aux crêtes sur l'image du gradient. Ces crêtes représentent les extréma locaux du gradient.

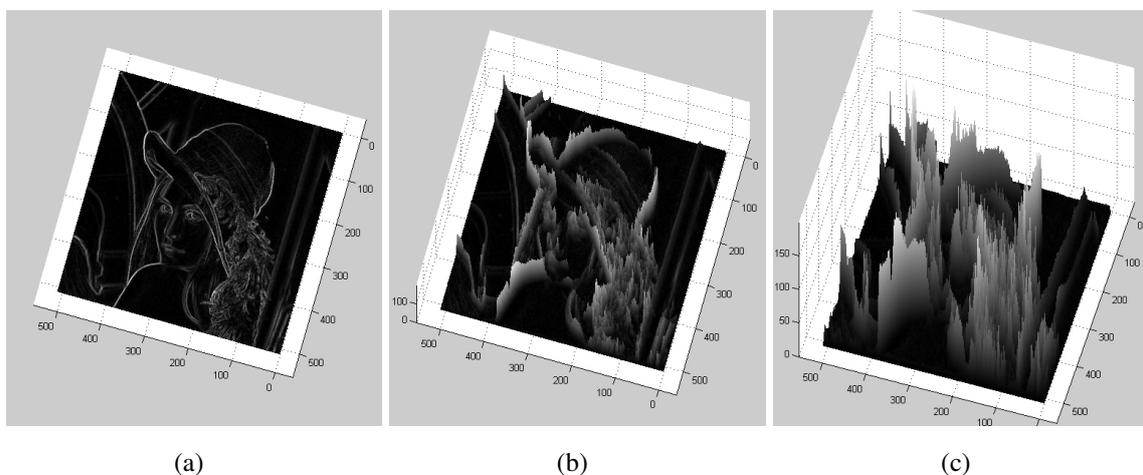


FIG. 4.2 – Vue 3D des contours présents dans l'image de Léna.

Pour déterminer efficacement un contour, l'estimation de gradient doit être suivie d'une étape d'extraction des maxima locaux de la norme du gradient. Cette étape permet d'affiner les contours détectés.

4.1.3 Maximisation de gradient

La détection de contour basée sur le gradient de l'image consiste à minimiser ou maximiser la dérivée première de cette fonction d'intensité. L'extraction des extréma locaux du gradient dans la direction du gradient revient à déterminer, pour un pixel $P_{i,j}$ donné, les valeurs du gradient sur la droite passant par $P_{i,j}$ dans la direction de son gradient (Figure 4.3). Il s'agit donc d'estimer la norme du gradient des deux voisins de $P_{i,j}$, P_A et P_B , situés dans la direction du gradient à une distance unitaire de part et d'autre de $P_{i,j}$, par interpolation linéaire par exemple. On vérifie ensuite que le gradient en $P_{i,j}$ est bien localement maximal sur cette droite, c'est à dire que :

$$\begin{cases} |P_{i,j}| > |P_A| \\ |P_{i,j}| > |P_B| \end{cases}$$

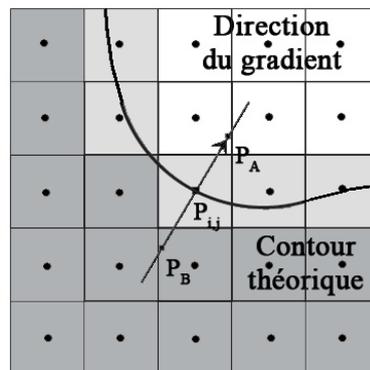


FIG. 4.3 – Maximisation de gradient : le pixel $P_{i,j}$ est un point de contour dans une image perspective si la norme de son gradient est maximum dans la direction du gradient.

Cependant, la qualité de la détection est très liée à la qualité des contours dans l'image. Souvent de bonne qualité, pour les applications de robotique en environnement artificiel ou les images à forte résolution, une dégradation est constatée pour les scènes d'extérieur ou les images médicales parfois très bruitée.

Ce bruit est lié le plus souvent à l'acquisition, la reconstruction, la numérisation ou aux textures de l'image. Il introduit des discontinuités indésirables dans la fonction d'intensité de l'image. Il est difficile de faire la différence entre ces contours liés au bruit et les contours utiles.

Reprenons l'image de Léna. Cette image contient du bruit provenant de son acquisition. Ce bruit a des répercussions lors de l'estimation des valeurs du gradient comme illustré Figure 4.4(a). Comme précédemment, les valeurs de gradient obtenues pour cette image ont été passées en 3D (Figure 4.4(b) et (c)). On peut remarquer la présence de nombreux pics dans la fonction d'intensité du gradient. Ces extrema parasitent la détection de contours. Pour supprimer ces valeurs liées au bruit, une solution consiste à utiliser un seuillage.

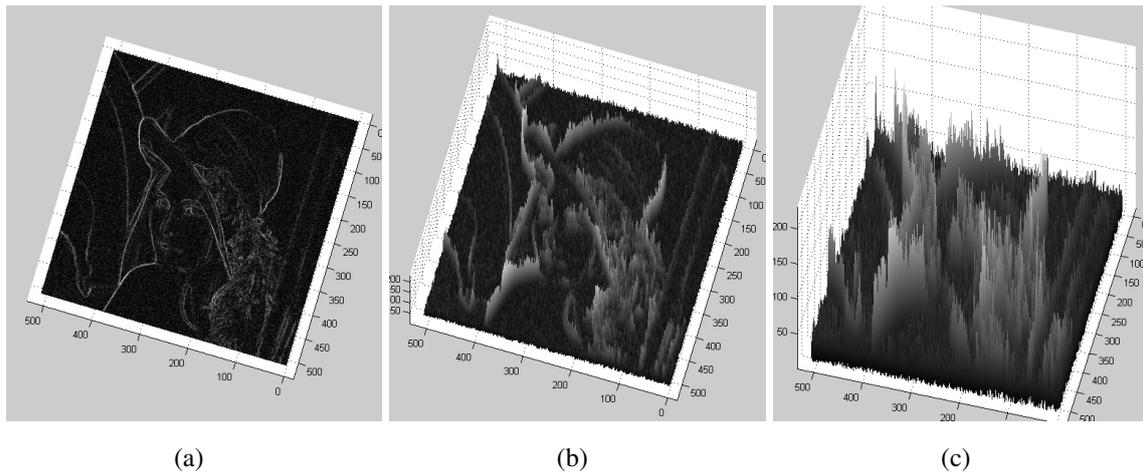


FIG. 4.4 – Vue 3D des contours présents dans l'image de Léna à laquelle on a ajouté un bruit gaussien.

4.1.4 Seuillage de l'image

Le seuillage est une étape fondamentale pour détecter les contours significatifs. Il est basé sur l'hypothèse que le bruit présent dans l'image est relativement faible par rapport au signal. L'opération dite de seuillage simple consiste à mettre à zéro tous les pixels ayant un niveau de gris inférieur à une certaine valeur et à la valeur maximale les pixels ayant une valeur supérieure. Ainsi le résultat du seuillage est une image binaire contenant des pixels noirs et blancs. En traitement d'images, on préfère souvent au seuillage simple, un seuillage par hystérésis utilisant deux valeurs de seuil. Le seuillage par hystérésis repose sur une hypothèse de connexité. Il permet d'éliminer les points de contours trop faibles et de ne conserver que les composantes connexes les plus importantes des contours. Le principe est d'utiliser deux seuils pour la norme du gradient : s_h et s_b . Les pixels dont la norme du gradient est supérieure à s_b sont sélectionnés. Si ces pixels sont connectés, par un chemin de pixels dont la norme est supérieure à s_b , à un pixel de la norme est supérieure à s_h , alors le contour est retenu.

Le choix des valeurs des seuils peut être automatisé, mais il est le plus souvent guidé par le jugement visuel. La Figure 4.5(a) représente les contours obtenus sur l'image de Léna avec un filtre de Deriche. Les valeurs des seuils ont été fixées de manière à obtenir le meilleur rapport entre les contours désirés et les contours non-désirés. Si les valeurs seuil choisies sont trop faibles, l'image contient un grand nombre de contours non-désirés (Figure 4.5(b)). Il s'agit en réalité de maxima locaux de faibles valeurs liés au bruit qui ont été assimilés à des contours. On obtient alors de faux-contours qui n'existent pas dans la scène réelle. Au contraire, lorsque les valeurs seuil sont trop importantes, certains contours présents dans l'image sont supprimés de l'image (Figure 4.5(c)). Plus le seuil est élevé, plus le risque de perdre les contours de l'image est important. De plus, pour des valeurs de seuil proches, il est souvent difficile de déterminer la meilleure détection (Figure 4.5(d)). La valeur de seuil des images est donc relativement difficile à estimer.

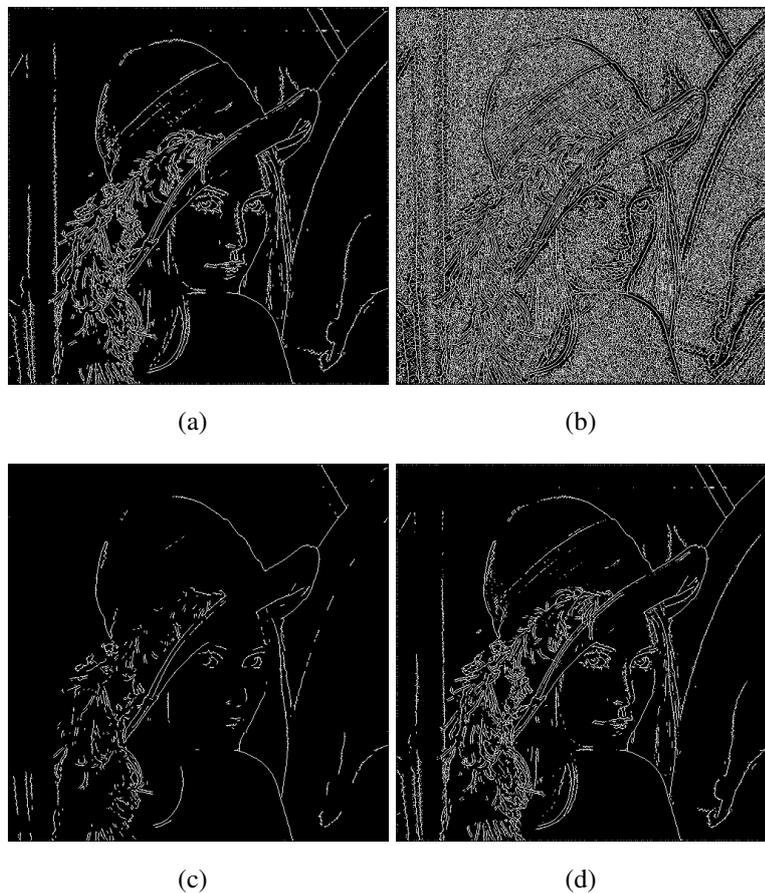


FIG. 4.5 – Detection de contours sur l'image de Léna pour différentes valeurs de seuil (a) $s_h = 29$ et $s_b = 26$, (b) $s_h = 0$ et $s_b = 0$, (c) $s_h = 60$ et $s_b = 55$, (d) $s_h = 32$ et $s_b = 30$.

Supposons que nous ayons réussi à trouver les meilleures valeurs de seuil pour l'image de Léna. Nous souhaitons à présent traiter d'autres images. Ces valeurs seront-elles également les meilleures valeurs de seuil pour ces autres images ? Et si non, existe-t-il des valeurs de seuil universelles qui

fonctionnent pour toutes les images ? Considérons une seconde image, celle des poivrons représentée Figure 4.6(a). Cette image est traitée comme l'image de Lena, avec un filtre de Deriche. La détection de contours est effectuée avec les mêmes seuils que l'image de Lena (Figure 4.5(a)). Sur cette image, nous souhaiterions détecter des contours fermés pour chaque poivrons et piments. Après traitement, l'image obtenue est représentée Figure 4.6(b). On remarque qu'une partie du contour du poivron central a été supprimé. Les valeurs de seuil sont donc trop élevées pour cette image. Le choix de valeurs plus faibles serait mieux adapté comme l'illustre la Figure 4.6(c). En effet, même si cette image contient quelques artéfacts, pour chaque végétal, on obtient un contour connexe et complet. Nous en concluons qu'il n'existe pas de bonnes valeurs de seuil, adaptées à toutes les images. Dans notre étude, nous avons constaté que les valeurs de seuil diffèrent d'une image à l'autre. Il est donc nécessaire de recalculer ces valeurs pour chaque image.

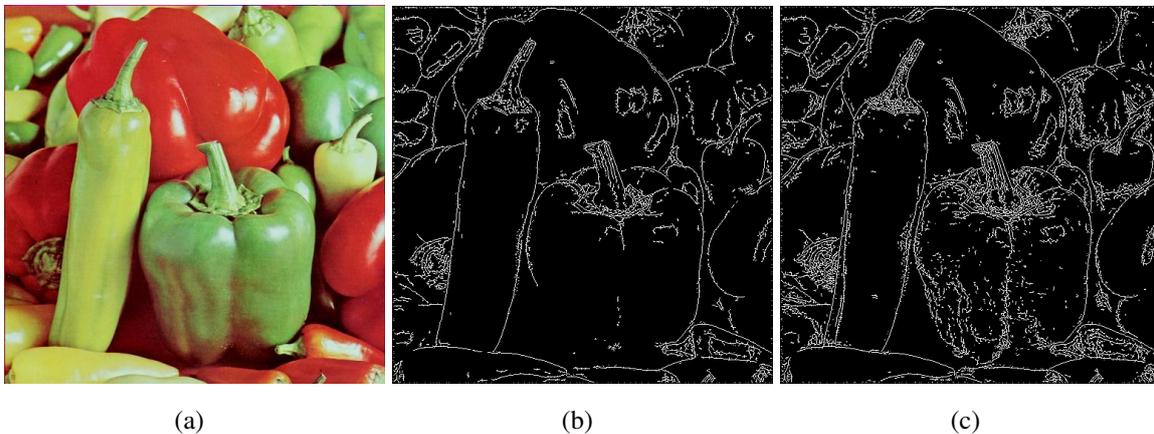


FIG. 4.6 – (a) Image originale des poivrons et détection de contours sur cette image avec (b) le même seuil que pour Léna $s_h = 29$ et $s_b = 26$ et (c) avec un seuil plus faible $s_h = 12$ et $s_b = 10$.

Un autre problème qui se pose, concerne le modèle de bruit présent dans les images. Habituellement, le bruit contenu dans les images est considéré comme uniforme. Ce bruit est indépendant de la position dans l'image et des valeurs des niveaux de gris des pixels. Cependant, dans certaines images, le bruit peut avoir un comportement non-uniforme. En imagerie médicale, par exemple, le bruit présent dans certaines image suit une distribution de Poisson. Lors de la détection de contours sur ces images, le seuillage par hystérésis échoue. Il est alors impossible de déterminer un seuil approprié à l'ensemble de l'image.

Afin d'illustrer nos propos, nous avons généré une image de synthèse présentée Figure 4.7(a). Cette image représente une rosace constituée de trois disques qui s'intersectent. Ces disques contiennent différents niveaux de gris. A cette image, nous avons ajouté un bruit de Poisson. Nous l'avons ensuite traité avec un filtre de Deriche et nous avons cherché une valeur de seuil permettant de détecter

complètement les contours de nos trois cercles. Notre objectif est atteint pour de faibles valeurs de seuil ($s_h = 10$ et $s_b = 8$). L'image (b) représente les contours obtenus pour ces valeurs de seuil. Cependant, si nos trois cercles sont correctement détectés, l'image contient également de nombreux contours non désirés. Ces contours sont liés uniquement au bruit de Poisson. Si on augmente les valeurs de seuil (Figures 4.7(c), (d) et (e)), ces artéfacts sont supprimés de l'image, mais les contours des disques le sont aussi. Le seuillage n'est donc pas efficace lorsque les images contiennent un bruit non-uniforme.

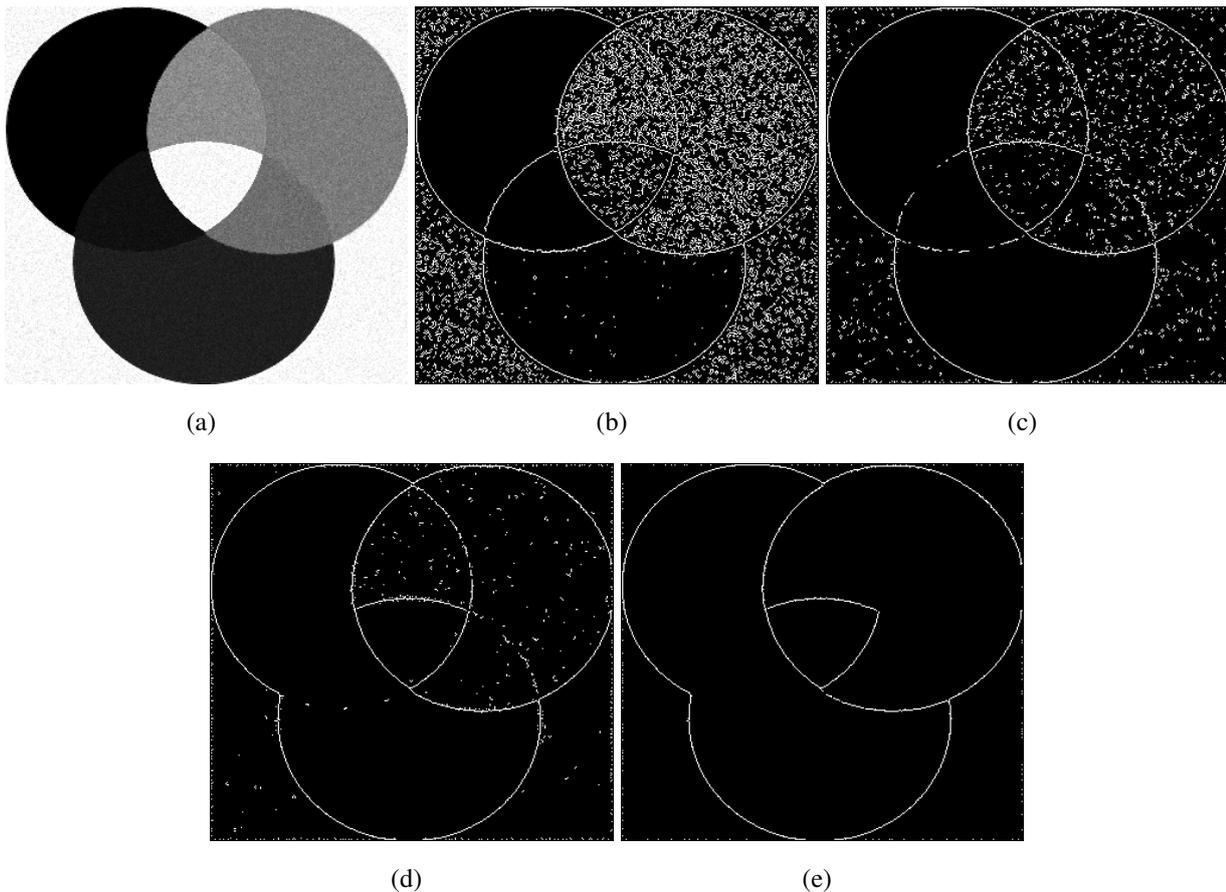


FIG. 4.7 – (a) Image originale bruitée avec un bruit de Poisson. (b) $s_h = 10$ et $s_b = 8$, (c) $s_h = 20$ et $s_b = 18$, (d) $s_h = 25$ et $s_b = 23$, (e) $s_h = 108$ et $s_b = 105$.

Nous avons pu observer à travers plusieurs exemples que les valeurs de seuil sont difficiles à définir. Elles diffèrent d'une image à l'autre et il n'y a pas d'approche universelle pour les fixer. De plus, il existe plusieurs sortes de bruits en fonction du capteur utilisé et pour les bruits non-uniformes, les méthodes de seuillage classiques échouent. Pour résoudre ces problèmes, nous avons donc proposé une nouvelle approche d'estimation de gradient permettant de seuiller l'image quelle que soit sa nature et le bruit qu'elle contient.

4.1.5 Variations entre les contours obtenus avec différentes approches

Si l'image traitée ne contient pas de bruit, les contours détectés devraient être les mêmes quelque soit le noyau sommatif utilisé. Cependant, la plupart des images réelles contiennent du bruit, en plus ou moins grande proportion. Lorsque l'on effectue une détection de contours sur la même image avec différents noyaux sommatifs, on peut constater certaines variations entre les contours obtenus. Ces variations dépendent uniquement du bruit présent dans les images. Considérons par exemple un zoom sur l'oeil de Léna représenté Figure 4.8(a). Cette image n'a pas été bruitée, elle contient uniquement les bruits d'acquisition initiaux. Les Figures (b), (c), (d) et (e) présentent les contours obtenus avec les filtres de Canny-Deriche, Shen-Castan, Sobel et Prewitt sur cette image après maximisation et seuillage arbitraire adapté à l'image (choisi de manière à obtenir le meilleur rapport entre les contours désirés et les contours indésirables dans l'image). On peut remarquer que tous les points détectés dans ces images ne sont pas identiques.

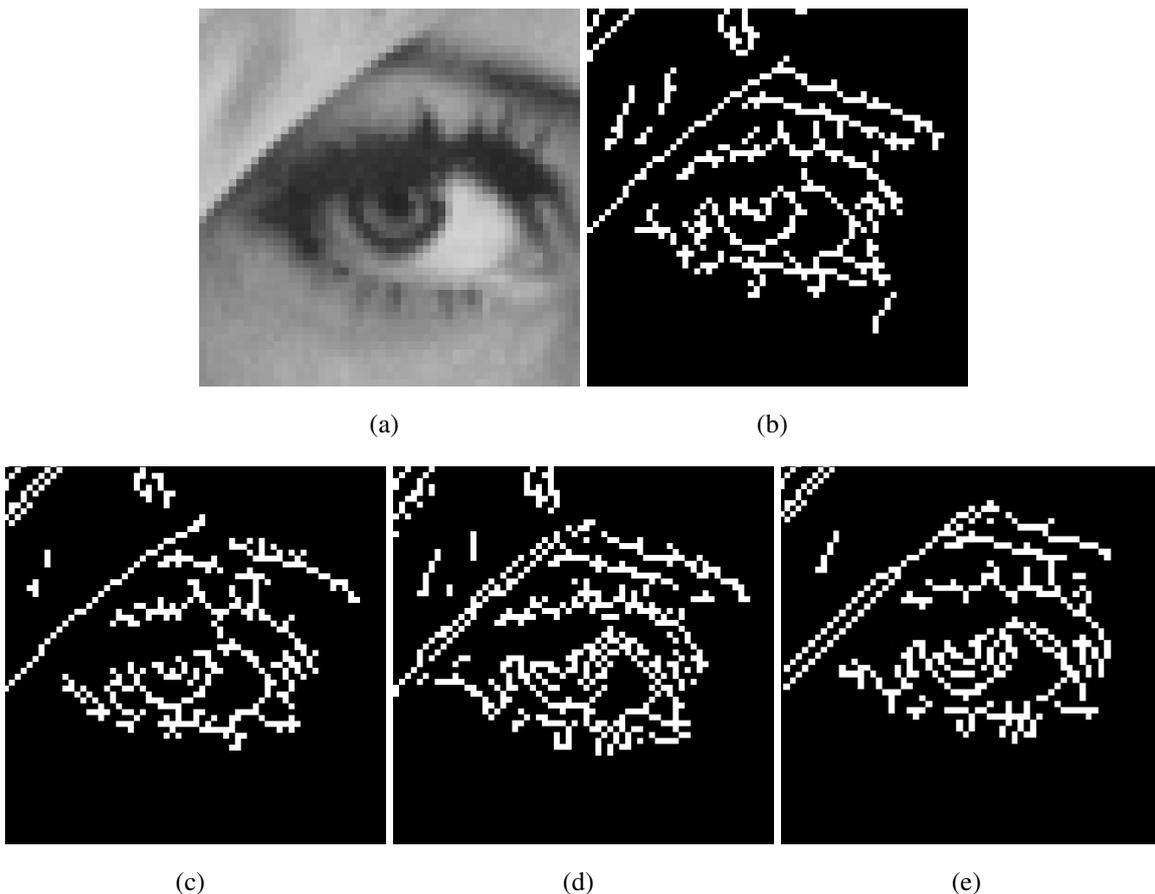


FIG. 4.8 – (a) Zoom sur l'oeil de Léna. (b) Contours obtenus avec le filtre de Canny-Deriche. (c) Contours obtenus avec le filtre de Shen-Castan. (d) Contours obtenus avec le filtre de Sobel. (e) Contours obtenus avec le filtre de Prewitt.

Sur la Figure 4.9, nous avons représenté les filtres les plus fréquemment utilisés (Canny-Deriche, Shen-Castan, gaussien, Prewitt et Sobel). Tous ces filtres sont basés sur des noyaux que nous avons

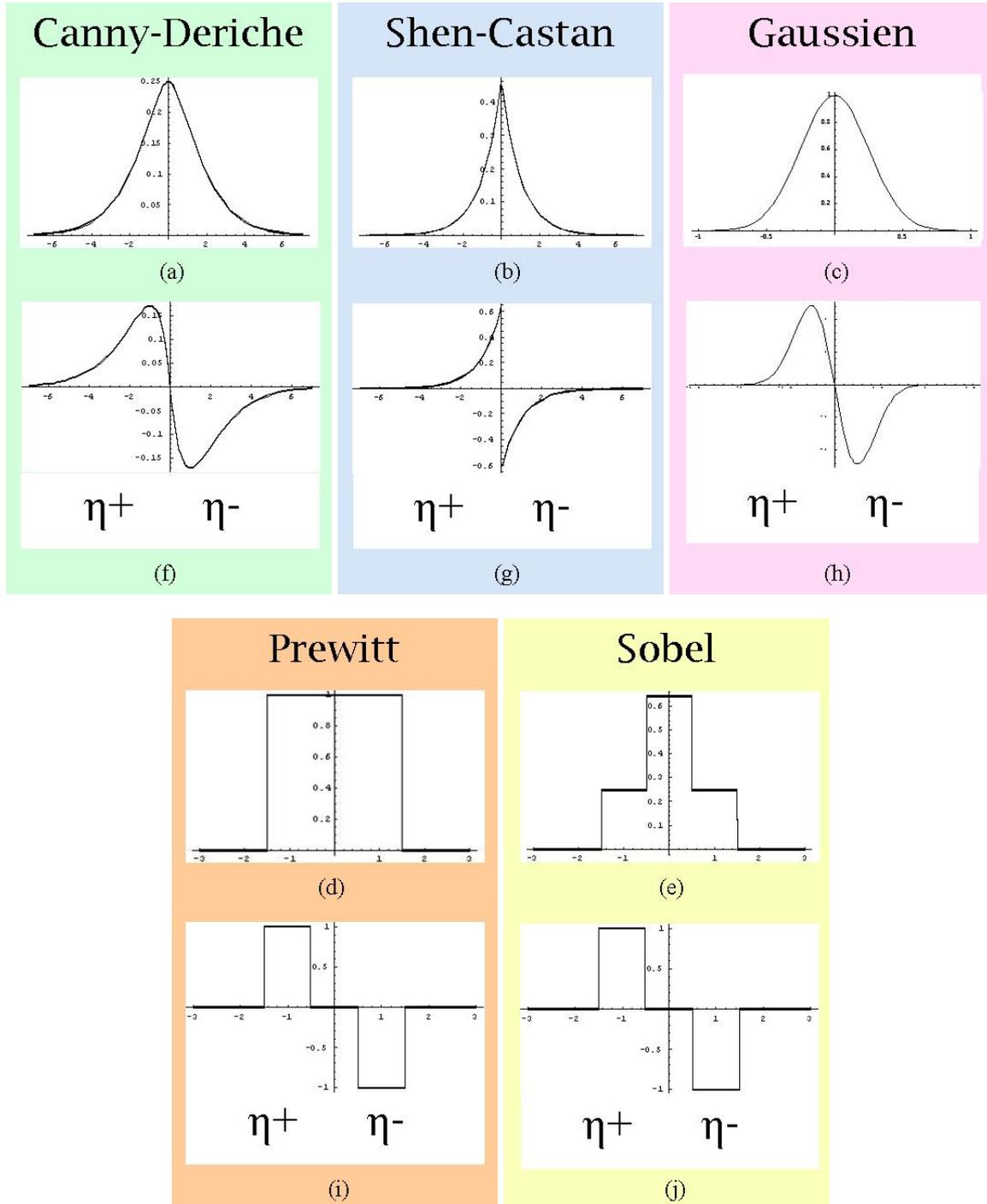


FIG. 4.9 – Présentation des noyaux de Canny-Deriche, Shen-Castan, gaussien, de Prewitt et de Sobel.

représentés en (a), (b), (c), (d) et (e). Ces noyaux ont été dérivés pour obtenir le filtre de dérivation

que nous connaissons et que nous avons illustré en (f), (g), (h), (i) et (j). On peut remarquer que tous ces noyaux dérivés ont des structures similaires : ils sont tous constitués la différence de deux noyaux usuels notés η^+ et η^- .

Les filtres gaussien, de Canny-Deriche et de Shen-Castan ont un support infini. Dans le cas de Canny-Deriche et Shen-Castan, ce support infini permet une implémentation récursive des filtres. Dans notre approche, nous nous intéresserons uniquement aux cas des noyaux bornés.

4.2 Méthode non-additive d'estimation de gradient dans les images

4.2.1 Dérivation d'un signal avec des noyaux non-sommatifs

L'estimation de la dérivée du signal continu s de l'image est interpolée à partir du signal discret S de l'image échantillonnée convolué avec un noyau sommatif dérivé $d\kappa$. Ce noyau est construit le plus souvent par soustraction de deux noyaux usuels.

$$d\kappa(x) = \eta^+(x - \epsilon_x) - \eta^-(x + \epsilon_x). \quad (4.5)$$

Convoluer un signal S avec le noyau dérivé $d\kappa(x)$ est équivalent à sommer la convolution de S par le noyau η^+ et la convolution de $-S$ par le noyau η^- . L'estimation de $DS(x)$, dérivée du signal discret S à la position x , est donnée par :

$$DS(x) = (\eta^+(x + \epsilon_x) - \eta^-(x - \epsilon_x)) * S(x) \quad (4.6)$$

$$= \eta^+(x + \epsilon_x) * S(x) + \eta^-(x - \epsilon_x) * (-S(x)) \quad (4.7)$$

Notons

$$\begin{cases} \kappa^+(x) = \eta^+(x + \epsilon_x) \\ \kappa^-(x) = \eta^-(x - \epsilon_x) \end{cases}$$

Dans ce cas l'expression (4.7) devient :

$$DS(x) = \mathbb{E}_{\kappa^+}(S) + \mathbb{E}_{\kappa^-}(-S) \quad (4.8)$$

Si κ^+ et κ^- sont des noyaux sommatifs, bornés, symétriques et positifs, alors nous avons vu en Section 2.2.3 que ces deux noyaux sommatifs sont dominés par deux noyaux non-sommatifs π^+ et π^- de même mode et même support. De plus, si π^+ et π^- sont triangulaires, alors ils dominent l'ensemble des noyaux sommatifs de même mode et même support que κ^+ et κ^- d'après le principe de domination. Ces deux noyaux sont représentés Figure 4.10.

En appliquant le théorème (2.30), on a :

$$\begin{cases} \mathbb{E}_{\kappa^+}(S) \in [-\mathbb{C}_{\nu_W^+}(-S), \mathbb{C}_{\nu_W^+}(S)] \\ \mathbb{E}_{\kappa^-}(-S) \in [-\mathbb{C}_{\nu_W^-}(S), \mathbb{C}_{\nu_W^-}(-S)] \end{cases}$$

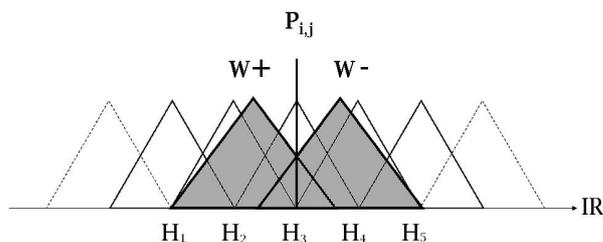


FIG. 4.10 – Noyau dérivé construit par soustraction de deux noyaux usuels.

Avec l'expression (4.8), on obtient alors une estimation imprécise de la dérivée de S :

$$[\underline{DS}, \overline{DS}] = [-\mathbb{C}_{\nu_W^+}(-S), \mathbb{C}_{\nu_W^+}(S)] + [-\mathbb{C}_{\nu_W^-}(S), \mathbb{C}_{\nu_W^-}(-S)] \quad (4.9)$$

$$= [-\mathbb{C}_{\nu_W^+}(-S) - \mathbb{C}_{\nu_W^-}(S), \mathbb{C}_{\nu_W^+}(S) + \mathbb{C}_{\nu_W^-}(-S)]. \quad (4.10)$$

Si la capacité ν_W est concave, cet intervalle contient toutes les valeurs possibles pour la dérivée.

Il est à noter que plus le niveau de bruit augmente dans les images, plus l'intervalle définissant DS s'étend, i.e. plus l'imprécision sur la valeur de la dérivée augmente.

Nous avons donc défini un noyau dérivé qui permet d'estimer, en chaque pixel de l'image, un intervalle de valeurs possibles de la dérivée en fonction du niveau de bruit de ce pixel. Rappelons que nous cherchions à détecter les points de contours de l'image. Si la dérivée de l'image est nulle cela signifie que la fonction d'intensité de l'image ne subit pas de variation, donc ce n'est pas un contour. Comment traduire cela avec notre représentation en intervalle de DS ? En fait on peut faire l'hypothèse que la dérivée peut être nulle si :

$$0 \in [\underline{DS}, \overline{DS}]. \quad (4.11)$$

En effet, si $0 \in [\underline{DS}, \overline{DS}]$, alors il existe deux noyaux sommatifs $\kappa^+(x)$ et $\kappa^-(x)$ tels que $\mathbb{E}_{\kappa^+}(S) - \mathbb{E}_{\kappa^-}(S) = 0$. Ce point n'est donc pas un maxima local du gradient. Il n'est pas considéré comme un point de contour. Au contraire, si $0 \notin [\underline{DS}, \overline{DS}]$, on ne peut trouver deux valeurs $\kappa^+(x)$ et $\kappa^-(x)$ telles que $\mathbb{E}_{\kappa^+}(S) - \mathbb{E}_{\kappa^-}(S) = 0$. Le point n'est donc pas négligeable, il s'agit bien d'un point de contour.

