

ECOLE NATIONALE SUPERIEURE POLYTECHNIQUE



DEPARTEMENT D'ELECTRONIQUE

THESE

DE DOCTORAT EN ELECTRONIQUE

Présentée par

AMMOUR Nassim

Chargé de cours au Département d'Electronique –USD-Blida

THEME

Méthode de segmentation d'image rapide basée sur
l'évolution probabiliste d'un graphe et de la
théorie de l'évidence

Devant le jury composé de :

M ^{me}	GUERTI Mhania	Professeur à l'ENP	Présidente
Mr	BENSEBTI Messaoud	Professeur à USD-Blida	Examineur
Mr	SALHI Hassan	Maître de Conférences à USD-Blida	Examineur
M ^{me}	HAMAMI Latifa	Maître de Conférences à l'ENP	Examinatrice
Mr	GUESSOUM Abderrezak	Professeur à USD-Blida	Rapporteur
Mr	BERKANI Daoud	Professeur à l'ENP	Rapporteur

Remerciements

EN premier lieu, je tiens à remercier Monsieur Guessoum Abderrezak, Professeur à l'USDB pour m'avoir proposé ce sujet de recherche. Je tiens à lui adresser mes sincères et chaleureux remerciements pour la confiance qu'il m'a témoigné, pour sa disponibilité de tous les instants, pour son aide précieuse et ses compétences scientifiques qui m'ont permis de mener à bien ce travail.

Mes remerciements s'adressent également à mon co-directeur de thèse, Monsieur Berkani Daoud, Professeur à l'ENP pour m'avoir accueilli au sein de son groupe des doctorants, pour m'avoir accordé toute sa confiance, et pour son aide inestimable pour l'aboutissement de ce travail.

Je remercie Madame Guerti Mhania, Professeur à l'ENP pour l'intérêt qu'elle a porté à mon travail, dans son rôle de présidente du jury de cette thèse.

Merci à Monsieur Salhi Hassane Maître de conférence à l'USDB, Monsieur Bensebti Messaoud Professeur à l'USDB et Madame Hamami Latifa Maître de conférence à l'ENP, d'avoir accepté d'être les membres de mon jury.

Je tiens à exprimer ma gratitude bien sûr à ceux qui m'ont permis d'exister : Mes Parents, qui m'ont toujours soutenu dans mes choix et m'ont permis d'en arriver là.

Je tiens à remercier tous mes amis qui m'ont soutenu, et toute personne qui a participé de près ou de loin à l'aboutissement de ce travail.

ملخص

في هذه الأطروحة نقدم طريقة جديدة تعالج تقسيم الصورة. تعتمد هذه الطريقة على إستراتيجية جديدة التي تقوم بتمديد مخطط متغير الذي يمثل تركيب الصورة لتقسيمها. قرار التمديد يتم إحصائيا بعد مسح منطقة من صورة. يتم بعد ذلك تعيين مصفوفة المناطق لكل قسم. نقوم بعد ذلك برسم خريطة مناطق الحدود. وأخيرا نقوم بأخر مرحلة في تقسيم والتي تتمثل في ترتيب المواقع التي لم يمكن ترتيبها خلال المرحلة الأولى والتي اجتنبت خلال مسح الصورة. تتم هذه العملية باستعمال نظرية البديهي.

كلمات المفتاح - مخطط متغير. انتقال. تقسيم. قسم. محيط. منطقة. نظرية البديهي.

Résumé - Dans cette thèse nous présentons une nouvelle méthode de segmentation non supervisée d'image. Elle repose sur une stratégie originale qui consiste à faire progresser un graphe évolutif de composition de l'image à segmenter. L'injonction d'évolution est établie statistiquement après la traversée d'une région. La matrice de composition spatiale des régions dans chaque classe est ensuite déterminée. Une carte de localisation spatiale des sites de conflit qui représentent les zones bruitées ou les contours est établie. Enfin, la phase finale de la segmentation est réalisée, elle consiste à classifier les pixels qui occupent les zones de conflit et qui sont omis durant la première phase. A cette fin la technique de classification basée sur la théorie de l'évidence de Dempster-Shafer est utilisée. La robustesse de la méthode est illustrée sur des images synthétiques et des images réelles.

Mots clés :- Graphe évolutif, transition, segmentation, classe, contour, région, théorie de l'évidence de Dempster-Shafer.

Abstract - In this thesis, we introduce a new color images segmentation algorithm. The color scene analytic method is based on the progress of a probabilistic evolutionary graph. The strategy consists in making grow an evolutionary graph, which presents the scene elements in an unsupervised segmented image. The graph evolution development is based on the computation of the belonging probabilities to the existing classes of the last built region. The space composition matrix of the areas in each class is then given. A space delimitation map of the regions is established by a new method of contour localization and refinement. At last, the final segmented image is established by classification of the pixels in the conflict region using the Dempster-Shafer evidence theory. The effectiveness of the method is demonstrated on synthetic and real images.

Key-Words: - Evolutionary graph, transition, segmentation, class, contour, region, Dempster-Shafer evidence theory.

Table des Matières

Remerciements	i
Résumé	ii
Table des matières	iii
Table des figures	vi
Introduction générale	1
Chapitre I : Le robot mobile	5
1. Introduction	5
2. Domaine d'application	5
3. Les architectures de contrôle	8
4 Les sources d'information	6
5 Les capteurs	6
6 Les caméras	8
6.1 Caméras simples.....	8
6.2 Caméras stéréoscopiques.....	8
6.3 Caméras panoramiques.....	8
7 Autres capteurs	9
7.1 Les capteurs tactiles.....	9
7.2 Les boussoles.....	9
7.3 Les balises.....	9
7.4 Le GPS.....	10
8. Navigation	10
9 Évitement d'obstacles	12
9.1 Méthode des champs de potentiel.....	12
9.2 Méthode Vector Field Histogram.....	13
9.3 Méthode de la fenêtre dynamique.....	14
10 Localisation	15
10.1 Localisation par les cartes topologiques.....	15
10.2 Localisation par les Cartes métriques.....	15
11 Planification	16
11.1 Espace des configurations.....	16
11.2 Discrétisation de l'espace de recherche.....	17
11.3 Recherche de chemin.....	18
12 Conclusion	18
Chapitre II : De l'image à la vision artificielle	
1 Introduction	20
2 Le système de vision	20
3 De la vision humaine à la vision artificielle	21
4 Les objets et la lumière	21

5 La vision des couleurs (chromatopsie)	22
6 Physiologie de l'œil	23
6.1. Généralités.....	23
6.2. La rétine.....	24
6.3. Les Photorécepteurs.....	25
6.4. Répartition des photorécepteurs sur la rétine.....	26
7 La vision artificielle	27
8 L'image et son traitement	28
8.1. Préambule.....	28
8.2. L'image.....	28
8.3. Les capteurs d'images.....	29
9 Traitement d'image	29
9.1. Le Prétraitement d'image.....	30
9.2. Segmentation.....	30
9.3. Analyse d'images.....	30
10 Conclusion	31

Chapitre III : Segmentation d'image état de l'art.....

1. Introduction	32
2 Définition de la segmentation	32
3 Les différentes approches de la segmentation d'image	34
3.1 Contour.....	34
3.2 Région.....	37
3.2.1 Classification déterministe supervisée : Réseaux de neurones.....	37
3.2.2 Classification déterministe non supervisée.....	38
3.2.3 Classification probabiliste paramétrique (Mélange de Lois).....	39
3.2.4 Classification probabiliste non-paramétrique (Champs Aléatoires de Markov)..	40
3.3 Approches structurales (Ligne de Partage des Eaux).....	42
3.4 Théorie des graphes et hypergraphe.....	43
4 Conclusion	44

Chapitre IV : Théorie des graphes et théorie de l'évidence

1 Théorie des graphes	45
1.1 Introduction.....	45
1.2 Définition d'un graphe.....	45
1.3 Graphes non orientés.....	45
1.4 Graphes orientés.....	46
1.5 Définitions.....	46
1.6 Graphes eulériens (Théorème d'Euler 1766)	47
1.7 Graphes hamiltoniens.....	47
1.8 Matrice d'adjacence.....	48
1.9 Distance.....	48
2 Théorie de l'évidence	49
2.1 Introduction	49
2.2 Eléments mathématiques.....	51
2.2.1 Représentation de l'information.....	51
2.2.2 Outils associés.....	53
5 Conclusion	54

Chapitre V : Segmentation d'image par graphe évolutif probabiliste avec raffinement par la théorie de l'évidence.....

1 Introduction.....	55
2 Méthodologie générale.....	55
3 Etablissement du graphe.....	56
3.1. Parcours de l'image.....	56
3.2. L'état initial.....	56
3.3. Détection d'une transition.....	56
3.4. Première transition.....	56
3.5. Deuxième transition.....	56
3.6. Création d'un nouveau nœud.....	57
3.7. Matrice des classes.....	57
3.8. Carte des contours.....	58
3.9. L'image segmentée.....	59
4 Raffinement des extrémités (Classification dans les régions de conflit).....	59
4.1. Théorie de l'évidence.....	59
4.2. Niveau crédal (Modélisation de la connaissance).....	59
4.3 Niveau pignistique (Prise de décision).....	60
5 Application à la classification dans les zones de conflit.....	61
6 Application de la méthode à des images à niveaux de gris.....	61
7 Application de l'algorithme à des images couleurs.....	62
7.1. Segmentation d'une image couleur RVB synthétique bruitée avec un bruit Gaussien $N(0,0.01)$	63
7.2. Segmentation d'une image couleur synthétique bruitée avec un bruit Gaussien $N(0,0.02)$	66
7.3 Segmentation d'image en couleurs réelle avec $ds=0.7$	69
7.4 Segmentation d'image en couleurs réelle avec $ds=0.8$	71
7.5 Segmentation d'image réelle avec $ds=0.9$	72
8 Etude comparative.....	72
9 Segmentation d'image réelle.....	74
10 Apport de la nouvelle méthode comparativement aux autres méthodes.....	76
11 Conclusion.....	76
Conclusions générales et perspectives.....	77
Références Bibliographiques.....	80

Table des figures

Fig.1.1 :	Schéma des interactions d'un robot avec son environnement.	5
Fig.1.2:	Illustration des architectures classiques des contrôleurs pour les robots mobiles	6
Fig.1.3:	Principe du télémètre à ultrasons et exemple de télémètre réel.	7
Fig.1.4:	Illustration d'un télémètre Laser.	8
Fig.1.5:	Principe des caméras panoramiques catadioptriques et exemple d'image obtenue.	9
Fig.1.6:	Action associée à un lieu.	10
Fig.1.7:	Navigation topologique.	11
Fig.1.8:	Navigation métrique.	12
Fig.1.9:	Illustration de la combinaison de différents potentiels primitifs.	12
Fig.1.10:	Illustration de la combinaison de différents potentiels primitifs dans l'espace relatif au robot.	13
Fig.1.11:	Partie supérieure : Grille d'occupation locale. Partie Inférieure : Utilisation de l'histogramme des obstacles pour déterminer la direction de déplacement du robot.	13
Fig.1.12:	Contrainte d'évitement d'obstacles pour la méthode de la fenêtre dynamique	14
Fig.1.13:	Fenêtre de sélection des vitesses tenant compte de la dynamique du robot.	15
Fig.1.14 :	Estimation de la position d'un robot par une comparaison de la carte locale avec la carte globale de l'environnement.	16
Fig.1.15:	Décompositions en cellules de l'espace libre dans les cartes métriques.	17
Fig.1.16:	Décompositions en chemins pré-calculés dans les cartes métriques.	18
Fig.2.1 :	Structure d'une onde électromagnétique	22
Fig.2.2 :	Spectre des ondes électromagnétiques	22
Fig.2.3 :	Spectre visible	23
Fig.2.4 :	Anatomie de l'œil	24
Fig.2.5 :	La rétine.	25
Fig.2.6:	Schéma global du système visuel humain.	25
Fig.2.7:	Sensibilité des différents types de photorécepteurs à la longueur d'onde de la lumière	26
Fig.2.8:	Concentration des photorécepteurs en fonction de l'excentricité	26
Fig.2.9 :	Processus de Vision Artificielle	27
Fig.2.10:	Schéma de la chaîne de traitement d'image	29
Fig.2.11:	Différents types de bruit	30
Fig.3.1 :	Image originale.	32
Fig.3.2 :	Image segmentée	32
Fig.3.3 :	Schéma d'un système imageur	33
Fig.3.4 :	Contours dans une image, a- Image originale, b- Contour des objets de l'image.	34
Fig.3.5 :	Modèles de contours	35
Fig.3.6 :	Opérateurs dérivatifs	35
Fig.3.7 :	Structure d'un neurone artificiel	38
Fig.3.8:	Principe de la Ligne de Partage des Eaux.	42
Fig.3.9:	Exemple d'hypergraphe par Claudio Rocchini	43
Fig.4.1 :	Exemple d'un graphe non orienté	45
Fig.4.2 :	Liens entre sommets et arêtes	46
Fig.4.3 :	Graphe orienté	46
Fig.4.4 :	G1 n'est pas un graphe eulérien, G2 est un graphe eulérien	47
Fig.4.5:	Le graphe G1 n'est pas hamiltonien (sommet de degré 1), le graphe G2 est hamiltonien	48
Fig.4.6 :	Graphe G et sa matrice d'adjacence	48

Fig.4.7:	le graphe G et sa matrice des distances.	49
Fig.5.1:	Balayage de l'image	56
Fig.5.2:	Matrice des régions d'une image synthétique.	58
Fig.5.3:	Choix des directions	58
Fig.5.4:	Carte des contours d'une image synthétique.	59
Fig.5.5:	Image réelle à segmenter	62
Fig.5.6:	L'image segmentée	62
Fig.5.7:	L'image synthétique pure à segmenter.	63
Fig.5.8:	L'image synthétique bruitée.	63
Fig.5.9:	L'image des zones de conflit.	63
Fig.5.10:	l'image des classes avant raffinement	64
Fig.5.11:	L'image des classes après raffinement.	64
Fig.5.12:	L'image segmentée sans raffinement.	65
Fig.5.13:	L'image segmentée avec raffinement.	65
Fig.5.14:	Image segmentée avec $d_s = 0.8$, $k=6$: (a) sans raffinement, (b) avec raffinement	65
Fig.5.15:	Image segmentée avec $d_s = 0.9$, $k = 9$: (a) sans raffinement, (b) avec raffinement	66
Fig.5.16:	L'image synthétique à segmenter avec $d_s = 0.7$: (a) L'image synthétique pure.	66
Fig.5.17:	L'image des zones de conflit.	67
Fig.5.18:	l'image des classes avant raffinement.	67
Fig.5.19:	L'image des classes après raffinement.	67
Fig.5.20:	L'image segmentée : (a) sans raffinement. (b) avec raffinement.	68
Fig.5.21:	Image segmentée avec $d_s = 0.8$, $k=9$: (a) sans raffinement. (b) avec raffinement.	68
Fig.5.22:	Image segmentée avec $d_s = 0.9$, $k=30$: (a) segmentation sans raffinement. (b) segmentation avec raffinement.	69
Fig.5.23 :	Image réelle d'un puzzle à segmenter	69
Fig.5.24:	L'image des zones de conflit.	70
Fig.5.25:	L'image segmentée avec $d_s = 0.7$: (a) Segmentation sans raffinement. (b) Segmentation avec raffinement	70
Fig.5.26:	L'image segmentée avec $d_s = 0.8$: (a) Segmentation sans raffinement. (b) Segmentation avec raffinement.	71
Fig.5.27:	L'image segmentée avec $d_s = 0.9$: (a) Segmentation sans raffinement. (b) Segmentation avec raffinement	72
Fig.5.28:	Image synthétique à segmenter.	72
Fig.5.29:	Image des zones de conflit.	73
Fig.5.30:	Image segmentée sans raffinement.	73
Fig.5.31:	Image segmentée avec raffinement.	73
Fig.5.32:	Image réelle à segmenter.	74
Fig.5.33:	Image des zones de conflit.	75
Fig.5.34:	Image segmentée sans raffinement.	75
Fig.5.35:	Image segmentée avec raffinement.	75

Introduction Générale

Le cadre général dans lequel s'inscrit cette thèse est celui de la segmentation d'image. Cette opération constitue la base de tous les processus d'interprétation d'images. Bien que le sujet ait déjà fait l'objet d'une vaste littérature [1], on n'arrive pas à trouver une méthode générale de segmentation applicable quelque soit le domaine d'intérêt [2]. Nos travaux sont orientés vers le problème de la navigation d'un robot mobile guidé par la vision dans un environnement partiellement ou totalement inconnu. Les recherches menées portent sur l'étude et la conception de machines autonomes, intégrant des capacités de perception, d'action et de raisonnement. Ces machines doivent interagir rationnellement avec un environnement variable et évolutif pour la réalisation de tâches diverses. Dans cette vision, on s'intéresse à doter un robot mobile d'un système visuel, lui permettant ainsi d'entrer dans le monde des voyants. Cette capacité, vue par les roboticiens comme précieuse et puissante, ouvre de nombreuses voies aux chercheurs pour la réalisation des tâches robotiques complexes en se basant notamment sur les images captées et traitées.