Cette règle simple est utilisée pour traiter les images. En annulant tous les pixels dont l'intervalle des valeurs de la dérivée contient la valeur 0, on effectue implicitement un seuillage naturel des valeurs estimées du gradient. Ce seuillage permet d'éviter le choix arbitraire de valeurs permettant de seuiller l'image.

4.2.2 Détection de contour dans une image perspective

Les noyaux non-sommatifs présentés dans la Section 2.2.2 sont étendus à en 2D (Section 2.2.3) pour effectuer une estimation imprécise du gradient d'une image. La propriété de domination est

toujours vérifiée. L'image étant bi-dimensionnelle, nous devons estimer la dérivée selon x et selon y . Par souci de simplicité, nous présenterons cette approche pour la dérivée selon x . L'estimation imprécise de la dérivée selon y s'effectue de la même manière.

Soit $\kappa(x, y) = \kappa_x(x)\kappa_y(y)$ un noyau de dérivation séparable, l'extension 2D de l'expression du gradient (4.5) est donnée par :

$$\frac{\delta}{\delta x}(\kappa(x, y)) = \frac{\delta}{\delta x}(\kappa_x(x)\kappa_y(y)) = \frac{\delta}{\delta x}(\kappa_x(x))\kappa_y(y) \quad (4.12)$$

$$= (\eta^+(x + \epsilon_x) - \eta^-(x - \epsilon_x))\kappa_y(y). \quad (4.13)$$

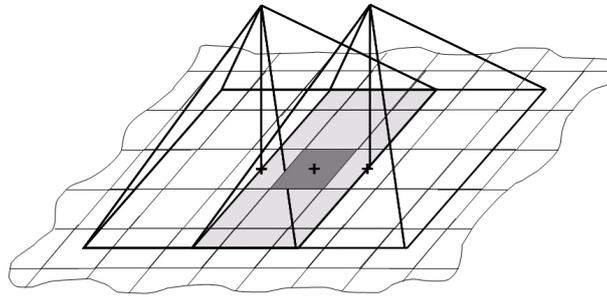


FIG. 4.11 – Extension en 2D du noyau dérivé.

Soit

$$\begin{cases} \kappa_x^+ = \eta^+(x + \epsilon_x)\kappa_y(y) \\ \kappa_x^- = \eta^-(x - \epsilon_x)\kappa_y(y) \end{cases}$$

Une estimation de la dérivée du signal 2D S suivant l'axe des x (GS_x) est évaluée en sommant la convolution de S avec le noyau κ_x^+ et la convolution de $(-S)$ avec le noyau sommatif κ_x^- :

$$GS_x = \mathbb{E}_{\kappa_x^+}(S) + \mathbb{E}_{\kappa_x^-}(-S). \quad (4.14)$$

Etant donné que $\mathbb{E}_{\kappa_x^+}(S)$ et $\mathbb{E}_{\kappa_x^-}(-S)$ sont des données imprécises, GS_x est également imprécise. Le principe de domination est étendu en 2D (cf. Section 2.2.3). Le noyau triangulaire 2D est estimé avec l'intégrale de Choquet (2.31). L'estimation imprécise de GS_x est une extension 2D de la Section 4.2.1. Deux noyaux non-sommatifs (représentés Figure 4.11) dominent une famille de noyaux κ_x^+ et κ_x^- .

Avec cette approche, la détection de contours est évaluée sous forme d'un intervalle de valeurs du gradient $[\underline{GS}_x, \overline{GS}_x] \times [\underline{GS}_y, \overline{GS}_y]$. Un point appartient au contour s'il s'agit d'un extremum de la dérivée de la fonction d'intensité de l'image. Un point n'est pas un extremum si

$$(0, 0) \in [\underline{GS}_x, \overline{GS}_x] \times [\underline{GS}_y, \overline{GS}_y]. \quad (4.15)$$

Un seuillage naturel de l'image est obtenu et nous permet d'éviter le choix arbitraire des valeurs de seuil.

4.2.3 Détection de contour dans une image omnidirectionnelle

Come nous l'avons vu Section 4.1.3 et 4.1.4, une détection de contours est systématiquement associée aux étapes de maximisation et seuillage. Dans les images perspectives, la scène projetée est échantillonnée régulièrement suivant une grille cartésienne de pixels. Le concept de voisinage est relativement intuitif et il est possible de définir les voisins d'un pixel dans une direction donnée pour estimer la norme du gradient. Dans les images omnidirectionnelles, la topologie du voisinage est modifiée par la forme du miroir. La définition des voisins dans la direction du gradient est beaucoup moins évidente. En réalité, la projection de la scène 3D sur le miroir provoque une concentration de l'information au centre de l'image. La résolution n'est pas invariante radialement. La définition de deux voisins à une distance unitaire d'un pixel n'a pas de sens dans l'espace de l'image. Pour les définir de manière géométriquement correcte, il est nécessaire d'utiliser un espace dans lequel les distances sont homogènes. Comme nous l'avons vu précédemment, une solution projective peut être utilisée pour définir les voisins de $P_{i,j}$ dans la direction du gradient. La Figure 4.12 illustre le calcul de deux "pixels" P_A et P_B projetés sur l'image omnidirectionnelle en utilisant le modèle de projection décrit dans [42] via le cylindre. Cependant, une telle approche est très coûteuse au niveau du temps de calcul (la direction du gradient, différente en chaque pixel, rend impossible une étape de pré-calculs).

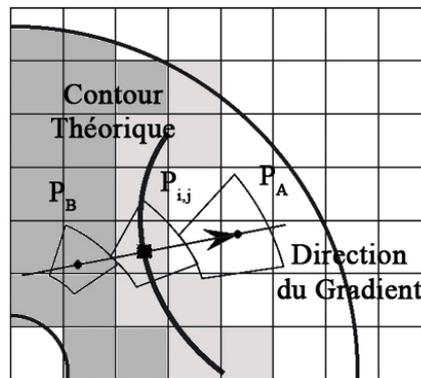


FIG. 4.12 – Maximisation de gradient dans une image omnidirectionnelle

Pour nous affranchir des étapes de maximisation et de seuillage dans les images omnidirectionnelles, nous avons donc adapté l'approche décrite Section 3.2 à l'aide de l'algorithme (2) de projection des opérateurs comme illustré Figure 4.13. Le centre de chaque pixel de l'image omnidirectionnelle est projeté sur le cylindre. Les coordonnées du noyau de dérivation sont calculées autour de ce point projeté. Le noyau est ensuite rétro-projeté sur l'image omnidirectionnelle, ainsi sa taille varie en fonction de sa position dans l'image. L'intervalle des valeurs du gradient du pixel est estimé à l'aide d'une intégrale de Choquet, en respectant la forme du noyau projeté.

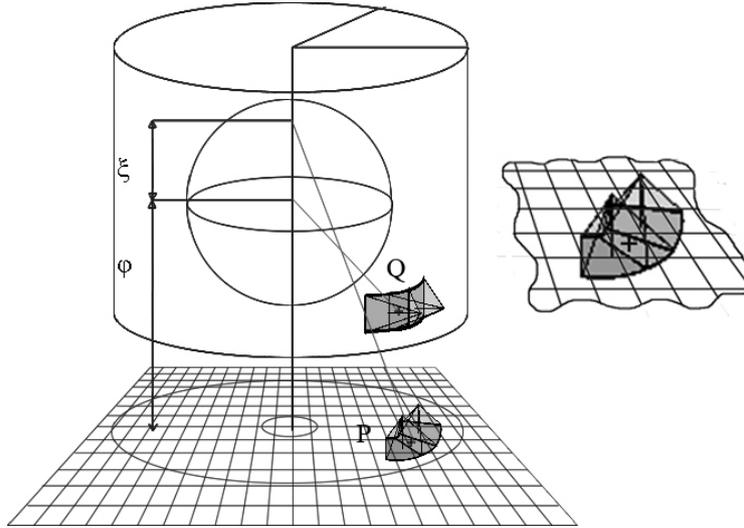


FIG. 4.13 – Projection des noyaux sur l'image omnidirectionnelle.

4.2.3.1 Modélisation du capteur

Notre algorithme utilise le modèle de projection unifié de Geyer et Daniilidis présenté Section 1.1.4.1.5. Ce modèle nous permet de relier un point $M(X, Y, Z)$ de l'espace 3D et sa projection $P(x, y)$ sur l'image, en utilisant une sphère de rayon R . Cette projection est décrite par les équations (4.16) et (4.17) :

$$x = \frac{(\varphi + \xi)X}{\xi R - Z} \quad (4.16)$$

$$y = \frac{(\varphi + \xi)Y}{\xi R - Z} \quad (4.17)$$

Ces équations de projection dépendent de deux paramètres seulement : φ , la distance entre l'image omnidirectionnelle et le centre de la sphère, et ξ , la distance entre le centre de la sphère et le point appartenant à l'axe de la sphère. De plus, on a $R^2 = X^2 + Y^2 + Z^2$.

L'échantillonnage réalisé par la rétine relie les coordonnées physiques de la projection de $M(X, Y, Z)$ à ses coordonnées discrètes (u, v) dans l'image :

$$u = k_u x + u_0 \quad (4.18)$$

$$v = k_v y + v_0 \quad (4.19)$$

avec (k_u, k_v, u_0, v_0) , des constantes liées à l'imageur. (u_0, v_0) représente les coordonnées de projection de l'axe optique dans l'image et (k_u, k_v) les rapports relatifs à la taille des pixels dans le plan (u, v) de l'image. En général, on a $k_u = k_v$.

en posant

$$d = \frac{R(\varphi + \xi)}{\xi \sqrt{1 + \frac{Z^2}{D^2} - R \frac{Z}{D}}} \quad (4.26)$$

Cette projection peut être faite dans le plan d'orientation θ et fait ainsi apparaître le rapport $\rho = \frac{Z}{D}$:

$$d = \frac{R(\varphi + \xi)}{\xi \sqrt{1 + \rho^2 - R\rho}} \quad (4.27)$$

On voit donc apparaître une relation non-linéaire entre (d, θ) la projection métrique et (ρ, θ) les coordonnées relatives sur le cylindre. Le rayon de la sphère est arbitraire. On considère généralement la sphère de rayon 1 mètre $R = 1$, ce qui nous donne la relation $d = f(\rho)$:

$$d = f(\rho) = \frac{\varphi + \xi}{\xi \sqrt{1 + \rho^2 - \rho}} \quad (4.28)$$

La relation inverse $\rho = f^{-1}(d)$ est obtenue de la manière suivante :

$$d(\xi \sqrt{1 + \rho^2} - \rho) = (\varphi + \xi) \quad (4.29)$$

$$d\xi \sqrt{1 + \rho^2} = d\rho + (\varphi + \xi) \quad (4.30)$$

$$d^2 \xi^2 (1 + \rho^2) = d^2 \rho^2 + (\varphi + \xi)^2 + 2d\rho(\varphi + \xi) \quad (4.31)$$

$$\rho^2 (d^2 \xi^2 - 2d\rho(\varphi + \xi) + d^2 \xi^2 - (\varphi + \xi)^2) = 0 \quad (4.32)$$

Le discriminant de cette expression est :

$$\Delta = 4(\varphi + \xi)^2 d^2 - 4(d^2 \xi^2 - (\varphi + \xi)^2)(d^2 \xi^2 - d^2) \quad (4.33)$$

$$= 4(d^2 \xi^2 (d^2 (1 - \xi^2) + (\varphi + \xi)^2)) \quad (4.34)$$

On obtient ainsi deux solutions :

$$\rho_{1,2} = \frac{(\varphi + \xi) \pm \xi \sqrt{d^2 (1 - \xi^2) + (\varphi + \xi)^2}}{d(1 - \xi^2)} \quad (4.35)$$

Pour discriminer ces deux solutions, il faut revenir aux projections.

Considérons les triangles OMZ et OSZ_S . On a :

$$\frac{Z_S}{D_S} = \frac{Z}{D} = \rho \quad (4.36)$$

De plus $D_S^2 + Z_S^2 = R^2$.

De même dans les triangles BPP_S et BSZ_S . On a :

$$\frac{\xi + Z_S}{D_S} = \frac{\varphi + \xi}{d} \quad (4.37)$$

Le dénominateur de D_S est positif. Sur la Figure 4.15, on voit que la solution correcte est toujours la plus grande (en valeur absolue) des deux solutions. Donc :

$$\rho = \frac{\xi}{D_S} - A \quad (4.43)$$

$$= \frac{\xi(A^2 + 1) - A(\xi A + \sqrt{R^2(A^2 + 1) - \xi^2})}{\xi A + \sqrt{R^2(A^2 + 1) - \xi^2}} \quad (4.44)$$

$$= \frac{\xi - A\sqrt{R^2(A^2 + 1) - \xi^2}}{\xi A + \sqrt{R^2(A^2 + 1) - \xi^2}} \quad (4.45)$$

$$= \frac{(\xi A - \sqrt{R^2(A^2 + 1) - \xi^2})(\xi - A\sqrt{R^2(A^2 + 1) - \xi^2})}{\xi^2 A^2 - (R^2(A^2 + 1) - \xi^2)} \quad (4.46)$$

$$= \frac{\xi^2 A - \xi\sqrt{R^2(A^2 + 1) - \xi^2} - \xi A^2\sqrt{R^2(A^2 + 1) - \xi^2} + A(R^2(A^2 + 1) - \xi^2)}{\xi^2 A^2 - (R^2(A^2 + 1) - \xi^2)} \quad (4.47)$$

$$= \frac{-\xi(A^2 + 1)\sqrt{R^2(A^2 + 1) - \xi^2} + AR^2(A^2 + 1)}{\xi^2(A^2 + 1) - R^2(A^2 + 1)} \quad (4.48)$$

$$= \frac{(A^2 + 1)(\sqrt{R^2(A^2 + 1) - \xi^2})}{(A^2 + 1)(\xi^2 - R^2)} \quad (4.49)$$

$$= \frac{AR^2 - \sqrt{R^2(A^2 + 1) - \xi^2}}{\xi^2 - R^2} \quad (4.50)$$

$$= \frac{(\frac{\xi+\varphi}{d})^2 R^2 - \xi\sqrt{R^2((\frac{\xi+\varphi}{d})^2 + 1) - \xi^2}}{\xi^2 - R^2} \quad (4.51)$$

$$= \frac{(\xi + \varphi)R^2 - \xi\sqrt{R^2((\xi + \varphi)^2 + d^2) - d^2\xi^2}}{d(\xi^2 - R^2)} \quad (4.52)$$

On obtient ainsi la valeur de ρ .

Remarque

Dans le cas des lentilles paraboliques, $\xi = 1$. Si on prend également $R = 1$, on a une singularité. Pour l'éviter, il faut faire un cas spécial pour la solution $\xi = 1$:

$$d = f(\rho) = \frac{\varphi + 1}{\sqrt{1 + \rho^2} - \rho} \quad (4.53)$$

dont l'inversion donne simplement :

$$\rho = f^{-1}(d) = \frac{d^2 - (\varphi + 1)^2}{2d(\varphi + 1)} \quad (4.54)$$

On en déduit les coordonnées (X, Y, Z) d'un point du cylindre associé au point (u, v) de l'image

omnidirectionnelle :

$$X = \frac{u - u_0}{k_u} \frac{\xi \sqrt{D^2 + Z^2} - Z}{\varphi + \xi} \quad (4.55)$$

$$Y = \frac{v - v_0}{k_v} \frac{\xi \sqrt{D^2 + Z^2} - Z}{\varphi + \xi} \quad (4.56)$$

$$Z = \rho d \quad (4.57)$$

4.2.3.3 Définition et projection du noyau

Connaissant les coordonnées du centre du pixel projeté, on peut souhaite définir géométriquement le noyau dérivatif sur le cylindre. Comme nous l'avons vu Section 4.2, ce noyau est constitué de deux noyaux usuels. La Figure 4.16 illustre le noyau dérivatif sur le cylindre dans chaque direction.

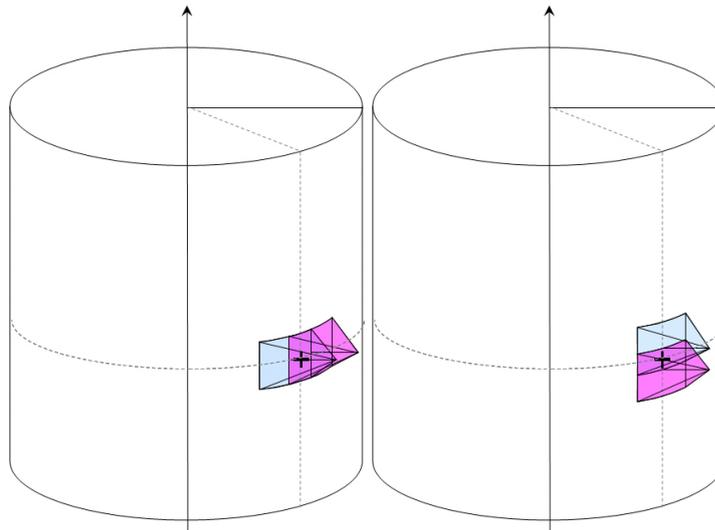


FIG. 4.16 – Définition des noyaux dérivatifs sur le cylindre dans chaque direction.

Ces noyaux sont définis géométriquement par 5 points : le centre du noyau et ses 4 coins (Figure 4.17). Ainsi, lorsque nous définissons un noyau dérivatif horizontal, nous devons calculer, en fonction du centre du pixel projeté M , les coordonnées des points M_1, A_1, B_1, C_1, D_1 et M_2, A_2, B_2, C_2, D_2 des deux noyaux qui le constituent. De même, lorsque nous définissons un noyau dérivatif vertical, nous devons calculer, les coordonnées des points M_3, A_3, B_3, C_3, D_3 et M_4, A_4, B_4, C_4, D_4 . Les coordonnées du point M sur le cylindre étant de la forme :

$$\begin{pmatrix} X_M \\ Y_M \\ Z_M \end{pmatrix} = \begin{pmatrix} D \cos(\theta) \\ D \sin(\theta) \\ Z \end{pmatrix},$$

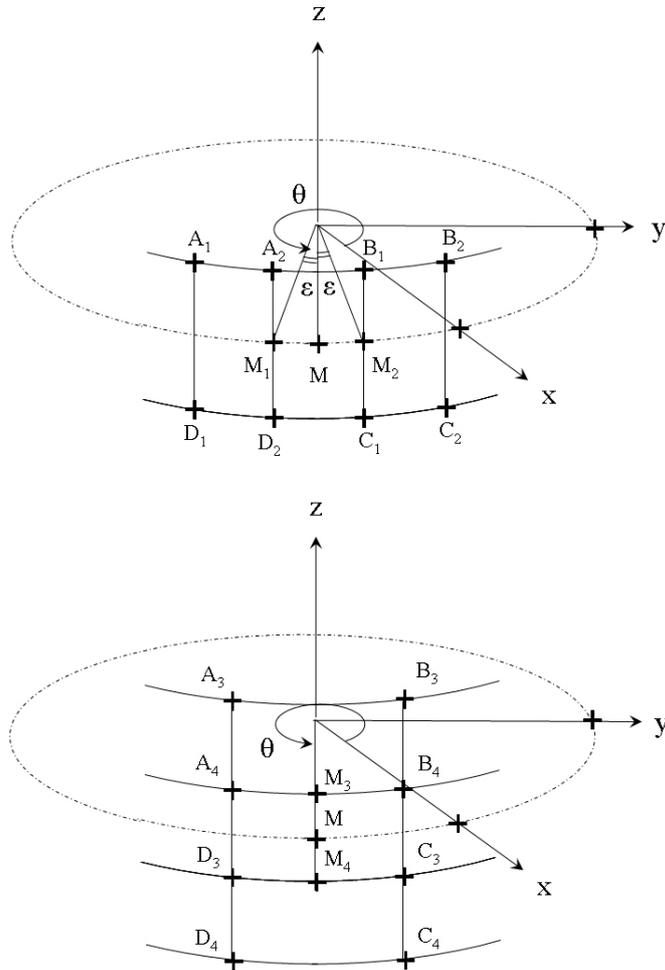


FIG. 4.17 – Calcul des coordonnées géométriques des noyaux sur le cylindre : (a) pour un noyau dérivatif horizontal, (b) pour un noyau dérivatif vertical.

les coordonnées des points M_1 et M_2 sont obtenues par variation angulaire :

$$\begin{pmatrix} X_{M_1} \\ Y_{M_1} \\ Z_{M_1} \end{pmatrix} = \begin{pmatrix} D \cos(\theta - \epsilon) \\ D \sin(\theta - \epsilon) \\ Z \end{pmatrix} \text{ et } \begin{pmatrix} X_{M_2} \\ Y_{M_2} \\ Z_{M_2} \end{pmatrix} = \begin{pmatrix} D \cos(\theta + \epsilon) \\ D \sin(\theta + \epsilon) \\ Z \end{pmatrix},$$

et les coordonnées des points M_3 et M_4 , par variation de la hauteur :

$$\begin{pmatrix} X_{M_3} \\ Y_{M_3} \\ Z_{M_3} \end{pmatrix} = \begin{pmatrix} D \cos(\theta) \\ D \sin(\theta) \\ Z - \delta \end{pmatrix} \text{ et } \begin{pmatrix} X_{M_4} \\ Y_{M_4} \\ Z_{M_4} \end{pmatrix} = \begin{pmatrix} D \cos(\theta) \\ D \sin(\theta) \\ Z + \delta \end{pmatrix},$$

On en déduit les coordonnées des autres points en fonction de δ et ϵ .

Ces coordonnées sont ensuite rétro-projetés sur l'image omnidirectionnelle à l'aide des équations de projection (4.16) et (4.16). Les noyaux projetés sont représentés Figure 4.18.

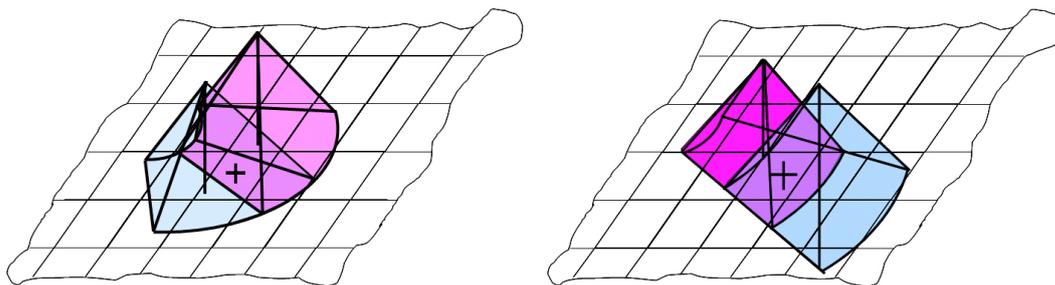


FIG. 4.18 – Noyaux dérivatifs projetés dans chaque direction.

4.2.3.4 Estimation du noyau dérivé dans chaque direction

Rappelons que l'image est considérée comme une partition floue du plan \mathbb{R}^2 (Section 2.1.2.2) : à chaque pixel est associé un sous-ensemble flou pyramidal de \mathbb{R}^2 . Pour estimer la valeur du noyau dérivé, il faut évaluer la valeur de chacun des deux noyaux usuels W^+ et W^- qui le composent, dans une direction donnée. Ces noyaux sont des noyaux non-sommatifs, triangulaires, bornés, symétriques, positifs. En utilisant le théorème (2.30) et l'intégrale de Choquet, on obtient une estimation de la convolution de l'image par chacun de ces noyaux. Ainsi, la convolution de l'image par le noyau W^+ appartient à l'intervalle $[-\mathbb{C}_{\nu_W^+}(-S), \mathbb{C}_{\nu_W^+}(S)]$ et la convolution de l'image par le noyau W^- appartient à l'intervalle $[-\mathbb{C}_{\nu_W^-}(S), \mathbb{C}_{\nu_W^-}(-S)]$. L'estimation de la dérivée suivant une direction donnée est obtenue par différence des intervalles de valeurs obtenus pour ces deux noyaux. On obtient ainsi une estimation du gradient sous la forme d'un intervalle $[\underline{DS}, \overline{DS}]$. Si $0 \in [\underline{DS}, \overline{DS}]$, alors il existe deux noyaux sommatifs dominés par les noyaux W^+ et W^- tels que l'estimation du gradient via ces noyaux est nulle. Le point considéré n'est donc pas un point de contour. Sa valeur est mise à 0. Dans le cas contraire, $0 \notin [\underline{DS}, \overline{DS}]$ il s'agit bien d'un point de contour. On donne la valeur $\frac{\underline{DS} + \overline{DS}}{2}$ au pixel de l'image.

Conclusion

Dans ce chapitre, nous avons présenté un nouvel opérateur de détection de contours basé sur les noyaux non-sommatifs et l'intégrale de Choquet. Cet opérateur permet d'obtenir une estimation du gradient sous la forme d'un intervalle. De plus il fournit un seuillage naturel de l'image. Cet opérateur a été développé de manière à pouvoir traiter n'importe quelle image indépendamment du bruit qu'elle contient. Il a ensuite été adapté aux images omnidirectionnelles. En effet, la géométrie de cette image rend les traitements de maximisation et seuillage délicats. Cette nouvelle approche permet de simplifier l'adaptation des traitements aux images omnidirectionnelles. De nombreuses expériences ont été réalisées dans le chapitre 5 pour montrer la validité de cette approche et mettre en évidence la robustesse de notre méthode en présence de bruit.

Chapitre 5

Expériences

Introduction

Dans ce chapitre nous proposons d'étudier la précision et la robustesse des approches présentées dans les Chapitres 3 et 4 sur des images artificielles et réelles. Nous avons cherché à comparer nos approches aux méthodes existantes afin de prouver leur efficacité. L'utilisation d'images de synthèse permet de contrôler tous les paramètres de l'imageur, ainsi que l'ensemble des éléments présents dans l'image. Pour estimer quantitativement et qualitativement nos approches, nous avons naturellement utilisé des images de synthèse. Cependant, dans la réalité, l'étalonnage n'est jamais parfait et l'acquisition d'images numériques introduit des bruits parasites. Des images réelles ont donc été utilisées pour montrer la pertinence de nos approches dans les conditions d'utilisation habituelles.

5.1 Quelques critères quantitatifs pour la détection de contours

Cette section présente succinctement des critères objectifs permettant d'estimer quantitativement la précision et la robustesse de nos approches. Le premier critère auquel nous nous sommes intéressé est la sensibilité au bruit de notre filtre. Pour la mesurer, nous avons utilisé les paramètres de Fram et Deutsch. Nous avons également étudié les critères de Canny de bonne détection, bonne localisation. Ces critères, présentés en Annexe C pour un filtre continu, ont été reformulés afin d'être appliqués aux images numériques, de nature discrète.

5.1.1 Paramètres de Fram et Deutsch

La comparaison entre différents opérateurs de détection de contours est souvent difficile. Les critères de détection de contours et de réduction de bruit dans les images sont souvent antagonistes. Canny a proposé des critères permettant de créer des filtres optimaux. Cependant, d'autres filtres peuvent être optimaux en respectant des critères différents. Une méthode d'évaluation quantitative

des performances des opérateurs d'extraction de contours a été proposée par Fram et Deutsch dans [16] et [29]. Les performances sont quantifiées par les valeurs de deux paramètres P_1 et P_2 à valeurs dans $[0, 1]$. P_1 mesure la robustesse du détecteur en présence de bruit : il est faible lorsqu'il y a beaucoup de fausses détections et vaut 1 quand il n'y a aucun pixel détecté dû au bruit. P_2 caractérise le comportement du détecteur vis-à-vis du signal seul : il vaut 1 lorsque tous les points de la transition réelle sont extraits sans bruit et qu'il n'y a pas de lacune.

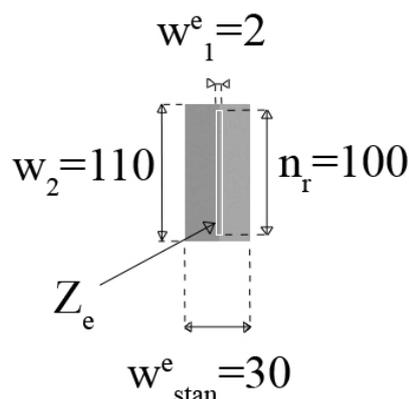


FIG. 5.1 – Modélisation d'un contour dans l'image.

Ces critères sont estimés en considérant une zone de l'image qui entoure le contour à évaluer (Figure 5.1). Soit n_{tot} le nombre total de points détectés de la région considérée, Z_e la zone de contour de cette région, n_{in} le nombre de points à l'intérieur de Z_e et n_{out} le nombre de points à l'extérieur ($n_{tot} = n_{in} + n_{out}$). n^0 est le nombre de points seuillés à l'extérieur du contour et n^e le nombre total de points seuillés. w_{stan}^e est le nombre de colonnes de la zone considérée et w_1^e est le nombre de colonnes de Z_e . De même w_2 est le nombre total de lignes de la zone considérée et n_r le nombre de lignes de Z_e . Un paramètre f est utilisé pour normaliser les proportions entre la zone de contour Z_e et la région considérée ($f = \frac{w_1^e}{w_{stan}^e}$). L'estimation du bruit dans Z_e vaut $n_{noise}^e = n^0 \frac{n_{in}}{n_{out}}$ et le nombre de points liés uniquement au signal n_{sig}^e (certains points peuvent appartenir au signal et au bruit) est donné par :

$$n_{sig}^e = \frac{n^e - n_{noise}^e}{1 - \frac{n_{noise}^e}{n_{in}}}$$

Les paramètres P_1 et P_2 sont estimés de la manière suivante :

$$P_1 = \frac{n_{sig}^e}{n_{sig}^e + (n_{noise}^e + n^0) \frac{n_{in}}{f n_{tot}}} \quad (5.1)$$

$$P_2 = \frac{\frac{n_r}{w_2} - \left\{ 1 - \left[1 - \frac{n_{noise}^e}{n_{in}} \right]^{w_1^e} \right\}}{\left[1 - \frac{n_{noise}^e}{n_{in}} \right]^{w_1^e}} \quad (5.2)$$

5.1.2 Critère de bonne localisation

Il s'agit de vérifier à quel point le contour détecté correspond au contour réel. Pour cela, il existe plusieurs techniques : soit on calcule une statistique de l'erreur de position entre les pixels détectés et le contour théorique (par mesure des écarts), soit on évalue à l'aide d'une estimation au sens des moindres carrés les paramètres du contour que l'on compare avec les valeurs théoriques (c'est la méthode de régression linéaire).

5.1.2.1 Mesure des écarts

Une technique pour évaluer la bonne localisation des contours est de calculer la moyenne et l'écart-type des distances entre les pixels du contour extrait et la transition théorique. L'écart mesuré est la distance perpendiculaire entre le centre de chaque pixel appartenant au contour et la transition réelle, comme illustré Figure 5.2.

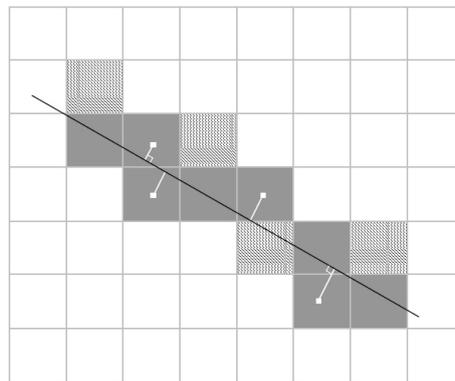


FIG. 5.2 – Mesure des écarts-types : les pixels grisés appartiennent au contour. Les pixels gris foncé ont été marqués par le détecteur de contours, les pixels gris clair n'ont pas été détectés.

Le centre d'un pixel p_i est exprimé par ses coordonnées cartésiennes (x_i, y_i) dans l'image et la transition théorique par son équation de la forme $ax + by + c = 0$. La distance d_i entre le centre du pixel et la droite est alors donnée par :

$$d_i = \left| \frac{ax_i + by_i + c}{\sqrt{a^2 + b^2}} \right| \quad (5.3)$$

Si N est le nombre de pixels détectés, alors la moyenne μ et l'écart-type λ sont respectivement donnés par :

$$\mu = \frac{\sum_i d_i}{N} \quad (5.4)$$

$$\lambda = \sqrt{\frac{\sum_i (d_i - \mu)^2}{N}} \quad (5.5)$$

5.1.2.2 Régression linéaire

Un autre moyen de quantifier les performances de l'opérateur est de chercher une corrélation entre le contour extrait et le contour théorique. Le contour le plus simple à modéliser étant la droite, notre choix s'est naturellement porté sur la régression linéaire. On cherche par la méthode des moindres carrés la droite qui minimise la somme des carrés des écarts entre les abscisses x_i des pixels extraits de la zone de contour et la droite théorique $y = ax + b$. Cette droite s'appelle la droite de régression linéaire de x en y . Elle est représentée Figure 5.3.

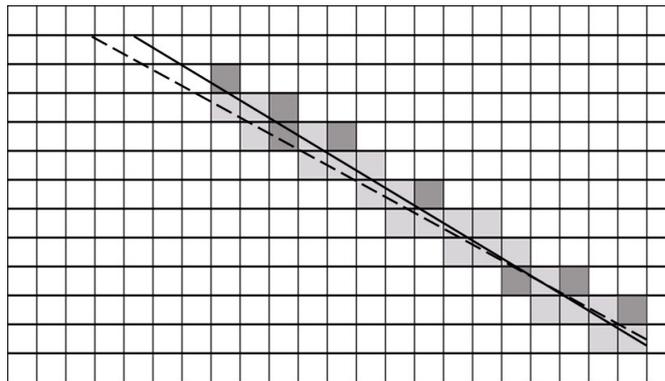


FIG. 5.3 – Droite de régression linéaire : les pixels grisés appartiennent à la zone de contour. Les pixels gris clair sont ceux détectés par l'opérateur et les pixels gros foncé ne le sont pas. Le contour théorique, représenté en pointillé est calculé à partir de l'ensemble des pixels appartenant au contour alors que la droite de régression linéaire représentée en noir est calculée uniquement à partir des pixels marqués par le détecteur de contour.