Les procédés d'asservissement visuel consistent à utiliser les informations fournies par un système visuel, qui regroupe le capteur de vision et la chaîne de traitement, pour contrôler les mouvements d'un système dynamique, ce système pouvant être réel dans le cadre de la robotique, ou bien virtuel dans le cadre de l'animation d'entités artificielles ou de la réalité augmentée. Ce thème de recherche se situe à l'intersection des domaines de la robotique, de l'automatique, et de la vision par ordinateur. A l'intérieur de ce spectre, nous nous focalisons sur l'interaction entre l'action et la perception visuelle.

Le traitement de l'image est considéré comme un sous-domaine du traitement du signal, science au carrefour des mathématiques, de la physique et de l'informatique. Les années vingt voient le balbutiement dans le domaine de l'image sans véritable évolution jusqu'à la fin de la seconde guerre mondiale. Les techniques de traitement du signal, et par conséquent celles de l'image, évoluèrent grâce à l'arrivée du radar ainsi que la prospection pétrolière. Dans les années quatre-vingt, le milieu médical s'intéresse aux images et particulièrement aux techniques d'imageries afin d'améliorer les diagnostics.

Les images numériques ont envahi notre quotidien, ainsi que celui de nombreux secteurs professionnels. On observe ainsi que les entreprises qui développent et/ou utilisent les techniques de traitement d'image au sens large sont extrêmement nombreuses et variées. Ainsi sans être exhaustif, les domaines d'applications sont le médical et la biologie, la sécurité et la surveillance, le multimédia, la photographie et la vidéo, la robotique, la vision industrielle, la biométrie ou encore la télédétection.

Dans le domaine de la robotique[3, 4, 5, 6], le problème du contrôle robotique par la vision artificielle réside dans l'association intelligente d'un système de perception et d'un système d'action en vue de l'exécution de tâches. Cette association est univoque. Le système sensoriel de vision artificielle fournit au système moteur robotique les informations utiles à la réalisation des tâches, mettant en évidence une articulation intelligente de la perception sur l'action. La synthèse de lois de commande permettant de piloter efficacement un robot mobile apparaît comme un enjeu important dans la définition d'une tâche robotique. En effet, il faut pouvoir asservir le système robotique de manière à réaliser les mouvements désirés, tout en tenant compte de la variabilité et de l'incertitude

que l'on peut avoir de l'environnement. La littérature regorge de méthodes concernant la commande des robots mobiles. Généralement, ces lois de commande sont posées en termes de régulation à zéro d'une erreur entre la situation du véhicule et une situation définie a priori par la trajectoire, le chemin ou la posture de référence [7, 8, 9, 10].

Dans le domaine médical, l'imagerie constitue aujourd'hui un outil très important dans l'aide au diagnostic pour le médecin. Les modalités d'acquisition de l'information image sont nombreuses et la quantité d'images manipulées s'accroît. La radiographie par rayon X constitue la part la plus importante du marché, l'imagerie ultrasonore, rapide, non invasive et peu onéreuse, gagne du terrain. L'imagerie médicale est certainement l'un des domaines de la médecine qui a le plus progressé ces vingt dernières années. Ces récentes découvertes permettent non seulement un meilleur diagnostic mais offrent aussi de nouveaux espoirs de traitement pour de nombreuses maladies. Cancer, épilepsie et d'autres maladies, l'identification précise de la lésion facilite déjà le recours à la chirurgie, seule solution thérapeutique pour certains malades. De telles techniques permettent également de mieux comprendre le fonctionnement de certains organes encore mystérieux, comme le cerveau.

Découverte il y a plus d'un siècle, la radiographie utilise les rayons X, capables de se jouer de la matière. Passant à travers une certaine partie du corps, ils impressionnent un film radiographique, plus ou moins noirci en fonction de l'organe traversé.

L'échographie est une technique d'exploration de l'intérieur du corps basée sur les ultrasons. Une sonde envoie un faisceau d'ultrasons dans la zone du corps à explorer. Selon la nature des tissus, ces ondes sonores sont réfléchies avec plus ou moins de puissance. Le traitement de ces échos permet une visualisation des organes observés. L'imagerie ultrasonore constitue la modalité principale dans certains domaines d'application. Elle peut également être utilisée en complément d'une autre modalité d'image où les autres modalités d'imagerie présentent un danger pour le patient. L'imagerie ultrasonore possède également l'avantage d'être très pratique et simple d'utilisation. Le médecin a la possibilité de déplacer la sonde échographique très librement pour explorer les régions suspectes. Par ailleurs, la rapidité d'obtention des images en temps réel fait de l'imagerie ultrasonore une modalité d'image très intéressante pour la visualisation et le diagnostic des tissus en mouvement. Elle est particulièrement bien adaptée à l'étude du cœur, l'échocardiographie. En effet, la dynamique temporelle ou la déformation des structures cardio-vasculaires est intimement liée à leur fonction. L'analyse des images ultrasonores permet de fournir au praticien des informations quantitatives fiables en 2D ou en 3D, informations qui complètent l'aspect visuel et qualitatif.

L'IRM étudie avec une grande précision de nombreux organes tels que le cerveau, la colonne vertébrale, les articulations et les tissus mous. Elle permet d'analyser à distance des organes de manière très précise. Elle est d'une grande utilité lorsqu'une analyse très fine est nécessaire et que certaines lésions ne sont pas visibles sur les radiographies standards, l'échographie ou le scanner. Elle permet de faire des images en coupes dans différents plans et de reconstruire en trois dimensions la structure analysée.

Elle utilise un champ magnétique (aimant) et des ondes radio. Aucune radiation ionisante n'est émise, le champ magnétique est sans danger. Son principe consiste à réaliser des images du

corps humain grâce aux nombreux atomes d'hydrogène qu'il contient. Placés dans un puissant champ magnétique, tous les atomes d'hydrogène s'orientent dans la même direction, ils sont alors excités par des ondes radio durant une très courte période (ils sont mis en résonance). A l'arrêt de cette stimulation, les atomes restituent l'énergie accumulée en produisant un signal qui est enregistré et traité sous forme d'image par un système informatique.

Dans le domaine de la radioactivité, les techniques de scintigraphie nucléaire reposent sur l'utilisation d'un traceur radioactif qui émet des rayonnements détectables par les appareils de mesure. Ces molécules radio-pharmaceutiques sont choisies pour se fixer préférentiellement sur certaines cellules selon le type de diagnostic voulu. Un traitement informatique des données permet ensuite de reconstituer l'origine spatiale de ces rayonnements et de déduire les régions du corps où le traceur s'est concentré. L'image obtenue est le plus souvent une projection mais on peut obtenir une coupe ou une reconstruction 3D de la répartition du traceur.

Tomographie d'émission mono-photonique (TEMP ou SPECT) : elle utilise l'émission de photons gamma par une molécule marquée par un isotope radioactif injecté dans l'organisme.

Tomographie à émission de positon (TEP ou PET) : elle utilise le plus souvent du sucre (un analogue du glucose) marqué par un corps radioactif émettant des positons (e.g., Fluor 18), et permet alors de voir les cellules à fort métabolisme (ex : cellules cancéreuses, infection...).

La TEP permet en général d'obtenir des images de meilleure qualité que la TEMP. Toutefois, le nombre et la disponibilité des radio-pharmaceutiques utilisables en TEMP ainsi que le coût modéré des gamma-caméras compensent ce défaut.

Dans le domaine de l'agriculture, au-delà de l'outil de comparaison visuelle on peut aussi exploiter l'image pour quelle donne des informations plus quantitatives, comme la taille, le nombre ou le poids de certains produits. Ainsi mesurer le diamètre d'un arbre ou d'un cep de vigne est possible par simple photographie. On peut aussi utiliser l'image pour comptabiliser les fruits dans les arbres. Des projets européens ont permis dans les années 1989 le développement du concept de la robotique agricole. Magali le robot cueilleur de pomme et Citrus son homologue pour les orangeries devaient permettre la récolte automatique des fruits sur l'arbre. Le robot était constitué d'une plate-forme automotrice qui se guidait seule en suivant le palissage des arbres. Sur cette plate-forme était montée un ou deux bras robotisés couplés à un système de vision couleur.

Les images en provenance des satellites permettent d'établir des cartes de rendement, d'effectuer des inventaires agricoles, de cartographier les calamités (incendie de forêt et inondation), de surveiller l'état phytosanitaire des cultures et de leurs stress hydriques sur des grandes surfaces notamment d'un département, d'une région, ou d'un pays. L'image satellite est par définition une image prise à très haute altitude, sa précision est assez faible et limite son usage à des grandes surfaces : par exemple les images SPOT couvrent des surfaces allant de 3600 km² à 4800 km² avec une résolution ne dépassant pas 2.5m.

De même nature que l'image satellite, l'imagerie aérienne est beaucoup plus précise, les 50cm voire moins sont souvent accessibles rendant l'outil potentiellement plus utile à l'agriculteur.

Pour l'instant, les applications les plus utilisées sont des applications de contrôle. On peut ainsi calculer avec un certain réalisme les surfaces au sol et vérifier que le champ est en jachère ou que l'arrachage a bien été effectué comme prévu et cela sans qu'on ait à se déplacer sur la parcelle. Les drones sont des objets volants sans pilotes volant généralement à basse altitude. Ils sont équipés de dispositif photographique radiocommandé. L'opérateur au sol est doté d'un retour vidéo pour vérifier qu'il prend bien la zone concernée. Une fois la campagne de photos effectuées il y a un long travail de correction géométrique et d'assemblage des photos. En effet l'avion n'est jamais vraiment parallèle au sol, il faut donc compenser les erreurs dues au roulis et au tangage.

Dans le cadre de nos travaux, le problème abordé dans cet ouvrage est celui du développement d'une nouvelle méthode de segmentation rapide d'images. L'objectif est de fournir à un robot mobile autonome, navigant dans un environnement incertain et dynamique, les informations nécessaires, issues de son système visuel, pour l'évolution dans son environnement sans collisions jusqu'à l'accomplissement de la tâche confiée. Ces informations visuelles seront à la base de la définition de schémas de commande en boucle fermée. L'aspect réactif vis-à-vis d'événements extérieurs perturbant le déroulement de la tâche fait partie de nos préoccupations. Notre objectif est de définir un ensemble de lois de commande référencées multi-capteurs permettant à un robot mobile de réaliser une tâche référencée vision dans un environnement encombré d'obstacles.

Cet ouvrage est organisé en cinq chapitres structurés de la manière suivante. Dans le premier chapitre, nous présentons le domaine d'application vers lequel nos travaux sont orientés, c'est la robotique mobile. Une brève étude sur la robotique mobile est effectuée dans le but de mettre en sailli l'importance que occupe la vision artificielle dans ce domaine. Ensuite, dans le deuxième chapitre, le sens visuel humain est étudié afin d'enlever le voile sur ce précieux phénomène. Cette étude est suivie dans le même chapitre par un bref aperçu sur la vision artificielle. Au cours du troisième chapitre, la segmentation d'image est traitée ainsi que les différentes techniques de la segmentation d'image qui existent sont citées. Une introduction à la théorie des graphes, ainsi qu'une brève étude sur la théorie de l'évidence font l'objet du quatrième chapitre. Le cinquième et dernier chapitre est consacré à l'étude de la méthode de segmentation proposée et aux résultats de la segmentation sur les images aux niveaux de gris et les images en couleurs. Enfin, la conclusion générale présente une synthèse de ce travail et dégage les nombreuses perspectives qui en découlent.

Chapitre I

De l'image segmentée à la navigation d'un robot mobile

1 Introduction

L'image est une source d'information fortement puissante, elle peut être considérée comme la lumière qui éclaire le robot dans son déplacement et dans sa prise de décision. Un robot est un système équipé de capacités de perception, de décision et d'action qui lui permettent d'agir de manière autonome dans son environnement. De manière générale, on regroupe sous l'appellation robots mobiles l'ensemble des robots à base mobile, par opposition notamment aux robots manipulateurs. L'usage veut néanmoins que l'on désigne le plus souvent par ce terme les robots mobiles à roues. Les autres robots mobiles sont en effet le plus souvent désignés par leur type de locomotion, qu'ils soient marcheurs, sous-marins ou aériens.

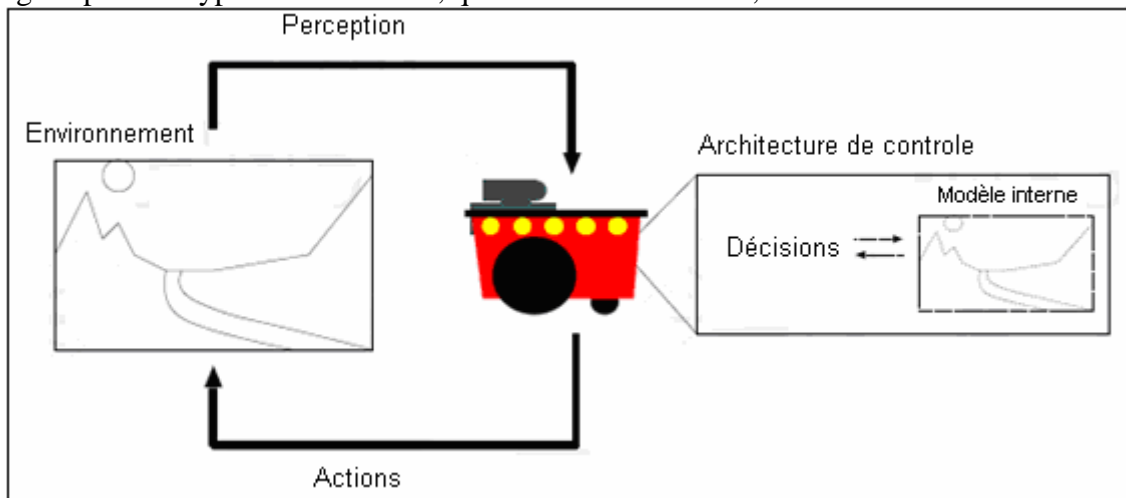


Fig.1.1 : Schéma des interactions d'un robot avec son environnement.

Cette définition s'illustre par un schéma classique des interactions d'un robot avec son environnement (Fig.1.1). Les différentes notions sont essentiellement issues de cette vision de la robotique, donnée par l'Intelligence Artificielle, qui place au centre des préoccupations l'enchaînement de ce cycle perception/Décision/Action. La manière dont un robot gère ces différents éléments est définie par son architecture de contrôle, qui peut éventuellement faire appel à un modèle interne de l'environnement pour lui permettre alors de planifier ses actions à long terme.

2 Domaines d'applications

Aujourd'hui, le marché commercial de la robotique mobile est toujours relativement restreint, mais il existe de nombreuses perspectives de développement qui en feront probablement un domaine important dans le futur. Les applications des robots peuvent se trouver dans de nombreuses activités ennuyeuses, salissantes ou dangereuses, mais également pour des applications ludiques ou de service, comme l'assistance aux personnes âgées ou handicapées.

3 Les architectures de contrôle

Un robot est un système complexe qui doit satisfaire à des exigences variées et parfois contradictoires. Un exemple typique pour un robot mobile est l'arbitrage qui doit être fait entre l'exécution la plus précise possible d'un plan préétabli pour atteindre un but et la prise en compte d'éléments imprévus, tels que les obstacles mobiles. Ces arbitrages, que ce soit au niveau de l'utilisation des capteurs, des effecteurs ou des ressources de calcul, sont réglés par

un ensemble logiciel appelé architecture de contrôle du robot. Cette architecture permet donc d'organiser les relations entre les trois grandes fonctions que sont la perception, la décision et l'action.

Bien que dans cette définition une architecture soit conçue pour un robot précis, ces architectures peuvent néanmoins être classées en trois grandes catégories, les contrôleurs hiérarchiques, les contrôleurs réactifs et les contrôleurs hybrides.

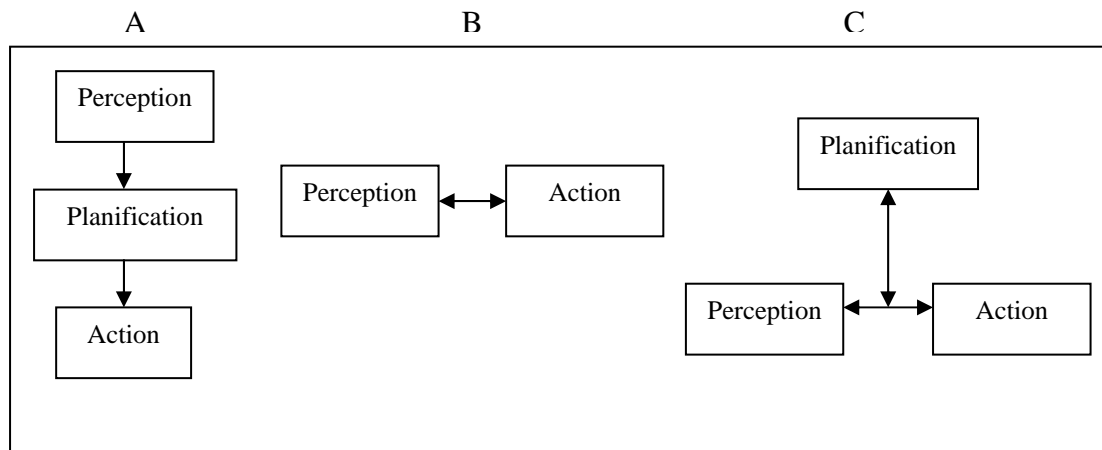


FIG. 1.2: Illustration des architectures classiques des contrôleurs pour les robots mobiles : Hiérarchique (A), Réactive (B) et Hybride (C).

4 Les sources d'information

La navigation repose sur deux types généraux d'informations différentes, les informations proprioceptives et les informations extéroceptives. Les informations proprioceptives sont des informations internes au robot qui le renseignent, dans le cas de la navigation, sur son déplacement dans l'espace. Les informations extéroceptives ou plus simplement les perceptions, sont des informations caractéristiques d'une position que le robot peut acquérir dans son environnement. Ces informations peuvent être de nature très variée. Par exemple, un robot peut mesurer la distance des obstacles avec des capteurs infrarouges ou utiliser une caméra. Les informations proprioceptives renseignent sur le déplacement du robot dans l'espace. Elles constituent donc une source d'information très importante pour la navigation. Cependant, la qualité de cette information se dégrade continuellement au cours du temps, la rendant inutilisable comme seule référence à long terme. Cette dégradation continue provient de l'intégration temporelle des mesures bruitées effectuées par les capteurs internes. Malgré ce défaut important, les informations proprioceptives ont l'avantage de dépendre assez peu des conditions environnementales qui perturbent fortement les informations perceptives. Les informations extéroceptives, ou plus simplement les perceptions, fournissent un lien beaucoup plus fort entre le robot et son environnement. En effet, les informations proprioceptives fournissent des informations sur le déplacement du robot, alors que les informations perceptives fournissent des informations directement sur la position du robot dans l'environnement.

5 Les capteurs

En robotique mobile, les capteurs sont des systèmes de perceptions nécessaires pour les besoins de la navigation du robot dans son environnement. Les capteurs proprioceptifs

permettent une mesure du déplacement du robot. L'odométrie permet d'estimer le déplacement à partir de la mesure de rotation des roues ou du déplacement des pattes. Le problème majeur de cette mesure est que l'estimation du déplacement fournie dépend très fortement de la qualité du contact entre la roue et le sol. Le système radar doppler est un petit radar pointé vers le sol qui permet de mesurer la vitesse du véhicule par effet Doppler. Ce système présente l'avantage d'être beaucoup plus précis que la mesure passant par les roues. Les systèmes inertiels sont potentiellement fiables car ils ne dépendent pas de la nature locale de l'environnement. L'accélération en translation de la plate-forme est mesurée par des accéléromètres linéaires. On dispose en général deux accéléromètres pour prendre des mesures dans deux directions perpendiculaires du plan de déplacement du robot. Un troisième peut être disposé verticalement afin de mesurer la position en trois dimensions. L'accélération angulaire est mesurée par des gyromètres. Les télémètres à ultrason utilisent la mesure du temps de retour d'une onde sonore réfléchie par les obstacles pour estimer la distance.

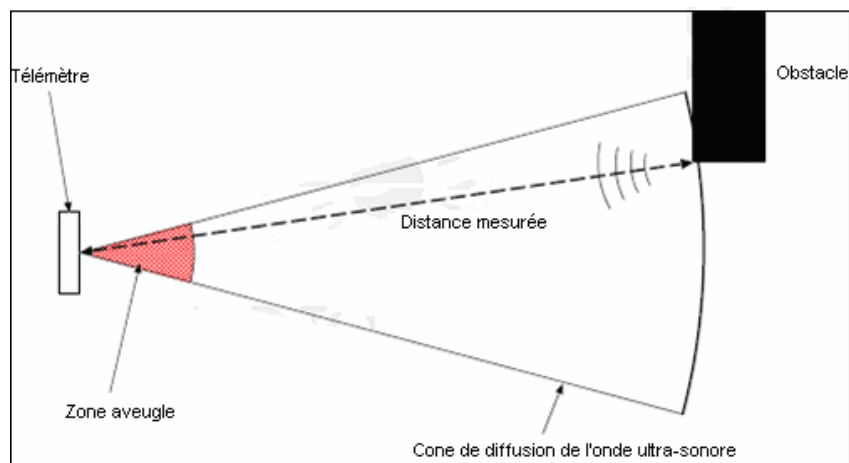


FIG. 1.3: Principe du télémètre à ultrasons et exemple de télémètre réel.

Les télémètres à infrarouge possèdent l'avantage d'avoir un cône de détection beaucoup plus restreint. Ils utilisent une lumière infrarouge pour la détection et peuvent être basés sur différentes techniques qui permettent de recueillir plus ou moins d'information. Il est possible de mesurer simplement le retour ou le non-retour d'une impulsion codée, ce qui permet de détecter la présence ou l'absence d'un obstacle dans une certaine portion de l'espace. Les télémètres laser à balayage sont les plus utilisés à l'heure actuelle pour des applications de cartographie et de localisation. Ils utilisent un faisceau laser mis en rotation afin de balayer un plan, en général horizontal, et qui permet de mesurer la distance des objets qui coupent ce plan.

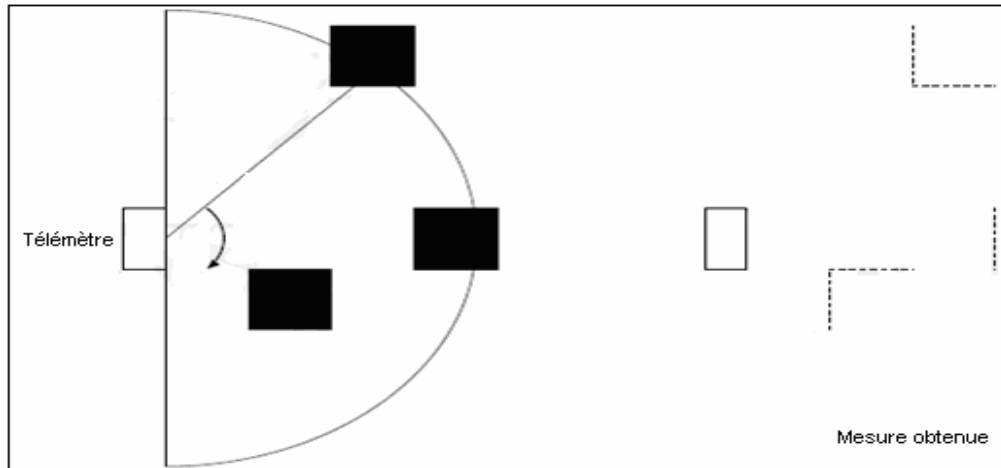


Fig.1.4: Illustration d'un télémètre Laser.

6. Les caméras

L'utilisation d'une caméra pour percevoir l'environnement est une méthode attractive car elle semble proche des méthodes utilisées par les humains. Le traitement des données volumineuses et complexes fournies par ces capteurs reste cependant difficile à l'heure actuelle, même si cela reste une voie de recherche très explorée.

6.1 Caméras simples

Une caméra standard peut être utilisée de différentes manières pour la navigation d'un robot mobile. Elle peut être utilisée pour détecter des amers à partir desquels il sera possible de calculer la position du robot. Si ces amers sont simplement ponctuels, ou de petite taille, il sera en général simplement possible d'estimer leur direction. Elle peut également être utilisée pour détecter des guides de navigation pour le robot, tels que des routes ou des couloirs. Il est également possible d'utiliser globalement une image pour caractériser une position ou un point de vue dans l'environnement. Il faudra alors comparer cette image aux nouvelles images acquises par le robot pour savoir si le robot est revenu à cette position.

6.2 Caméras stéréoscopiques

Lorsque l'on dispose de deux caméras observant la même partie de l'environnement à partir de deux points de vue différents, il est possible d'estimer la distance des objets et d'avoir ainsi une image de profondeur, qui peut être utilisée pour l'évitement d'obstacles ou la cartographie.

6.3 Caméras panoramiques

Les caméras panoramiques (catadioptriques) sont constituées d'une caméra standard pointant vers un miroir de révolution (par exemple un simple cône, ou un profil plus complexe qui peut s'adapter à la résolution exacte que l'on veut obtenir sur le panorama). L'image recueillie permet d'avoir une vision de l'environnement sur 360 degrés autour de la camera. Le secteur angulaire vertical observé dépend de la forme du miroir et peut être adapté aux besoins de chaque application.

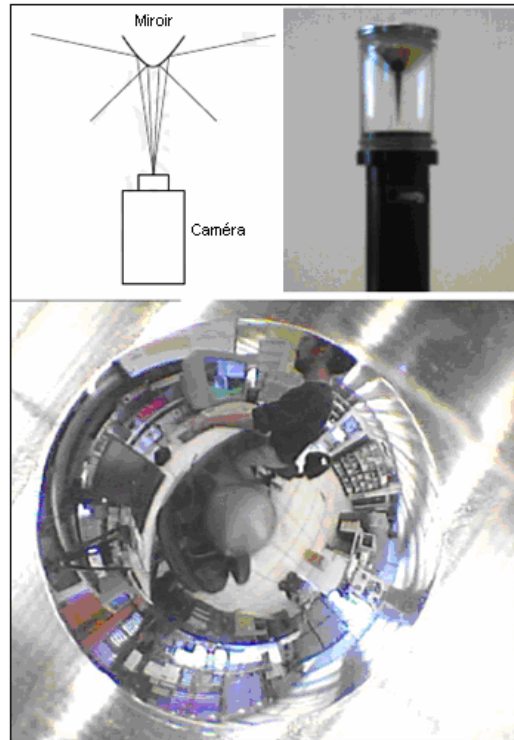


Fig.1.5: Principe des caméras panoramiques catadioptriques et exemple d'image obtenue.

7. Autres capteurs

7.1 Les capteurs tactiles

Les robots peuvent être équipés de capteurs tactiles, qui sont le plus souvent utilisés pour des arrêts d'urgence lorsqu'il rencontre un obstacle qui n'avait pas été détecté par le reste du système de perception. Ces capteurs peuvent être de simples contacteurs répartis sur le pourtour du robot. Ils ne détectent alors le contact qu'au dernier moment. Il est également possible d'utiliser des petites tiges arquées autour du robot pour servir d'intermédiaire à ces contacteurs, ce qui permet une détection un peu plus précoce et donne ainsi plus de marge pour arrêter le robot.

7.2 Les boussoles

Les boussoles permettent, par la mesure du champ magnétique terrestre, de déduire la direction du nord. Ces capteurs peuvent utiliser différentes technologies et ont l'avantage de fournir une direction de référence stable au cours du temps. En pratique, on les utilise principalement en extérieur en apportant le plus grand soin à leur positionnement sur le robot pour éviter les influences des composants du robot, notamment les moteurs électriques.