La droite de régression, comme toute droite, est donnée par son équation $x = cy + d$. Si N est le nombre de pixels détectés, alors les paramètres c et d sont donnés respectivement par :

$$c = \frac{n \sum_i (x_i y_i) - \sum_i x_i \sum_i y_i}{N \sum (y^2) - \sum (y)^2} \quad (5.6)$$

$$d = \frac{\sum_i x_i - m \sum_i y_i}{N} \quad (5.7)$$

5.1.3 Critère de bonne détection

Nous avons également étudié le rapport signal sur bruit SNR ("Signal-to-Noise Ratio"). Un SNR élevé correspond à une faible probabilité de ne pas détecter un vrai point de contour ou de marquer de faux points de contours. Cet opérateur mesure le rapport entre l'information et le bruit dans une image. Il est calculé à partir d'une image initiale $Im1$ qui contient le signal et le bruit et d'une version restaurée de cette image $Im2$ de même taille. Par conséquent, plus le SNR est élevé,

meilleur est le signal. Le SNR est défini par :

$$\sigma = 10 \times \log_{10} \left(\frac{\sum Im1^2}{\sum (Im2 - Im1)^2} \right) \quad (5.8)$$

5.2 Présentation des approches comparées

Dans la suite de ce document, nous présentons une comparaison des différentes approches de détection de contours. Pour l'approche présentée Chapitre 3, le masque de convolution que nous avons choisi pour estimer la qualité de l'Algorithme (2) est le masque de Prewitt, présenté Annexe B (appliqué à l'image dans les 4 directions). Pour des raisons d'espace dans les tableaux de données, les méthodes seront respectivement mentionnées par les noms suivants : Prewitt, Désanamorphose, Proportionnelle, Floue sans interactions, Floue avec interactions, Deriche, Shen et Gradient maxitif. Nous avons choisi tous ces filtres de manière à ce qu'ils aient la même granulosité.

1. **Prewitt** : Le masque perspective est convolué à l'image omnidirectionnelle.
2. **Désanamorphose** : L'image omnidirectionnelle est désanamorphosée sur le cylindre. Le masque perspective est convolué avec l'image reconstruite sur le cylindre. Puis, l'image filtrée est retro-projetée sur le plan omnidirectionnel pour être comparée avec les autres méthodes.
3. **Proportionnelle** : L'image omnidirectionnelle est traitée avec l'approche proportionnelle présentée Section 3.3.

Les deux approches suivantes présentent deux extensions possibles de l'approche floue décrite Section 3.4. Ces variations dépendent de la manière de considérer le masque.

4. **Floue sans interactions** : Le masque est considéré comme une quantité précise. Le chevauchement entre les éléments du masque n'est donc pas pris en compte pour estimer la valeur du pixel filtré.
5. **Floue avec interactions** : Le masque est une quantité imprécise. Les éléments du masque se chevauchent. L'intégrale de Choquet est utilisée pour prendre en compte les interactions entre les éléments du masque et calculer la valeur du pixel filtré.

Deux filtres optimaux, références en détection de contour, sont également présentés à titre indicatif. Ces filtres n'ont pas été adaptés à la géométrie des images omnidirectionnelles.

6. **Deriche** : Le filtre de Canny-Deriche (Annexe D).
7. **Shen** : Le filtre de Shen-Castan (Annexe E).

Enfin, l'approche présentée Chapitre 4 :

8. **Gradient maxitif** : Le noyau dérivatif défini est appliqué à l'image.

Cette approche sera validée dans une première partie sur des images perspectives avant d'être appliquée aux images omnidirectionnelle.

Remarque :

Afin d'obtenir les contours proprement dits, les images subissent les post traitement de maximisation et seuillage présentés Section 4.1.3 et 4.1.4. Les seuils ont été choisis dans le but d'extraire le maximum de pixels de contours et le minimum de pixels de bruit. Il y a cependant un compromis à faire entre sur-détection (extraire plus de bruit) et sous-détection (laisser des « trous » dans le contour). Dans notre étude, le choix a été fait d'extraire le contour le plus complet possible, même s'il faut détecter plus de bruit. Cela revient à privilégier le critère P_2 par rapport à P_1 . Le choix des seuils est guidé par le jugement (visuel) de l'opérateur.

L'approche du gradient maxitif fournit un seuillage naturel de l'image. L'étape de seuillage n'a donc pas été réalisée avec cette approche.

5.3 Validation de l'approche Gradient maxitif avec des images perspectives

5.3.1 Evaluation quantitative de l'approche du Gradient maxitif

5.3.1.1 Application de l'approche du Gradient maxitif sur des images de synthèse

Ayant créé un nouvel opérateur de détection de contours, nous avons souhaité étudier le comportement de cet opérateur sur des images perspectives avant de l'appliquer aux images omnidirectionnelles. Cet opérateur fournit un seuillage naturel des valeurs estimées du gradient. Notre objectif était de créer un opérateur applicable à toutes les images, indépendamment de leur nature et du bruit qu'elles contiennent.

Notre approche est comparée avec les filtres de Prewitt, Canny-Deriche et Shen-Castan. La comparaison est effectuée sur l'image de synthèse présentée Figure 5.4(a). Cette image est bruitée avec deux types de bruit afin de tester la robustesse de notre détecteur de contour. Dans un premier temps, nous avons appliqué un bruit gaussien de variance σ croissante. Ce bruit Gaussien est obtenu en ajoutant à chaque pixel une valeur aléatoire suivant une loi de probabilité Gaussienne de variance σ et de moyenne μ :

$$G_{\sigma,\mu}(S) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(S-\mu)^2}{2\sigma^2}} \quad (5.9)$$

La Figure 5.4(b) montre l'image originale obtenue pour un bruit gaussien de variance $\sigma = 10$ et la Figure 5.4(c) pour $\sigma = 20$. Dans un second temps, nous avons considéré un bruit non-uniforme, le bruit de Poisson, présent notamment en imagerie médicale, pour estimer les performances de notre

approche (Figure 5.4(d)). Notre image est bruitée par un bruit de Poisson dont la loi de probabilité s'écrit en chaque pixel :

$$P_{\lambda}(S) = \frac{\lambda^S}{S!} e^{-\lambda} \quad (5.10)$$

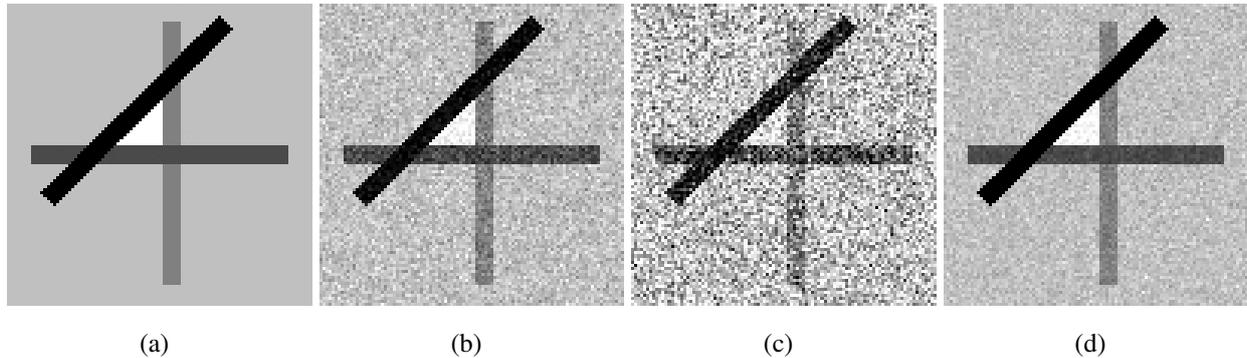


FIG. 5.4 – (a) Image de synthèse originale non bruitée. (b) Ajout d'un bruit gaussien de variance $\sigma = 10$. (c) Ajout d'un bruit gaussien de variance $\sigma = 20$. (d) Ajout d'un bruit de Poisson.

La Figure 5.5 présente un comparatif des résultats obtenus avec le bruit gaussien croissant. Lorsque l'image ne contient pas de bruit, toutes les approches sont capables, sans seuillage, de détecter correctement les contours de l'image (première ligne de la Figure 5.5). Ajoutons donc un bruit blanc gaussien (représentant le bruit d'acquisition) dans nos images. On peut observer en comparant toutes ces images que, sans seuillage ((j) et (r) pour Deriche, (l) et (t) pour Shen, (n) et (v) pour Prewitt, (p) et (x) pour notre approche), seule notre approche est capable de détecter correctement les contours dans l'image. En effet, pour éliminer le bruit présent dans les images, les approches classiques font l'hypothèse que le bruit est faible comparé au signal. Le choix d'un seuillage arbitraire permet de supprimer toutes ces valeurs parasites inférieures aux valeurs du signal. Ces approches s'avèrent relativement efficaces lorsque le niveau de bruit est bien inférieur au niveau du signal ((k) pour Deriche, (m) pour Shen, (o) pour Prewitt). Cependant, lorsque le bruit est ramené au même niveau que le signal ((s) pour Deriche, (u) pour Shen, (w) pour Prewitt), elles détectent autant de bruit que de signal. Notre approche est capable de discerner le signal du bruit. Ainsi la plupart des points de bruit sont supprimés afin de ne conserver que les contours (p). Lorsque le bruit et le signal sont indissociables, notre approche ne détectera que des points de bruit qui seront supprimés (x). Avec un bruit de Poisson, la plupart des approches sont incapables de discerner le signal et le bruit comme nous pouvons le voir sur la Figure 5.6. Nous pouvons constater que notre approche en revanche fournit des contours tout à fait pertinents et ne détecte pas ou très peu ce bruit.

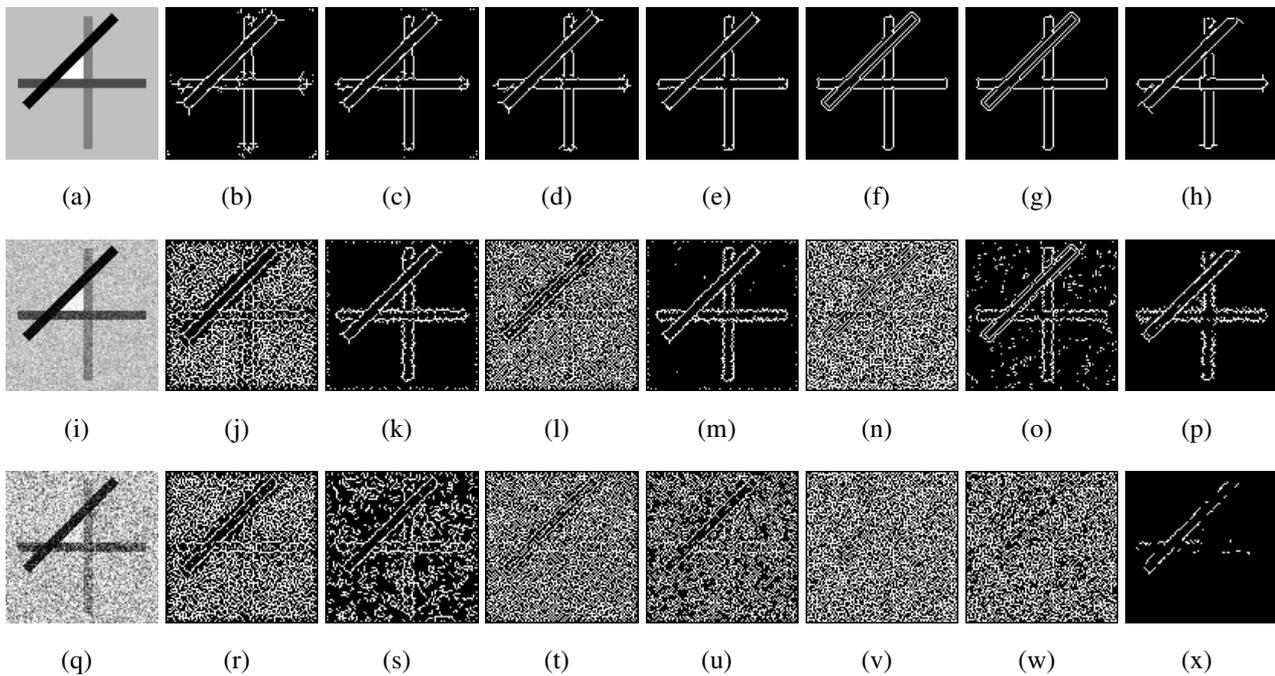


FIG. 5.5 – Image originale (a) et les contours obtenus avec un filtre de (b) Deriche sans seuillage, (c) Deriche avec un seuillage arbitraire adapté $s_h = 35$ et $s_b = 30$. (d) Shen sans seuillage, (e) Shen avec un seuillage arbitraire adapté $s_h = 24$ et $s_b = 20$. (f) Prewitt sans seuillage, (g) Prewitt avec un seuillage arbitraire adapté $s_h = 120$ et $s_b = 115$. (h) avec notre approche sans seuillage. (i) et (q) représentent l'image originale avec ajout d'un bruit gaussien tel que $\sigma = 10$ et $\sigma = 20$. Les mêmes estimations que (b), (c), (d), (e), (f), (g) et (h) sont fournies pour l'image (i) en (j), (k), (l), (m), (n), (o) et (p) et pour l'image (q) en (r), (s), (t), (u), (v), (w) et (x).

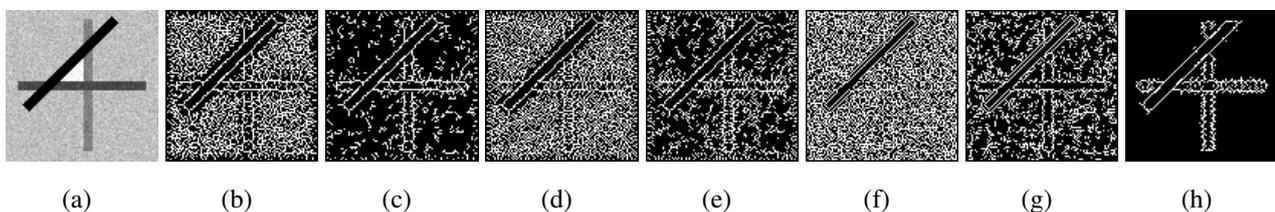
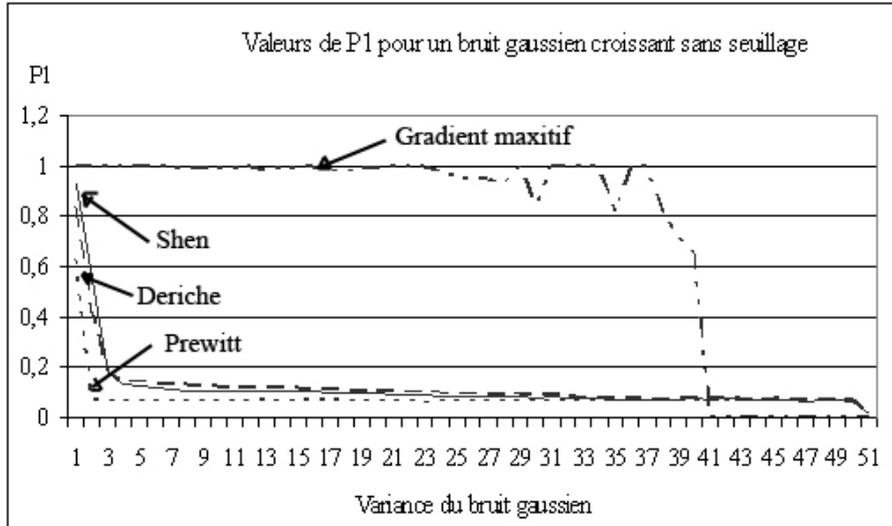


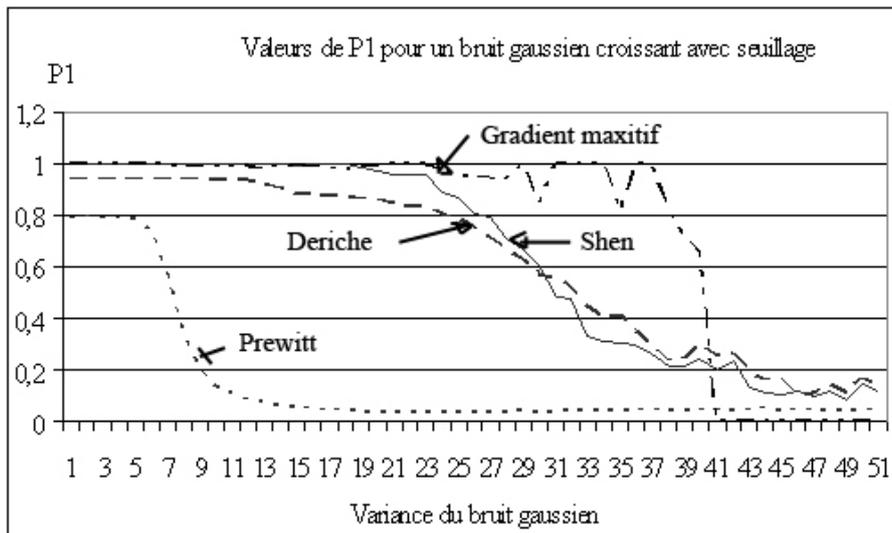
FIG. 5.6 – (a) représente l'image originale 5.4(a) avec ajout d'un bruit de poisson et les contours obtenus avec un filtre de Deriche sans seuillage, (b) Deriche avec un seuillage arbitraire adapté $s_h = 12$ et $s_b = 10$. (c) Shen sans seuillage, (d) Shen avec un seuillage arbitraire adapté $s_h = 8$ et $s_b = 6$. (e) Prewitt sans seuillage, (f) Prewitt avec un seuillage arbitraire adapté $s_h = 60$ et $s_b = 55$. (g) avec notre approche sans seuillage.

5.3.1.2 Evolution du paramètre P_1 en présence de bruit

Pour estimer quantitativement la robustesse au bruit sur les images présentées ci-dessus, nous avons utilisé le paramètre P_1 de Fram et Deutsch présenté Section 5.1.1. Pour la première expérience, avec un bruit gaussien de variance croissante appliqué à la Figure 5.4(a), on n'effectue pas de seuillage. La Figure 5.7(a) montre que toutes les approches échouent à part la nôtre pour $\sigma > 3$.



(a)



(b)

FIG. 5.7 – Valeurs comparatives de P_1 pour un bruit gaussien. (a) Sans seuillage, (b) Avec seuillage (pour toutes les approches excepté l'approche du Gradient maxitif qui fournit un seuillage naturel).

Nous avons ensuite appliqué un seuillage arbitraire de manière à obtenir les meilleurs résultats possibles pour chaque approche. La Figure 5.7(b) montre que le filtre de Prewitt est le plus sensible

au bruit. Les approches de Canny-Deriche et Shen-Castan ont des comportements similaires : pour $\sigma < 25$, un contour optimal est détecté. Ensuite les courbes déclinent : les contours sont de plus en plus bruités. Pour $\sigma > 40$, les contours et le bruit sont indissociables. Avec notre approche, un contour optimal est détecté pour $\sigma < 30$. Lorsque $\sigma > 40$, toutes les valeurs estimées, en chaque point, pour le gradient sont liées au bruit, donc considérées comme nulles.

Dans un second temps, un bruit de Poisson a été appliqué à l'image. Nous avons traité cette image et reporté les valeurs obtenues dans la Table 5.1. Ces résultats montrent que notre approche est 8 fois moins sensible au bruit que le filtre de Prewitt et 5 fois moins que les approches de Canny-Deriche et Shen-Castan.

Approche	P_1
Prewitt	0.089676
Canny-Deriche	0.14975
Shen-Castan	0.141
Gradient maxitif	0.69881

TAB. 5.1 – Valeurs de P_1 avec un bruit de Poisson.

Dans le Chapitre 4, nous avons expliqué, avec l'exemple de la Figure 4.7, que pour des images contenant un bruit non-uniforme, il était impossible de trouver de bonnes valeurs de seuil qui préservent les contours et suppriment le bruit. Des résultats complémentaires sont présentés Figure 5.8. Nous avons présentés les résultats obtenus avec les filtres de Deriche, Shen et Prewitt pour différentes valeurs de seuil. Les meilleures valeurs de seuil sont celles permettant de visualiser entièrement les 3 disques de la rosace. Nous les avons estimées et avons ainsi obtenus les images (c), (g) et (k), dans lesquelles les 3 disques sont totalement détectés. En augmentant ces valeurs de seuil, la quantité de bruit présente dans l'image diminue, mais des contours sont également supprimés. La Figure 5.8(b) illustre les résultats obtenus en traitant l'image (a) avec l'approche du Gradient maxitif. On peut remarquer que dans cette image, les 3 disques sont correctement détectés et que, de plus, l'image contient très peu de contours indésirables.

5.3.1.3 Estimation des paramètres P_1 en fonction de l'intensité du contour

Pour cette expérience, nous avons repris l'image test présentée Figure 5.4(a) et nous l'avons légèrement modifiée. Cette image ne contient plus que deux valeurs de niveaux de gris, le premier pour le fond et le second pour la croix. La valeur du fond est de 147 niveaux de gris. La croix est définie pour des sauts de 3, 6, 12, 20 et 100 niveaux de gris. Dans la réalité, l'information peut subir des dégradations venant de la chaîne d'acquisition ou de la scène elle-même. Tout opérateur de détection de contours doit donc faire preuve de robustesse en présence du bruit. Afin d'étudier cela,

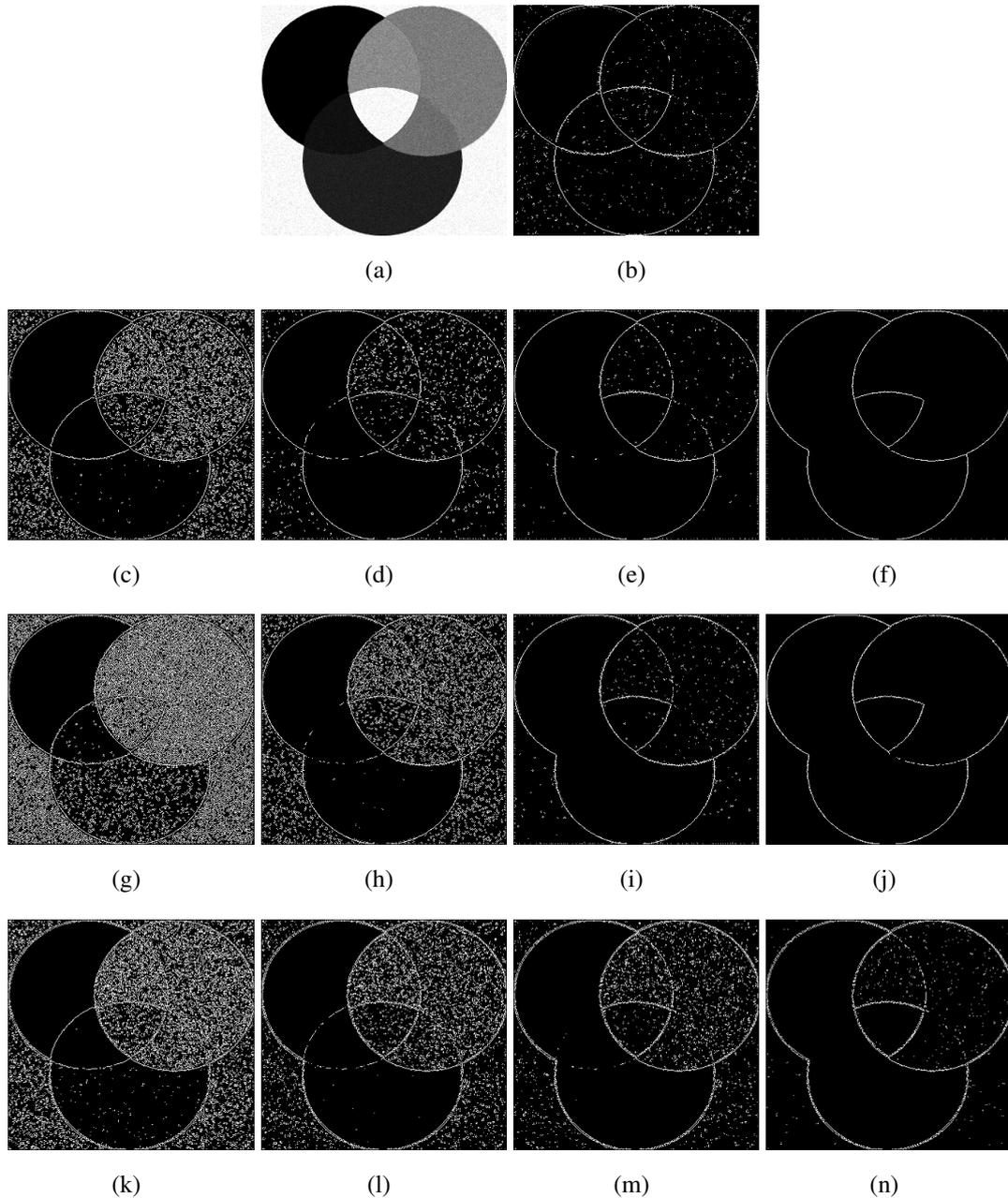


FIG. 5.8 – (a) Image originale bruitée avec un bruit de Poisson. (b) Résultat obtenu avec notre approche sans seuillage. Les résultats obtenus avec un filtre de Deriche sont présentés pour différentes valeurs de seuil : (c) $s_h = 10$ et $s_b = 8$, (d) $s_h = 20$ et $s_b = 18$, (e) $s_h = 25$ et $s_b = 23$, (f) $s_h = 108$ et $s_b = 105$. Les résultats obtenus avec un filtre de Shen sont présentés pour différentes valeurs de seuil : (g) $s_h = 1$ et $s_b = 0$, (h) $s_h = 3$ et $s_b = 2$, (i) $s_h = 10$ et $s_b = 8$, (j) $s_h = 14$ et $s_b = 13$. Les résultats obtenus avec un filtre de Prewitt sont présentés pour différentes valeurs de seuil : (k) $s_h = 34$ et $s_b = 33$, (l) $s_h = 48$ et $s_b = 46$, (m) $s_h = 60$ et $s_b = 55$, (n) $s_h = 83$ et $s_b = 79$.

les images ont été bruitées par un bruit gaussien de moyenne nulle et d'écart-type égal à 2 niveaux de gris. La Figure 5.9 illustre les images que nous avons utilisées.

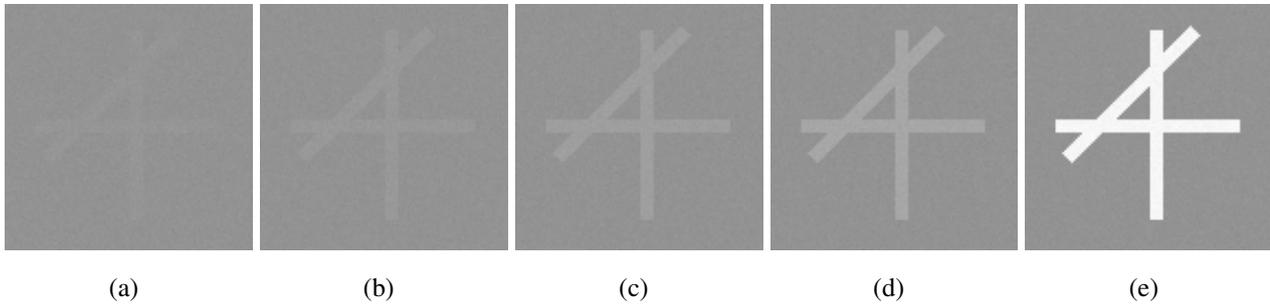
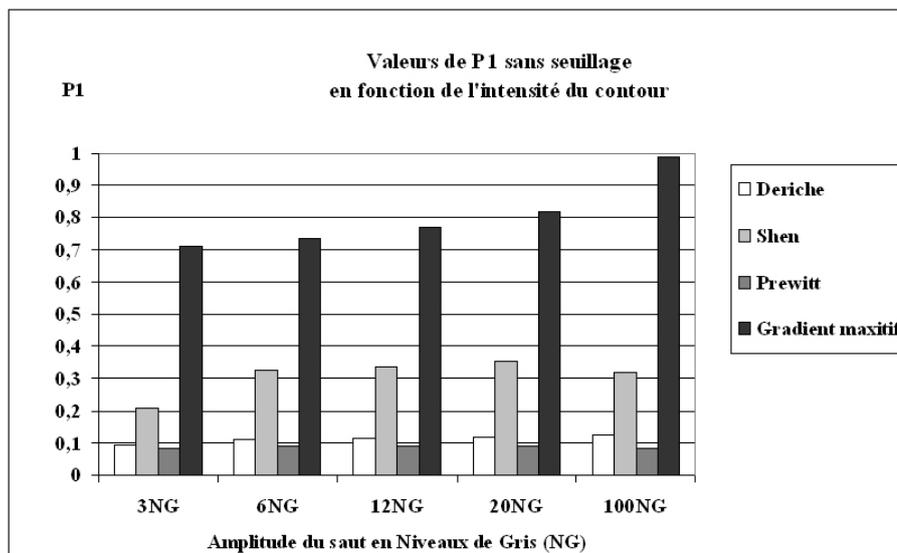


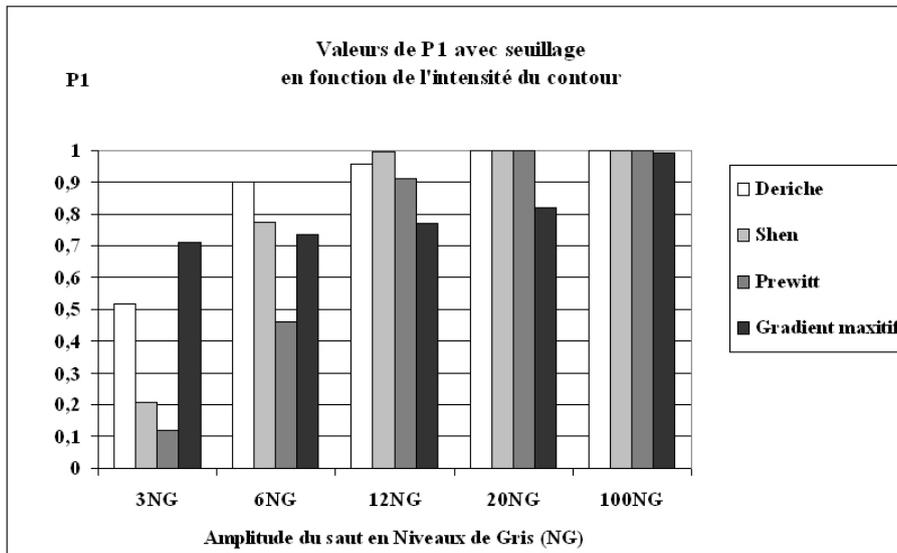
FIG. 5.9 – Image de la croix pour des sauts de (a) 3 niveaux de gris, (b) 6 niveaux de gris, (c) 12 niveaux de gris, (d) 20 niveaux de gris et (e) 100 niveaux de gris.

Si aucun seuillage n'est appliqué, seule notre approche réussit à différencier les points de contours des points de bruit (Figure 5.10(a)). Nous obtenons donc de meilleurs résultats que l'ensemble des méthodes classiques. Les traitements classiques nécessitent une étape de seuillage arbitraire. La Figure 5.10(b) montre les résultats obtenus avec le seuillage arbitraire permettant de détecter le contour le plus complet possible. Avec un saut de niveaux de gris de faible amplitude, notre approche élimine la plus grande partie du bruit, ce qui lui permet de fournir de très bons résultats, même si le contour détecté est incomplet. Lorsque l'amplitude du saut augmente, on peut remarquer que les approches classiques sont plus efficaces que notre approche. Ce n'est que pour un saut supérieur ou égal à 20 niveaux de gris que les approches fournissent des résultats similaires.



(a)

FIG. 5.10 – (a) Valeurs de P_1 obtenues sans seuillage des images.

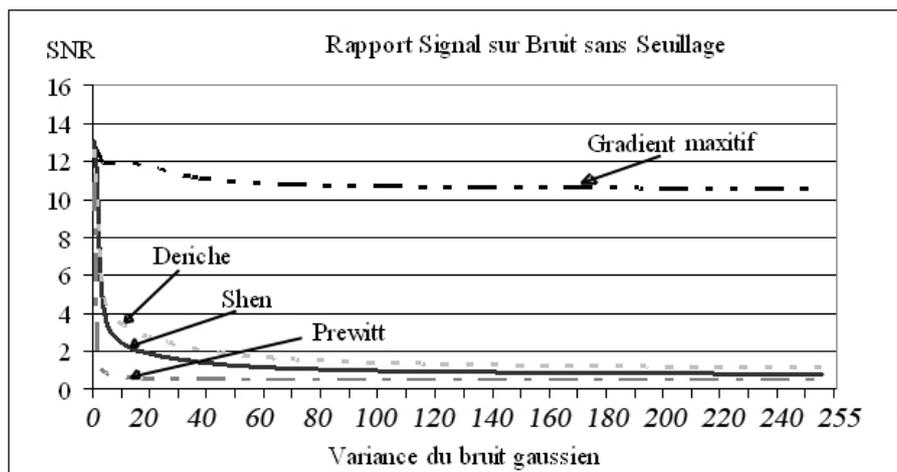


(a)

FIG. 5.10 – (b) Valeurs de P_1 obtenues avec seuillage des images (sauf pour l'approche Gradient maxitif).

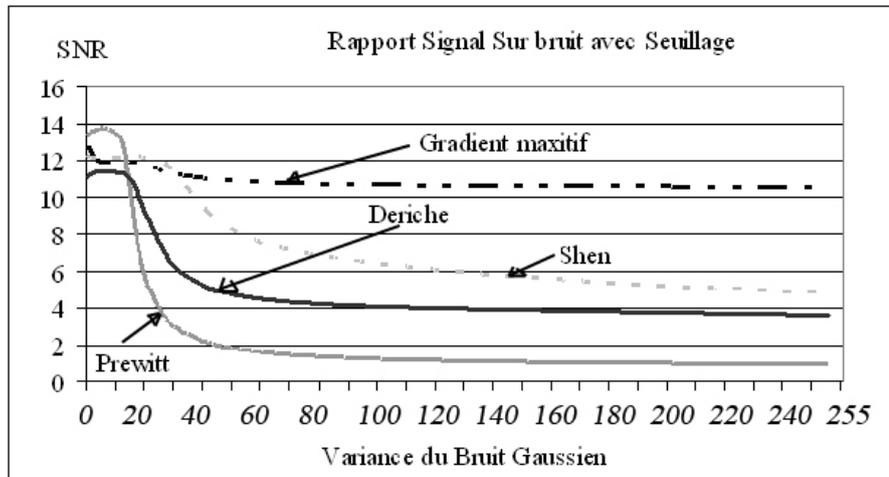
5.3.1.4 Critère de bonne détection

Toujours en utilisant la Figure 5.4(a), nous avons estimé le rapport signal sur bruit obtenu d'abord sans seuillage (Figure 5.11(a)), puis en appliquant un seuillage arbitraire (Figure 5.11(b)). Sans seuillage, nous pouvons observer que l'ensemble des approches classiques donne un très mauvais rapport signal sur bruit pour $\sigma \geq 2$. Seule notre approche avec son seuillage naturel est capable de détecter les contours présents dans l'image et fournit de très bons résultats. En seuillant l'image, les



(a)

FIG. 5.11 – (a) Rapports signal sur bruit obtenus sans seuillage.