7.3 Les balises

Dans certaines applications, il est également possible d'utiliser des balises dont on connaît la position, et qui pourront être facilement détectées par le robot, afin de faciliter sa localisation. On peut également utiliser des codes couleurs ou des codes des barres qui pourront être détectés par une caméra.

7.4 Le GPS

Les besoins de localisation étant omniprésents dans de très nombreux secteurs de la vie actuelle, l'idée d'avoir un système de localisation le plus universel possible a donné lieu à l'apparition du "Global Positioning System". C'est un système de balises dont on a placé les balises sur des satellites en orbite terrestre et qui est par conséquent accessible de quasiment partout à la surface du globe. Ce système permet donc d'avoir une mesure de sa position dans un repère global couvrant la terre avec une précision variant de quelques dizaines de mètres à quelques centimètres suivant les équipements. En pratique, il est souvent couplé à un système inertiel qui permet de palier aux pertes du signal GPS et il ne remplace de toute façon pas les capteurs du robot qui lui permettent de percevoir son environnement immédiat, qui constitue la source d'information principale pour la navigation à court terme (par exemple l'évitement d'obstacles, par opposition à la navigation à long terme qui consiste à rejoindre un but distant).

8 Navigation

Les stratégies de navigation permettant à un robot mobile de se déplacer pour rejoindre un but sont extrêmement diverses, de même que les classifications qui peuvent en être faites. L'approche d'un objet permet de se diriger vers un objet visible depuis la position courante du robot. Elle est en général réalisée par une remontée de gradient basée sur la perception de l'objet. C'est une stratégie locale, c'est-à-dire fonctionnelle uniquement dans la zone de l'environnement pour laquelle le but est visible. Le Guidage permet d'atteindre un but qui n'est pas un objet matériel directement visible, mais un point de l'espace caractérisé par la configuration spatiale d'un ensemble d'objets remarquables, ou amers, qui l'entourent ou qui en sont voisins. La stratégie de navigation, souvent une descente de gradient également, consiste alors à se diriger dans la direction qui permet de reproduire cette configuration.

L'action associée à un lieu est la première capacité réalisant une navigation globale, c'est-à-dire qui permette de rejoindre un but depuis des positions pour lesquelles ce but ou les amers qui caractérisent son emplacement sont invisibles. Elle requiert une représentation interne de l'environnement qui consiste à définir des lieux comme des zones de l'espace dans lesquelles les perceptions restent similaires, et à associer une action à effectuer à chacun de ces lieux (Fig.1.6). L'enchaînement des actions associées à chacun des lieux reconnus définit une route qui permet de rejoindre le but. Ces modèles permettent donc une autonomie plus importante mais sont limités à un but fixé. Une route qui permet de rejoindre un but ne pourra en effet pas être utilisée pour rejoindre un but différent. Changer de but entraînera l'apprentissage d'une nouvelle route, indépendante des routes permettant de rejoindre les autres buts.

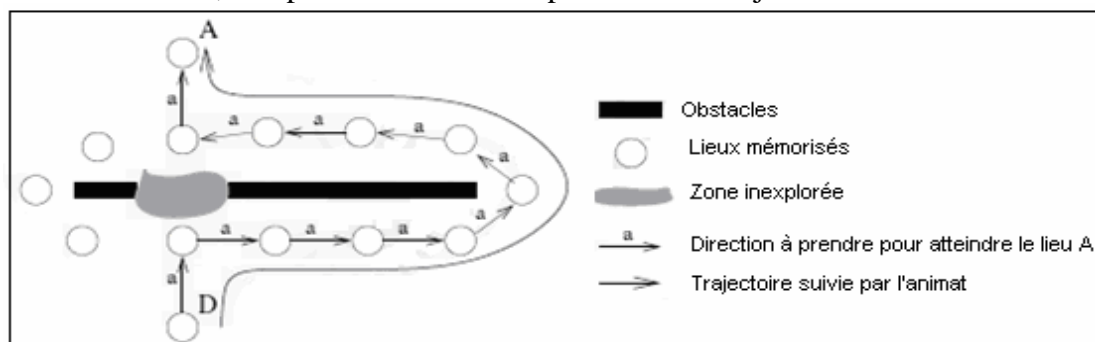


Fig.1.6 : Action associée à un lieu.

En chaque lieu, représenté par un cercle, l'action à accomplir pour rejoindre le but A est représentée par une flèche indiquant la direction à suivre à partir de ce lieu. Cette stratégie permet de rejoindre un but distant dans l'environnement mais repose sur des chemins figés. Dans cet exemple, le chemin joignant le lieu D au lieu A et passant par la droite de l'obstacle a été appris. Rejoindre le lieu A depuis le lieu D ne pourra alors être réalisé que par ce chemin. Le raccourci empruntant le chemin de gauche, par exemple, est inutilisable.

La navigation topologique est une extension de la précédente qui mémorise dans le modèle interne les relations spatiales entre les différents lieux. Ces relations indiquent la possibilité de se déplacer d'un lieu à un autre, mais ne sont plus associées à un but particulier. Ainsi le modèle interne est un graphe qui permet de calculer différents chemins entre deux lieux arbitraires. Ce modèle ne permet toutefois que la planification de déplacements parmi les lieux connus et suivant les chemins connus (Fig.1.7).

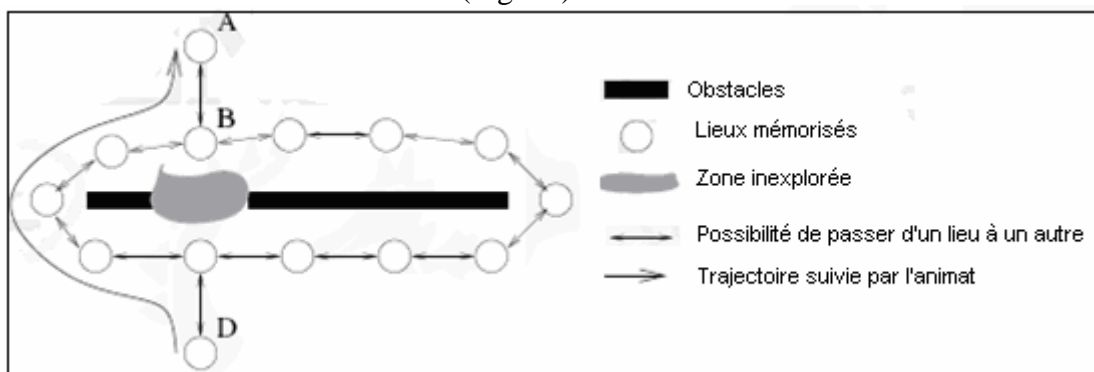


Fig.1.7: Navigation topologique.

Cette stratégie permet de mémoriser un ensemble de lieux et les possibilités de passer de l'un à l'autre, indépendamment de tout but. Pour rejoindre un but, il faut alors une étape de planification qui permet de rechercher, parmi tous les chemins possibles, le chemin rejoignant le but. Dans notre exemple, le chemin le plus court entre D et A peut alors être calculé, mais uniquement parmi les lieux et les chemins déjà connus. Cette stratégie permet, par exemple, de contourner l'obstacle par la gauche mais ne permet pas de le traverser en ligne droite de D à A.

La navigation métrique est une extension de la précédente car elle permet au robot de planifier des chemins au sein de zones inexplorées de son environnement. Elle mémorise pour cela les positions métriques relatives des différents lieux, en plus de la possibilité de passer de l'un à l'autre. Ces positions relatives permettent, par simple composition de vecteurs, de calculer une trajectoire allant d'un lieu à un autre, même si la possibilité de ce déplacement n'a pas été mémorisée sous forme d'un lien.

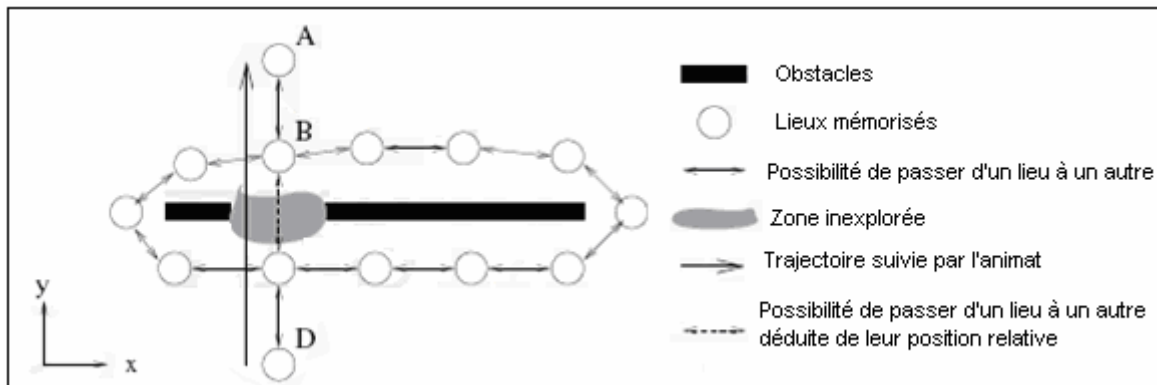


Fig.1.8: Navigation métrique.

9 Évitement d'obstacles

9.1 Méthode des champs de potentiel

Dans la méthode d'évitement d'obstacles par champs de potentiels, on assimile le robot à une particule se déplaçant suivant les lignes de courant d'un potentiel créé en fonction de l'environnement perçu par le robot. Ce potentiel traduit différents objectifs tels que l'évitement d'obstacles ou une direction de déplacement préférée. Le gradient de ce potentiel donne, en chaque point de l'espace, la direction de déplacement du robot (Fig.1.9).

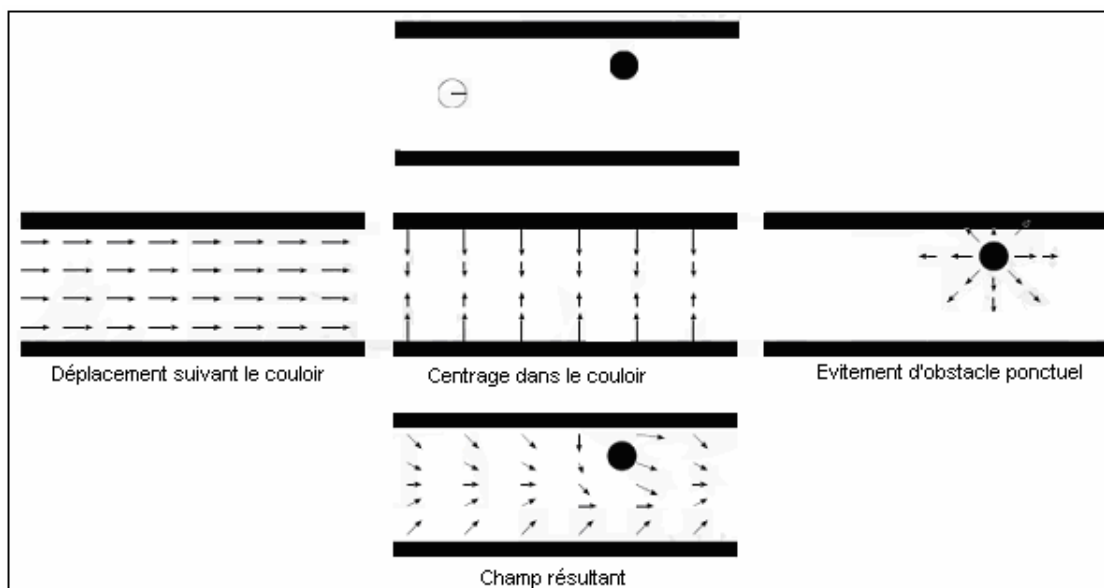


Fig.1.9: Illustration de la combinaison de différents potentiels primitifs.

Dans la pratique, pour l'évitement d'obstacles, le potentiel est en général calculé dans l'espace relatif au robot et ne sert qu'à décider de la vitesse et de la direction courante, en sommant simplement la contribution des différents éléments perçus (Fig.1.10).

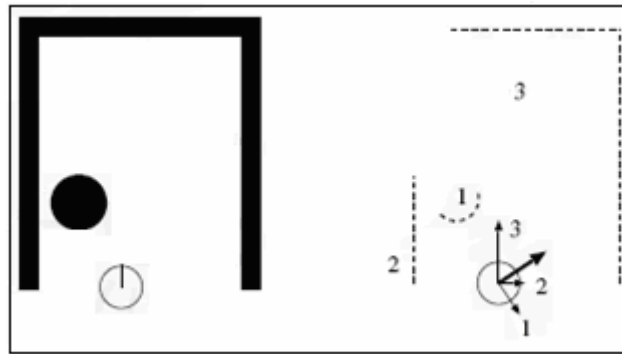


Fig.1.10: Illustration de la combinaison de différents potentiels primitifs dans l'espace relatif au robot.

9.2 Méthode Vector Field Histogram

Une grille d'occupation locale est construite à partir de capteurs à ultrasons de manière très rapide par la méthode "Histogrammic in motion mapping". Chaque cellule de la grille contient un nombre d'autant plus élevé qu'elle a souvent été perçue comme contenant un obstacle (Fig.1.11). Un histogramme représentant l'occupation de l'environnement autour du robot est ensuite construit à partir de cette grille d'occupation locale. Pour cela, l'environnement est discrétisé en secteurs angulaires pour lesquels la somme des valeurs des cellules est calculée (Fig.1.11). Un seuil permettant d'inférer en tolérant un certain bruit est ensuite utilisé pour déterminer les directions possibles pour le robot. Le choix de la direction est finalement réalisé parmi les directions possibles, dont la valeur est inférieure au seuil, en fonction de contraintes externes (par exemple la direction la plus proche de la direction du but).

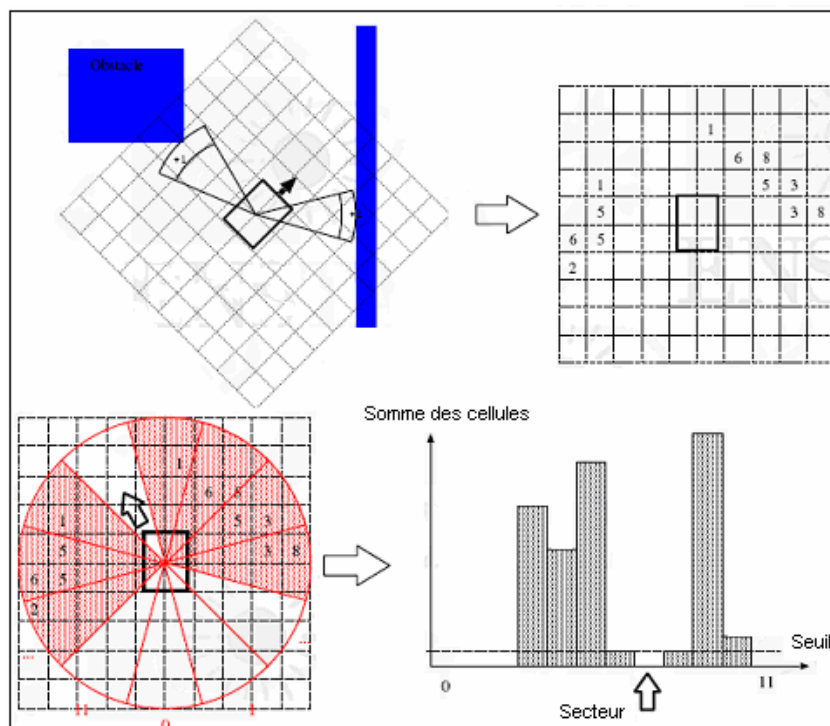


Fig.1.11: Partie supérieure : Grille d'occupation locale. Partie Inférieure : Utilisation de l'histogramme des obstacles pour déterminer la direction de déplacement du robot.

9.3 Méthode de la fenêtre dynamique

La méthode de la fenêtre dynamique permet, à partir de la perception locale de l'environnement, de sélectionner un couple (v, ω) de vitesses de translation et de rotation du robot qui répond à différentes contraintes, dont celle d'éviter les obstacles. Un tel couple de vitesses, lorsqu'il est appliqué au robot, produit une trajectoire circulaire, pour laquelle la satisfaction des différentes contraintes peut être évaluée. A l'issue de l'évaluation de toutes les contraintes pour tous les couples de vitesses possibles, la méthode de la fenêtre dynamique permet de sélectionner le couple le plus pertinent (qui répond le mieux aux contraintes).

La contrainte d'évitement d'obstacles est une contrainte dure (succès / échec) et doit obligatoirement être satisfaite. Elle est évaluée pour chacune des trajectoires possibles à partir de la perception locale de l'environnement à un instant donné et de la position estimée du robot à un pas de temps fixé dans le futur pour la trajectoire courante. Si le robot n'a pas rencontré d'obstacles à cet horizon, la contrainte est respectée, dans le cas contraire, elle ne l'est pas (Fig.1.12).

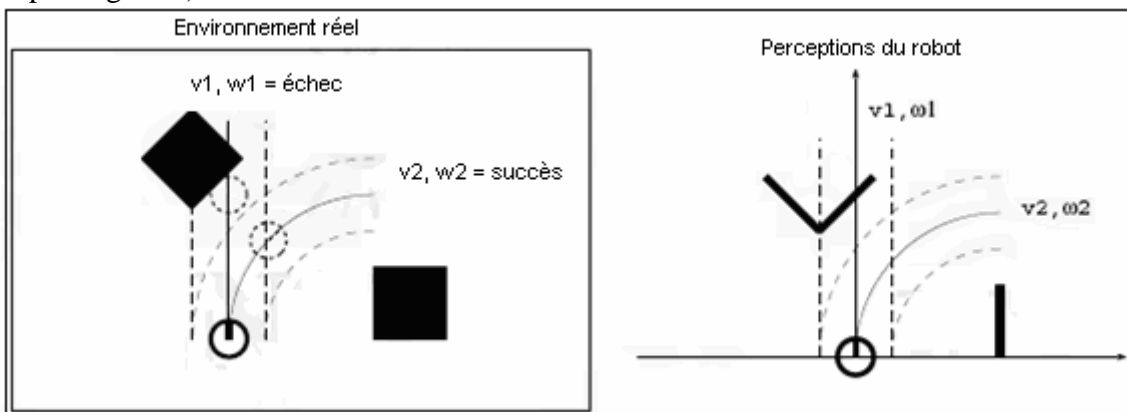


Fig.1.12: Contrainte d'évitement d'obstacles pour la méthode de la fenêtre dynamique.

Le respect ou le non respect de cette contrainte est reporté dans un graphe des vitesses qui indique, pour chaque couple de vitesses possible (donc chaque trajectoire), si le robot va ou ne va pas rencontrer un obstacle (Fig.1.13). Dans ce graphe, il est alors possible de tracer la fenêtre des vitesses accessibles au prochain pas de temps à partir des vitesses courantes du robot et des valeurs d'accélération et décélération maximales. Il reste alors à choisir, au sein de cette fenêtre, un couple de vitesses qui ne conduise pas à percuter un obstacle pour garantir un déplacement sûr du robot. Pour faire le choix parmi toutes les vitesses possibles au sein de cette fenêtre, il est possible d'utiliser des contraintes "souples" supplémentaires pour exprimer des préférences par une fonction de coût $G(v, \omega)$ au sein de cet espace des vitesses accessibles.

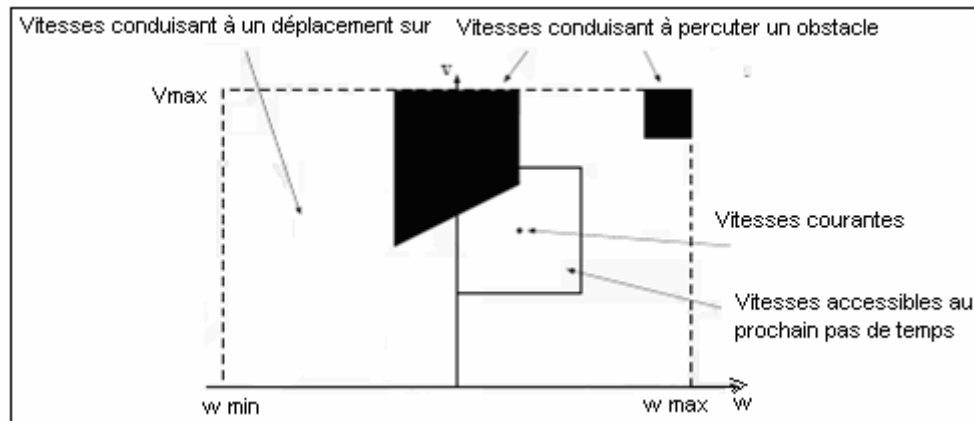


Fig.1.13: Fenêtre de sélection des vitesses tenant compte de la dynamique du robot.

10 Localisation

Pour reprendre pied dans le monde réel, pour décider des prochaines actions à accomplir, nous avons besoin d'informations sur ce qui nous entoure et en particulier, nous ressentons le besoin de savoir où nous nous trouvons. De même, comme nous allons le voir, pour accomplir seul des tâches de complexité variable, pour planifier ses mouvements et ses actes, un robot a besoin de se situer, même de manière locale ou qualitative, dans son environnement.

Il existe plusieurs méthodes de localisation d'un robot dans un environnement [16], parmi elles la localisation globale permet de retrouver la position du robot dans son environnement sans qu'aucune estimation initiale ne soit fournie en fusionnant les informations proprioceptives et les perceptions. Cette capacité est très importante du point de vue de l'autonomie, car elle permet au robot de trouver sa position initiale, dans toutes les conditions, sans intervention extérieure.

10.1 Localisation par les cartes topologiques

Dans le cas des cartes topologiques, estimer la position à partir des seules perceptions est extrêmement simple. En effet, parmi tous les lieux représentés dans la carte, la position du robot est celle d'un des nœuds qui correspond le mieux aux perceptions courantes. La recherche de ce nœud passe donc par la comparaison des perceptions du robot avec les perceptions mémorisées dans chacun des nœuds de la carte. Les nœuds qui sont identiques ou suffisamment similaires sont alors reconnus comme positions possibles du robot. En l'absence de perceptual aliasing, tous les nœuds de la carte correspondent à des situations différentes. Cette étape est alors suffisante pour la localisation complète du robot car le nœud reconnu est unique. Différents systèmes perceptifs ont été utilisés pour implanter de tels modèles.

Certains auteurs utilisent des images panoramiques de l'environnement pour définir les nœuds de la carte [17]. D'autres modèles utilisent les directions ou les distances d'amers ponctuels tous discernables, soit en simulation [9, 18], soit sur des robots réels [6].

10.2 Localisation par les Cartes métriques

Dans le cas des cartes métriques, diverses méthodes d'estimation de la position existent. Lorsque les perceptions sont constituées d'amers ponctuels, une méthode de triangulation peut être utilisée [19, 9, 20]. Cette méthode repose sur la mesure de la direction et de la

distance d'amers ponctuels connus. La perception de trois amers de ce type permet en effet de définir la position du robot de manière unique. Un simple calcul mathématique permet donc de calculer cette position à partir des positions des amers. Ce calcul peut également être approché par des réseaux de neurones, ou par des méthodes heuristiques qui permettent une meilleure résistance au bruit. Lorsque cette méthode est utilisée avec des cartes ne comportant pas de perceptual aliasing, chaque amer est unique et cette méthode permet d'estimer directement de manière non ambiguë la position du robot. En cas de perceptual aliasing, certains amers ne peuvent être distingués et il faut tenir compte de l'estimation précédente de la position afin de pouvoir identifier correctement les différents amers et estimer correctement la position.

Lorsque les objets mémorisés dans la carte ont une certaine étendue spatiale en deux dimensions, il est par contre possible d'utiliser la perception d'un seul objet afin d'estimer directement la position du robot. Les amers utilisés peuvent alors être des objets tridimensionnels détectés par une caméra, les angles des obstacles détectés par un télémètre laser [21] ou un capteur à ultrason [22], des segments détectés en utilisant une caméra ou un télémètre laser [23].

D'autres modèles n'estiment pas directement la position du robot au vu des perceptions, mais reposent sur la comparaison d'une carte métrique locale avec la carte métrique globale (Fig.1.14). La carte métrique locale est construite soit à partir des seules perceptions courantes, soit à partir des données proprioceptives et des perceptions recueillies sur un court laps de temps. Le problème est alors de trouver la portion de carte globale qui correspond le mieux à la carte locale. Cette méthode est très souvent utilisée avec les grilles d'occupation [24, 11], ainsi qu'avec des données brutes de télémètres laser.

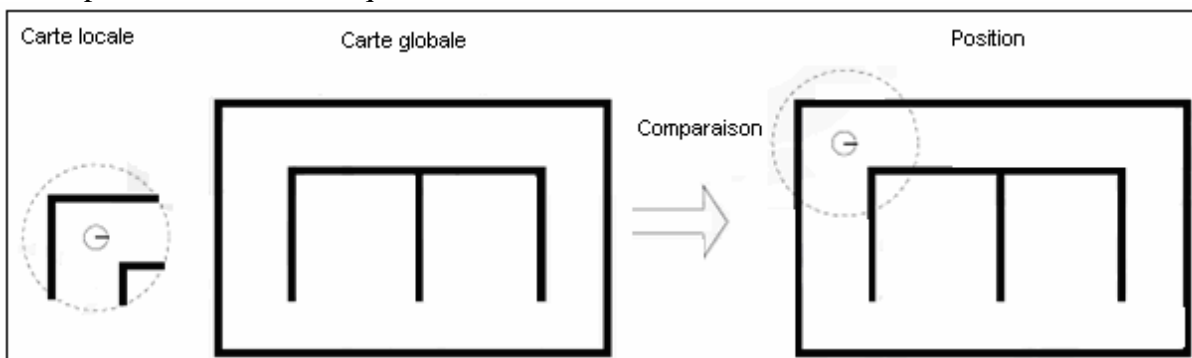


Fig.1.14: Estimation de la position d'un robot par une comparaison de la carte locale avec la carte globale de l'environnement.

11 Planification

Connaissant une carte de l'environnement et la position du robot au sein de cette carte, il est possible de calculer une trajectoire pour rejoindre un but.

11.1 Espace des configurations

Dans le cas 2D pour un robot holonome, l'espace des configurations est simplement l'espace de travail auquel nous enlèverons toutes les positions conduisant à percuter un obstacle, c'est à dire les obstacles eux mêmes, plus une marge de sécurité autour des obstacles correspondant

au rayon du robot. En pratique l'espace des déplacements et l'espace des configurations sont confondus.

11.2 Discrétisation de l'espace de recherche

Les cartes topologiques fournissent directement une discrétisation de l'environnement réel utilisable par les techniques de planification. Dans le cas des cartes métriques, qui représentent l'espace de manière continue, ces techniques ne sont utilisables qu'après discrétisation de l'espace libre représenté dans la carte. Pour ce faire, certains modèles intègrent directement cette décomposition au niveau de la cartographie, en construisant une carte topologique parallèlement à la carte métrique [11]. D'autres modèles font appel à des décompositions de l'espace libre spécifiques à la planification. Cependant, différentes techniques, tels les champs de potentiel [25, 2], permettent de calculer des chemins directement dans le domaine continu, sans phase préalable de discrétisation.

Il existe deux catégories de méthodes pour discrétiser l'espace de recherche des cartes métriques. Les méthodes de la première catégorie font appel à des décompositions en cellules, de différents types, qui permettent de reproduire la topologie de l'espace libre [25, 2] (Fig.1.15). Les cellules obtenues sont alors utilisées de manière similaire aux nœuds des cartes topologiques dans le processus de planification, les cellules adjacentes étant considérées comme reliées par une arête.

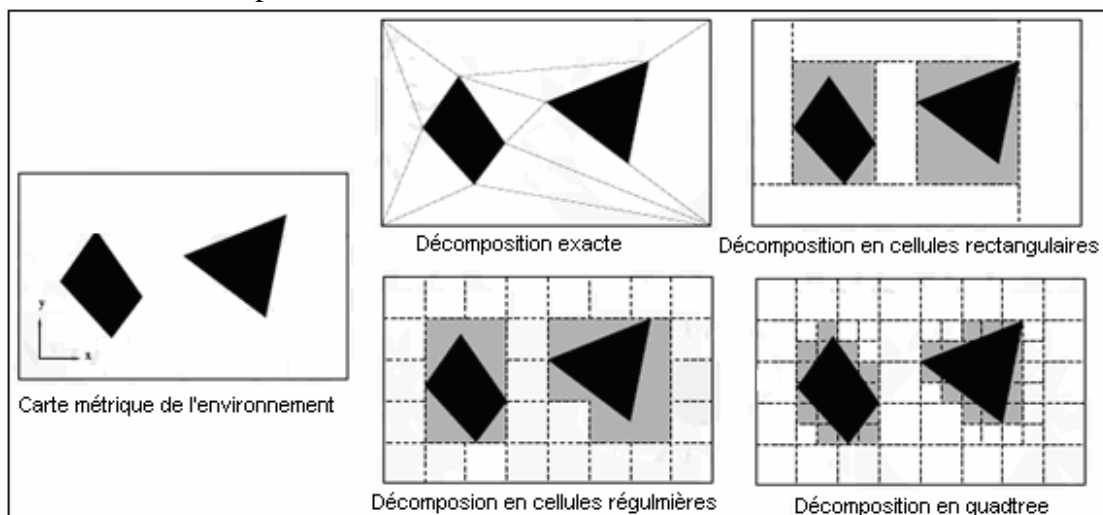


Fig.1.15: Décompositions en cellules de l'espace libre dans les cartes métriques

La décomposition exacte permet de représenter l'ensemble de l'espace libre, à l'aide de cellules de formes irrégulières. La décomposition en cellules rectangulaires ne représente qu'un sous-ensemble de l'espace libre suffisant pour la planification. Ces cellules peuvent être de taille régulière ou non. Enfin une représentation hiérarchique telle que le «quadtree» permet d'utiliser des cellules de taille variable en fonction de la complexité locale de l'environnement.