(a)

FIG. 5.11 – (b) Rapports signal sur bruit obtenus avec seuillage.

résultats obtenus avec les approches classiques sont meilleurs. Cependant, on peut remarquer que leur rapport signal sur bruit reste toujours inférieur à celui obtenu avec notre approche.

5.3.1.5 Mesure des écarts

Pour mesurer les écarts des pixels à la droite théorique, nous avons défini une image très simple constitué d'un unique contour d'équation $y = x\sqrt{3}$.

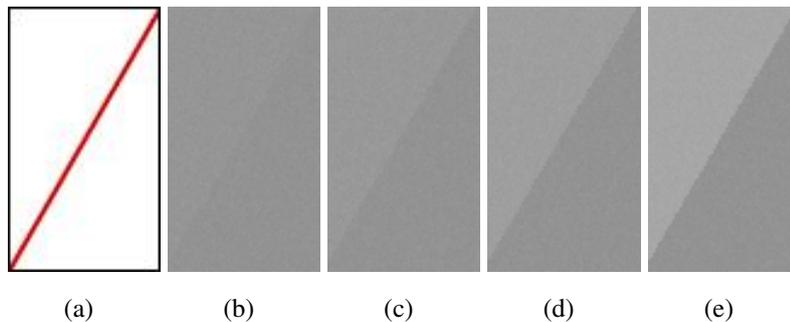
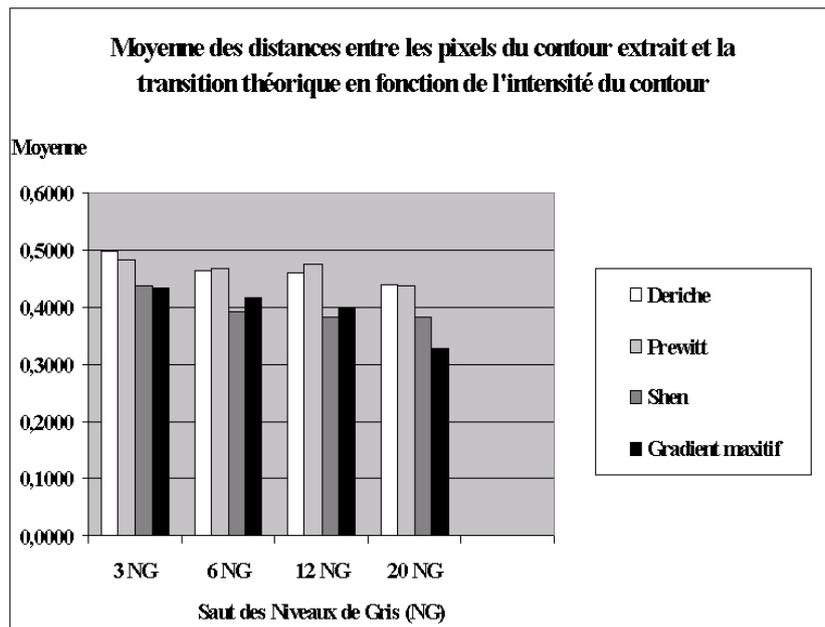


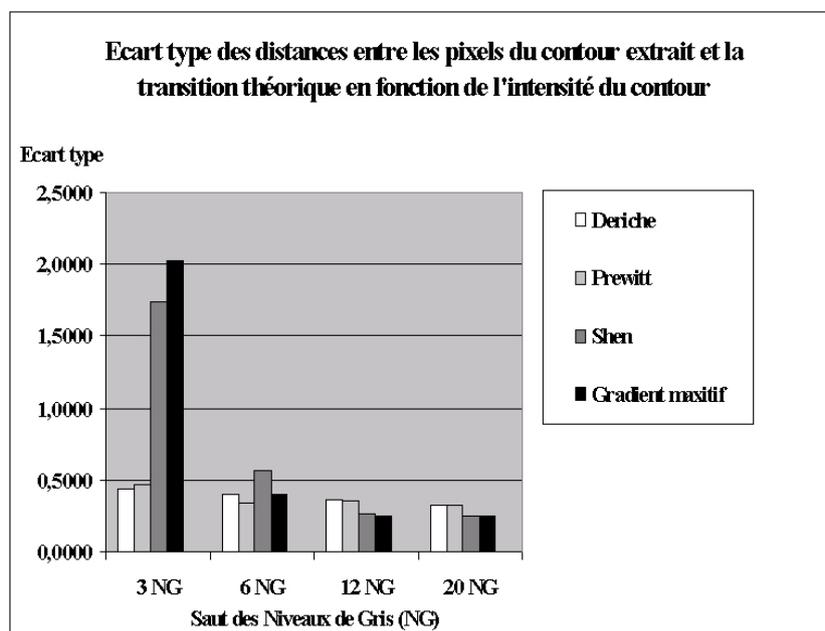
FIG. 5.12 – (a) Contour théorique. Les images obtenues pour des sauts de (b) 3 Niveaux de gris, (c) 6 Niveaux de gris, (d) 12 Niveaux de gris, (e) 20 Niveaux de gris.

La Figure 5.12 illustre les images utilisées pour cette expérience. Ces images ont des sauts d'amplitude de 3, 6, 12 et 20 niveaux de gris. Comme précédemment, nous avons ajouté à ces images un bruit gaussien de variance $\sigma = 2$. Nous avons étudié les moyennes et écarts-types obtenus entre le contour théorique et le contour détecté dans ces images pour différentes approches. Les résultats sont présentés sur les Figures 5.13(a) et (b). On peut observer qu'avec un saut d'amplitude faible, l'approche maxitive n'est pas très performante. L'écart-type des pixels avec le contour théorique est



(a)

FIG. 5.13 – (a) Moyennes obtenues pour des sauts d'amplitude 3, 6, 12 et 20 niveaux de gris pour chaque approches.



(a)

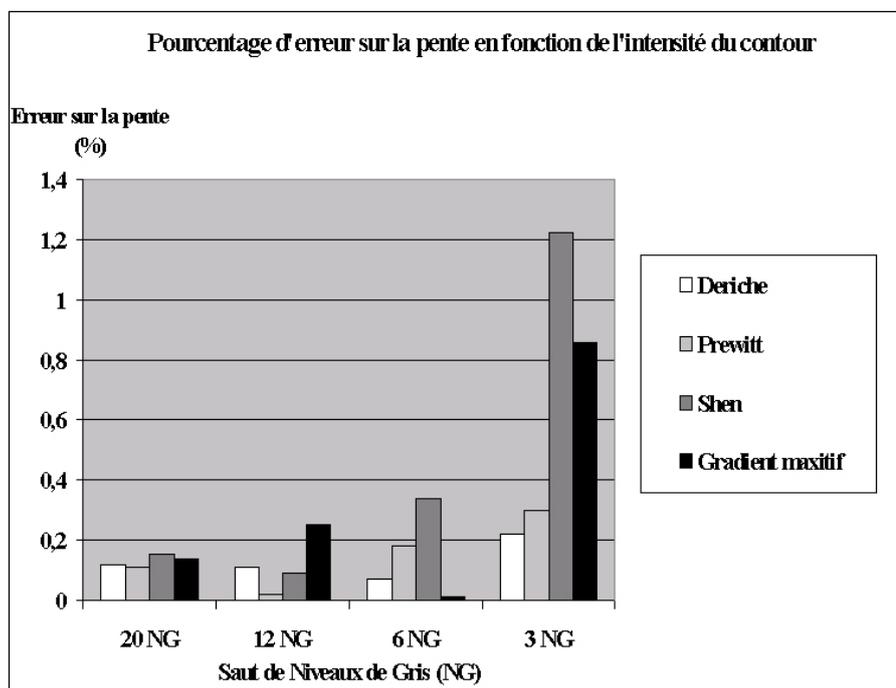
FIG. 5.13 – (b) Ecart-types entre les pixels du contour extrait et la transition théorique obtenus pour des sauts d'amplitude 3, 6, 12 et 20 niveaux de gris pour chaque approche.

très important. Avec des sauts inférieurs à 5 niveaux de gris, les pixels détectés sont très éloignés du contour théorique. En effet, l'utilisation de noyaux maxitifs introduit une sorte de lissage qui rend la détection de contours difficile lorsque l'intensité du contour est trop faible. Cependant, dès que le contour est légèrement plus contrasté, on peut remarquer que cette méthode fournit de meilleurs résultats que les approches classiques. Le contour obtenu est plus proche du contour théorique.

5.3.1.6 Régression linéaire

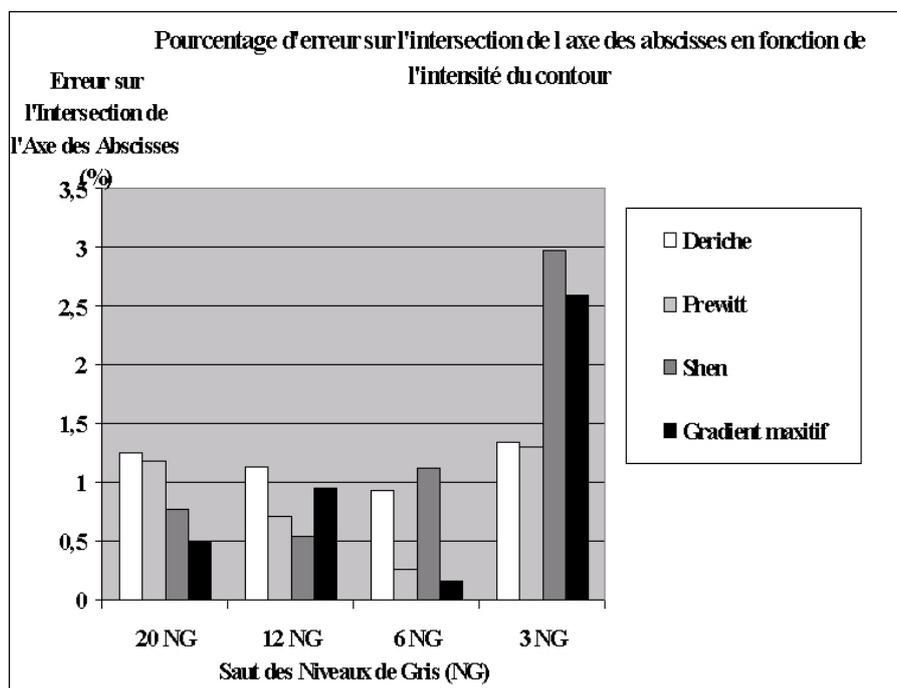
Pour estimer la bonne localisation du contour, nous avons estimé le pourcentage d'erreur entre la droite théorique et la droite détectée. L'erreur effectuée sur la pente de la droite est donnée par $|\frac{a-c}{a}|$ et celle effectuée sur l'ordonnée à l'origine par $|\frac{b-d}{b}|$.

On peut constater le même phénomène que pour la mesure des écarts. Pour de faibles niveaux de gris, l'approche maxitive fournit des résultats médiocres. Cependant, dès que l'amplitude du saut de niveaux de gris est suffisamment importante, nous obtenons de meilleurs résultats qu'avec les approches classiques.



(a)

FIG. 5.14 – (a) Pourcentage d'erreur sur la pente de la droite détectée pour des sauts d'amplitude 20, 12, 6 et 3 niveaux de gris pour chaque approche.



(a)

FIG. 5.14 – (b) Pourcentage d'erreur sur l'ordonnée à l'origine pour des sauts d'amplitude 20, 12, 6 et 3 niveaux de gris pour chaque approche.

5.3.2 Robustesse de l'approche du Gradient maxitif sur des images réelles

Sur les images artificielles, nous avons obtenu des résultats très prometteurs pour cette nouvelle approche. Dans cette section, notre approche est appliquée à une image réelle présentée Figure 5.15. Pour comparer qualitativement nos résultats, nous avons zoomé sur cinq régions de l'image.

Les résultats ont été présentés Figure 5.16. La première ligne représente les 5 parties sur lesquelles nous avons zoomé. Afin de fournir une meilleure visualisation au lecteur des contours que nous cherchons à détecter, la luminosité et le contraste ont été modifiés.

La seconde ligne présente les contours obtenus avec l'approche du Gradient maxitif. Rappelons que cette approche fournit un seuillage naturel de l'image. Les trois dernières lignes montrent les contours obtenus avec un filtre de Deriche pour trois seuillages différents. La troisième ligne est celle qui donne le meilleur ratio entre les contours désirables et les contours indésirables dans l'image. C'est donc cette ligne qui va être la ligne de comparaison avec notre approche.

La première image (a) représente une grille régulière. Les contours de cette grille semblent correctement détectés par notre approche (f) et par l'approche de Deriche (k). La seconde image (b) montre la même grille mais moins contrastée. De plus dans le bas de l'image, on peut observer une zone très contrastée. La totalité des contours de cette image est détectée par notre approche caméra (g). Avec l'approche de Deriche (l), la zone la plus contrastée est bien détectée, mais la grille est relativement

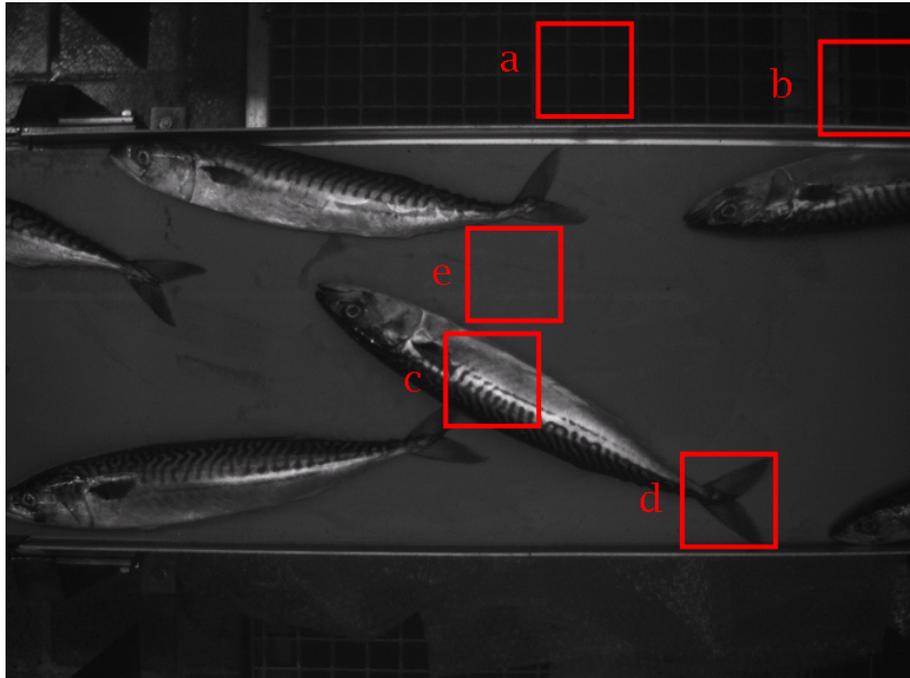


FIG. 5.15 – Image réelle originale.

incomplète. La troisième image (c) représente une partie du corps du poisson sur laquelle on peut distinguer de nombreux détails. Notre approche (h) permet de distinguer la majeure partie de ces détails. En revanche, l'approche de Deriche (m) détecte de nombreux détails supplémentaires qui ne sont pas présents sur l'image originale. La quatrième image (d), qui est également très sombre sur l'image originale, se focalise sur la queue du poisson. Celle-ci est correctement identifiée avec notre approche (i) et avec l'approche de Deriche (n). Enfin la dernière région (e) ne contient rien, si ce n'est du bruit. Notre approche (j) détecte juste les deux petites taches sombres présentes dans cette image, alors que l'approche de Deriche (o) détecte une partie des variations lumineuses présentes dans le bruit. L'image obtenue contient donc de nombreux artéfacts.

A présent, supposons que l'on souhaite détecter entièrement la grille (b), il est alors nécessaire de trouver une valeur de seuil plus basse, préservant plus de détails dans l'image. Nous avons trouvé un tel seuil comme l'illustre (r); cependant, les autres images, (p), (q), (s) et (t), contiennent de nombreux contours indésirables.

Si l'on cherche à détecter correctement les détails présents dans le corps du poisson, des valeurs de seuil plus hautes doivent être utilisées. Les contours obtenus pour l'image (c) sont alors beaucoup moins bruités et on peut distinguer de nombreux détails sur le poisson (w), cependant, sur les autres images, (u), (v), (x) et (y), de nombreux contours ont été supprimés. La détection de contour obtenue sur cette image est alors très incomplète.

Pour traiter correctement cette image avec les approches classiques, il faudrait la découper en petites imquettes et chercher pour chacune de ces imquettes les meilleures valeurs de seuil. Cette opération

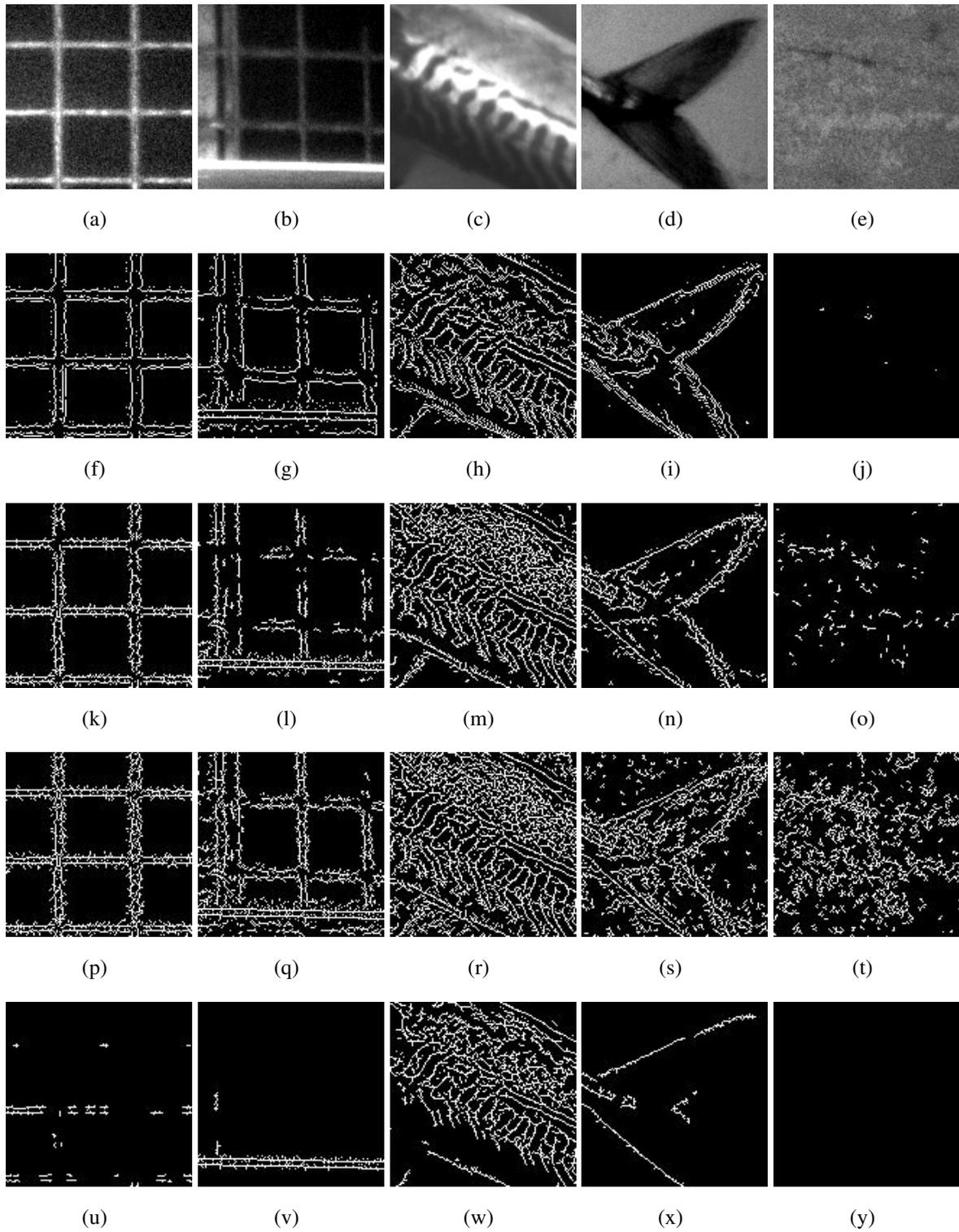


FIG. 5.16 – (a), (b), (c), (d) et (e) représentent l'image originale après réhaussement de la luminosité et du contraste afin de mieux visualiser les contours.

FIG. 5.16 – (f), (g), (h), (i) et (j) montrent les contours obtenus avec notre approche. (k), (l), (m), (n) et (o) montrent les contours obtenus avec le filtre de Deriche pour les valeurs de seuil donnant le meilleur rapport entre contours désirés et contours non-désirés $s_h = 16$ et $s_b = 13$; (p), (q), (r), (s) et (t) les images obtenues avec des valeurs de seuil plus élevées permettant de détecter correctement la grille $s_h = 30$ et $s_b = 28$; (u), (v), (w), (x) et (y) les images obtenues avec des valeurs de seuil plus faibles permettant de détecter correctement le corps du poisson $s_h = 6$ et $s_b = 4$.

n'est évidemment pas envisageable. Notre approche en revanche fournit donc de très bons résultats sur des images réelles et artificielles, quelle que soit la nature du bruit contenu dans l'image.

5.4 Comparaison des différentes approches sur des images omnidirectionnelles

5.4.1 Influence de l'étalonnage de la caméra sur les traitements

Rappelons que l'image omnidirectionnelle est très difficilement interprétable pour l'homme. Les déformations générées par le miroir amènent les utilisateurs de ces capteurs à chercher des caractéristiques particulières de l'environnement dans l'image comme les droites radiales qui correspondent aux amers verticaux. Dans la Section 1.1.4, nous avons noté les difficultés rencontrées pour étalonner correctement les capteurs omnidirectionnels. En effet, la coïncidence de l'axe de révolution du miroir avec l'axe optique de la caméra est une opération très minutieuse et complexe à réaliser. Cependant, l'influence d'un léger non-alignement entre ces axes introduit une distorsion minime dans la détection et la localisation des droites radiales dans l'image. Dans de nombreuses applications ([68], [7], [33], [67]), l'étalonnage du capteur n'a pas été mis en œuvre. Avec notre approche, il est impossible d'ignorer les paramètres d'étalonnage de l'imageur. Notre méthode est dépendante du bon étalonnage de la caméra. En effet, si un léger décalage entre les axes de la caméra et du miroir aura une incidence moindre sur notre algorithme, un décalage plus important modifiera la taille des noyaux projetés sur l'image omnidirectionnelle.

Dans cette partie, nous avons généré une image de synthèse constituée de lignes radiales avec le logiciel POVRay. Nous avons modifié les paramètres φ , ξ , k_u , k_v , u_0 et v_0 de la caméra présentés Section 4.2.3 afin de déterminer leur influence sur les traitements que nous avons proposé. Les paramètres d'étalonnage de la caméra que nous avons définie sont les suivants :

$$\left\{ \begin{array}{l} (k_u = k_v = 163.100 \\ (u_0, v_0) = (300, 400) \\ \xi = 0.94 \\ \varphi = 1.2811 \end{array} \right.$$

Avec notre algorithme, nous effectuons une rétro-projection. Ainsi, certains paramètres n'ont aucune influence sur la taille ou la position des noyaux dans l'image. φ , par exemple, fournit la distance

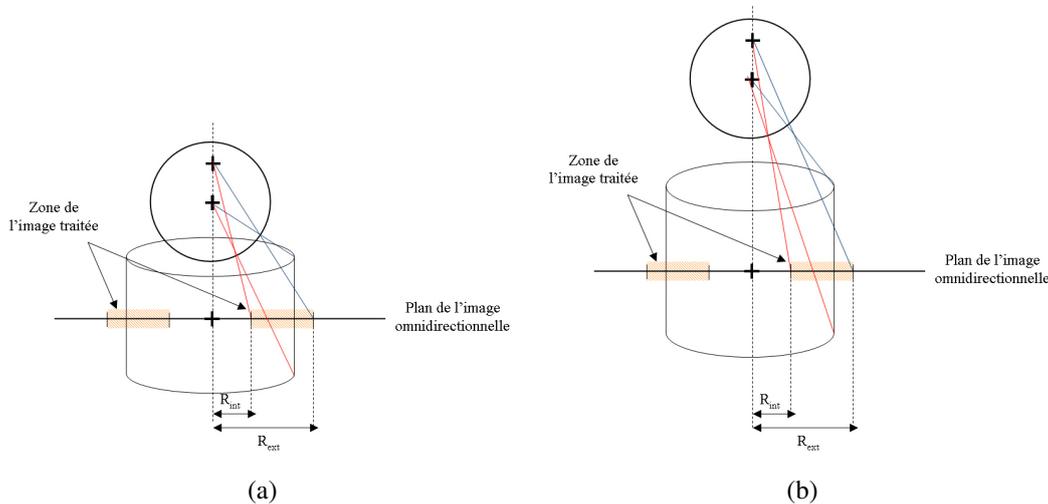


FIG. 5.17 – Influence de φ sur le système de projection.

entre l'image omnidirectionnelle et le centre de sphère. Les variations de ce paramètre engendrent une modification de la taille et de la position du cylindre dans l'espace 3D. La Figure 5.17 montre la position et la hauteur du cylindre obtenu pour deux valeurs différentes de φ . Cette modification a des répercussions au niveau du calcul des coordonnées des points projetés sur le cylindre. Cependant, la hauteur des noyaux étant définie en fonction de la hauteur du cylindre, les noyaux seront toujours adaptés à l'image quelle que soit la valeur du paramètre φ . Notre détection de contours ne sera pas affectée par une erreur sur ce paramètre.

ξ , au contraire, représente la distance entre le centre de la sphère et le point appartenant à l'axe de la sphère. C'est elle qui détermine le type de capteur catadioptrique central utilisé, en fonction de sa position par rapport au centre de la sphère. La modification de ξ introduira donc des distorsions importantes dans l'image. En effet, en le modifiant au hasard, nous risquons de transformer notre miroir hyperbolique en miroir parabolique ou elliptique. Nous avons représenté sur la Figure 5.18 les images obtenues pour différentes valeurs de ξ .

k_u et k_v indiquent les rapports relatifs à la taille des pixels dans le plan. Nos pixels étant carrés, on a $k_u = k_v$. Tant que cette égalité est conservée, nous restons cohérent avec notre image à un facteur d'échelle près. La détection de contours ne sera donc pas modifiée tant que $k_u = k_v$. En revanche, en prenant $k_u \neq k_v$, le rapport relatif à la taille des pixels dans le plan n'est plus conservé, on va donc avoir une déformation horizontale ou verticale sur les noyaux. Un exemple des déformations verticales obtenues par modification de ces paramètres est présenté Figure 5.19.

(u_0, v_0) fournissent les coordonnées de projection de l'axe optique dans l'image. Tous les noyaux

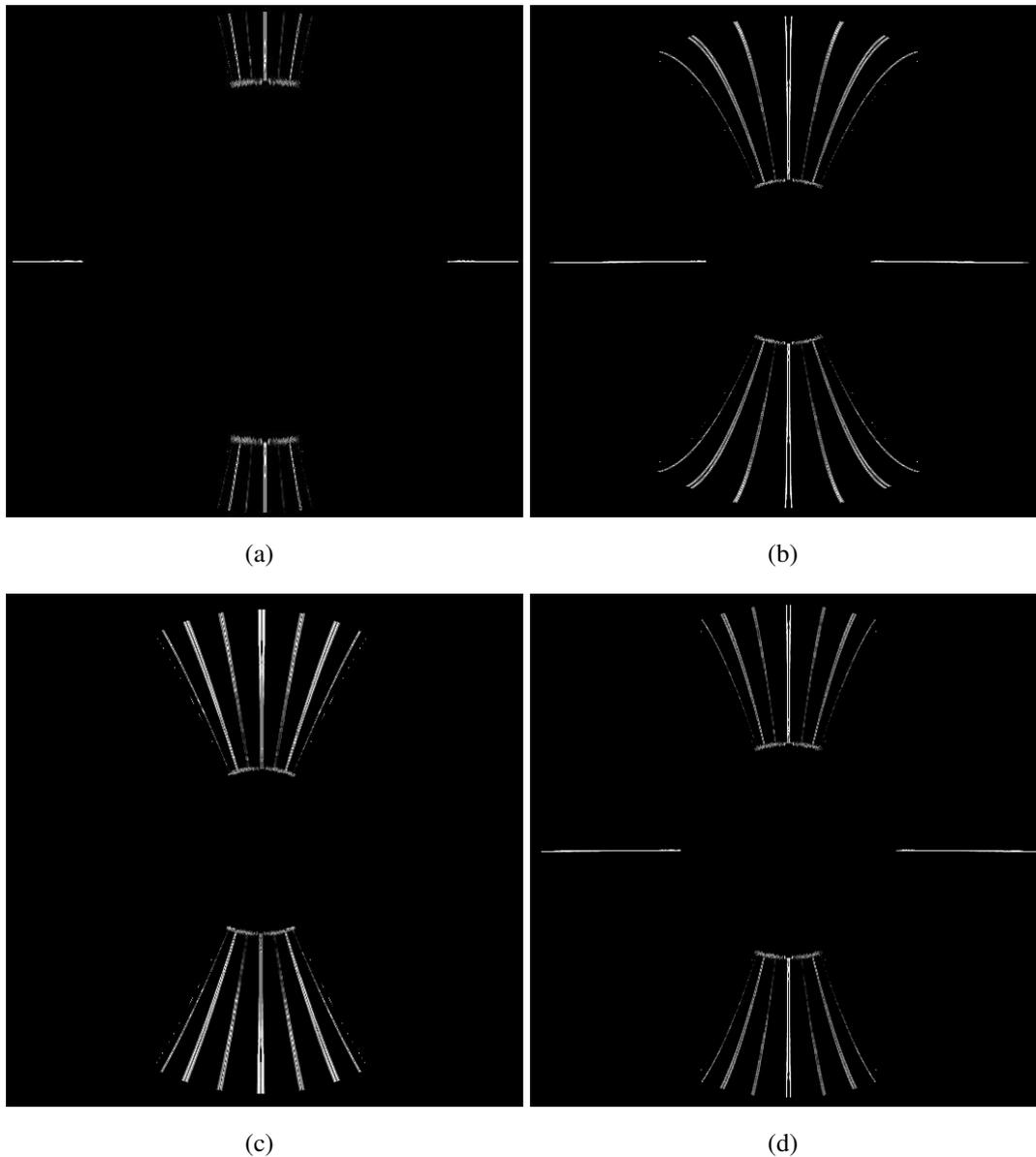


FIG. 5.18 – Influence de ξ sur l'image pour (a) $\xi = 1.54$, (b) $\xi = 1.88$, (c) $\xi = 2$, et (d) $\xi = 2.35$.

projetés sont calculés en fonction de ces coordonnées, comme le montre la Figure 5.20(a). Une modification de ces coordonnées impliquerait une mauvaise connaissance de la position de la caméra par rapport à l'image. Tous nos noyaux seraient alors légèrement décalés par rapport à l'image (Figure 5.20(b)). En pratique, la taille des noyaux étant relativement petite, il est difficile de voir, à l'oeil nu, les variations sur les contours obtenus, lorsque l'on fait une erreur de quelques pixels sur u_0 et/ou v_0 comme l'illustre la Figure 5.21(a). En revanche, si cette erreur est suffisamment importante, elle est très rapidement repérable. En effet, l'image omnidirectionnelle formant une couronne, une erreur sur les coordonnées de u_0 et/ou v_0 implique une translation de la couronne considérée par

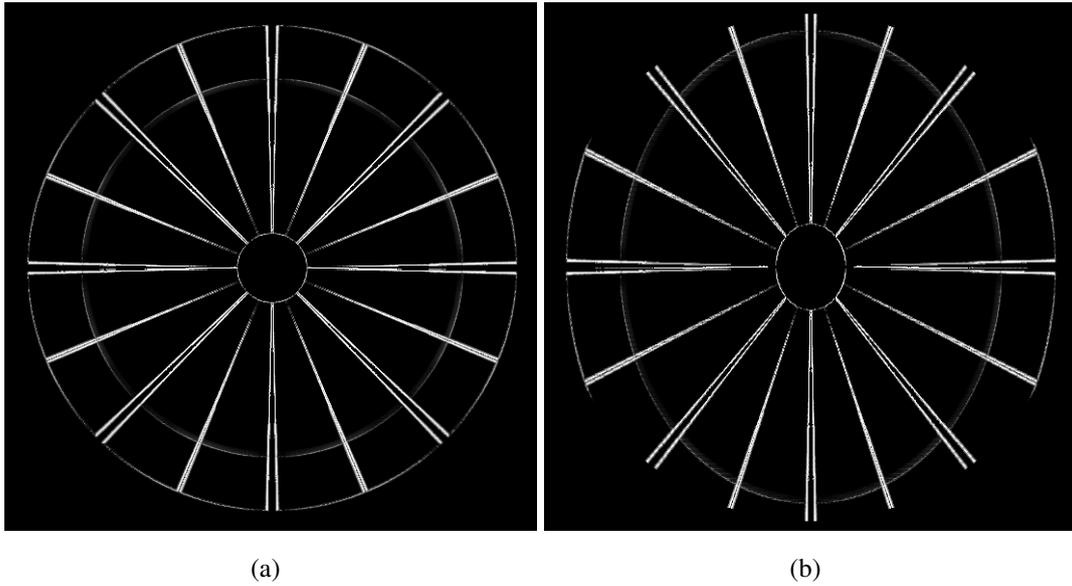


FIG. 5.19 – Détection de contour obtenue avec (a) les paramètres de calibration correctes de la caméra (b) avec $k_u = 163.100$ et $k_v = 203.875$.

rapport à l'image. Or sur notre image omnidirectionnelle, on peut voir clairement la position de la caméra. Ce décalage sera donc visible directement sur l'image. Par exemple, pour la Figure 5.21(b), un décalage de 50 pixels a été effectué sur l'image.