Les méthodes de la seconde catégorie font appel au pré-calcul de chemins entre des points répartis dans l'environnement [25] (Fig.1.16). Les points seront utilisés comme les nœuds d'une carte topologique, tandis que les chemins pré-calculés reliant les nœuds seront utilisés comme les arêtes de cette carte.

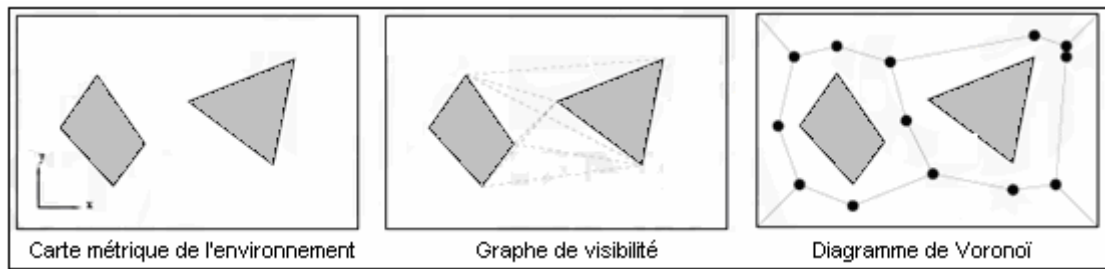


FIG.1.16: Décompositions en chemins pré-calculés dans les cartes métriques.

Ces chemins relient des points particuliers utiles pour la navigation et répartis dans l'environnement. Différents choix de points sont possibles. Le graphe de visibilité utilise les angles d'obstacles qui sont les points que le robot devra contourner pour éviter ces obstacles. Le diagramme de Voronoï utilise les points équidistants de plusieurs obstacles qui permettent de générer des chemins passant le plus loin possible des obstacles.

La planification du chemin entre deux points de l'environnement se réalise alors en deux étapes. La première étape permet de calculer un chemin direct entre, d'une part, le point de départ et le point le plus proche dans l'espace discrétisé et, d'autre part, le point de l'espace discrétisé le plus proche du but et le but en question. La seconde étape permet ensuite de calculer un chemin entre ces deux points de l'espace discrétisé. Ces trois parties de trajectoires sont ensuite assemblées pour obtenir le chemin reliant le point de départ au but. Une phase d'optimisation supplémentaire peut être utilisée pour limiter les effets de la discrétisation et lisser la trajectoire.

11.3 Recherche de chemin

A partir d'une carte métrique discrétisée ou d'une carte topologique, il existe différentes méthodes pour calculer un chemin entre la cellule de départ et la cellule but. Un plan peut être généré qui contient une suite d'actions à effectuer par le robot, ou une suite de points à atteindre afin de rejoindre le but [9, 26]. Un second type de plan, qui associe à chacune des positions possibles du robot au sein de la carte l'action qu'il doit effectuer pour atteindre son but, peut être utilisé. Ce type de plan est appelé politique ou plan universel. L'enchaînement de reconnaissances de positions et de réalisations des actions associées à ces positions permet donc de générer une route joignant le but. Ce type de plan présente l'avantage de permettre au robot d'atteindre le but, aussi longtemps qu'il possède une estimation correcte de sa position.

12 Conclusion

Nous nous sommes penchés, au cours de ce chapitre, sur les concepts et outils permettant à un robot mobile, évoluant dans un environnement dynamique incertain, de se déplacer d'une façon autonome vers un objectif.

Pour arriver au bout de du travail confié, le robot doit généralement exécuter deux tâches, le calcul préalable d'une trajectoire de référence à l'aide de méthodes de planification, puis l'exécution de cette trajectoire à l'aide d'une méthode réactive qui permet de prendre en compte les petits imprévus et de rattraper les erreurs de suivi. Disposer d'une trajectoire de référence ne suffit pas à garantir la sécurité du robot, il doit la suivre correctement. Ce problème devient crucial lorsque l'ensemble des paramètres de contrôle ne sont pas maîtrisés.

L'absence de connaissance complète a priori sur l'environnement rend la trajectoire calculée incertaine sur un temps indéfini, elle doit être recalculée régulièrement. L'aspect dynamique de l'environnement oblige le robot à rester réactif et à adapter son évolution à la dynamique de son environnement.

Equiper un robot mobile d'un système de vision artificiel ouvre une fenêtre sur la simplification de la tâche de la navigation en évitant les difficultés de la navigation d'un système aveugle.

Chapitre II

L'image et la vision artificielle

1 Introduction

Chez tous les êtres vivants, la perception visuelle joue un rôle fondamental. Elle est à la base de la caractéristique essentielle de tout être vivant qui est de réagir dynamiquement à des variations de son environnement après la perception et l'analyse des données sensorielles. Un des buts de la vision artificielle est d'aider à la construction de systèmes robotiques capables d'évoluer et d'agir dans des environnements non contrôlés et d'effectuer ainsi des tâches complexes de manière plus ou moins autonome. L'information visuelle est sans doute la plus riche des différentes sources d'informations disponibles. De ce fait, la conception des systèmes de vision pour l'interprétation automatique de scènes suscite un intérêt sans cesse croissant.

De nombreux chercheurs se sont penchés sur ce problème et ont mis en œuvre plusieurs approches en vue de la conception d'un système de vision complet. L'objectif des premières étapes de la plupart des systèmes de vision est de décrire l'importante quantité d'informations contenues dans l'image en recherchant des indices visuels ou des primitives pertinents permettant de la représenter sous une forme plus condensée et facilement exploitable. La performance des systèmes de vision artificielle est tributaire de la qualité de cette représentation.

2 Le système de vision

L'espace qui nous entoure a une structure tridimensionnelle. Lorsque l'on demande à une personne de décrire ce qu'elle voit, elle n'éprouve aucune difficulté à nommer les objets qui l'entourent (Le téléphone, la table ou le livre ... etc.). Et pourtant l'information qui est réellement disponible sur la rétine de ses yeux n'est ni plus ni moins, une collection de points. En chaque point ou pixel il y a tout simplement une information qui donne une indication quant à la quantité de lumière et la couleur qui proviennent de l'espace environnant et qui ont été projetées à cet endroit de la rétine. Le téléphone, la table ou le livre n'existent pas sur la rétine. Guidé à la fois par l'information codée dans l'image (ou la rétine) et par ses propres connaissances, le processus visuel construit des percepts. Le téléphone ou le livre sont les réponses finales, résultant d'un processus d'interprétation qui fait partie intégrante du système de vision. De plus, il n'y a pas de correspondance terme à terme entre l'information sensorielle (la lumière et la couleur) et la réponse finale (des objets 3D). Le système de vision doit fournir les connaissances nécessaires afin de permettre une interprétation non ambiguë.

La vision a suscité l'intérêt de nombreux scientifiques et philosophes depuis déjà très longtemps. Parmi ceux-ci, les neurobiologistes mènent des recherches théoriques et expérimentales afin d'essayer de comprendre l'anatomie et le fonctionnement du cerveau dans son ensemble. Ils ont découvert une structure très complexe qui est loin de leur avoir révélé tous ses secrets. La tâche des neurobiologistes semble être à la fois grandiose et illusoire. Grandiose, parce que le cerveau est une des plus complexes inventions de la nature. Il reste et restera pour longtemps le bastion encore inconnu que les sciences humaines se proposent de conquérir. Illusoire, car on ne connaît pas ses limites. Ces limites ne sont-elles pas repoussées à chaque découverte ? David Hubel a merveilleusement bien exprimé ce paradoxe : Le cerveau peut-il comprendre le cerveau ? Avec la naissance de machines de calcul de plus en plus sophistiquées, un certain nombre de scientifiques se sont attaqués au problème de la vision d'un point de vue quantitatif afin de construire un modèle de perception visuelle.

3 De la vision humaine à la vision artificielle

Il existe de nombreuses méthodes qui visent à reproduire la vision humaine. Mais à l'heure actuelle, aucun système n'est en mesure de remplacer les niveaux perceptifs et cognitifs dans leur totalité. On ne peut en effet concevoir de vision artificielle globale sans intelligence artificielle de très haute qualité. Il est par contre possible d'obtenir une intelligence efficace pour des domaines bien ciblés (environnement et connaissances limitées). C'est le principe des systèmes experts et cela permet la réalisation de visions artificielles très spécialisées qui sont efficaces. On pourra par exemple citer les systèmes de reconnaissance d'empreintes digitales ou ceux qui peuvent garder votre voiture à bonne distance de celles qui la précèdent.

Une méthode courante de vision artificielle consiste en une succession de traitements : on commence par des traitements numériques sur l'image et l'on en extrait finalement une information sémantique. Mais des recherches considèrent plutôt la vision comme un système complexe où toutes les étapes seraient plus ou moins liées. Chaque traitement serait alors exécuté en fonction du précédent et le résultat serait alors plus efficace. On comprend dans cette méthode que la vision artificielle se doit d'être active (tout comme la vision humaine) et c'est ainsi qu'apparaissent des techniques de prédiction et vérification. La stratégie de perception est alors d'explorer la scène puis de traiter les objets séparément de manière adéquate (tout comme l'homme le fait).

On voit bien que l'élaboration de système de vision artificielle nous amène directement vers la psychologie de la vision humaine. Et c'est donc tout naturellement que des recherches se sont tournées vers ce qu'on appelle l'attention visuelle. On procède là à une analyse sérielle de zones restreintes de l'image. Le système localise ainsi plus rapidement les informations pertinentes.

Loin d'atteindre des résultats satisfaisants dans un environnement quelconque, ces techniques sont efficaces dans des domaines tels que la robotique et elles montrent bien que les recherches sont à mener autant (si ce n'est plus) en psychologie de la vision qu'en algorithmique. Plus le modèle de perception sera proche de l'homme, plus les résultats devraient être encourageants.

Une des voies prometteuses est celle de la vision stéréoscopique qui permet de reconstruire une scène en trois dimensions. On peut alors détacher les objets les uns des autres beaucoup plus efficacement. Mais le problème majeur réside dans la reproduction du processus cérébral qui nous permet d'obtenir une image en relief à partir de deux images en deux dimensions.

4 Les objets et la lumière

D'une manière générale, La lumière est une onde électromagnétique, se propageant dans l'espace et le temps. Les ondes électromagnétiques sont caractérisées par leur longueur d'onde, trajet parcouru par l'onde pendant une période (fig.2.1). Elles correspondent à la variation concomitante du champ électrique noté E et du champ magnétique noté B suivant des plans orthogonaux.

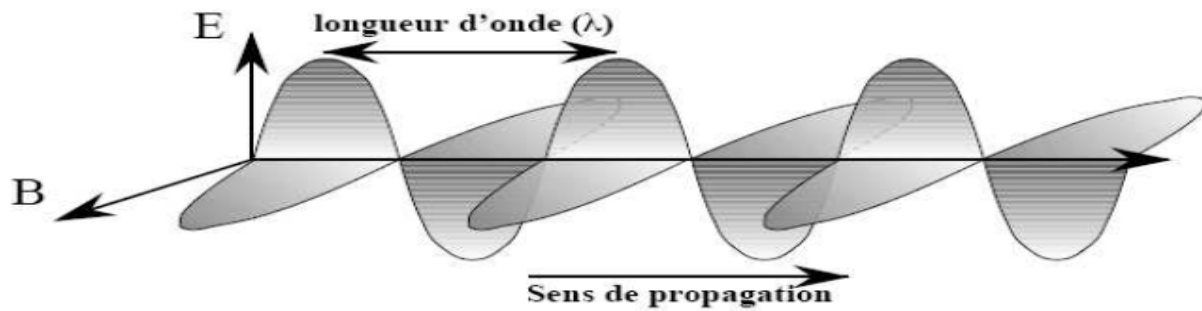


Fig.2.1 : Structure d'une onde électromagnétique.

Dès que la lumière interagit avec la matière, elle n'est plus considérée comme une onde mais comme un flux de particules discrètes dotées d'une énergie cinétique inversement proportionnelle à la longueur d'onde. Ces corpuscules sont des photons.

L'ensemble des ondes électromagnétiques comprend les rayons gamma, X, ultraviolets, infrarouges, les ondes radios, etc, (fig.2.2).

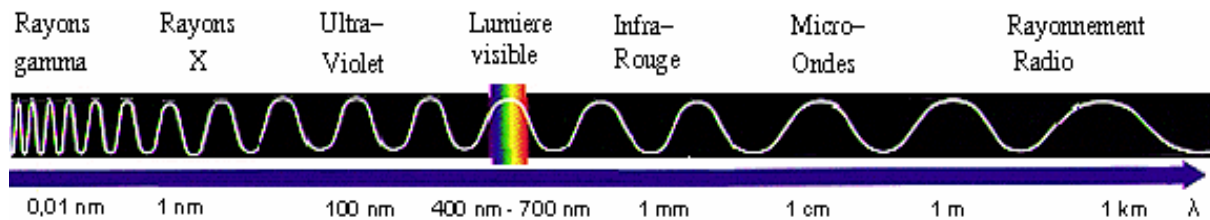


Fig.2.2 : Spectre des ondes électromagnétiques.

Les travaux de Max Planck et Albert Einstein sur la lumière quantique ont montré que l'énergie de la lumière est aussi en quelque sorte granuleuse, ce grain d'énergie est appelé un photon. Chaque photon d'un rayonnement (lumière, rayons X ...) porte une quantité d'énergie caractéristique de sa fréquence. Le photon est une particule élémentaire de la famille des bosons, et explique les échanges d'énergie entre la lumière et la matière. Ces échanges avec la matière expliquent en effet comment on peut voir certains objets. Il existe deux types d'objets :

- Les objets qui produisent de la lumière, comme le Soleil, les flammes, les lampes à incandescence, etc. Ceux-ci produisent souvent de la lumière par incandescence, le mouvement perpétuel d'agitation de la matière émettant des ondes électromagnétiques.
- Les objets qui ne sont visibles que s'ils sont éclairés (invisibles dans l'obscurité). Ils diffusent dans toutes les directions la lumière qu'ils reçoivent, ce qui est appelé l'émission atomique ou moléculaire.

5 La vision des couleurs (chromatopsie)

Le système visuel humain ne peut détecter, dans le spectre de la lumière, que des longueurs d'ondes comprises entre environ 400 et 700 nanomètres. En dessous de ces limites on parle de l'ultraviolet, au dessus de l'infrarouge. Notre système visuel perçoit cet intervalle de fréquences d'ondes lumineuses comme un arc-en-ciel de couleurs variant progressivement. On appelle cet intervalle de fréquences d'ondes lumineuses le spectre visible. L'illustration suivante montre approximativement le spectre visible (Fig.2.3).