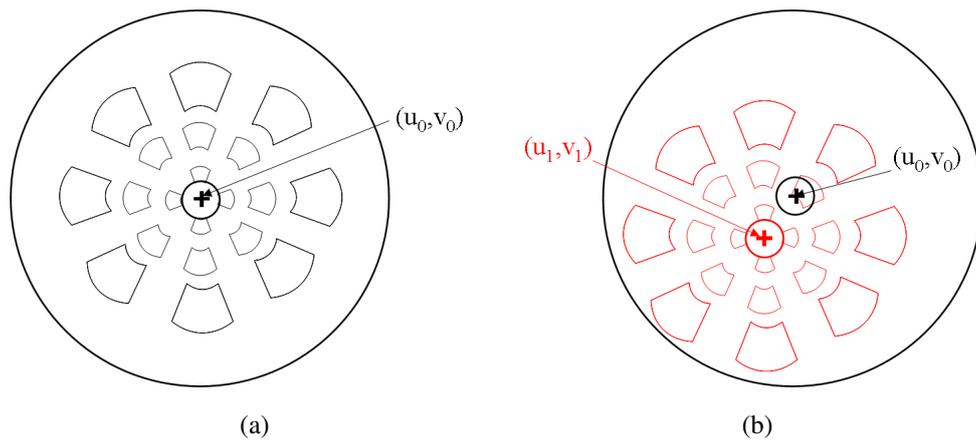


FIG. 5.20 – Exemples de noyaux projetés (a) avec les bonnes coordonnées de projection de l'axe optique dans l'image et (b) avec des coordonnées erronées.

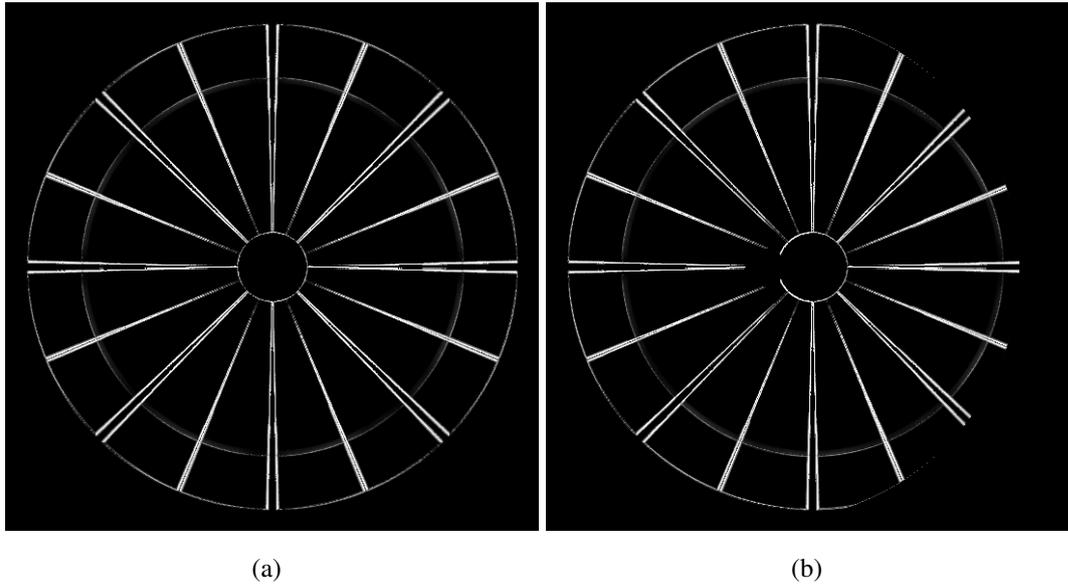


FIG. 5.21 – Exemples de noyaux projetés (a) avec $(u_0, v_0) = (300, 390)$ et (b) avec $(u_0, v_0) = (300, 350)$.

5.4.2 Influence de la taille du noyau sur les traitements

Avec notre approche projective, le masque est défini sur le cylindre. Lors de sa projection, il subit les déformations liées au miroir. Il recouvrira donc plusieurs pixels de l'image omnidirectionnelle. Une question importante réside dans le choix de la taille du masque sur le cylindre. Ce facteur joue sur la qualité de l'image résultante.

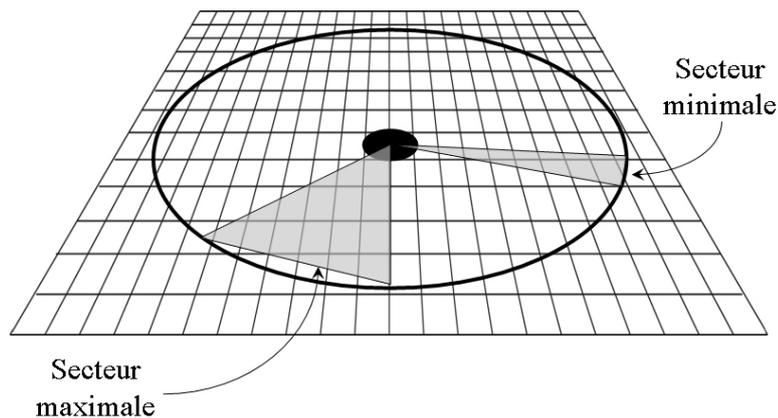


FIG. 5.22 – Secteur angulaire minimal et maximal du noyau.

La taille angulaire du noyau dépend de la partie de l'image que l'on veut privilégier. Si on privilégie la région centrale de l'image, on prendra un large secteur angulaire de manière à ce qu'un pixel

proche du centre soit entièrement inclus dans ce secteur. Si au contraire on privilégie la périphérie de l'image, le secteur permettant d'inclure un pixel sera beaucoup plus étroit (Figure 5.22).

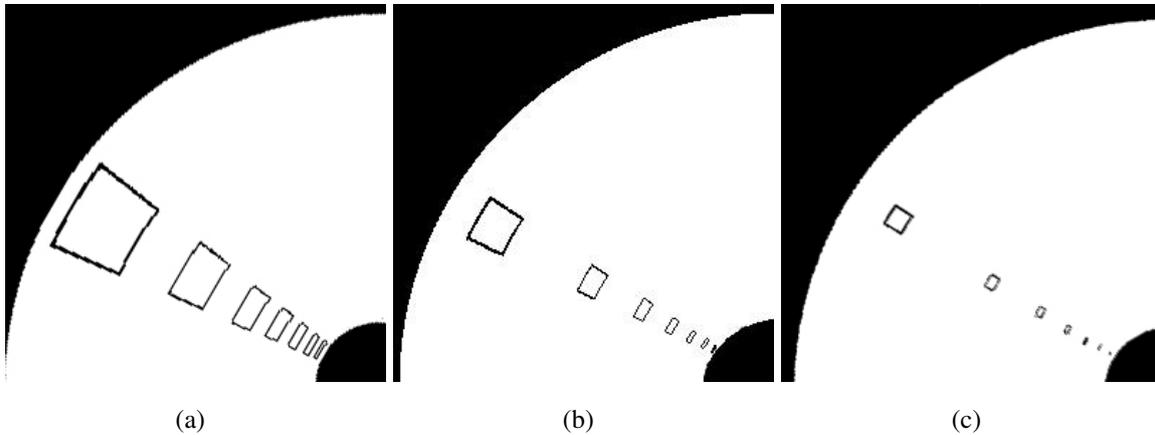


FIG. 5.23 – Projections de noyaux de différentes tailles sur l'image omnidirectionnelle : (a) Noyaux de grande taille (40 pixels) (b) Noyaux de taille moyenne (20 pixels) (c) Noyaux de petite taille (10 pixels).

La Figure 5.23 montre les projections obtenues avec trois tailles de noyaux différentes. Dans la première (a), la taille des noyaux couvre 40 fois la taille d'un pixel sur le cylindre ; dans la seconde (b), 20 fois la taille d'un pixel et dans la troisième (c), seulement 10 fois la taille d'un pixel. Afin de bien visualiser le phénomène de projections, les noyaux ont volontairement été choisis très gros comparés aux noyaux réellement utilisés qui ne recouvrent que quelques pixels.

Lorsque la taille des noyaux augmente, ils recouvrent un plus grand nombre de pixels dans l'image omnidirectionnelle. L'estimation de la valeur du pixel filtré contiendra moins de bruit, mais le signal obtenu sera flou, lissé et le niveau de détail réduit. Au contraire, en utilisant des noyaux de plus petite taille, on obtiendra un signal filtré plus détaillé mais aussi plus bruité. Il faut également noter qu'en prenant un masque de taille extrêmement petite, ce masque risque d'être entièrement inclus dans un pixel, n'apportant alors aucune information sur le voisinage et sur les contours dans l'image.

La Figure 5.24 illustre ces phénomènes sur une image de synthèse (a). Si la taille du masque est trop petite (b), un masque associé à un pixel proche du centre de l'image risque d'être entièrement contenu dans ce pixel. Aucun filtrage ne sera alors possible au centre de l'image. En revanche, si le masque est trop large (c), les contours détectés à la périphérie de l'image subiront un lissage très important. Ils seront donc plus difficiles à identifier et plusieurs contours proches risquent de fusionner entre eux et d'être perdus. Nous avons donc déterminé la taille du masque arbitrairement en fonction des contours recherchés (d). Sa largeur sur le cylindre est choisie proportionnelle au périmètre intérieur de l'image catadioptrique (exprimé en pixels). Sa hauteur fonction de la largeur (en pixels) de la couronne qui constitue l'image omnidirectionnelle ($Rayon_{exterieur} - Rayon_{interieur}$) dans l'image omnidirectionnelle et de la hauteur du cylindre.

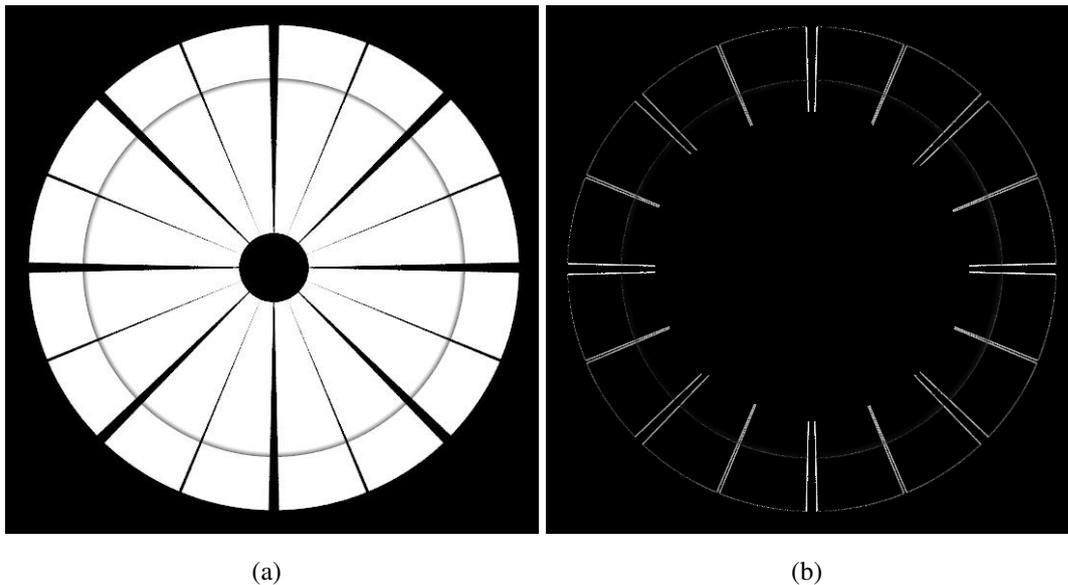


FIG. 5.24 – (a) Image originale. (b) Détection de contours obtenue avec un masque quatre fois plus petit que celui que nous avons utilisé.

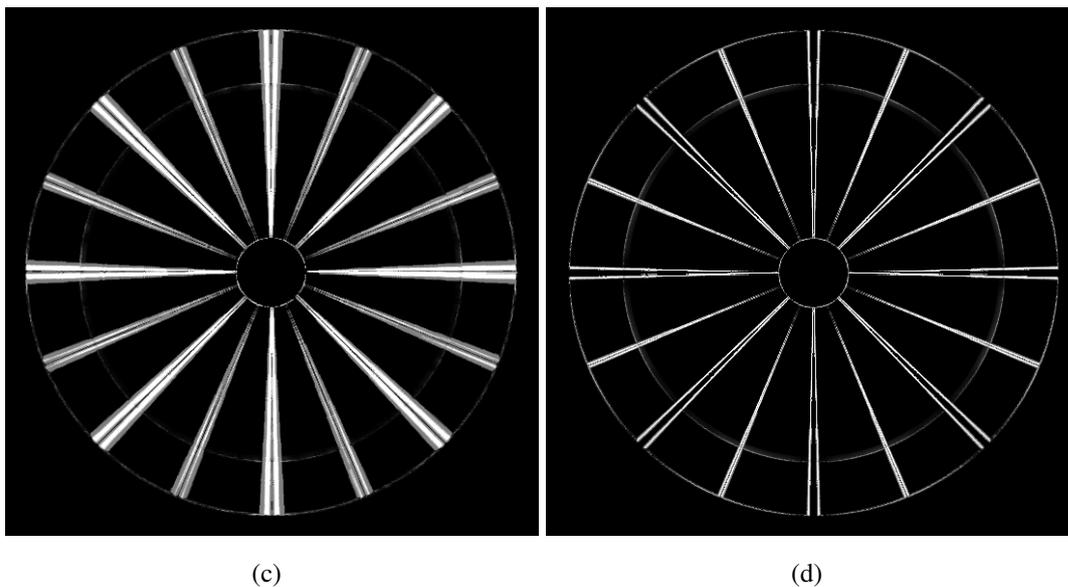


FIG. 5.24 – (c) Détection de contours obtenue avec un masque quatre fois plus grand que celui que nous avons utilisé, (d) la taille du masque que nous avons utilisé, résulte d'un bon compromis entre le lissage et la précision.

Sur l'image originale, on peut également remarquer la présence d'une ligne circulaire (horizontale dans la scène observée). Les valeurs des niveaux de gris de cette ligne ont volontairement été choisies en dégradé du blanc au gris au blanc. Lorsque la taille des noyaux est adaptée à l'image, cette ligne est correctement détectée. Cependant, avec des noyaux trop gros ou trop petits, celle-ci est

difficilement détectable car dans le premier cas, elle a été lissée et dans le second, les noyaux ne sont pas assez informatifs sur le voisinage des pixels.

5.4.3 Comparaison quantitative sur des images artificielles

La suite de cette section présente les performances des détecteurs de contours sur des images omnidirectionnelles. Dans un premier temps, nous avons testé nos approches sur des images artificielles. Cinq images de dimensions 600×800 pixels ont été générées à l'aide du logiciel POV-Ray, pour cinq valeurs de saut d'amplitude : 3, 6, 12, 20 et 100 niveaux de gris. L'intensité de la zone la plus sombre est de 147 et reste invariante. La Figure 5.25 représente cette image artificielle pour un saut de 20 niveaux de gris.

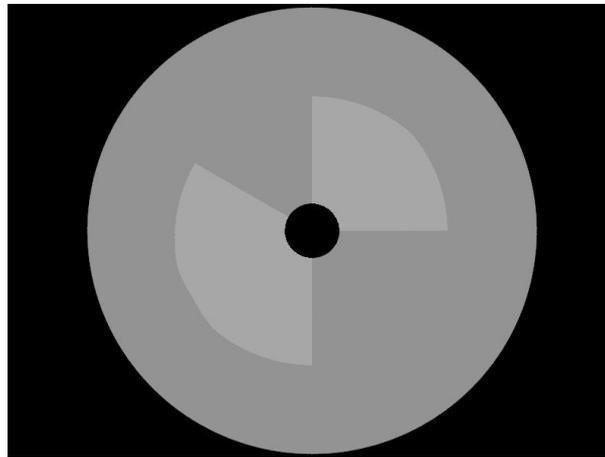
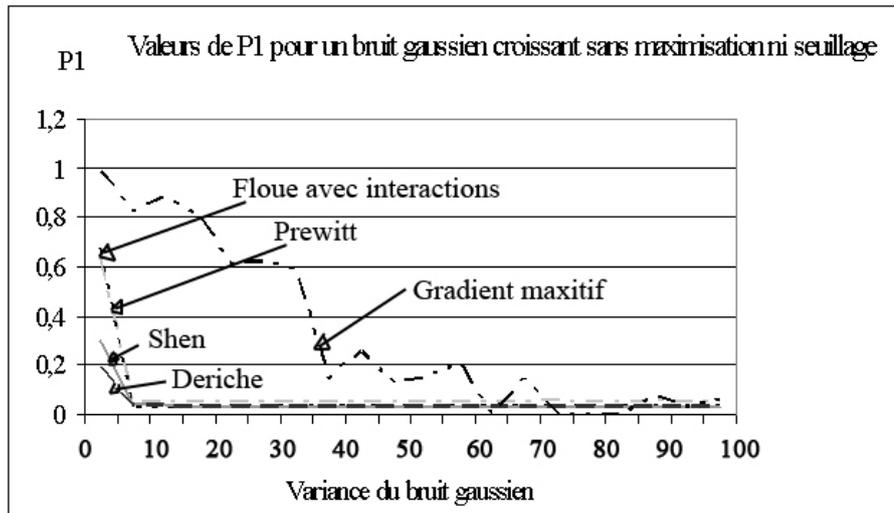


FIG. 5.25 – Image artificielle bruitée avec un saut de 20 niveau de gris.

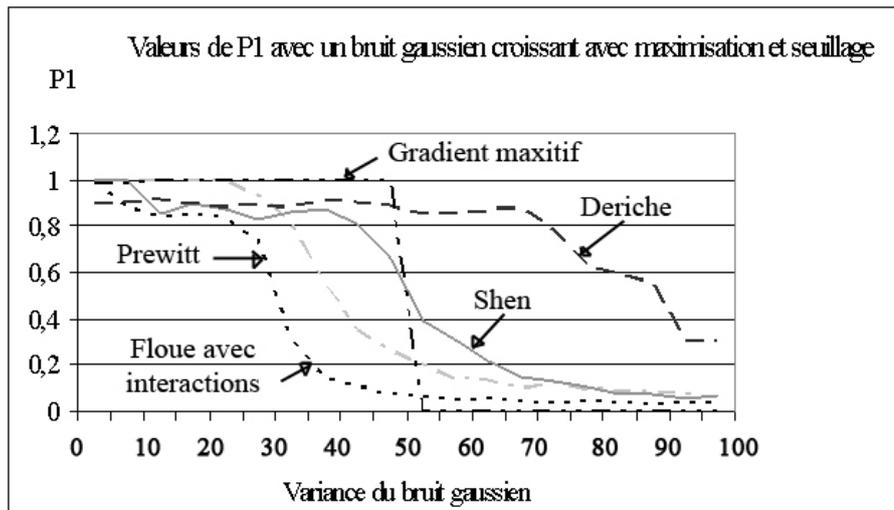
5.4.3.1 Evolution du paramètre P_1 en présence de bruit

Dans cette partie, nous utiliserons uniquement l'image artificielle 5.25 possédant un saut de 100 niveaux de gris. De la même manière que dans la Section 5.3.1, nous avons comparé le critère P_1 pour différentes approches appliquées aux images omnidirectionnelles : le filtre de Prewitt, son adaptation Floue présentée Chapitre 3, les approches optimales de Shen-Castan et Canny-Derliche et l'approche du Gradient maxitif présentée Chapitre 4. Dans un premier temps on applique un bruit gaussien de variance σ croissante. Pour la première expérience, on n'effectue pas de maximisation de gradient, ni de seuillage. La Figure 5.26(a) montre que toutes les approches échouent à part la nôtre pour $\sigma > 2$. Dans la seconde expérience, nous avons maximisé le gradient et appliqué un seuillage arbitraire de manière à obtenir les contours les plus optimaux possibles pour chaque approche. La Figure 5.26(b) montre que le filtre de Prewitt est le plus sensible au bruit. Son adaptation présentée dans [18] est également assez sensible au bruit. Notre approche donne de meilleurs résultats que la plupart des



(a)

FIG. 5.26 – (a) Valeurs comparatives de P_1 pour un bruit gaussien sans maximisation ni seuillage.



(a)

FIG. 5.26 – (b) Valeurs comparatives de P_1 pour un bruit gaussien avec maximisation et seuillage.

filtres. Les contours sont détectés correctement pour $\sigma < 45$. Si $\sigma > 45$, toutes les valeurs estimées sont bruitées donc nulles. Seul le filtre de Canny-Deriche détecte un contour optimal pour $\sigma < 70$. Ensuite, les valeurs diminuent jusqu'à ce que le contour soit complètement dissimulé par le bruit.

5.4.3.2 Estimation des paramètres P_1 et P_2 en fonction de l'intensité du contour

Notre étude se portera sur 2 transitions : une transition verticale entre les colonnes 399 et 400 et une transition oblique à 60° de la transition verticale. Pour le contour oblique, nous souhaitons

éviter d'effectuer une rotation sur notre image et d'introduire des valeurs interpolées. Connaissant

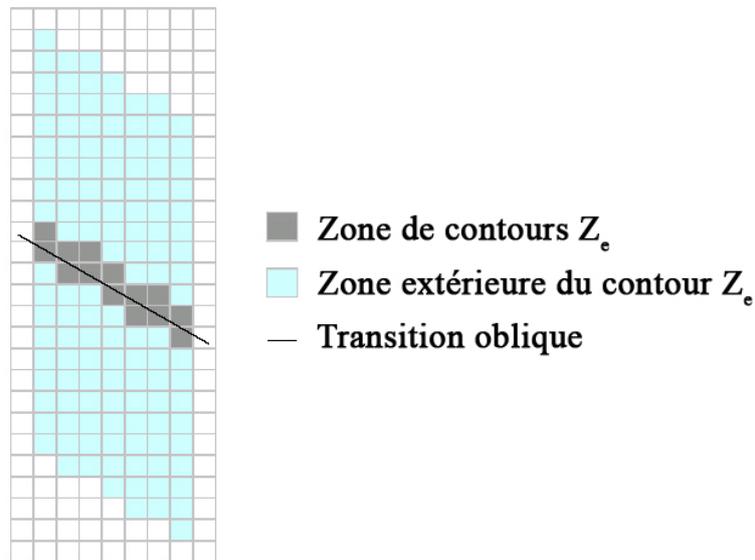


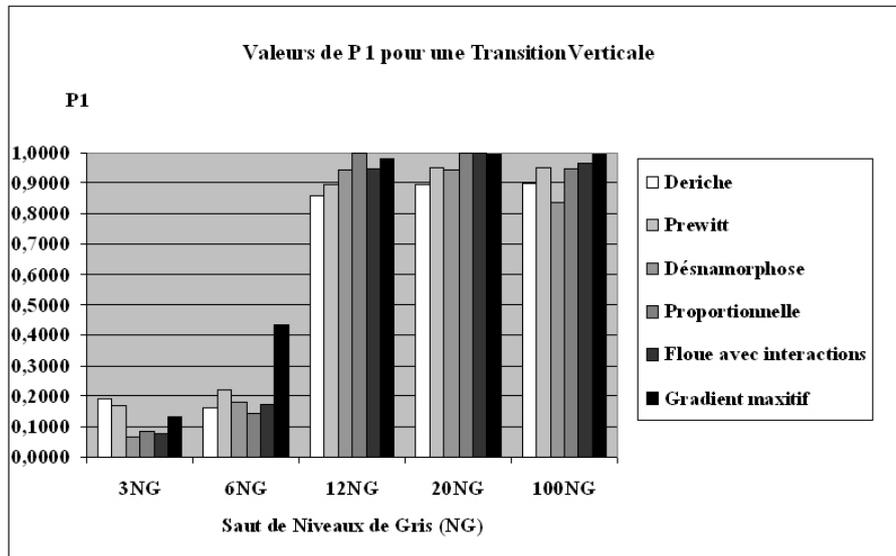
FIG. 5.27 – Adaptation de la zone de contour pour une transition oblique.

précisément la transition oblique dans nos images, nous avons transposé nos calculs en colonnes : au lieu de faire 2 pixels d'épaisseur par ligne, Z_e est maintenant formée par 2 pixels par colonne comme illustré Figure 5.27. Par conséquent, la zone extérieure à Z_e sera constituée de 14 pixels au-dessus et en dessous de Z_e .

Pour un saut de 3 niveaux de gris par exemple, les contours extraits par la méthode Désanamorphose sont indiscernables du bruit quels que soient les seuils appliqués. Cette perte d'information peut être expliquée par les effets d'interpolation intervenant lors du dépliage de l'image. De plus, le bruit gaussien ajouté rend plus difficile l'extraction.

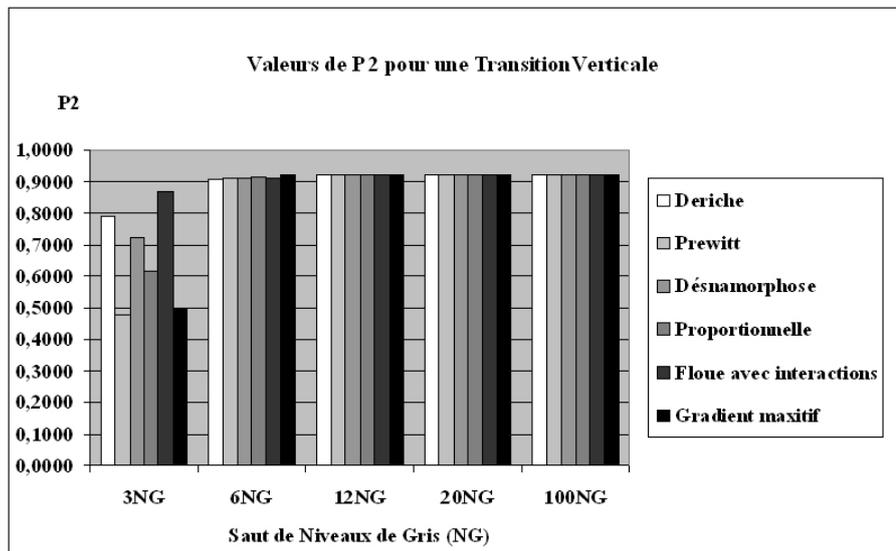
Les résultats pour les paramètres P_1 et P_2 obtenus avec une transition verticale sont présentés Figure 5.28 et ceux obtenus avec une transition oblique Figure 5.29. Ces figures nous montrent que nos approches fournissent des résultats au moins aussi bons que les approches optimales.

Avec un faible saut de niveau de gris, les contours sont difficilement détectables quelle que soit l'approche utilisée. Avec les approches à noyaux variables, le noyau projeté introduit une sorte de lissage. Ce lissage rend quasiment impossible la détection de contours lorsque le saut de niveaux de gris est trop faible. On peut voir sur les Figures 5.28(a) et 5.29(a) que pour un saut inférieur à 6 niveaux de gris, les valeurs obtenues sont très faibles pour les approches Floue, Proportionnelle et Gradient maxitif. On peut remarquer que deux approches se démarquent à partir d'un saut de 6 niveaux de gris : l'approche du Gradient maxitif pour la transition verticale et Deriche pour la transition oblique.



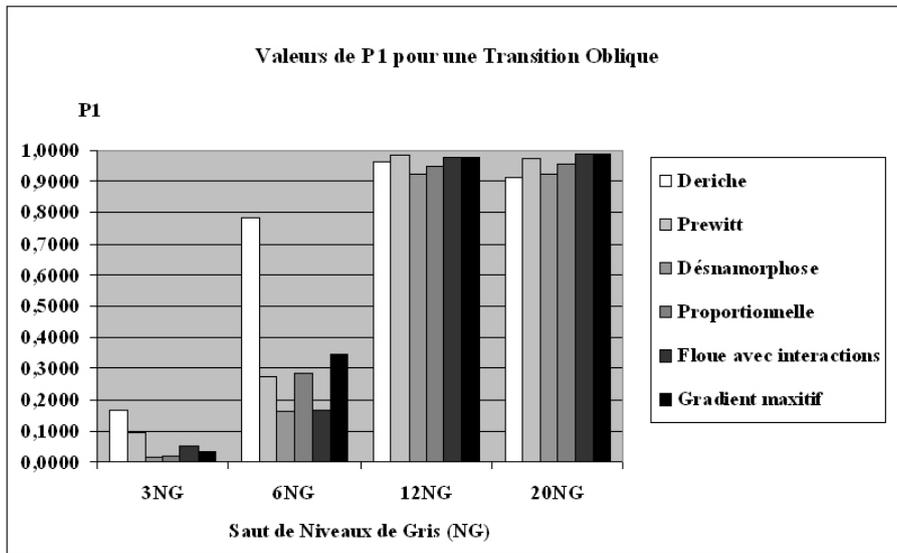
(a)

FIG. 5.28 – (a) Paramètres P_1 obtenus pour une transition verticale.



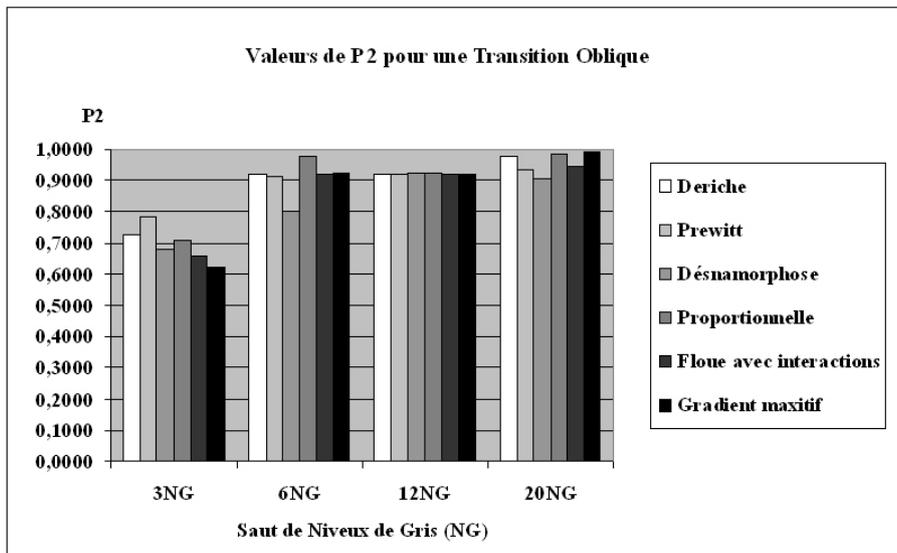
(b)

FIG. 5.28 – Paramètres P_2 obtenus pour une transition verticale.



(a)

FIG. 5.29 – (a) Paramètres P_1 obtenus pour une transition oblique.



(b)

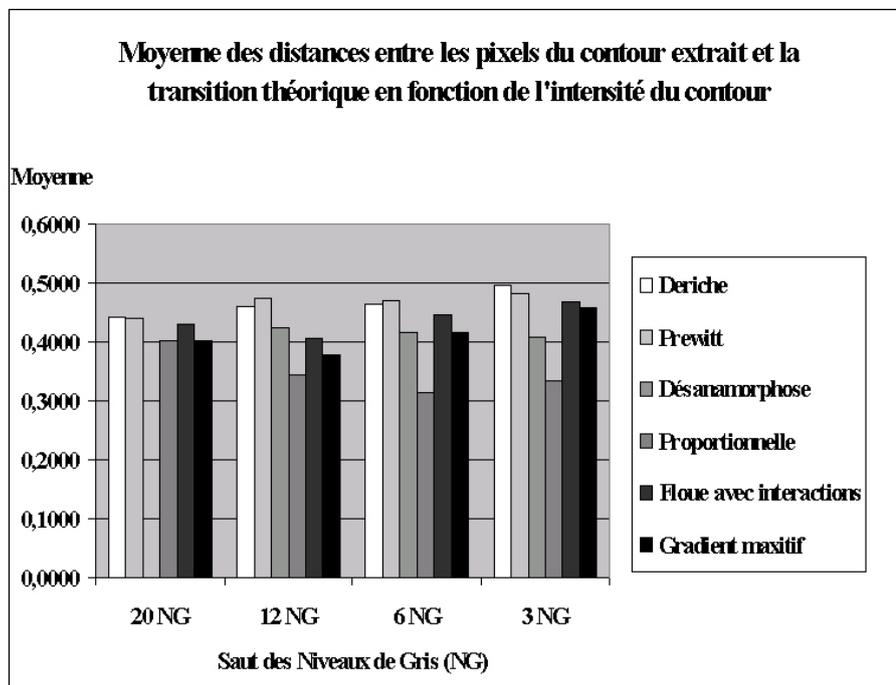
FIG. 5.29 – (b) Paramètres P_2 obtenus pour une transition oblique.

5.4.3.3 Mesure des écarts

Comme nous l'avons mentionné, le filtre de Prewitt est très sensible au bruit. Pour un saut d'amplitude très faible dans une image bruitée, ce filtre ne sera pas en mesure de détecter correctement le contour. Le contour est noyé dans le bruit. Nous n'avons donc pas pu estimer la moyenne et l'écart-type de l'erreur entre la droite théorique et la droite détectée pour ce filtre sur l'image pour un saut de 3 niveaux de gris.

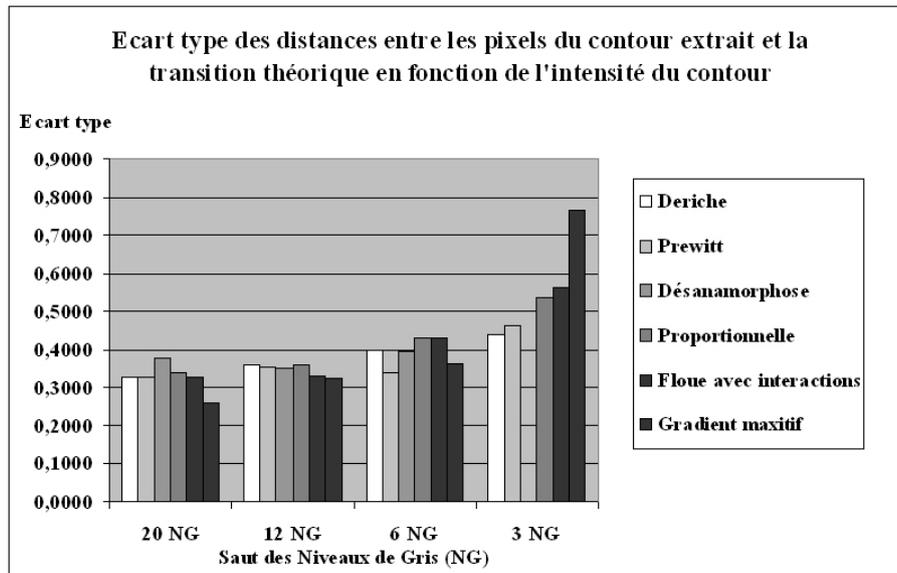
Sur la Figure 5.30(a), on peut observer que la moyenne de l'erreur entre la droite théorique et la droite détectée varie peu quelque soit l'approche. Les techniques d'extraction à noyaux variables donnent des moyennes sensiblement inférieures à celles des méthodes dérivatives classiques. On peut penser que cette performance est due à l'adaptation de la taille du noyau à la position dans l'image omnidirectionnelle.