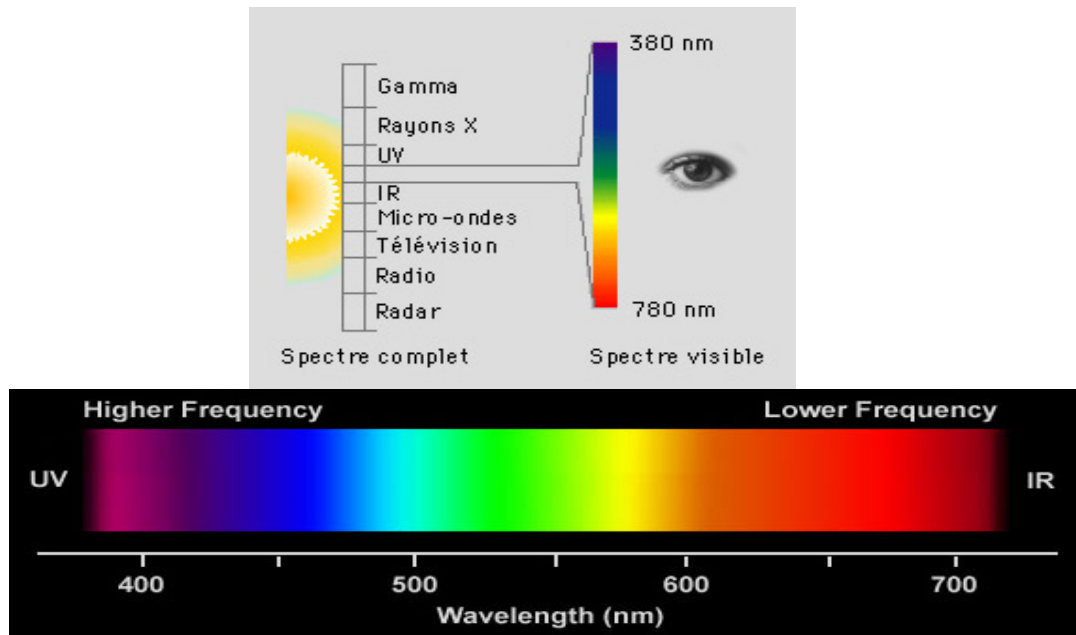


Fig.2.3 : Spectre visible

Une couleur est donc définie par sa longueur d'onde, ou par un mélange de longueurs d'onde. Par exemple, un vert pur est une radiation monochromatique de longueur d'onde 530 nm, tandis que la lumière blanche est un spectre continu contenant toutes les longueurs d'onde du domaine du visible.

6 Physiologie de l'œil

6.1 Généralités

Parmi les cinq sens intervenant dans la perception humaine, la vue est sans aucun doute le sens le plus étudié. On ne compte plus les ouvrages traitant de la vision, et la vue semble être le sens qui permet le mieux de percevoir le monde.

La perception visuelle est un traitement actif de l'information qui fait appel aux capacités de construction de notre cerveau. L'information visuelle est tout d'abord traitée de façon sensorielle, l'œil transmet les stimuli au cerveau via le nerf optique. Ensuite de façon perceptive, c'est à ce stade qu'on perçoit les formes, les couleurs et les mouvements. Et enfin de façon cognitive, on identifie alors les objets en faisant intervenir nos représentations et nos connaissances. Ce sont précisément ces deux derniers stades que l'on essaie de reproduire sur des machines afin de concevoir une vision artificielle.

De forme approximativement sphérique, l'œil est l'organe de base de la vision. Il comporte un ensemble d'éléments destinés à capter l'image de l'environnement vers où il est orienté. L'œil constitue le récepteur des phénomènes lumineux et les focalise pour en projeter l'image sur la rétine. Cette dernière analyse l'image, commence à la coder et envoie le message visuel au cerveau via le nerf optique et les voies optiques intracrâniennes.

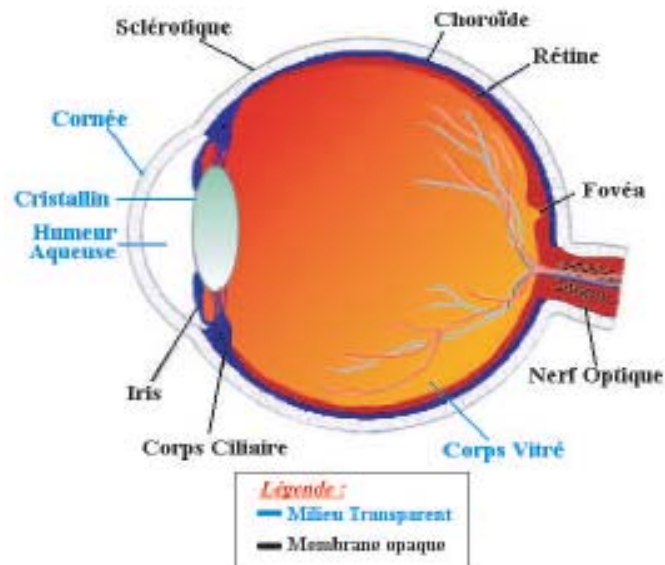


Fig.2.4 : Anatomie de l'œil

6.2 La rétine

Tapissée de photorécepteurs, elle est le siège de réception de l'information lumineuse. La rétine est constituée d'une couche fine de cellules nerveuses dont une grande partie est sensible à la lumière. Les cellules sensibles à la luminosité et au mouvement sont appelées les bâtonnets. Les cellules sensibles à la couleur sont appelées les cônes (Fig.2.5). Les bâtonnets sont responsables de la vision nocturne (vision scotopique) et possèdent un maximum de sensibilité vers 510nm. Leur sensibilité est liée à un colorant, la rhodopsine, qui blanchit à la lumière du jour, expliquant par là leur insensibilité la journée. Les bâtonnets ne fournissent qu'une réponse photométrique et ne permettent donc pas de déterminer les couleurs ("la nuit, tous les chats sont gris"). Les cônes (on parle aussi de cônes S M L) fournissent une réponse photométrique et chromatique, grâce à des pigments dont les maximums d'absorption se situent dans le bleu (450nm, Short : sensible aux longueurs d'onde visibles courtes), le vert (540nm, Medium : sensible aux longueurs d'onde visibles moyennes) ou le rouge (580nm, Large : sensible aux grandes longueurs d'onde visibles). Une lumière, à n'importe quelle longueur d'onde du spectre visible entre 400 et 700 nm, va donc exciter plus ou moins les cellules de ces 3 types. Notre perception de la couleur dépend donc de cette combinaison. C'est là la base de la vision des couleurs et son aspect trichromatique.

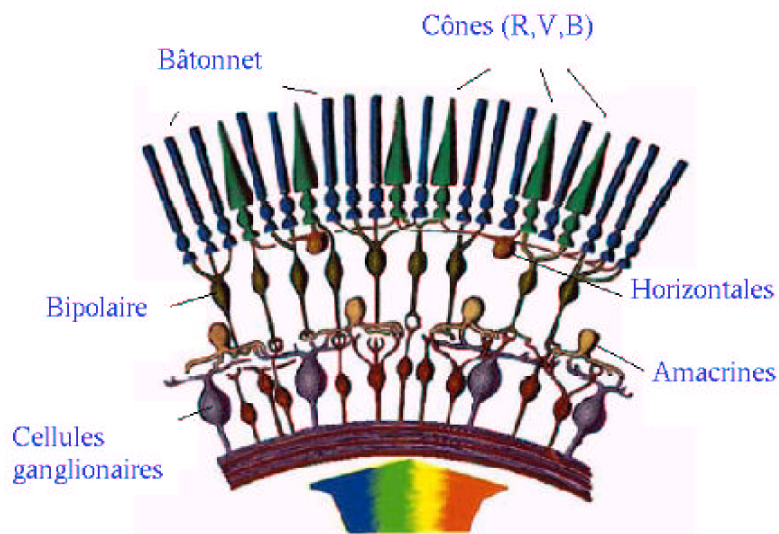


Fig.2.5 : La rétine.

L'information visuelle est reçue par l'œil puis transmise au cerveau après différents traitements préliminaires réalisés par la rétine. Les traitements de plus haut niveau sont réalisés dans le cerveau au niveau des aires visuelles que l'on trouve dans la partie arrière des deux hémisphères cérébraux (Fig.2.6). D'après les études physiologiques, les zones corticales dédiées au traitement de l'information visuelle occupent une place importante. De plus, les connexions avec les autres aires du cerveau font du système visuel un ensemble très complexe.

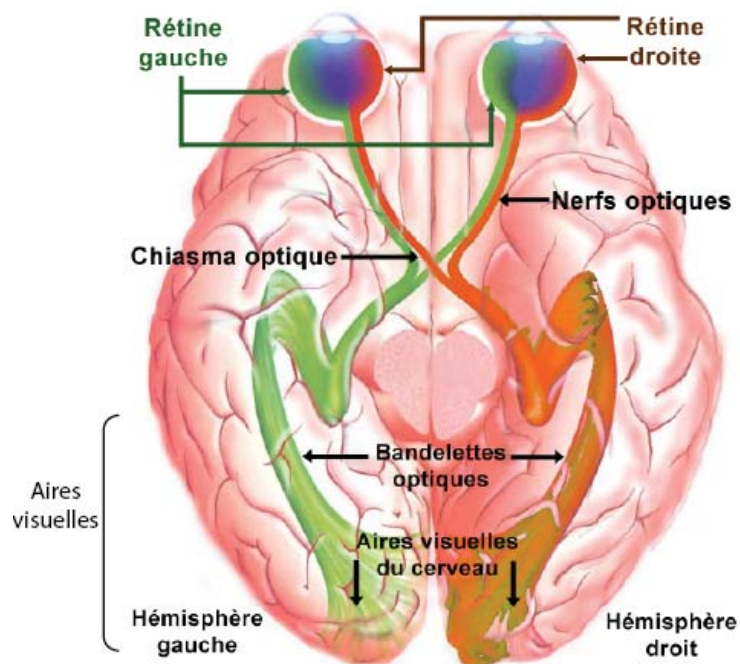


Fig.2.6: Schéma global du système visuel humain.

6.3 Les Photorécepteurs

La lumière traverse les couches neuronales de la rétine et arrive au niveau des photorécepteurs (cônes et bâtonnets). Ces cellules ont pour but d'effectuer la transduction du stimulus lumineux en un potentiel de membrane. Les photorécepteurs présentent des sensibilités spectrales (couleur) et des sensibilités en amplitude (quantité de lumière) différentes (Fig.2.7).

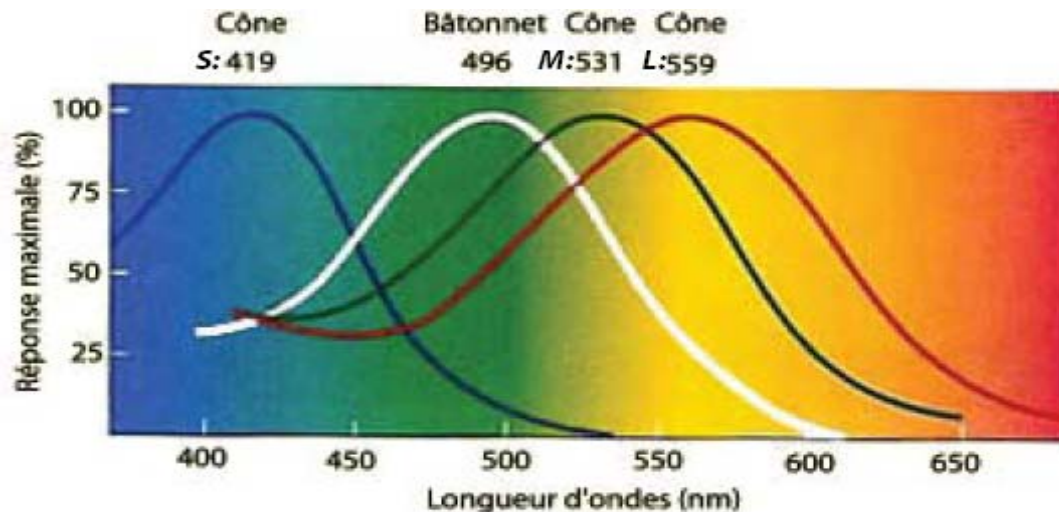


Fig.2.7: Sensibilité des différents types de photorécepteurs à la longueur d'onde de la lumière.

6.4 Répartition des photorécepteurs sur la rétine

La population des bâtonnets est nettement plus importante que celle des cônes. Les bâtonnets seraient au nombre de 120 millions contre 5 millions pour les cônes. La répartition des cônes et des bâtonnets n'est pas uniforme sur la rétine, les bâtonnets étant beaucoup plus nombreux à la périphérie et les cônes plus nombreux près de l'axe optique (à faible excentricité). Cette variation de la concentration des photorécepteurs est représentée sur la (Fig.2.8). La zone centrale (faible excentricité) ne contient que des cônes, elle constitue la fovéa. On y trouve un pas d'échantillonnage des photorécepteurs très faible d'où une vision de haute précision de jour et une vision de très faible sensibilité dans l'obscurité du fait de l'absence de bâtonnets. La zone de moyenne excentricité (parafovéa) comporte une faible proportion de cônes et une population de bâtonnets croissante. Dans cette zone la vision de jour est moins précise, mais la vision en obscurité est performante. Enfin, pour les fortes excentricités, les cônes sont presque absents et laissent la place aux bâtonnets. On remarque aussi le disque optique. C'est une zone de la rétine dépourvue de capteur, car c'est à cet endroit que les axones des neurones qui constituent le nerf optique se rejoignent pour partir en direction du cerveau. Cette zone est aussi appelée tâche aveugle.

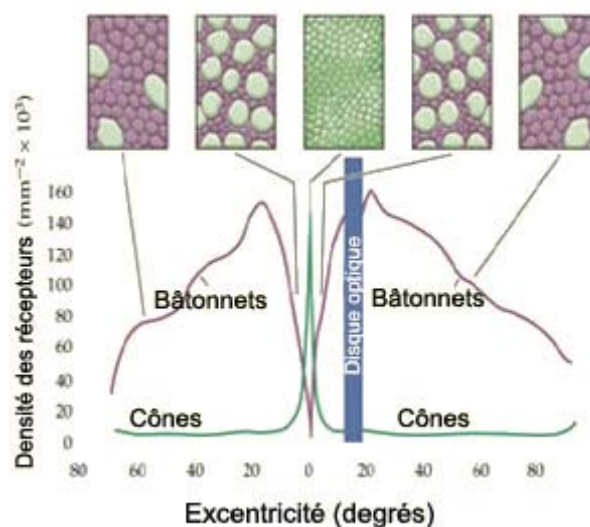


Fig.2.8: Concentration des photorécepteurs en fonction de l'excentricité.

7 La vision artificielle

La vision est le processus de découvrir avec des images ce qui est présent dans le monde réel, et où il est localisé. Il s'agit d'extraire ou de découvrir des structures sémantiques à partir d'une série d'images qui ont des correspondances avec les structures réelles du monde. L'étude de la vision artificielle a pour but de reproduire le sens de la vue. Dans le processus de vision artificielle existent généralement cinq étapes Fig.2.9. Chaque étape est chargée d'effectuer une tâche très spécifique.