La mesure de l'écart-type donne une idée de la dispersion des pixels sur la droite détectée. Les valeurs obtenues sont présentées Figure 5.30(b). Les histogrammes montrent qu'il n'y a pas de grande différence entre les écarts-types pour les sauts de 12 et 20 niveaux de gris. L'allure générale montre une forte augmentation des écarts lorsque l'amplitude des sauts diminue. Les écarts-types des méthodes Floue et Proportionnelle augmentent plus rapidement que ceux des autres méthodes. L'étude des moyennes et des écarts-types de l'erreur entre la droite théorique et la droite détectée



(a)

FIG. 5.30 – (a) Moyennes obtenues pour des sauts d'amplitude 20, 12, 6 et 3 niveaux de gris pour chaque approche.



(e)

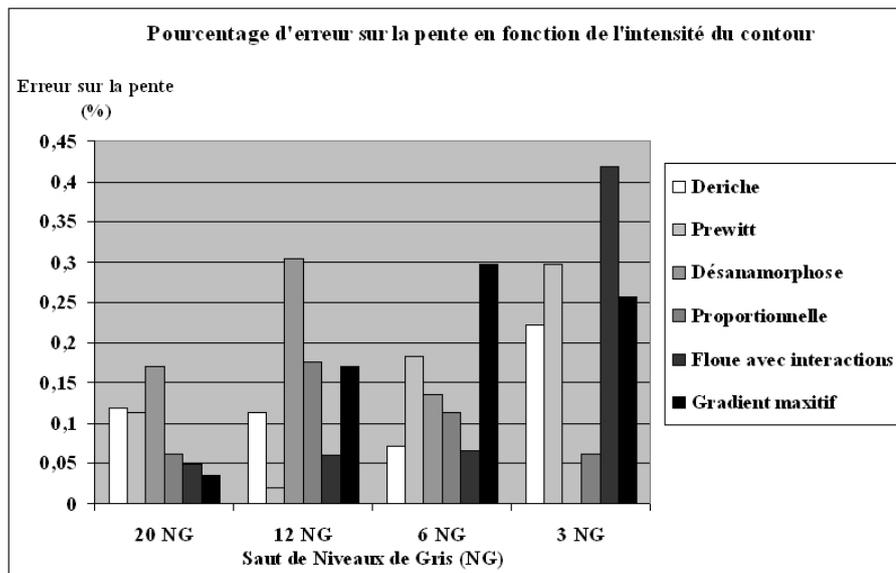
FIG. 5.30 – (b) Ecart-types entre les pixels du contour extrait et la transition théorique obtenus pour des sauts d'amplitude 20, 12, 6 et 3 niveaux de gris pour chaque approche.

permet de dire que les méthodes à noyaux variables extraient des contours plus proches de la transition réelle, mais plus étalés. Ce phénomène s'explique par la manière dont opèrent ces techniques : le noyau est plus large sur le bord de l'image et plus petit vers le centre, contrairement à ceux des méthodes classiques de taille constante en tout point de l'image.

5.4.3.4 Régression linéaire

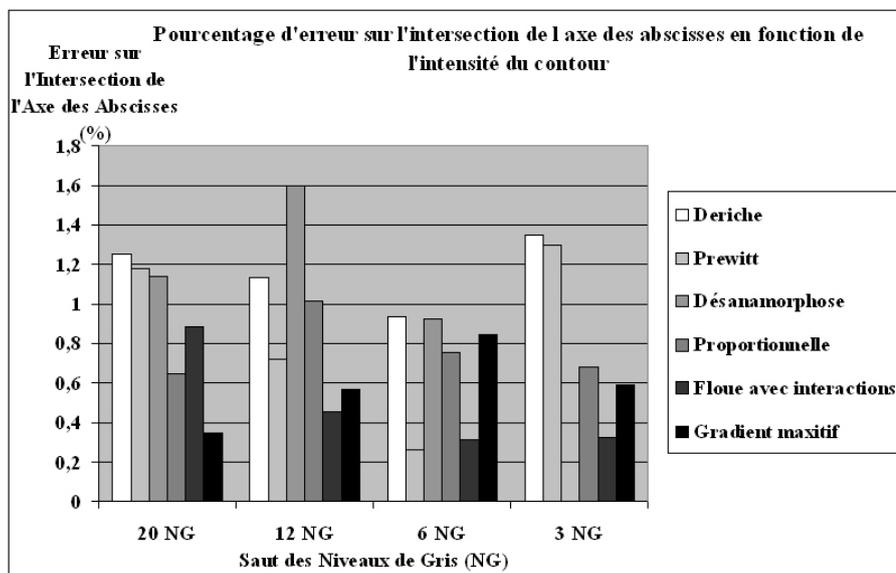
Pour estimer l'erreur entre le contour théorique et le contour détecté sur l'image omnidirectionnelle, nous utilisons la droite de régression linéaire présentée Section 5.4.3.4. Le contour théorique est défini par la droite $y = ax + b$ avec $a = \tan 60 = 1.7321$ and $b = u_0 - (v_0 * a) = -119.6152$. On estime la droite de régression linéaire en fonction des points détectés à l'aide des équations (5.6) et (5.7). Les résultats obtenus sont présentés Figures 5.31 (a) et (b).

Avec un faible saut de niveau de gris (inférieur à 5), aucun algorithme ne détecte de contours. Celui-ci est totalement perdu dans le bruit. De plus, pour les approches à noyaux variables, le noyau projeté introduit une sorte de lissage. Ce lissage a des effets catastrophiques lorsque les variations du signal sont faibles face au bruit. Avec une amplitude supérieure à 10 niveaux de gris, le pourcentage d'erreur décroît. Les différentes approches donnent de bons résultats. On peut remarquer que plus l'amplitude augmente, meilleurs sont les résultats donnés par les approches à noyaux variables. La méthode Désanamorphose donne les plus mauvais résultats à cause du bruit introduit par la rétro-projection.



(a)

FIG. 5.31 – (a) Pourcentage d’erreur sur la pente de la droite détectée pour des sauts d’amplitude 20, 12, 6 et 3 niveaux de gris pour chaque approche.



(b)

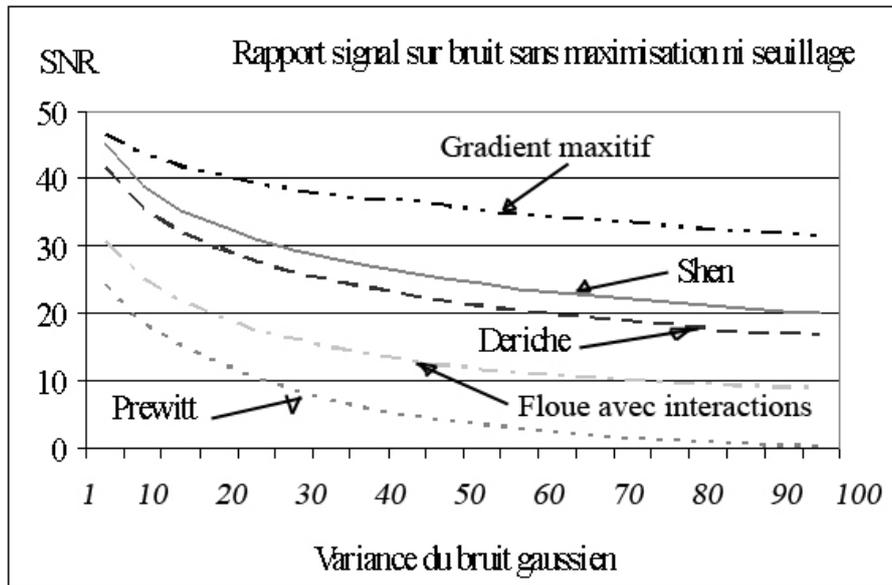
FIG. 5.31 – (b) Pourcentage d’erreur sur l’ordonnée à l’origine pour des sauts d’amplitude 20, 12, 6 et 3 niveaux de gris pour chaque approche.

5.4.3.5 Critère de bonne détection

Dans ce paragraphe, nous avons estimé le rapport signal sur bruit présenté Section 5.1.3 pour chacune des différentes approches. Ce rapport permet de tester la bonne détection du contour. L’image

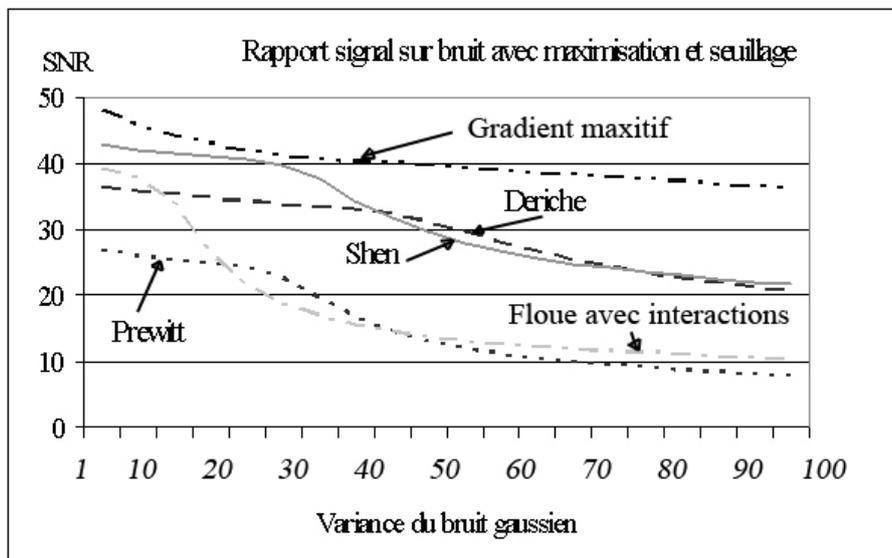
5.4. Comparaison des différentes approches sur des images omnidirectionnelles

artificielle avec un saut d'amplitude de 100 niveaux de gris et un bruit gaussien d'intensité croissante a été utilisée.



(a)

FIG. 5.32 – (a) Rapport signal sur bruit obtenu sur des images sans maximisation ni seuillage.



(e)

FIG. 5.32 – (b) Rapport signal sur bruit obtenu sur des images après maximisation et seuillage (sauf l'approche maxitive qui fournit un seuillage naturel et n'a donc pas été seuillée).

On peut observer que le meilleur rapport signal sur bruit est donné par l'approche du Gradient

maxitif loin devant les autres. Remarquons également que l'adaptation du filtre de Prewitt donne de meilleurs résultats que le filtre de Prewitt appliqué directement à l'image omnidirectionnelle. Les approches de Canny-Deriche et Shen-Castan ont des résultats relativement similaires. Ceux-ci sont plus robustes au bruit que de simples noyaux de convolution mais moins robustes que l'approche du Gradient maxitif.

5.4.4 Expériences sur des images réelles

Pour ces expérimentations, nous avons utilisé un capteur composé d'une caméra numérique munie d'un miroir hyperbolique illustré Figure 5.33. Une procédure d'étalonnage, décrite dans [17], nous a permis de déterminer les paramètres intrinsèques de la caméra.



FIG. 5.33 – Notre capteur.

Cette section présente les principaux résultats obtenus avec des images réelles. Sur l'image 5.34(a), la partie du bas est une mire constituée de droites parallèles à l'axe optique. L'effet de distorsion sur ces droites est le plus visible. La pertinence de notre opérateur est ainsi facilement illustrée. On peut observer une alternance de lignes radiales noires fines et épaisses. En raison de la projection sur le miroir, les lignes semblent plus fines à mesure que l'on se rapproche du centre de l'image. Un bon détecteur de contours doit mettre en évidence les transitions de l'image : passage d'un pixel noir à un pixel blanc ou d'un pixel blanc à un pixel noir. Ainsi sur notre image, pour chaque ligne noire sur fond blanc, on devrait détecter deux contours caractéristiques de la ligne de la périphérie au centre de l'image.

La Figure 5.35 montre l'ensemble des contours détectés pour chaque approche. La Figure 5.36 représente un zoom sur une partie de l'image afin de mieux visualiser les détails des contours détectés. Comme nous pouvons voir sur la Figure 5.36(a) le détecteur de contours classique de Prewitt n'est pas adapté aux images omnidirectionnelles. Le masque fournit un contour double pour les lignes minces. Pour les lignes épaisses, ces contours fusionnent à proximité du centre de l'image. Ainsi, les effets de détecteur de contours ne sont pas identiques selon la position du contour sur l'image.

Les détecteurs de Canny-Deriche (Figure 5.36(g)) et Shen-Castan (Figure 5.36(h)) présentent de très bons résultats. Le double contour est détecté correctement de la périphérie au centre de l'image.

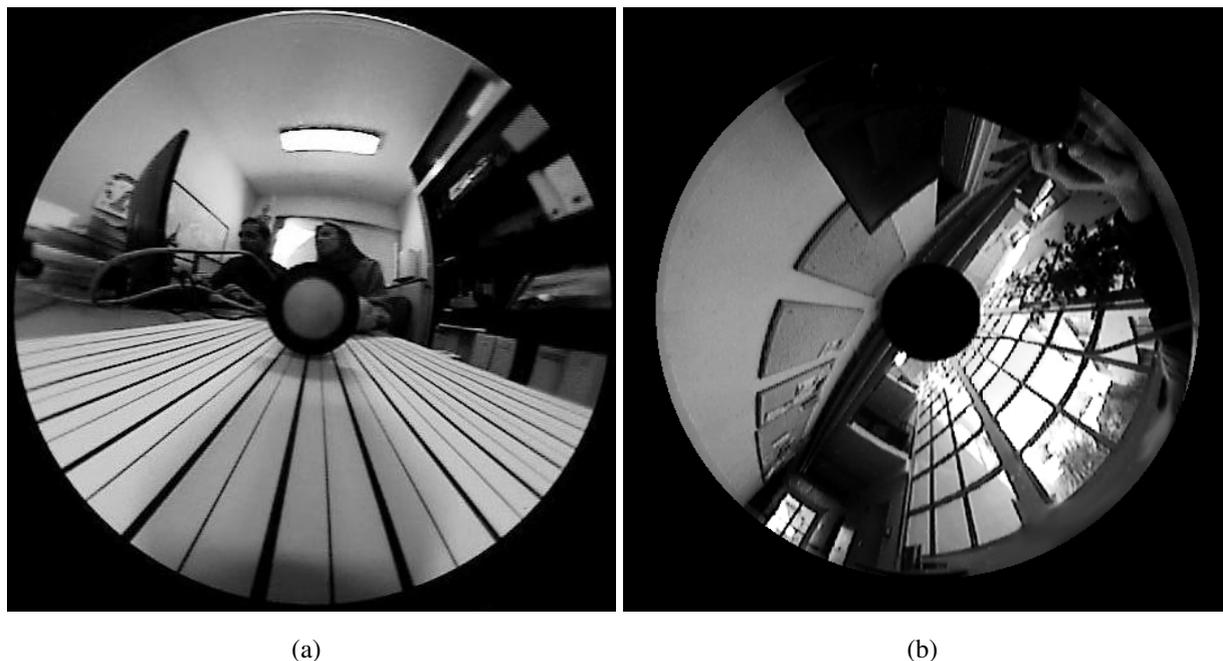


FIG. 5.34 – Deux images réelles

Cependant, la taille de chaque contour détecté est la même quelle que soit sa position dans l'image. Cet opérateur n'est donc pas adapté pour détecter des contours correspondant à des éléments réels de la scène 3D capturée par la caméra.

L'approche par Désanamorphose (Figure 5.36 (b)) présente de bons résultats. Les lignes minces sont correctement détectées comme dans l'approche classique. Cependant, pour les lignes épaisses, les deux frontières sont détectées de la périphérie au centre de l'image. En raison des effets d'interpolation dus à la projection et à la rétro-projection, les contours sont bruités.

L'approche Proportionnelle (Figure 5.36(c)) présentée dans la Section 3.3, offre des résultats légèrement meilleurs que ceux obtenus avec l'approche par Désanamorphose. Les résultats sont moins bruités, particulièrement au centre de l'image. Cependant, la largeur du contour change en fonction de sa localisation sur l'image. Cet effet est dû à l'imprécision des cellules du masque de Prewitt. En effet, la projection d'une cellule du masque à la périphérie de l'image couvre beaucoup de pixels, alors qu'elle ne couvre que peu de pixels à proximité du centre de l'image.

L'approche Floue avec interactions, illustrée Figure 5.36(e), présente les meilleurs résultats. Les contours sont correctement identifiés sur l'image omnidirectionnelle. Les deux contours des lignes épaisses sont bien détectés et restent distincts de la périphérie au centre de l'image. Comme avec l'approche proportionnelle, l'épaisseur du contour dépend de la position sur l'image en raison de l'imprécision du masque. On peut remarquer que l'approche Floue sans interactions (Figure 5.36(d)) fournit aussi de bons résultats. Cependant, vers la périphérie de l'image, le double contour a tendance à se chevaucher.

L'approche du Gradient maxitif fournit, elle aussi, de très bons résultats comme nous pouvons le voir sur la Figure 5.36(f). Comme pour l'approche floue et l'approche proportionnelle, la taille des masques dépend de leur position dans l'image. On obtient donc des contours plus fins au centre de l'image qu'à la périphérie. De plus cette approche fournit un seuillage naturel de l'image. Tous les points détectés sont donc des points associés aux contours présents dans l'image.

La Figure 5.34(b) représente un couloir avec un quadrillage de fenêtres et de l'autre côté, des panneaux d'affichages. Cette image contient toute sorte de contours radiaux, concentriques et obliques et nous permet de valider nos algorithmes sans fournir de mire de départ. Cette image ne contient que des contours physiques. Ces contours sont par conséquent plus bruités et d'autant plus difficile à détecter correctement.

Comme pour la Figure 5.34(a), nous avons représenté Figure 5.37 l'ensemble des contours détectés avec chaque approche. La Figure 5.38 représente un zoom sur une partie des fenêtres de l'image afin de mieux visualiser les détails des contours détectés.

Sur la Figure 5.38(a), nous pouvons constater que le filtre de Prewitt s'avère incapable de détecter les contours des fenêtres correctement : il ne détecte qu'un seul contour pour deux fenêtres mitoyennes. L'utilisation d'images réelles contenant une quantité de bruit non-négligeable a des conséquences désastreuses sur la détection de contours .

Les filtres de Canny-Deriche (Figure 5.38(g)) et Shen-Castan (Figure 5.38(h)) fournissent comme précédemment de très bons résultats. Le contour de chaque fenêtre est correctement détecté.

L'approche par Désanamorphose (Figure 5.38 (b)) est elle aussi très sensible au bruit. Les fenêtres sont correctement détectées à l'extérieur de l'image, cependant, les contours obtenus sont relativement bruités et fusionnent en se rapprochant du centre de l'image.

L'approche Proportionnelle (Figure 5.38(c)) présentée dans la Section 3.3, offre des résultats à peu près similaires aux résultats obtenus avec l'approche par Désanamorphose. Les noyaux variables fournissent des contours plus fins au centre qu'à la périphérie de l'image, cependant ces contours sont noyés dans le bruit, ce qui les rend difficilement interprétables.

Les approches Floues avec (Figure 5.38(f)) et sans (Figure 5.38(e)) interactions fournissent toutes deux de bons résultats (d'autant plus si on les compare aux autres images obtenues en utilisant le filtre de Prewitt). Les contours sont correctement détectés pour chaque fenêtre du centre à la périphérie de l'image. Ces contours sont adaptés à la géométrie de l'image omnidirectionnelle et, sur ces figures, on peut visualiser très nettement la variation de la taille des contours en fonction de la position dans l'image (les contours sont plus fin au centre de l'image et plus large à la périphérie). L'approche du Gradient maxitif (Figure 5.38(g)) est également une approche à noyaux variables. Les contours sont correctement détectés sur l'ensemble de l'image. Cependant, notre image étant relativement bruitée, certains contours apparaissent moins nettement que d'autres. En effet, notre approche a supprimé la plupart des points apparentés au bruit.

Les opérateurs adaptés aux images perspectives se comportent de la même manière quelque soit leur position dans l'image. En effet, la convolution discrète est un outil permettant la construction de filtres linéaires ou de filtres de déplacements invariants. Les variations de résolution présentes dans l'image omnidirectionnelle ne sont pas prises en compte. L'intérêt de notre approche est de conserver le même comportement sur l'espace projectif du cylindre, relativement proche de l'espace de l'image perspective. Cette approche permet de détecter des contours dans la scène réelle. La largeur du contour dépend de sa position radiale sur l'image omnidirectionnelle.

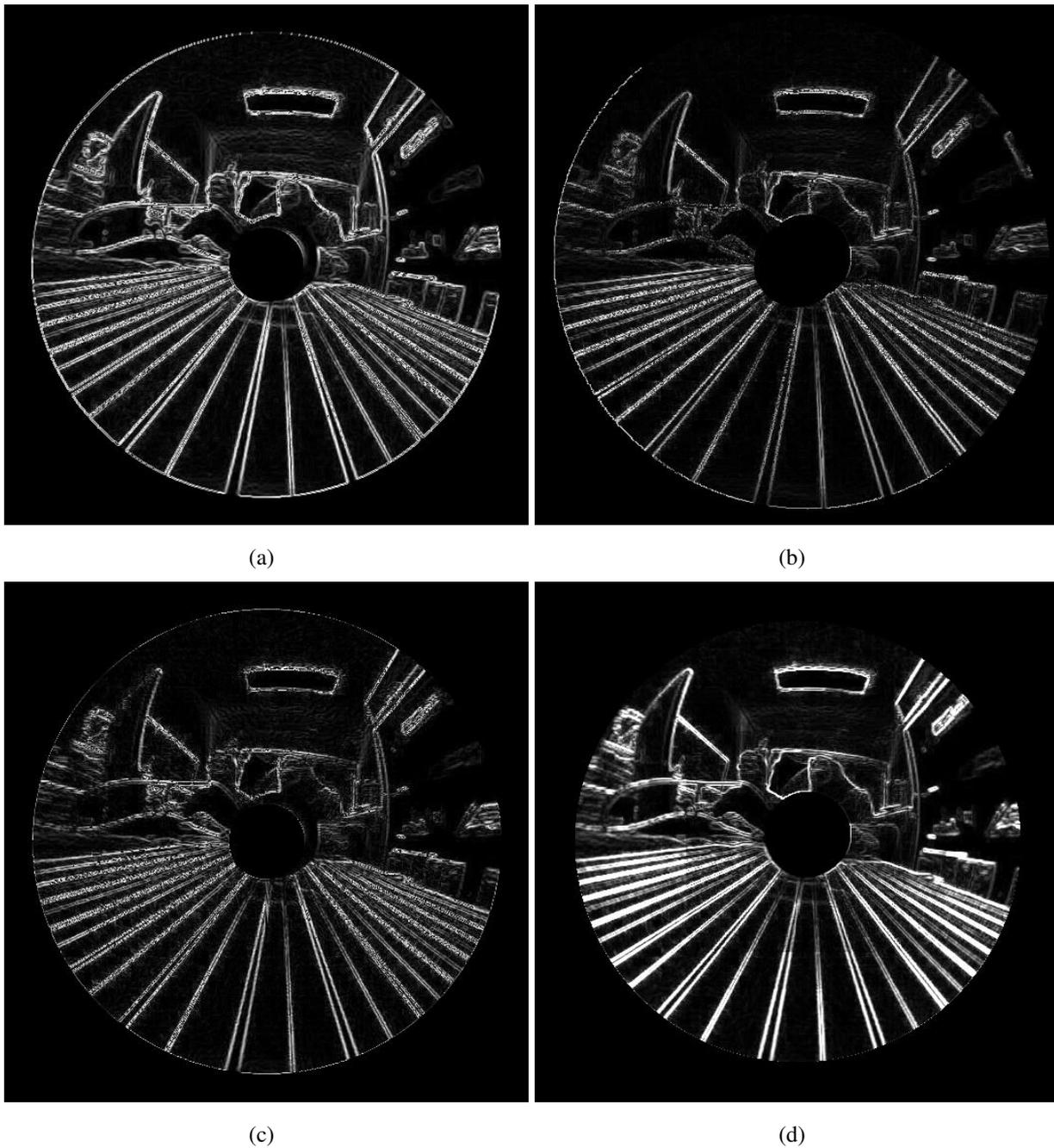


FIG. 5.35 – (a) Filtre de Prewitt classique appliqué à l’image omnidirectionnelle, (b) Filtre de Prewitt classique appliqué à l’image désanamorphosée, (c) Filtre de Prewitt utilisant l’approche proportionnelle présentée dans la Section 3.3, (d) Filtre de Prewitt utilisant l’approche floue présentée dans la Section 3.4 sans prendre en compte les interactions entre les éléments du masque.

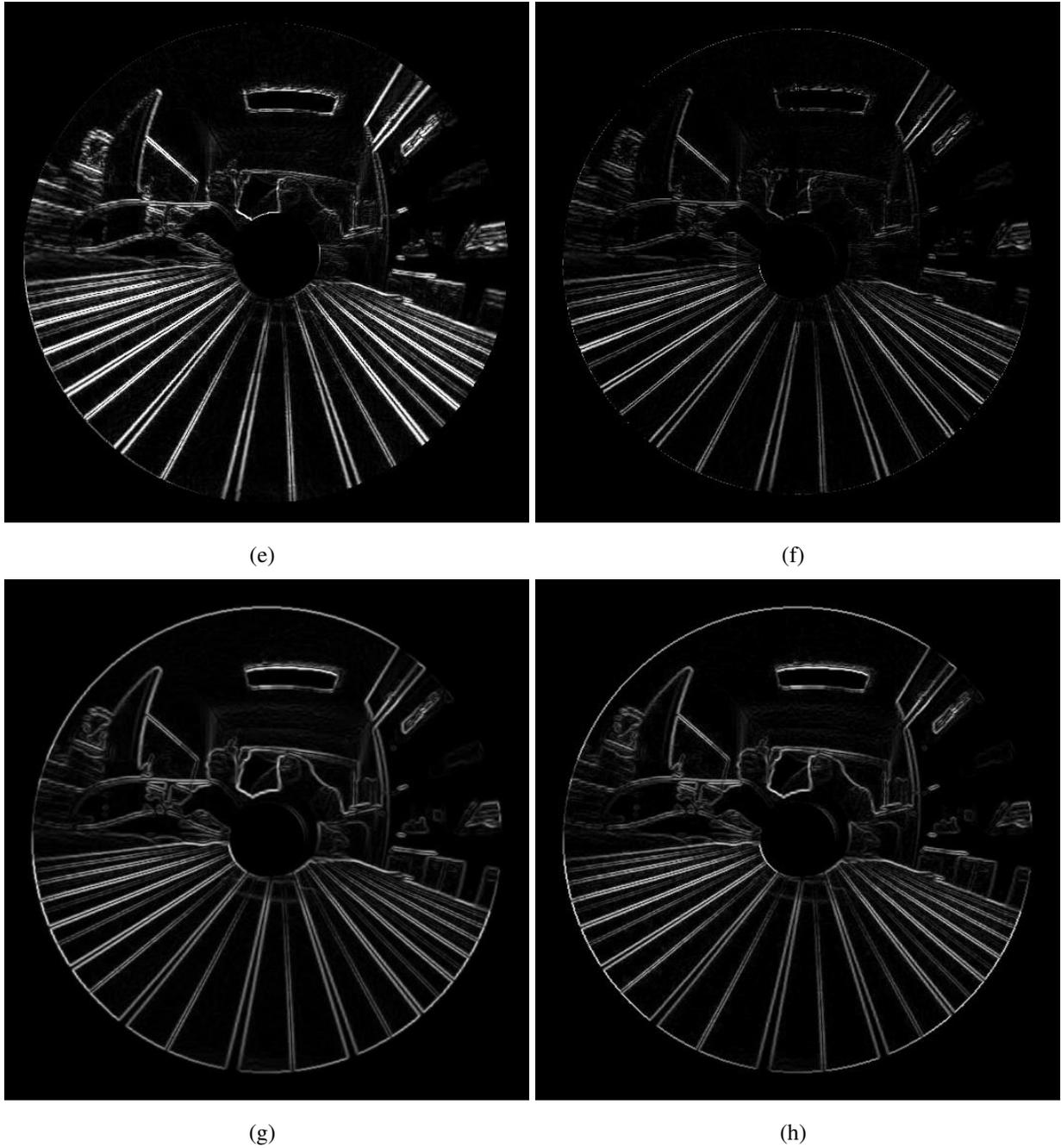


FIG. 5.35 – (e) Filtre de Prewitt utilisant l’approche floue avec interactions présentée dans la section 3.4, (f) Estimation de gradient maxitif présentée dans la section 4.2, (g) Filtre de Canny-Deriche appliqué à l’image omnidirectionnelle, (h) Filtre de Shen-Castan appliqué à l’image omnidirectionnelle.

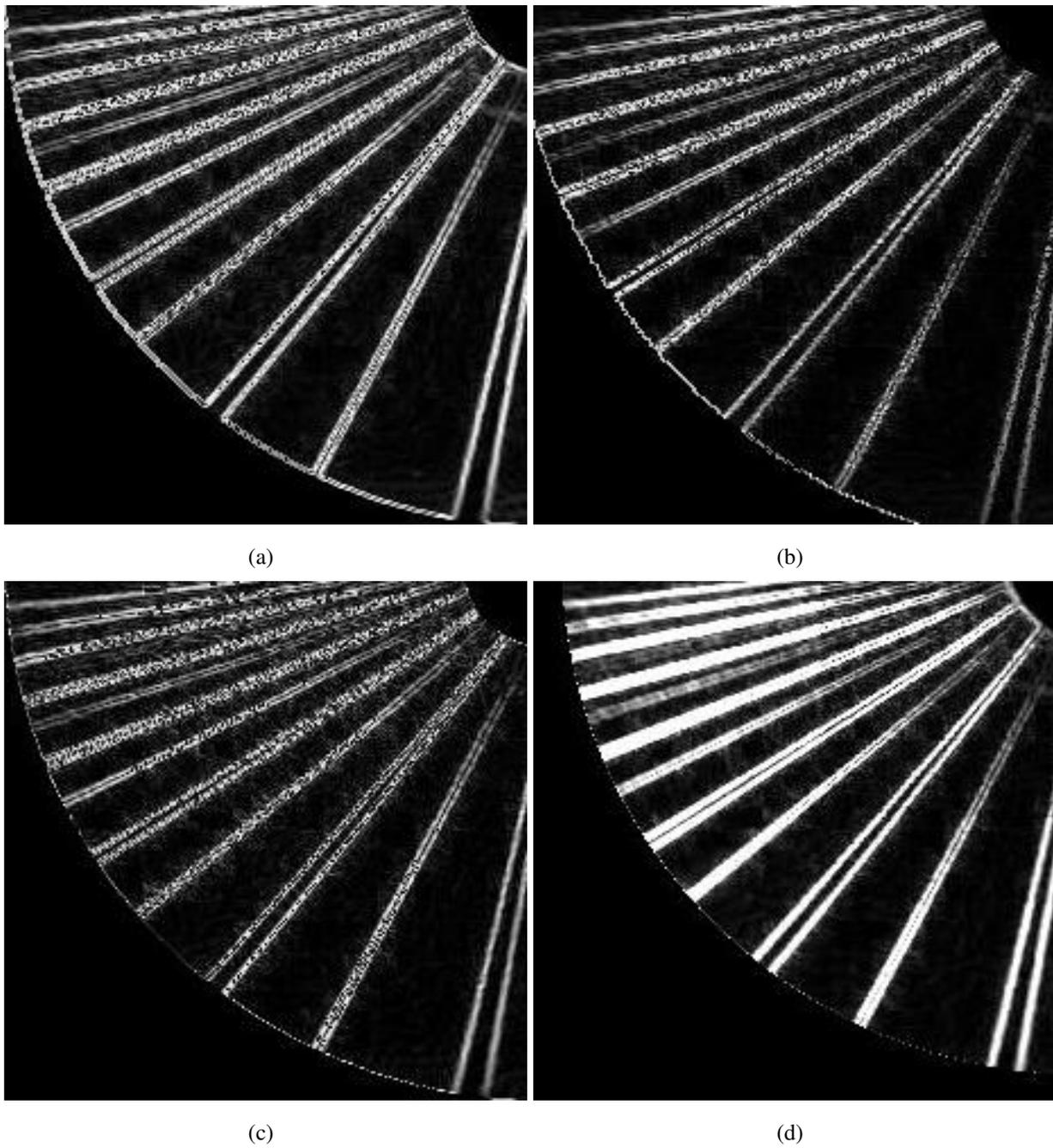


FIG. 5.36 – Zoom sur l’image précédente : (a) Filtre de Prewitt classique appliqué à l’image omnidirectionnelle, (b) Filtre de Prewitt classique appliqué à l’image désanamorphosée, (c) Filtre de Prewitt utilisant l’approche proportionnelle présentée dans la section 3.3, (d) Filtre de Prewitt utilisant l’approche floue présentée dans la section 3.4 ans prendre en compte les interactions entre les éléments du masque.

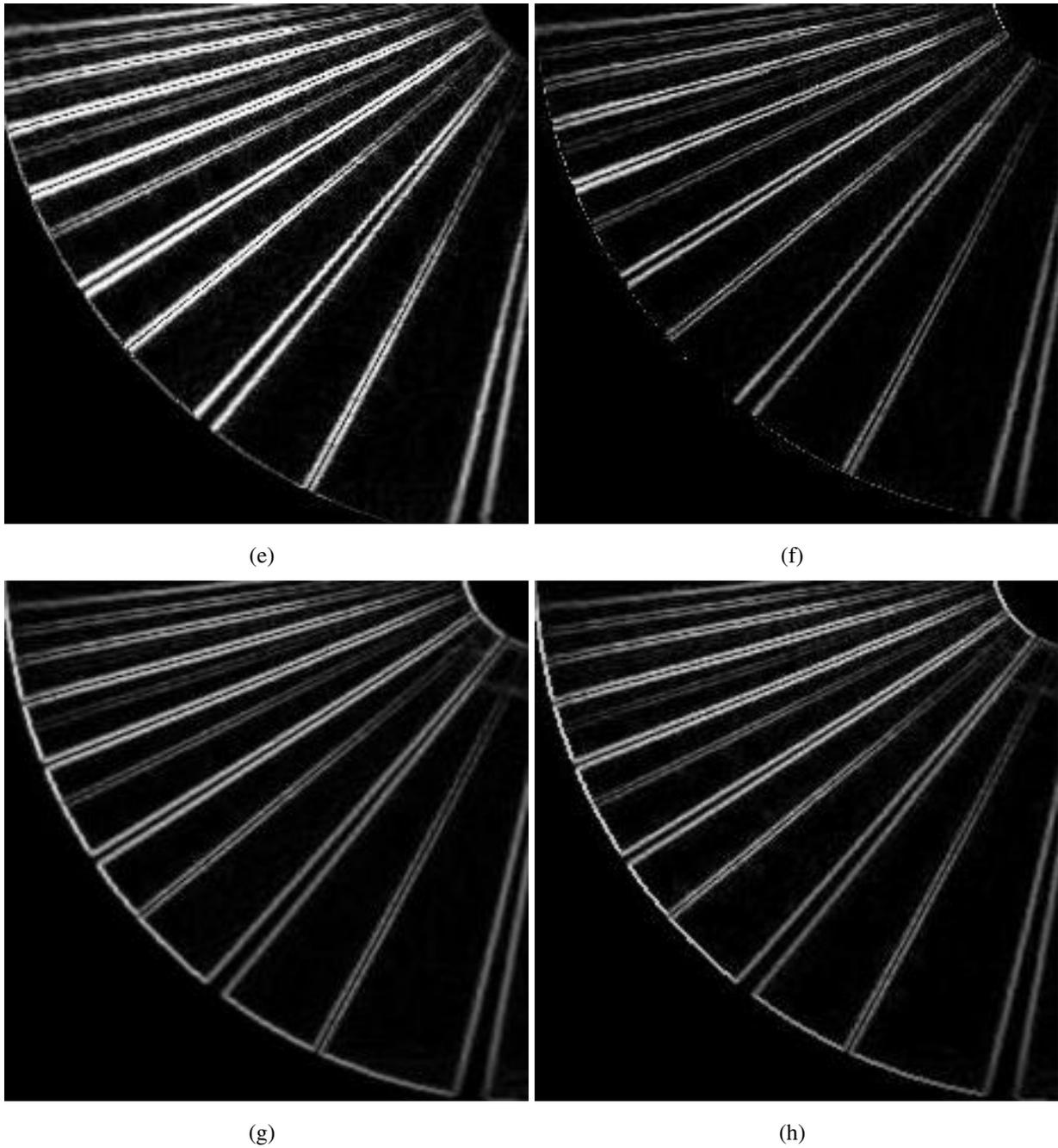


FIG. 5.36 – Zoom sur l’image précédente : (e) Filtre de Prewitt utilisant l’approche floue avec interactions présentée dans la section 3.4, (f) Estimation de gradient maxitif présentée dans la section 4.2, (g) Filtre de Canny-Deriche appliqué à l’image omnidirectionnelle, (h) Filtre de Shen-Castan appliqué à l’image omnidirectionnelle.

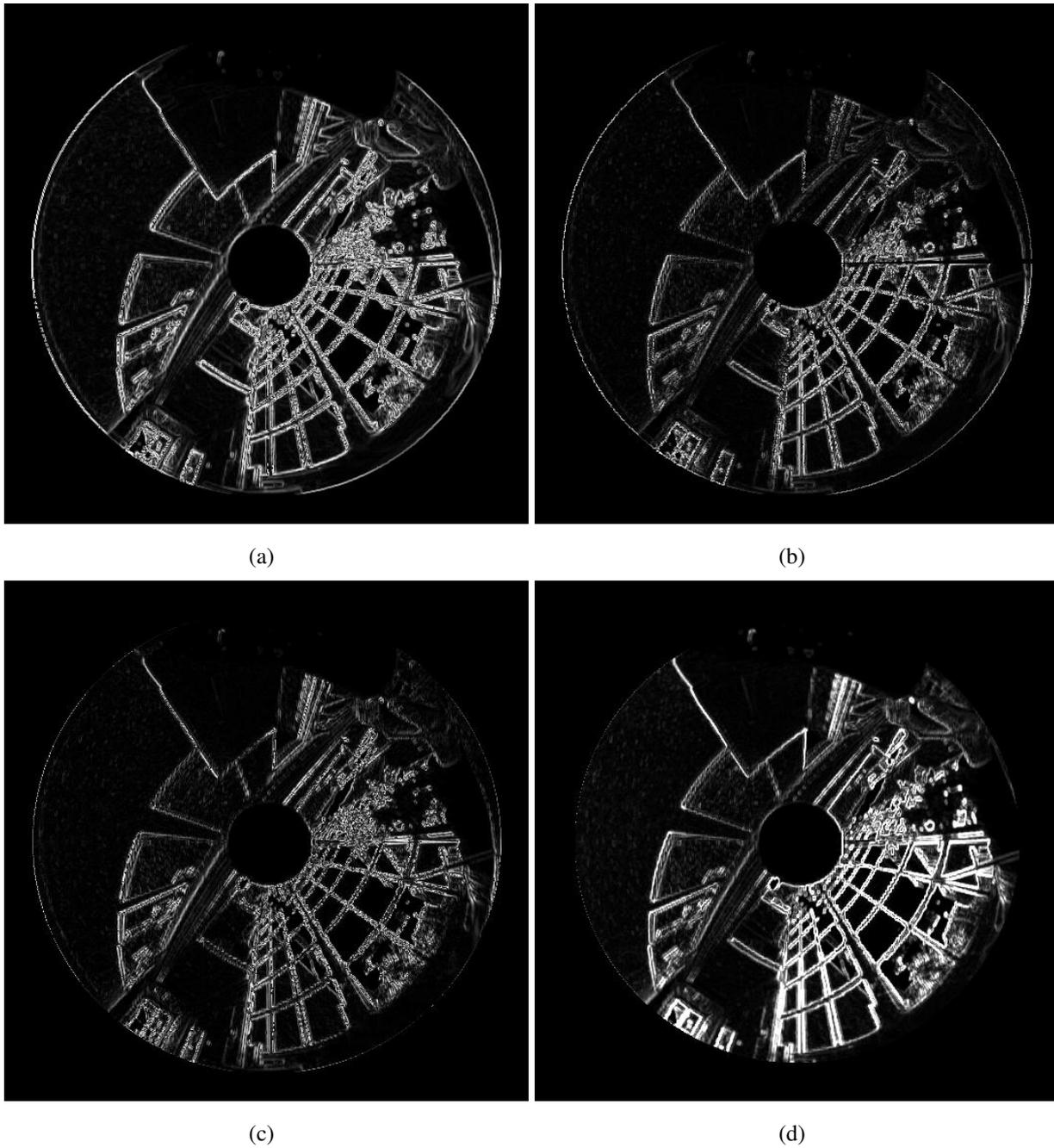


FIG. 5.37 – (a) Filtre de Prewitt classique appliqué à l’image omnidirectionnelle, (b) Filtre de Prewitt classique appliqué à l’image désanamorphosée, (c) Filtre de Prewitt utilisant l’approche proportionnelle présentée dans la section 3.3, (d) Filtre de Prewitt utilisant l’approche floue présentée dans la section 3.4 sans prendre en compte les interactions entre les éléments du masque.

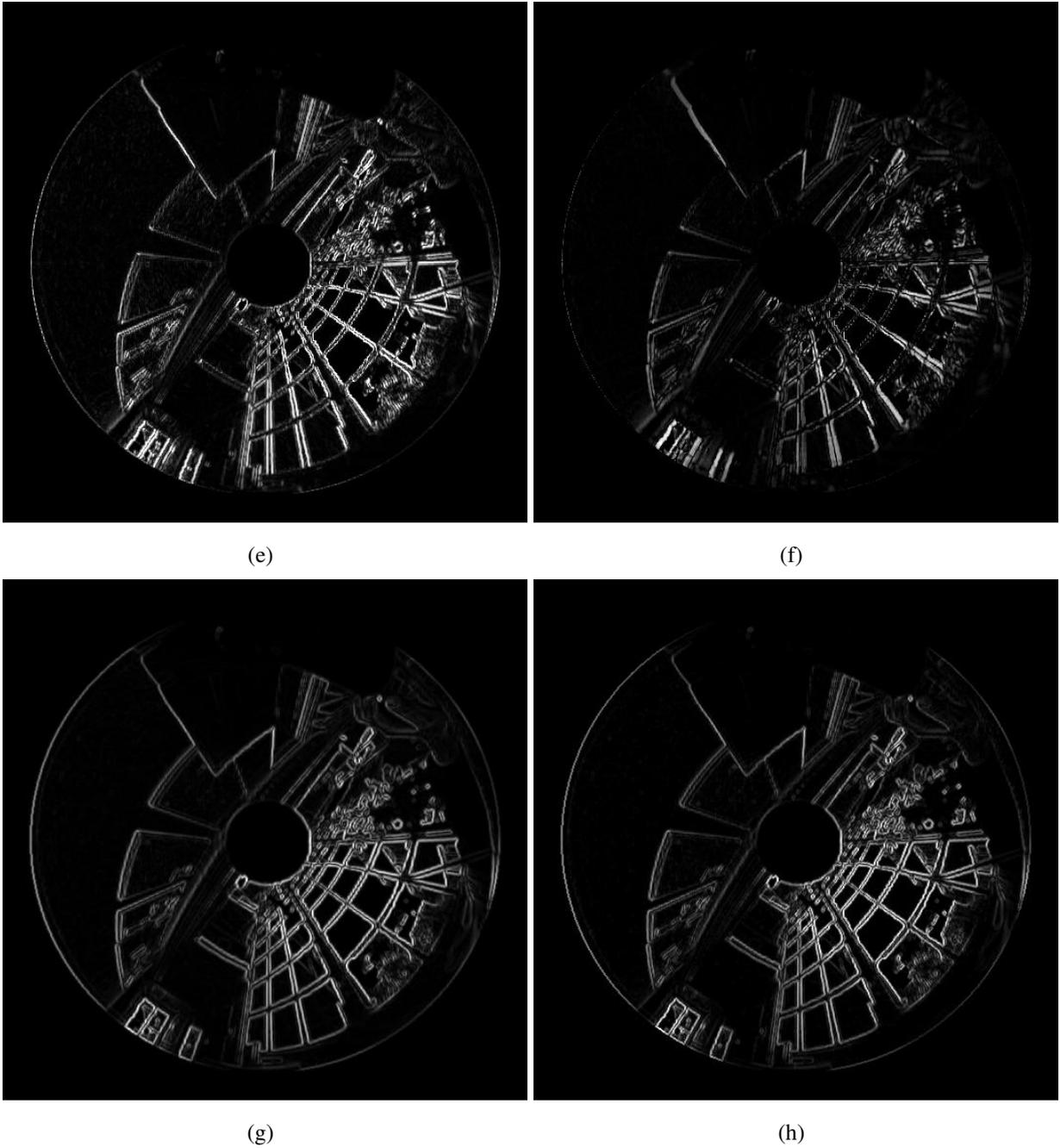


FIG. 5.37 – (e) Filtre de Prewitt utilisant l'approche floue avec interactions présentée dans la section 3.4, (f) Estimation de gradient maxitif présentée dans la section 4.2, (g) Filtre de Canny-Deriche appliqué à l'image omnidirectionnelle, (h) Filtre de Shen-Castan appliqué à l'image omnidirectionnelle.

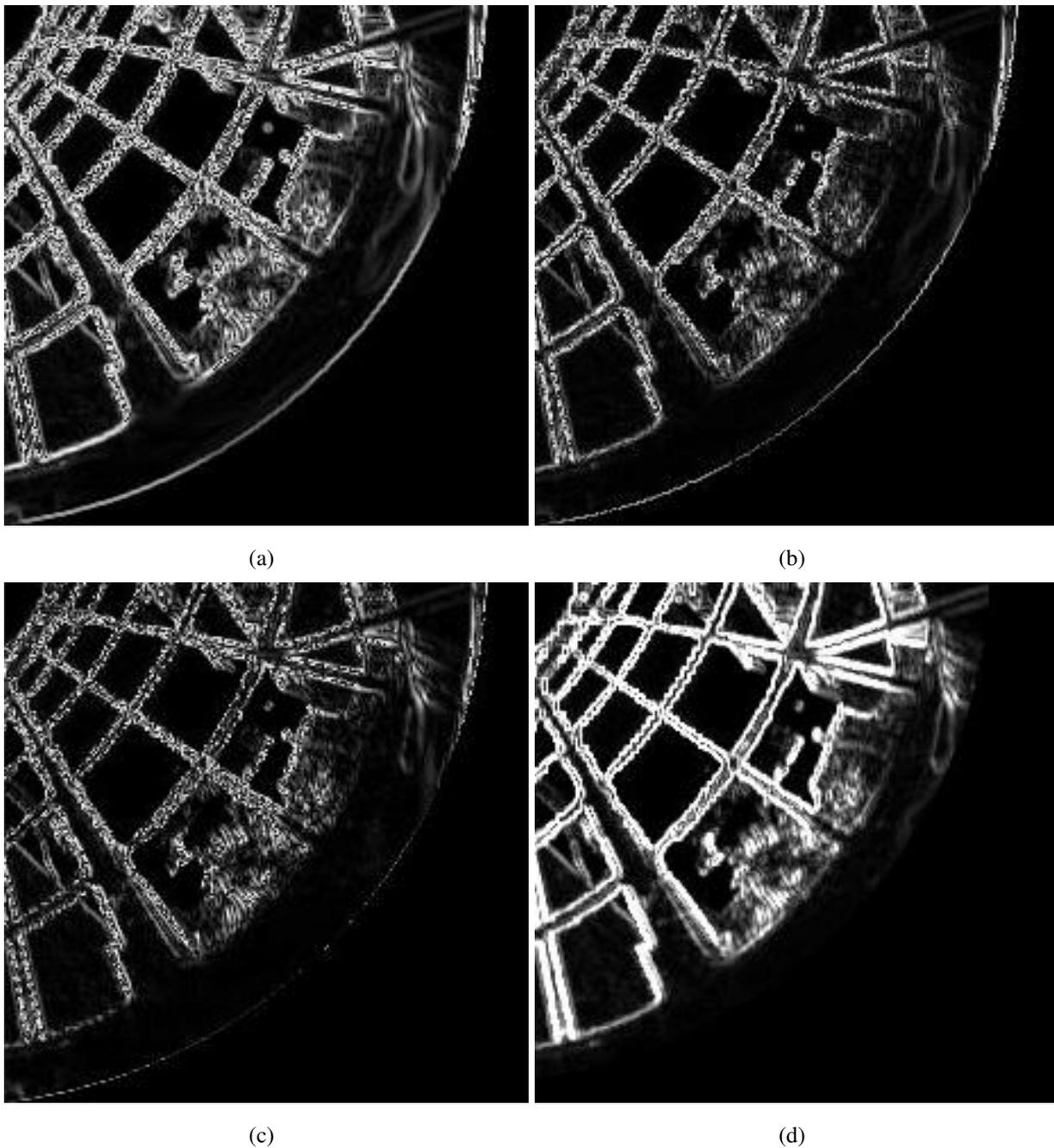


FIG. 5.38 – Zoom sur l’image précédente : (a) Filtre de Prewitt classique appliqué à l’image omnidirectionnelle, (b) Filtre de Prewitt classique appliqué à l’image désanamorphosée, (c) Filtre de Prewitt utilisant l’approche proportionnelle présentée dans la section 3.3, (d) Filtre de Prewitt utilisant l’approche floue présentée dans la section 3.4 sans prendre en compte les interactions entre les éléments du masque.

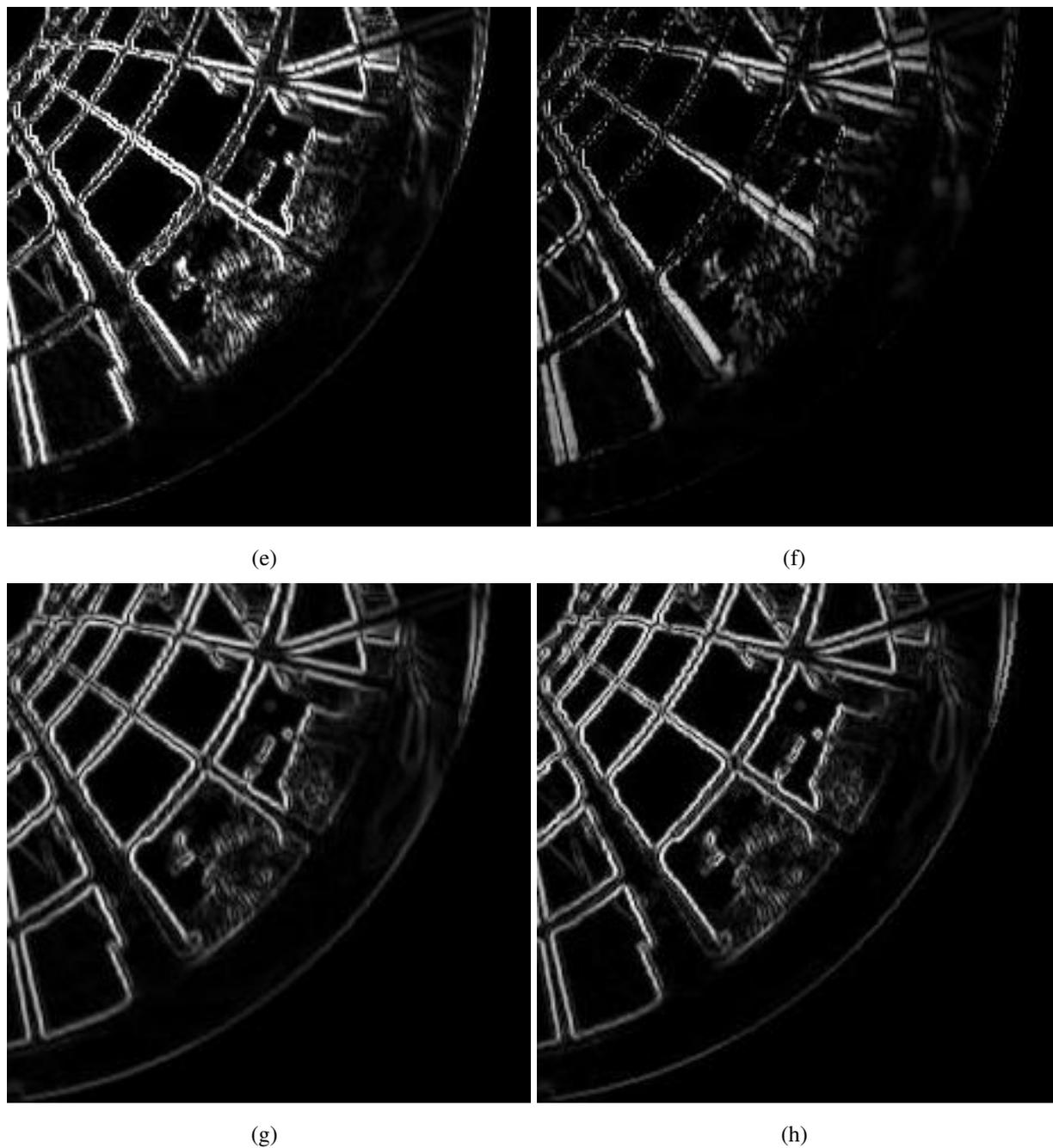


FIG. 5.38 – Zoom sur l’image précédente : (e) Filtre de Prewitt utilisant l’approche floue avec interactions présentée dans la section 3.4, (f) Estimation de gradient maxitif présentée dans la section 4.2, (g) Filtre de Canny-Deriche appliqué à l’image omnidirectionnelle, (h) Filtre de Shen-Castan appliqué à l’image omnidirectionnelle.

Conclusion

Dans ce chapitre, nous avons effectué des tests quantitatifs et qualitatifs sur des images réelles et artificielles afin de prouver la fiabilité de nos algorithmes. L'approche du Gradient maxitif a été validée sur des images perspectives et fournit de très bons résultats par rapport aux approches basées sur des noyaux sommatifs. Fournissant un seuillage naturel de l'image, elle permet d'effectuer les traitements indépendamment de l'image ou du bruit qu'elle contient. Nous évitons ainsi le choix laborieux et parfois impossible de valeurs de seuil arbitraires adaptées à l'image. Nous avons ensuite proposé une série de tests sur des images omnidirectionnelles afin de tester la bonne détection, la bonne localisation et la robustesse au bruit de nos détecteurs de contours. La comparaison effectuée nous a permis de montrer que nos approches sont au moins équivalentes et souvent meilleures que les approches existantes. En particulier, nous avons mis en évidence la robustesse de nos opérateurs en présence de bruit.

$$dS * \kappa = S * d\kappa \quad (5.11)$$

- (i) La condition de cohérence : $\forall A \subseteq \Omega, P(A) \leq \Pi(A), .$
(ii) Le principe de préservation de la préférence : $\forall (a, b) \in \Omega^2,$

$$p(a) > p(b) \Rightarrow \pi(a) > \pi(b)$$

$$p(a) = p(b) \Rightarrow \pi(a) = \pi(b)$$

- (iii) Le principe du maximum de spécificité : $\forall A \subseteq \Omega,$

$$N_1 \leq N_2$$

$$\Pi_2 \leq \Pi_1$$

$$\pi_1 \geq \kappa_1$$

$$\pi_2 \geq \kappa_2$$

$$\overline{I}_1 = \mathbb{C}_{\pi_1}(I) \geq \mathbb{E}_{\kappa_1}(I)$$

$$\underline{I}_1 = \mathbb{C}_{\pi_1^c}(I) \leq \mathbb{E}_{\kappa_1}(I)$$

$$\overline{I}_2 = \mathbb{C}_{\pi_2}(I) \geq \mathbb{E}_{\kappa_2}(I)$$

$$\underline{I}_2 = \mathbb{C}_{\pi_2^c}(I) \leq \mathbb{E}_{\kappa_2}(I)$$

$$\widehat{I}_1 - \widehat{I}_2 \in [\underline{I}_1 - \overline{I}_2, \overline{I}_1 - \underline{I}_2]$$

$$d\kappa(x) = \eta^+(x - \epsilon_x) - \eta^-(x + \epsilon_x)$$

$$\overline{I} = \mathbb{C}_{\pi}(I) \geq \mathbb{E}_{\kappa}(I)$$

$$\underline{I} = \mathbb{C}_{\pi^c}(I) \leq \mathbb{E}_{\kappa}(I)$$

$$\underline{I} \leq \widehat{I} \leq \overline{I}$$

$$\mathbb{C}_{\pi^c}(I) \leq \mathbb{E}_{\kappa}(I) \leq \mathbb{C}_{\pi}(I)$$

$$\widehat{I} = \mathbb{E}_{\kappa}(I)$$

$$\kappa \preceq \pi$$

$$\kappa_1 \preceq \pi_1$$

$$\kappa_2 \preceq \pi_2$$

$$\overline{I}_1 = \mathbb{C}_{\pi_1}(I) \geq \mathbb{E}_{\kappa_1}(I)$$

$$\underline{I}_1 = \mathbb{C}_{\pi_1^c}(I) \leq \mathbb{E}_{\kappa_1}(I) \leq \mathbb{C}_{\pi_1}(I) = \overline{I}_1$$

$$\underline{I}_2 = \mathbb{C}_{\pi_2^c}(I) \leq \mathbb{E}_{\kappa_2}(I) \leq \mathbb{C}_{\pi_2}(I) = \overline{I}_2$$

$$\mathbb{E}_{\kappa_1}(I) - \mathbb{E}_{\kappa_2}(I) = 0$$

Conclusion Générale

Associant une caméra perspective et un miroir de révolution, les caméras catadioptriques permettent d'élargir le champ de vision des images à 360°. Les images omnidirectionnelles fournissent une vue globale de l'environnement. Cependant, les miroirs utilisés pour obtenir ces images introduisent des déformations et modifient la notion de voisinage et la résolution de l'image. Les opérateurs de traitement d'images classiques ont été développés pour des images perspectives et ne sont pas satisfaisant pour les images omnidirectionnelles.

Le travail de thèse que nous avons présenté dans ce manuscrit, s'inscrit dans la problématique du traitement des images omnidirectionnelles.

Nous avons présenté un large éventail des systèmes d'acquisition d'images réalisés pour étendre le champ de vision. Parmi ces systèmes d'acquisition, nous nous sommes intéressés aux capteurs catadioptriques centraux. Dans la littérature, nous avons référencé de nombreux articles dans lesquels ces capteurs étaient utilisés pour leurs propriétés géométriques et leur champ de vision à 360°. Nous avons également pu remarquer que souvent les auteurs traitaient ces images avec des méthodes classiques et soulignaient que des outils adaptés à la géométrie de l'image donneraient de meilleurs résultats. Déplier l'image ou la projeter sur un autre support impliquerait l'interpolation des valeurs des niveaux de gris des pixels et génèrerait des erreurs répercutées lors des traitements. Pour minimiser l'impact de ces erreurs, nous avons proposé de repousser, le plus tard possible dans la chaîne de calculs, l'étape de modification ou d'ajout d'informations arbitraires dans l'image originale. Sous cette contrainte de préservation des informations contenues dans l'image originale, nous avons développé de nouvelles approches de traitement d'images, applicables directement sur l'image omnidirectionnelle et prenant en compte de sa géométrie.

Afin de prendre en compte l'imprécision de la mesure de luminance introduite dans l'image lors de l'échantillonnage spatial, nous avons choisi une modélisation par une partition floue. Pour traiter les images, nous avons utilisé un raisonnement basé sur la théorie des possibilités. Cet outil nous a permis de remplacer les notions probabilistes par des notions de possibilistes.

Notre objectif principal, dans ce travail sur les images omnidirectionnelles, a été de réaliser des outils qui prennent en compte les déformations introduites par les miroirs et la résolution non-uniforme.

Pour traiter ces images, nous avons défini un modèle projectif nous permettant de créer nos opéra-

teurs dans l'espace régulier du cylindre englobant le miroir de révolution. Nous avons proposé de projeter le centre de chaque pixel de l'image omnidirectionnelle sur le cylindre, de définir la géométrie du noyau autour du point projeté et de les rétro-projeter sur l'image omnidirectionnelle. Avec cette approche, nous associons à chaque pixel un noyau projeté qui nous permet d'estimer la valeur du niveau de gris du pixel filtré en respectant la géométrie de l'image omnidirectionnelle. Cette adaptation prend en compte la non-uniformité de l'image liée aux distorsions du miroir.

L'intérêt de cette approche comparée aux approches existantes est sa simplicité. Avec une caméra perspective, la scène acquise est supposée invariante par translation. Cela implique que la distance entre la scène et la caméra est constante. Avec une caméra omnidirectionnelle, la scène observée est à une distance constante d'une caméra ayant tourné sur elle-même. Par analogie, on retrouve l'espace du cylindre.

Dans la première partie de nos travaux, nous nous sommes intéressés à l'adaptation des masques de convolution aux images omnidirectionnelles. Ces opérateurs sont très simples à appliquer aux images (une sommation de multiplications) et permettent de réaliser de nombreux traitements tels que le lissage, la suppression de bruit ou la détection de contours.

En appliquant notre algorithme aux masques de convolution, chaque élément du masque projeté chevauche plusieurs pixels de l'image omnidirectionnelle. Pour transférer l'information des pixels intersectés de l'image omnidirectionnelle vers un élément du masque projeté, nous avons proposé intuitivement une solution inspirée du transfert pignistique. Le poids est proportionnel à l'aire d'intersection du pixel de l'image omnidirectionnelle avec l'élément du masque projeté.

Pour prendre en compte les imprécisions de luminance induites par l'échantillonnage spatial, nous avons ensuite proposé une seconde solution, basée sur une modélisation floue des données de l'image. Les pixels sont alors assimilés à des nombres triangulaires flous 2D. L'imprécision de la localisation des niveaux de gris des pixels induit une imprécision sur la localisation des éléments du masque. Nous avons donc modélisé notre masque de convolution avec des nombres triangulaires flous 2D.

Contrairement aux approches par interpolation dans le plan de l'image (interpolation bilinéaire par exemple), notre approche présente l'avantage d'estimation du gradient sous la forme d'un intervalle.

Dans la suite de nos travaux, nous nous sommes ensuite focalisés plus particulièrement à la détection de contours dans les images. Nous avons étudié les nombreux opérateurs existants (filtres basés sur l'approximation de dérivée de la fonction d'intensité de l'image par différences finies, filtres à réponses impulsionnelles finies ou infinies, filtres optimaux implémentés récursivement, etc.) et nous avons pu constater que tous ces filtres présentent la même faiblesse : ils sont incapables de discerner le signal du bruit présent dans une image. L'application d'un seuillage s'avère donc nécessaire pour faire la distinction entre les contours pertinents et les contours liés au bruit. Cependant, le choix de ces valeurs de seuil n'est pas aisé. Elles dépendent de l'image, diffèrent d'une image à l'autre et échouent lorsque le bruit présent dans l'image n'est pas uniforme. Pour pallier à ces problèmes, nous avons développé une nouvelle approche non-additive d'estimation de gradient.

Nous avons remarqué que les opérateurs de dérivation étaient généralement constitués de la soustraction de deux noyaux usuels. Nous avons proposé de remplacer ces noyaux probabilistes par des noyaux possibilistes. Dubois et Prade [12] ont montré qu'un noyau possibiliste triangulaire de mode et de support donné domine l'ensemble des noyaux probabiliste ayant même mode et même support. Nous avons donc choisi d'utiliser des noyaux possibilistes triangulaires.

Pour estimer l'espérance de l'image via chacun de ces noyau, nous avons utilisé l'intégrale de Choquet. Cette intégrale fournit pour chacun des noyaux un intervalles de valeurs. Nous avons ainsi obtenu une estimation du gradient de l'image sous la forme d'un intervalle. De cet intervalle, nous avons déduit l'appartenance ou non d'un point à un contour. En ce sens, notre approche fournit un seuillage naturel des valeurs du gradient. Un des avantages de cette approche est qu'elle ne dépend pas de l'image traitée ni du bruit qu'elle contient. Nous avons ensuite appliqué cette méthode aux images omnidirectionnelles.

La validité des approches proposées a été établie à l'aide d'expérimentations sur des images de synthèses et des images réelles. De nombreuses expériences quantitatives basées sur des critères objectifs ont été effectuées sur des images artificielles afin d'estimer les performances de ces approches comparées aux approches de traitement d'images classiques.

L'approche Gradient a été validée dans un premier temps sur des images perspectives avant d'être appliquée à des images omnidirectionnelles. Sur les images perspectives, nous avons pu mettre en évidence la robustesse au bruit de cette approche. Avec son seuillage naturel, cette approche permet de traiter toutes les images quel que soit le bruit qu'elles contiennent. Pour des contours de faible intensité, cette approche introduit un phénomène de lissage et les contours ne sont pas détectés. Cependant, dès que l'amplitude du saut du contour dépasse 5 niveaux de gris, on obtient une meilleure détection et une meilleure localisation des contours de l'image qu'avec les approches classiques.

Nous avons ensuite comparé les résultats obtenus sur des images omnidirectionnelles avec différents filtres présentant la même granulosité. Nous avons étudié la robustesse au bruit, ainsi que la bonne détection, bonne localisation de nos détecteurs de contours : Les résultats obtenus prouvent que pour chaque critère, nos approches sont au moins équivalentes et le plus souvent meilleures. Nous avons constaté en particulier que l'utilisation d'une modélisation floue rend les filtres moins sensibles au bruit présent dans les images.

Avec la méthode d'adaptation d'un masque de convolution de Prewitt, nous avons pu observer une meilleure détection, une meilleure localisation et une moins importante sensibilité au bruit qu'en appliquant directement le filtre de Prewitt classique. Nous avons également pu remarquer les très bons résultats fournis par l'approche Gradient. Nous obtenons des contours dont la taille varie en fonction de leur position dans l'image. Ces contours sont cohérents avec la scène 3D observée.

Il est vrai que certaines approches classiques comme les filtres optimaux fournissent de bons résultats sur les images omnidirectionnelles. Cependant, ces détecteurs ont le même comportement quelle que soit leur position dans l'image et les contours détectés ont donc tous la même largeur. Ces contours

correspondent aux discontinuités de la fonction d'intensité de l'image, mais pas nécessairement à des contours de la scène 3D observée. L'utilisation de méthodes à noyaux variables pour les images omnidirectionnelles nous a permis de prendre en compte les variations de résolution présentes dans l'image et de détecter des contours cohérents qui respectent la scène observée. Dans notre approche, la largeur des contours dépend de la position radiale dans l'image omnidirectionnelle.

Les perspectives de ce travail sont multiples.

Tout d'abord, une question récurrente se pose : pourquoi créer de nouveaux opérateurs adaptés aux images omnidirectionnelles qui mettent en œuvre des méthodes complexes, parfois coûteuses en temps de calcul, si les opérateurs existants donnent de bons résultats ? Nous avons souligné le fait que nos algorithmes permettent de détecter des contours qui ont une signification au sens de la scène 3D observée. L'utilisation de détecteurs de contours adaptés aux images omnidirectionnelles a été envisagé dans de nombreux travaux concernant la détection de mouvements ou la détections de points d'intérêts dans l'image entre autre. Il serait intéressant de coupler nos détecteurs de contours aux algorithmes existants en matière de traitement et d'interprétations d'images omnidirectionnelles et d'estimer les performances obtenues avec une approche adaptée par rapport à celles obtenues précédemment avec les traitements classiques.

L'estimation des performances d'un détecteur de contour est une opération difficile. Nous avons appliqués les critères de Canny et les critères de Fram et Deutsch, aux images omnidirectionnelles, comme s'il s'agissait d'images perspectives. Ces critères ne sont pas adaptés à la géométrie de l'image. Peut-être serait-il nécessaire de définir de nouveaux critères quantitatifs (ou d'adapter les critères existants) appropriés aux images omnidirectionnelles.

Les noyaux maxitifs ont fournit de très bons résultats dans les images perspectives et dans les images omnidirectionnelles. Des travaux sont actuellement en cours au LIRMM concernant l'estimation de bruit dans les images via les noyaux maxitifs.

Annexe A

Ensembles flous

La théorie des ensembles flous a été proposée par Zadeh [70] en 1965. Associée à la théorie des possibilités, elle permet de représenter mathématiquement l'imprécision relative de certaines informations. Dans cette théorie interviennent les notions de mesures de possibilité et de nécessité présentées par Dubois et Prade [11].

De même qu'un sous-ensemble classique E de X est défini par sa fonction caractéristique $\chi_E : \Omega \rightarrow [0, 1]$, un sous-ensemble flou A de Ω est défini par une fonction $\mu_A : \Omega \rightarrow [0, 1]$ qui associe à chaque élément ω de Ω son degré d'appartenance $\mu_A(\omega)$ à A .

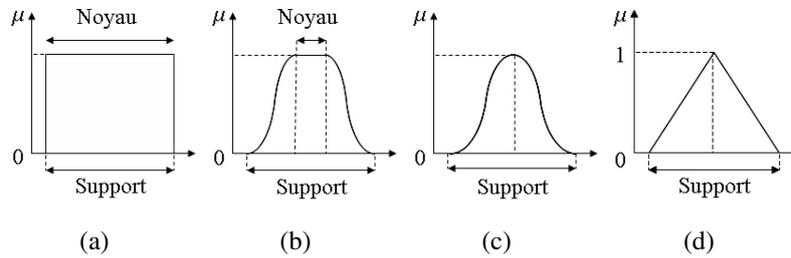


FIG. A.1 – (a), (b) Intervalle flou. (c) Nombre flou (d) Nombre flou triangulaire.

Un sous-ensemble flou est dit normalisé si $\sup_{\omega \in \Omega} \mu_A(\omega) = 1$. Il est caractérisé par son support et son noyau. Le support de A noté $Supp(A)$ est la partie de Ω sur laquelle la fonction d'appartenance à A n'est pas nulle et son noyau $Noy(A)$ est non vide si le sous-ensemble flou A est normalisé :

$$\begin{cases} Supp(A) = \{\omega \in \Omega | \mu_A(\omega) \neq 0\}; \\ Noy(A) = \{\omega \in \Omega | \mu_A(\omega) = 1\}. \end{cases}$$

De plus, pour toute valeur α de $[0, 1]$, on définit l' α -coupe stricte A_α ou coupe de niveau α d'un sous-ensemble flou A de Ω comme le sous ensemble $A_\alpha = \{\omega \in \Omega | \mu_A(\omega) > \alpha\}$.

Un sous-ensemble flou normalisé de \mathbb{R} est appelé quantité floue. Une valeur modale d'une quantité floue Q est un élément ω_m de \mathbb{R} telle que $\mu_Q(\omega_m) = 1$. Un sous-ensemble flou F de X est convexe si et seulement si toute α -coupe de F est une partie convexe de X . Un intervalle flou I est une quantité

floue convexe. On dit qu'un intervalle flou est un nombre flou (Figure A.1(c) et (d)) si sa fonction d'appartenance est semi-continue supérieurement, de support $\{x \in \mathbb{R} \mid \mu_I(x) > 0\}$ borné, admettant une unique valeur modale ω_m .

Annexe B

Dérivation par différences finies

Les premières approches de filtrage consistent à approximer les dérivées de la fonction d'intensité de l'image par différences finies :

$$\nabla_x I(x, y) = I(x + n, y) - I(x - n, y) \quad (\text{B.1})$$

avec en général $n = 1$.

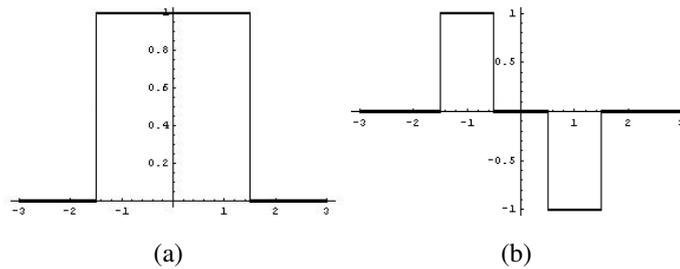


FIG. B.1 – Filtre de Prewitt (a) Filtre passe-bas effectuant le lissage, (b) Filtre passe haut effectuant la dérivation.

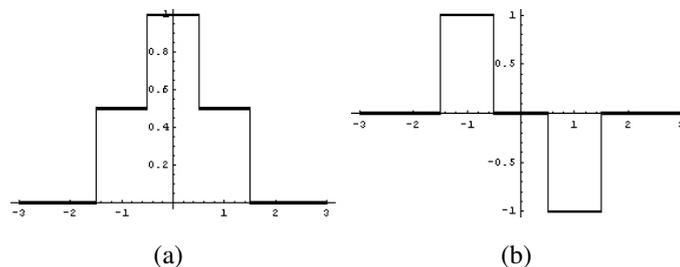


FIG. B.2 – Filtre de Sobel (a) Filtre passe-bas effectuant le lissage, (b) Filtre passe haut effectuant la dérivation.

Cette approche donne naissance à des masques de différences que l'on convolue à l'image. Elle donne de bons résultats sur des images faiblement bruitées. Plus le masque est grand, plus larges

sont les contours. Les opérateurs les plus connus sont les filtres de Prewitt et Sobel représentés respectivement Figures B.1 et E.1. Il s'agit de masques directionnels, séparables qui peuvent être vu comme la convolution d'un filtre de lissage et d'un filtre de dérivation comme explicité Equation 3.3. La direction du gradient est celle du masque donnant la réponse maximale.

Annexe C

Critères de Canny

L'approche de Canny [8] consiste à définir un filtre optimal satisfaisant trois critères : bonne détection, bonne localisation de la réponse et unicité de la réponse à un contour unique. Ces trois critères s'expriment par l'optimisation conjointe de trois fonctionnelles qui permettent de définir le filtre linéaire optimal pour la détection d'une marche d'escalier sous l'hypothèse d'un bruit additif indépendant du signal.

C.1 Critère de bonne détection

Il s'agit de minimiser la probabilité de ne pas détecter un vrai point de contour ou de marquer de faux points de contours. Ce critère correspond à la maximisation du rapport signal sur bruit SNR .

$$\Sigma = \frac{\int_0^{\infty} \psi(x) dx}{\sqrt{\int_{-\infty}^{\infty} \psi^2(x) dx}} \quad (\text{C.1})$$

où ψ est un filtre linéaire.

C.2 Critère de bonne localisation

Les points marqués comme contours par le détecteur doivent être aussi près que possible du centre du contour véritable. Ce critère revient à maximiser l'écart type de la position des passages par zéro, c'est à dire à l'inverse de l'espérance de la distance entre le vrai point de contour et le point de contour détecté.

$$\Lambda = \frac{|\psi'(0)|}{\sqrt{\int_{-\infty}^{\infty} \psi'^2(x) dx}} \quad (\text{C.2})$$

C.3 Critère d'unicité de la réponse

Le détecteur ne doit pas fournir de multiples réponses à un seul contour. Ce critère est inclus implicitement dans le critère de bonne détection puisque si l'on détecte deux contours alors qu'il n'y en a qu'un, l'une des deux réponses doit être considérée comme fautive. Cependant, la forme mathématique choisie ne met pas en évidence cette condition : elle ne mesure que la capacité du filtre à distinguer le signal du bruit au centre du contour. Ce critère permet de prendre en compte le comportement du filtre dans son voisinage et d'écartier les points de bruit.

$$\Gamma = \frac{|\psi'(0)|}{\sqrt{\int_{-\infty}^{\infty} \psi'^2(x) dx}} = k \times \frac{\int_{-\infty}^0 \psi(x) dx}{\sqrt{\int_{-\infty}^{\infty} \psi^2(x) dx}} \quad (\text{C.3})$$

La prise en compte de ces trois critères pour l'opérateur idéal de détection fournit une première approximation par la dérivée de la fonction de Gauss $g(i)$, proposée par Canny :

$$\phi(x) = k(\alpha|x| + 1)e^{-\frac{x^2}{2\sigma^2}} \quad (\text{C.4})$$

C.4 Résolution du système

Canny souhaite obtenir un filtre mono-dimensionnel à réponse impulsionnelle finie. Les critères de détection et de localisation étant antinomiques, on les combine en maximisant le produit $\Sigma\Lambda$ sous la contrainte du troisième critère. On obtient une équation différentielle dont la solution est :

$$\psi(x) = a_1 e^{\alpha x} \sin \omega x + a_2 e^{\alpha x} \cos \omega x + a_3 e^{-\alpha x} \sin \omega x + a_4 e^{-\alpha x} \cos \omega x + c \quad (\text{C.5})$$

Annexe D

Approche de Deriche

Au filtre de Canny, on préfère souvent le détecteur de Deriche [14], qui vérifie les mêmes critères, mais qui possède une réponse impulsionnelle infinie. Ce filtre est constitué d'un filtre de lissage et d'un filtre de dérivation. Le filtre de lissage préalable permet de s'immuniser en partie du bruit. Ce filtre s'écrit de façon récursive en fonction des valeurs de l'image, par l'intermédiaire de deux filtres l'un décrivant l'image de gauche à droite et l'autre de droite à gauche. Les images créées ont subi un lissage vertical (resp. horizontal) et une dérivation horizontale (resp. vertical). L'image finale est le maximum des valeurs des pixels des images obtenues.

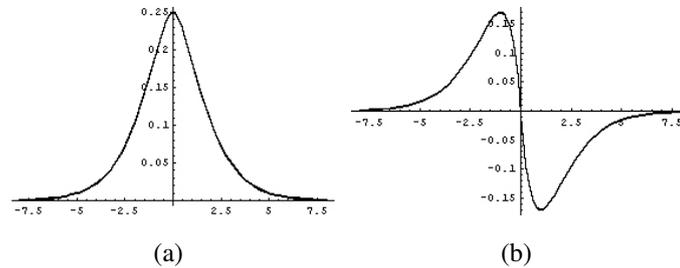


FIG. D.1 – (a) Filtre passe-bas de Canny-Deriche, (b) Filtre passe haut de Canny-Deriche.

Ce filtre est illustré en deux dimensions. Les calculs sont explicités pour estimer un gradient selon la direction verticale. L'estimation du gradient suivant une direction horizontale peut être simplement obtenu en inversant les positions des indices. $x[i, j]$ représente les valeurs des pixels de l'image à la position (i, j) , $y(i, n)$ l'image lissée et $z(i, j)$ le gradient de l'image.

Le filtre de lissage est de la forme $\phi(x) = k(\alpha|x| + 1)e^{-\alpha|x|}$ avec $k = \frac{(1-e^{-\alpha})^2}{1+2\alpha e^{-\alpha}-e^{-2\alpha}}$. Son implémentation récursive s'écrit $y(i, j) = y^+(i, j) + y^-(i, j)$

avec :

$$\begin{cases} y^+(i, j) = kx[i, j] + ke^{-\alpha}(\alpha - 1)x[i - 1, j] + 2e^{-\alpha}y^+(i - 1, j) - e^{-2\alpha}y^+(i - 2, j); \\ y^-(i, j) = ke^{-\alpha}(\alpha + 1)x[i + 1, j] - ke^{-2\alpha}x[i + 2, j] + 2e^{-\alpha}y^-(i + 1, j) - e^{-2\alpha}y^-(i + 2, j) \end{cases}$$

Le filtre de dérivation est de la forme $\psi(x) = -cxe^{-\alpha|x|}$ avec $c = \frac{(1-e^{-\alpha})^2}{e^{-\alpha}}$. Ce qui conduit à $z(i, j) = z^+(i, j) + z^-(i, j)$

où

$$\begin{cases} z^+(i, j) = ce^{-\alpha}y[i, j] + 2e^{-\alpha}z^+(i, j - 1) - e^{-2\alpha}z^+(i, j - 2) \\ z^-(i, j) = -ce^{-\alpha}y[i, j] + 2e^{-\alpha}z^-(i, j + 1) - e^{-2\alpha}z^-(i, j + 2) \end{cases}$$

Annexe E

Approche de Shen-Castan

Shen et Castan [56] ont également proposé un filtre optimal récursif minimisant les critères de Canny. Comme avec l'ensemble des détecteurs de contours, l'image est lissée par des filtres passe-bas, puis on lui applique un opérateur différentiel pour mettre en évidence les transitions.

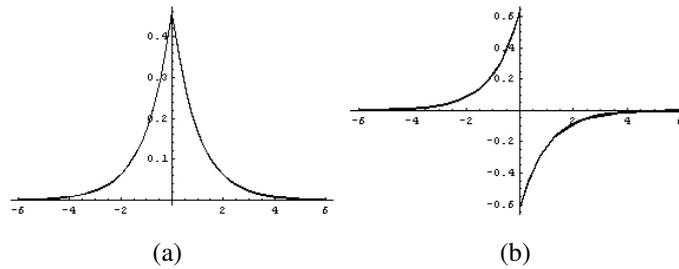


FIG. E.1 – (a) Filtre passe-bas de Shen-Castan, (b) Filtre passe haut de Shen-Castan.

Le filtre de lissage est de la forme $\phi(x) = ce^{-\alpha|x|}$ avec $c = \frac{(1-e^{-\alpha})}{(1+e^{-\alpha})}$. Son implémentation récursive s'écrit $y(i, j) = y^+(i, j) + y^-(i, j)$

avec :

$$\begin{cases} y^+(i, j) = cx[i, j] + e^{-\alpha}y^+(i - 1, j); \\ y^-(i, j) = cx[i + 1, j] - e^{-\alpha}y^-(i + 1, j) \end{cases}$$

Le filtre de dérivation est de la forme $\psi(x) = \begin{cases} de^{-\alpha x} & \text{si } x > 0 \\ de^{\alpha x} & \text{sinon.} \end{cases}$ avec $d = 1 - e^{-\alpha}$.

ce qui conduit à $z(i, j) = z^+(i, j) + z^-(i, j)$

où

$$\begin{cases} z^+(i, j) = dy[i, j] + e^{-\alpha}z^+(i, j - 1) \\ z^-(i, j) = -dy[i, j] + e^{-\alpha}z^-(i, j + 1) \end{cases}$$

Ce filtre présente une discontinuité d'ordre 1 au point 0 du filtre de Shen.

Publications de l'auteur

Journal international en science et technologie de l'information

- Fuzzy Edge Detection for Omnidirectional Images, JACQUEY Florence, COMBY Frédéric, STRAUSS Olivier, *FSS07 Fuzzy Sets and Systems, Elsevier*.

Congrès internationaux avec comité de lecture

- Non-additive approach for image gradient estimation, JACQUEY Florence, LOQUIN Kévin, COMBY Frédéric, STRAUSS Olivier, *ICIP07, IEEE International Conference on Image Processing*, présenté à San Antonio, Texas, USA, 16-19/09/2007.
- Non-additive Approach for Omnidirectional Image Gradient Estimation, JACQUEY Florence, COMBY Frédéric, STRAUSS Olivier, *ICCV'07 International Conference on Computer Vision, OMNIVIS'2007 the 7th Workshop on Omnidirectional Vision*, poster présenté à Rio de Janeiro, Brazil, 14-20 Novembre 2007.

Workshop

- Non-additive approach for omnidirectional image gradient estimation, JACQUEY Florence, COMBY Frédéric, STRAUSS Olivier, *OMNIVIS07, the 7th Workshop on Omnidirectional Vision at ICCV07, the IEEE International Conference on Computer Vision, Rio de Janeiro, Brazil*.

Congrès nationaux avec comité de lecture

- Détection de Contours Floue pour les Images Omnidirectionnelles, JACQUEY Florence, COMBY Frédéric, STRAUSS Olivier, *LFA06 Rencontres francophones sur la Logique Floue et ses Applications, Cépaduès édition, pages 379-386, présenté le 19/10/2006 à Toulouse, France*.

Colloques nationaux et groupes de recherche

- Nouvelle approche d'estimation de gradient pour les Images Omnidirectionnelles, JACQUEY Florence, COMBY Frédéric, *Journée thématique "Vision omnidirectionnelle", GDR/ISIS, 27/03/2007, Amiens.*
- Détection de Contours pour les Images Omnidirectionnelles, JACQUEY Florence, *DOCTISS07, La quatorzième session des journées des doctorants de l'école doctorale "Information, Structures et Systèmes", poster présenté le 19/04/2007 à Montpellier.*

Bibliographie

- [1] K. Daniilidis A. Makadia. Direct 3d-rotation estimation from spherical images via a generalized shift theorem. *CVPR'03 : Proceedings of the IEEE conference on computer vision and pattern recognition*, 2 :217, 2003.
- [2] C. Drocourt P. Vasseur C. Cauchois, E. Brassart. Calibration of the omnidirectional vision sensor : Syclop. pages 1287–1292, May 1999.
- [3] D. Kachi-Akkouche C. Demonceaux. Optical flow estimation in omnidirectional images using wavelet approach. volume 7, 2003.
- [4] P. Vasseur C. Demonceaux. Markov random fields for catadioptric image processing. *Pattern Pattern Recognition Letters*, 27(16) :1957–1967, 2006.
- [5] K. Daniilidis C. Geyer. Catadioptric projective geometry. *International Journal of Computer Vision*, 43 :223–243, 2001.
- [6] S. Sastry C. Geyer, R. Bajcsy. Euclid meets fourier : Applying harmonic analysis to essential matrix estimation in omnidirectional cameras. *OMNIVIS'04 : Proceedings of Omnidirectional Vision, Camera Networks and Non-classical Cameras in conjunction with ECCV'04 : European Conference on Computer Vision*, June 2004.
- [7] E.M. Mouaddib C. Pégard. A mobile robot using panoramic view. pages 89–94, April 1996.
- [8] J. Canny. A computational approach to edge detection. *Transactions on Pattern Analysis and Machine Intelligence*, 8(6) :679–698, November 1986.
- [9] G. Mauris H. Prade D. Dubois, L. Foulloy. Probability-possibility transformations, triangular fuzzy sets, and probabilistic inequalities. *Reliable Computing*, 10(4) :273–297, 2004.
- [10] H. Prade D. Dubois. Unfair coins and necessity measures : toward a possibilistic interpretation of histograms. *Fuzzy Sets and Systems*, 10 :15–20, 1983.
- [11] H. Prade D. Dubois. *Possibility Theory An Approach to Computerized Processing of Uncertainty*. Plenum Press, New York, 1988.
- [12] L. Foulloy G. Mauris D. Dubois, H. Prade. Probability-possibility transformations, triangular fuzzy sets, and probabilistic inequalities. *Reliable Computing*, 10 :273–297, 2004.
- [13] D. Denneberg. *Non Additive Measure and Integral*. Kluwer Academic Publishers, 1994.

- [14] R. Deriche. Fast algorithms for low-level vision. *Transactions on Pattern Analysis and Machine Intelligence*, 12(1), 1990.
- [15] C. Drocourt. *Localisation et modélisation de l'environnement d'un robot mobile par coopération de deux capteurs omnidirectionnels*. PhD thesis, Université de Technologie de Compiègne, 2002.
- [16] J.R. Fram E.S. Deutsch. A quantitative study of the orientational bias of some edge detector schemes. *IEEE Transactions on Computers*, March 1978.
- [17] O. Strauss F. Comby, C. Caderas De Kerleau. Étalonage de caméras catadioptriques hyperboloïdes.
- [18] O. Strauss F. Jacquy, F. Comby. Fuzzy edge detection for omnidirectional images. *LFA'06 : Rencontres francophones sur la Logique Floue et ses Applications*, pages 379–386, 2006.
- [19] J.J. Gonzalez-Barbosa. *Vision panoramique pour la robotique mobile : stéréovision et localisation par indexation d'images*. PhD thesis, Université de Toulouse III, 2004.
- [20] P. Greguss. *PAL-optic based instruments for space research and robotics*. 1996.
- [21] T. Pajdla H. Bakstein. Panoramic mosaicing with a 180° field of view lens. pages 60–67, Copenhagen, Denmark, June 2002.
- [22] R. Cutler H. Nanda. Practical calibrations for a real-time digital omnidirectional camera. *CV-PR'01 : Computer Vision and Pattern Recognition*, December 2001.
- [23] J. Thiran P. Vanderghyest I. Bogdanova, X. Bresson. Scale-space analysis and active contours for omnidirectional images. *IEEE Transactions on Image Processing*, pages 1888–1901, July 2007.
- [24] P. Frossard P. Vanderghyest I. Tomic, I. Bogdanova. Multiresolution motion estimation for omnidirectional images. 2005.
- [25] N. Ayache I.L. Herlin. A new methodology to analyse time sequences of ultrasound images. Rapport de Recherche 1390, Rocquencourt, 1991.
- [26] H. Ishiguro. Development of low-cost compact omnidirectional vision sensors and their applications. In *ISAS'98 : International Conference on Information Systems, Analysis and Synthesis*, pages 433–439, 1998.
- [27] S. Lacroix J-J. Gonzalez-Barbosa. Un algorithme rapide de stéréovision panoramique dense. 2003.
- [28] P. Bolon J.M. Chassery D. Demigny C. Graffigne A. Montanvert R. Zéboudj J. Zérubia J.P. Cocquerez, S. Philipp. *Analyse d'images : filtrage et segmentation*. Masson, 1995.
- [29] E.S. Deutsch J.R. Fram. On the quantitative evaluation of edge detection schemes and their comparisons with human performance. *IEEE Transaction on Computers*, C-24(6) :616–627, 1975.

- [30] T. Bulow K. Daniilidis, A. Makadia. Image processing in catadioptric planes : Spatiotemporal derivatives and optical flow computation. *OMNIVIS'02 : Workshop on Omnidirectional Vision*, pages 3–10, June 2002.
- [31] M. Yachida K. Yamazawa, Y. Yagi. Omnidirectional imaging with hyperboloidal projection. In *IROS'93 : Proceedings IEEE/RSJ International Conference on Intelligent Robots and Systems*, volume 2, pages 1029–1034. IEEE Computer Society, July 1993.
- [32] N. Ahuja K.H. Tan, H. Hua. Multiview panoramic cameras using a mirror pyramid. *OMNIVIS'02 : Workshop on Omnidirectional Vision*, page 87, June 2002.
- [33] B. Marhic P. Vasseur L. Delahoche, C. Pegard. A navigation system based on an omnidirectional vision sensor. 2 :718–724, September 1997.
- [34] V. Lasserre. *Modelisation Floue des Incertitudes de Mesures de Capteurs*. PhD thesis, Université de Savoie, 1999.
- [35] C. Barrows M. Barth. A fast panoramic imaging system and intelligent imaging technique for mobile robots. *IROS'96 : Proceedings of the 1996 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2 :626–633, November 1996.
- [36] J.L. Marichal. *Aggregation Operations for Multicriteria Decision Aid*. PhD thesis, Department of Mathematics, University of Liège, 1998.
- [37] I. Hara H. Asoh T. Matsui N. Vlassis, Y. Motomura. Edge-based features from omnidirectional images for robot localization. volume 2, pages 1579– 1584, Seoul, Korea, 2001.
- [38] V. Nalwa. A true omnidirectional viewer. 1996.
- [39] S.K. Nayar. Omnidirectional video camera. *Proceedings of DARPA Image Understanding Workshop*, May 1997.
- [40] C. Gastaud P. Garda O. Romain, T. EA. Un capteur multi-spectral de vision panoramique 3d. *ORASIS'01 : Congrès francophone de vision*, pages 359–366, June 2001.
- [41] S. Sastry O. Shakernia, R. Vidal. Omnidirectional egomotion estimation from back-projection flow. Madison Wisconsin, USA, June 2003.
- [42] F. Comby O. Strauss. Fuzzy morphology for omnidirectional images. volume 2, pages 141–144, Genova, Italy, September 2005.
- [43] Y. Aloimonos R. Pless P. Baker, C. Fermüller. A spherical eye from multiple cameras (makes better model of the world). volume 1, pages 576–583, Kauaii, USA, December 2001.
- [44] B. Krose R. Bunschoten. Robust scene reconstruction from an omnidirectional vision system, 2003.
- [45] J. Salvi R. Orghidan, E. Mouaddib. Système de vision 3d omnidirectionnelle. In *9èmes journées ORASIS*, Puy-de-Dôme, France, May 2005.

- [46] S.K. Nayar R. Swaminathan. Non-metric calibration of wide-angle lenses and polycameras. 2 :413–419, June 2000.
- [47] S.K. Nayar R. Swaminathan, M.D. Grossberg. Non-Single Viewpoint Catadioptric Cameras : Geometry and Analysis. *International Journal of Computer Vision*, 66(3) :211–229, Mars 2006.
- [48] C. Rabaud. *Une Nouvelle Approche de Mise en Correspondance Stéréoscopique Dense par Méthodes Possibilistes*. PhD thesis, Université de Montpellier II, 2005.
- [49] Srikumar Ramalingam, Peter Sturm, and Suresh Lodha. Generic calibration of axial cameras. Rapport de Recherche 5827, INRIA, December 2005.
- [50] D. Rees. Panoramic television viewing system. 3,505,465, 1970.
- [51] S.K. Nayar S. Baker. A theory of single-viewpoint catadioptric image formation. *International Journal of Computer Vision*, 35(2) :175–196, 1999.
- [52] S.K. Nayar S. Baker. *Single Viewpoint Catadioptric Cameras*. Springer-Verlag, 2001.
- [53] S. Durand E. Mouaddib S. Bigot, D. Kachi. Spherical image denoising and its application to omnidirectional imaging. pages 8–11, March 2007.
- [54] J. Devars S. Ieng, R. Benosman. An efficient dynamic multi-angular feature points matcher for catadioptric views. *CVPRW'03 : Conference on Computer Vision and Pattern Recognition Workshop OMNIVIS'03 : Omnidirectional Vision and Camera Networks*, 7, June 2003.
- [55] M. Ban-Ezra S. Peleg. Stereo panorama with a single camera. pages 395–401, June 1999.
- [56] J. Shen and S. Castan. An optimal linear operator for step edge detection. *CVGIP'92 : Computer Vision Graphics and Image Processing*, 54(2) :112–133, 1992.
- [57] A. Karmarkar S.K. Nayar. 360*360 mosaics. volume 2, pages 388–395, Hilton head, South Carolina, USA, June 2000.
- [58] P. Smets. Constructing the pignistic probability function in a context of uncertainty. In M. Henrion, R. D. Schachter, L. N. Kanal, and J. F. Lemmer, editors, *Uncertainty in Artificial Intelligence 5*, pages 29–39. Elsevier Science Publishers, North-Holland, Amsterdam, 1990.
- [59] L. Spacek. Omnidirectional catadioptric vision sensor with conical mirrors. *TIMR'03 : Towards Intelligent Mobile Robotics*, August 2003.
- [60] R. Bajcsy S.S. Lin. True single view point cone mirror omni-directional catadioptric system. 2 :102–107, July 2001.
- [61] C. Gastaud P. Garda T. EA, O. Romain. Un capteur de sphereo-vision stereoscopique couleur. *ORASIS'01 : Congrès francophone de Vision par Ordinateur*, June 2001.
- [62] S. Vedula H. Saito T. Kanade, P. Rander. Virtualized reality : Digitizing a 3d time-varying event as is and in real time. In Hideyuki Tamura Yuichi Ohta, editor, *Mixed Reality, Merging Real and Virtual Worlds*, pages 41–57. Springer-Verlag, 1999.

- [63] S. Thompson T. Matsui, H. Asoh. Mobile robot localization using circular correlations of panoramic images. *Proceedings of the 2000 IEEE/RSJ International Conference on Intelligent Robots and Systems IROS2000*, November 2000.
- [64] V. Hlavàc T. Pajdla. Image-based self-localization by means of zero phase representation in panoramic images. In *Proceedings of the Second International Conference on Advances in Pattern Recognition*, pages 24–31, London, UK, 2001. Springer-Verlag.
- [65] V. Hlavac T. Svoboda, T. Pajdla. Motion estimation using central panoramic cameras. *International Conference on Intelligent Vehicles*, pages 335–340, October 1998.
- [66] N. Winters and J. Santos-Victor. Omni-directional visual navigation, July.
- [67] H. Tsuji Y. Yagi, S. Kawato. Real-time omnidirectional image sensor (copis) for vision-guided navigation. *IEEE Transactions Robotics and Automation*, 10 :11–22, February 1994.
- [68] Y. Yagi and S. Kawato. Panorama scene analysis with conic projection. In *IROS'90 : Proceedings IEEE/RSJ International Workshop on Intelligent Robots and Systems*, pages 181–187, Tsuchiura, July 1990.
- [69] G. Xu X. Lin D. Shi Z. Zhu, S. Yang. Fast road clasification and orientation estimation using omni-view image an neural networks. *Transactions on Image Processing*, 7 :1182–1197, 1998.
- [70] L.A. Zadeh. Knowledge representation in fuzzy logic. In R. R. Yager and L. A. Zadeh, editors, *An Introduction to Fuzzy Logic Applications in Intelligent Systems*, pages 1–25. Kluwer, Boston, USA, 1992.

Index

- α -coupe, 54, 147
- Acquisition, 36
- Algorithme de projection, 59
- Angle solide, 24
- Approche floue, 65
- Approche proportionnelle, 64
- Astigmatisme, 18
- Bruit de quantification, 37
- Caméra à balayage, 39
- Caméra matricielle, 39
- Coma, 24
- Condition de cohérence, 46, 141
- COPIS, 21
- Critère d'unicité de la réponse, 152
- Critère de bonne détection, 96, 105, 126, 151
- Critère de bonne localisation, 95, 151
- Critères de Canny, 151
- Deriche, 153
- Dichotomie, 55
- Différences finies, 149
- Distribution de possibilité, 45
- Distribution de probabilité, 45
- Echantillonnage, 36
- Ensembles flous, 147
- Erreurs de quantification, 37
- Estimation du gradient, 72
- Filtre Gaussien, 58
- Filtre moyennneur, 58
- Fish-eye, 8
- Fonction caractéristique, 147
- Fréquence d'échantillonnage, 38
- Humanoid Machine Vision System, 9
- Hypothèse de continuité de la matière, 40
- Intégrale de Choquet, 49, 50, 67
- Intervalle flou, 147
- Laplacien, 58
- Lentille télécentrique, 19
- Masques de convolution, 57
- Maximisation de gradient, 74
- Mesure de nécessité, 46
- Mesure de probabilité, 45
- Mesure des écarts, 95, 106, 124
- Mesure des possibilités, 45
- Mesures de confiance, 46
- Miroir conique, 20
- Miroir elliptique, 17
- Miroir hyperbolique, 18
- Miroir parabolique, 18
- Miroir plan, 16
- Miroir sphérique, 21
- Modèle de projection unifié, 19
- Nombre flou, 148
- Nombres flous triangulaires, 47
- Noyau, 147
- Noyaux non-sommatifs, 45
- Noyaux sommatifs, 44
- Numérisation, 36

-
- Panoramic Annular Lens, 9
Paramètres de Fram et Deutsch, 93
Partition couvrante, 43
Partition floue, 49
Point de vue unique, 15
Prewitt, 58
Principe de domination, 46, 48
Principe de préservation de la préférence, 47, 141
Principe du maximum de spécificité, 47, 141
Probabilités, 44
Produit de convolution, 57
Projection orthographique, 18
Propriété de maximisation normalisée, 45
Propriété de sommativité normalisée, 44
- Qualité, 37
Quantification, 36
Quantité floue, 147
- Régression linéaire, 95, 96, 108, 125
Rapport signal sur bruit, 96, 105, 126, 151
- Seuillage par hystérésis, 75
Shen-Castan, 155
Signal analogique, 35
Signal numérique, 35
Sobel, 58
Sous-ensembles flous, 147
Sous-ensembles flous normalisé, 147
Support, 147
SYCLOP, 21
Système de caméras multiples, 11
Système rotatif, 10
Systèmes catadioptrique axial, 22
Systèmes catadioptrique caustiques, 23
Systèmes catadioptriques, 13
- Théorème de Nyquist-Shannon, 38
Transfert de croyance, 64
Transfert pignistique, 51, 64
Transformation Probabilité-Possibilité, 46
- Valeur modale, 147
Vision horizontale, 54
Vision verticale, 51

Traitement d'images omnidirectionnelles

Résumé : Ce manuscrit présente une approche novatrice au problème du traitement d'images omnidirectionnelles. Elle se démarque des autres approches par le fait que les traitements sont adaptés à la géométrie de l'image et effectués directement dans l'espace de l'image afin d'interpoler le moins possible les données originales de l'image. L'originalité de ce travail repose sur l'utilisation d'un ensemble de techniques de représentation de l'imprécision de la mesure de luminance induite par l'échantillonnage spatial de l'image. Ces techniques s'appuient sur une représentation par partition floue de l'image.

Afin de traiter les images issues de capteurs omnidirectionnels catadioptriques, nous proposons une approche basée sur l'utilisation d'un espace projectif virtuel appelé le cylindre englobant. Ce cylindre nous permet de définir les différents opérateurs de convolution dans un espace quasi identique à celui de l'image perspective. Les opérateurs ainsi définis sont ensuite projetés sur l'image omnidirectionnelle pour tenir compte de la géométrie du capteur. Dans un premier temps, nous avons proposé une adaptation des masques de convolution aux images omnidirectionnelles prenant en compte les imprécisions sur la localisation du niveau de gris des pixels. Dans un second temps, nous nous sommes intéressés plus précisément aux détecteurs de contours. Nous avons développé une nouvelle approche d'estimation de gradient utilisant des noyaux possibilistes. L'estimation du gradient est fournie sous la forme d'un intervalle de valeurs. Cet intervalle permet de seuiller naturellement les valeurs de gradient obtenues.

La validité de nos approches est établie par des expérimentations sur des images omnidirectionnelles synthétiques et réelles. Les images synthétiques sont utilisées afin de tester la robustesse de nos détecteurs de contours en présence de bruit et d'étudier la bonne détection, bonne localisation et réponse unique du contour. Les images réelles permettent d'observer le comportement de nos méthodes vis à vis des variations d'illumination et des bruits d'acquisition.

Mots clefs : Traitement d'images omnidirectionnelles, détection de contours, imprécisions, théorie des possibilités, intégrale de Choquet.

Omnidirectional image processing

Abstract : This manuscript presents new approaches to process omnidirectional images. These methods differ from classical approaches related to image processing because they are adapted to the omnidirectional image geometry and applied directly in the image space to avoid interpolated data. One of the main originality of this work is related to the use of fuzzy sets theory to represent the imprecisions of the luminance measure due to the image sampling. This modeling associates to each pixel of the image, a fuzzy domain.

To process omnidirectional catadioptric images, a projective space, constituted by the unified projection model and the surrounding cylinder, is defined. The cylinder space is locally equivalent to the space of all perspective images acquired by a rotating camera. The operators are geometrically defined on the cylinder and projected on the omnidirectional image. First, we have adapted the convolution masks by taking into account the imprecision on the grey level localization. Next, we have studied edge detection and developed a new non-additive gradient based approach. The gradient value in each pixel is estimated by an interval which provide a natural threshold.

The validity of our approach is illustrated with experimentations on synthetic and real omnidirectional images. The synthetic images are first used to test the sensibility of our detectors to noise and then to study the good detection and good localization of the obtained edges. The real images allow to test the method's robustness towards classical disruptions in image processing such as brightness variation and acquisition noise.

Keywords : Omnidirectional images processing, edge detection, imprecisions, possibility theory, Choquet integral.