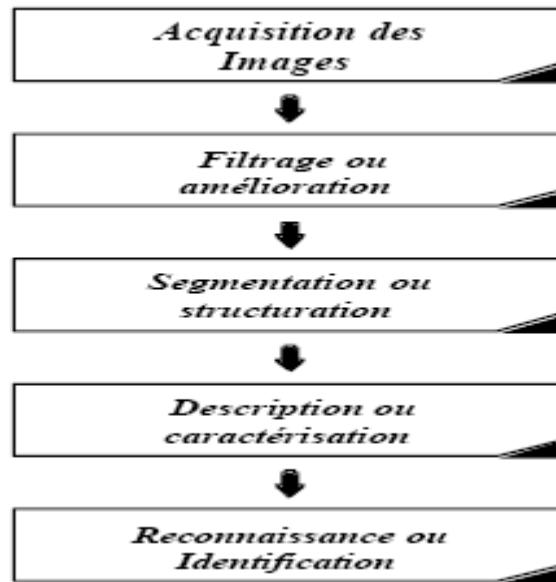


Fig.2.9. Processus de Vision Artificielle

La vision artificielle ou vision par ordinateur est une discipline qui consiste à convertir une image en données objets ou, plus explicitement, à identifier les objets contenus dans l'image par l'extraction et l'analyse de caractéristiques abstraites à partir des pixels, suivant un processus de reconnaissance de forme similaire à celui opéré par l'humain. Le But de la vision artificielle est de doter les machines d'un système visuel leur permettant, à partir d'images d'effectuer des tâches complexes comme, par exemple, la photo-interprétation, la télésurveillance, la navigation ou la manipulation d'objets par des robots. Parmi les autres applications dans ce domaine nous pouvons citer l'analyse d'images aériennes ou satellitaires (surveillance des cultures ou des forêts, prévisions météorologiques, ...) ; l'identification d'un individu par biométrie (par exemple l'analyse automatisée des empreintes digitales); l'analyse d'images médicales ; l'analyse de documents ; le contrôle de qualité ; ... etc.

Un système de vision peut être décrit comme un ensemble de processus fortement liés opérant des transformations progressives des informations sensorielles. Ces transformations assurent le passage de l'image physique saisie à une description structurelle puis sémantique, en passant par la segmentation et l'identification d'attributs caractéristiques de la scène et des objets qui la composent. D'une manière générale, deux éléments dans la définition d'un système de vision peuvent être retenus [28] :

- L'espace des représentations des données utilisées par le processus d'interprétation d'image.
- Les connaissances qui définissent le sens d'une interprétation. Elles représentent l'ensemble des relations existant au sein de l'espace des représentations.

8 L'image et son traitement

8.1 Préambule

La vision artificielle par ordinateur est la science qui développe les bases algorithmiques et théoriques par lesquelles l'information utile relative à l'environnement peut être automatiquement extraite et analysée à partir d'une image, d'un ensemble d'images ou d'une séquence d'images. Une telle information peut référer à la reconnaissance d'un objet générique, à la description tridimensionnelle d'un objet inconnu, à la position et l'orientation d'un objet observé, ou à la mesure de toute propriété spatiale d'un objet telle que la distance entre deux de ces points distincts ou le diamètre d'une section circulaire.

L'acquisition d'images est réalisée par l'intermédiaire d'une caméra vidéo dont le signal est numérisé. Une caméra vidéo comprend un système optique composé de lentilles qui forme une projection bidimensionnelle d'une scène de l'environnement observée sur un plan image photosensible (ou rétine). Chaque caméra délivre, à cadence vidéo, une copie de l'image optique reçue par sa rétine électronique. La rétine d'une caméra est divisée, en abscisse et ordonnée, en un certain nombre de points appelés pixels. Au cours du traitement de l'image, l'usage de l'informatique implique cette division en un nombre défini mais programmable de pixels.

8.2 L'image

Une image est une représentation spatiale d'un objet, d'une scène bidimensionnelle ou tridimensionnelle, ou d'une image elle-même. Elle peut être réelle ou virtuelle telle qu'en optique expérimentale. En vision par ordinateur, le terme image réfère usuellement à une image enregistrée, telle que l'image vidéo, l'image digitale ou la photographie.

Une image optique monochrome est abstraitement assimilable à une fonction continue I de deux variables réelles u et v définie sur une région rectangulaire bornée d'un plan. La valeur de la fonction image I au point de coordonnées (u,v) du plan est dénotée $I(u,v)$. Pour les capteurs optiques ou photographiques, la valeur $I(u,v)$ est typiquement proportionnelle à l'énergie lumineuse radiante reçue dans la gamme de fréquences électromagnétiques à laquelle le capteur d'image est sensible, au voisinage du point de coordonnées (u,v) . Dans ce cas, l'image est appelée image d'intensités ou de luminances.

La radiance dénote la quantité de lumière émise par une source en terme de puissance par unité de surface comprise dans un angle solide unité. La fonction d'intensités de l'image I s'apparente à la radiance de la scène et dépend de la quantité de lumière incidente, de la fraction de la lumière incidente réfléchie ainsi que de la géométrie de la réflexion de la lumière, c'est-à-dire des directions d'illumination et d'observation. Ainsi, différents points d'objets de l'environnement situés devant le système optique d'acquisition d'images présentent différentes valeurs d'intensités sur l'image, dépendant de la radiance incidente ainsi que des manières dont ils sont illuminés, dont ils réfléchissent la lumière, dont la lumière réfléchie est collectée par le système optique et enfin dont le capteur d'images répond à l'excitation lumineuse [31].

Une image digitale ou numérique est représentée par une matrice de valeurs numériques correspondant à des valeurs d'intensités quantifiées discrètes. Dans ce cas, I forme une matrice et $I(u,v)$ représente la valeur d'intensité à la position de l'image matérialisée par la ligne u et la colonne v de la matrice. L'unité de l'image digitale ou

numérique est le pixel qui associe une position (u,v) dans le plan image et une valeur d'intensité correspondante $I(u,v)$ [32].

Dans le contexte des images monochromes noir et blanc, les valeurs d'intensités associées aux pixels sont appelées niveaux de gris. Le niveau de gris d'un pixel représente une mesure de l'accumulation de la lumière collectée dans la région du plan occupée par ce pixel. Dans le contexte des images trichromes couleur, la représentation fonctionnelle de l'image fait intervenir trois fonctions d'intensités $I_R(u,v)$, $I_V(u,v)$, $I_B(u,v)$ relatives aux trois composantes primaires rouge, verte et bleue constitutives de la couleur [33].

8.3 Les capteurs d'images

Une image lumineuse est la représentation d'un objet, d'une scène ou d'une image elle-même obtenue par des procédés optiques. L'image est caractérisée par la distribution spatiale de ses éclaircissements et c'est précisément à cette distribution qu'est liée l'information spécifique qu'elle porte. Un capteur d'images est formé d'une association ordonnée d'un grand nombre de capteurs optiques sur l'ensemble desquels se trouve projetée l'image. Les capteurs d'images délivrent des signaux électriques dans lesquels se trouve transposée cette information recherchée dans l'image lumineuse originelle. Les signaux électriques sont ultérieurement susceptibles de traitements électroniques, qui permettent en particulier la transmission et la reconstitution de l'image sur un écran cathodique mais également leur numérisation qui conduit à une représentation digitale des images.

9 Traitement d'image

Le traitement d'images vise à mettre en exergue certaines particularités de l'image pour en faciliter l'analyse et l'interprétation. Le schéma de la figure 2.10 résume la chaîne de traitements effectués sur l'image.

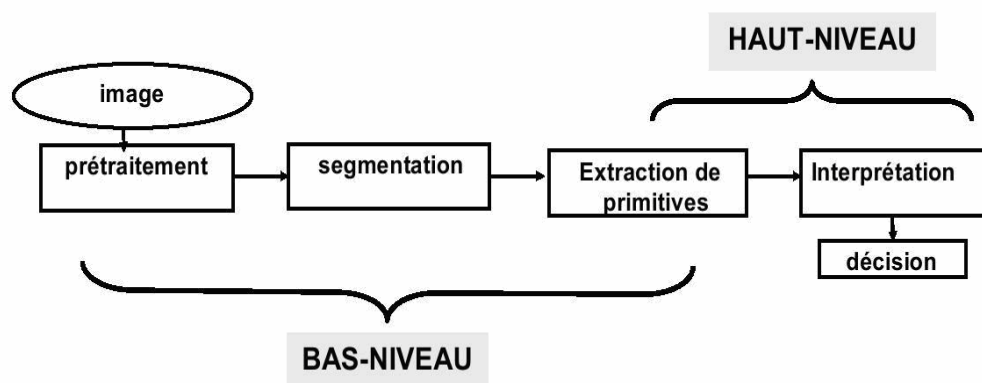


Fig.2.10 : Schéma de la chaîne de traitement d'image

Les traitements de bas niveau nécessitent très peu d'informations sur le contenu des Images, il s'agit ici de processus de prétraitement (des filtrages de lissage, d'amélioration, et de restauration d'images), segmentation,...etc. Les traitements de hauts niveaux se situent en aval du bas niveau, et peuvent nécessiter des informations sur le contenu de l'image. Il s'agit de la reconnaissance des formes,...

9.1 Le Prétraitement d'image

Le premier objectif du traitement est l'élimination des informations non pertinentes pour faciliter l'extraction des informations utiles à l'analyse. Toute acquisition est bruitée, le bruit dépendant de la technologie (capteur, échantillonnage); mais aussi de la scène, du contexte (plein soleil, ombres), ou de la nature des objets eux-mêmes (spécularités, transparence). Selon sa nature, le bruit induit un effet qui peut être localisé sur un pixel ou sur un voisinage de pixels spatial ou temporel, par exemple le bruit dit «poivre-et-sel» très localisé, le bruit de speckle, provoquant un chatolement, bruit granulaire de type multiplicatif ou encore le bruit additif gaussien, introduisant un flou (Fig.2.11).

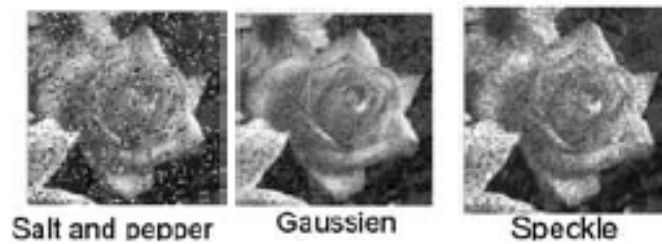


Fig.2.11 : Différents types de bruit

Les traitements préliminaires visent à améliorer la qualité d'une image ; cependant cette notion est très subjective et souvent assujettie à un objectif. Or cet objectif peut différer entre un système automatisé et un humain. Dans tous les cas, pour les images couleur deux approches sont possibles : soit un traitement marginal, qui est le traitement séparé et indépendant sur chacun des axes de représentation, de chacune des trois composantes ; soit directement un traitement vectoriel sur les pixels auxquels sont associés non plus un scalaire mais un vecteur de dimension trois. Une quantification sur une palette de couleurs réduite peut permettre d'appliquer des traitements de type niveaux de gris. Des changements d'espace de représentation sont toujours bien sûr envisageables avant d'opérer les traitements.

9.2 Segmentation

Le cœur d'un système d'analyse automatique d'images est l'étape de la segmentation ou d'extraction des caractéristiques. La segmentation des images consiste à regrouper les pixels de ces images qui partagent une même propriété pour former des régions connexes. Une image est donc une mosaïque de régions homogènes séparées par des zones de transitions.

9.3 Analyse d'images

L'objectif de l'analyse est de permettre l'exploitation du contenu de l'image à des fins d'interprétation (aide au diagnostic en imagerie aérienne, satellitaire, médicale par exemple), de localisation et de reconnaissance (vidéosurveillance, contrôle robotique) ou de mesure des évolutions (contrôle qualité, suivi thérapeutique, etc.).

L'espace qui nous entoure présente une structure tridimensionnelle. Lorsqu'on demande à une personne de décrire ce qu'elle voit, celle-ci n'éprouve aucune difficulté à nommer les objets qui l'entourent. Et pourtant, l'information disponible sur la rétine des yeux (ou image) n'est ni plus ni moins qu'une collection de points (environ 1 million). En chacun de ces points (ou pixels) se trouve codée une information qui fournit une indication quant à la quantité de lumière, provenant de l'espace environnant, projetée localement sur la rétine.

Il n'existe pas de correspondance terme à terme entre l'information sensorielle délivrée par les images (lumière et couleur) et le résultat du processus d'interprétation (reconnaissance des objets). Le système de vision doit fournir les connaissances nécessaires afin de permettre une interprétation sans ambiguïté [35].

Les attributs caractéristiques d'un objet observé présentent une organisation spatiale à la fois sur l'image et sur cet objet dans l'environnement tridimensionnel. Aussi, la reconnaissance d'un objet particulier est réalisée en établissant la correspondance entre la configuration spatiale de ses attributs caractéristiques observée dans les images, et leur configuration spatiale correspondante attendue dans l'environnement 3D. Comme le système de vision (tout comme l'homme) ne peut reconnaître un objet que s'il l'a déjà vu (s'il le connaît), celui-ci lui est appris en le lui montrant préalablement. Cette phase "d'apprentissage" en quelque sorte est destinée à définir ou désigner l'objet par des attributs de référence qui lui sont propres [32].

La nature de l'information qui peut être extraite est diverse, des caractéristiques couleur ou luminance, des caractéristiques d'aspect de surface (rugosité, texture, brillance), des caractéristiques de forme ou de construction (courbe, segments, courbure), des caractéristiques d'organisation (voisinage, inclusion, hiérarchie).

À partir de ces informations, l'analyse permet de mesurer des ressemblances ou des dissemblances entre deux images ou parties d'images. Dans le cadre générique de la reconnaissance des formes (RDF), l'enjeu est d'identifier un objet.

10 Conclusion

La vue est l'un des sens privilégié de l'être humain. Pour passer de la sensation, stimulation des organes des sens, à la perception, impression obtenue à partir des sens et appréhendé par l'esprit, de nombreux traitements sont effectués dans notre cerveau. La perception fait ainsi suite à la sensation, en incluant des traitements des stimuli perçus consciemment ou inconsciemment et en prenant en considération des informations contextuelles et des expériences passées, stockées dans une bibliothèque de taille impressionnante. Cette bibliothèque inclut des images, des sons, des odeurs. De plus, les interactions entre les différentes aires cérébrales motrices, sensitive, visuelle et auditive sont encore mal connues.

Automatiser le traitement et l'analyse de scènes est une mission particulièrement difficile. Si de nombreuses applications permettent d'automatiser des tâches liées à la perception humaine, comme le contrôle qualité sur des objets manufacturés ou la conduite autonome de véhicules, ceci ne peut être réalisé que dans des cas bien particuliers, avec des contraintes très fortes quant à l'environnement et la nature des objets eux-mêmes.

Chapitre III

Segmentation d'image état de l'art

1 Introduction

Le processus d'analyse d'images peut être défini comme l'ensemble de méthodes et d'outils permettant de décrire quantitativement le contenu d'une image. Il est généralement décomposé en plusieurs étapes : acquisition, numérisation, prétraitement, segmentation, interprétation.

L'étape de segmentation d'image est une étape très importante dans cette chaîne d'analyse car c'est à partir de l'image segmentée, que les mesures sont effectuées pour l'extraction des paramètres discriminants en vue de la classification ou de l'interprétation. La tâche de segmentation peut se résumer de la manière suivante : Etant donnée une image, l'objectif de la segmentation est d'établir une description compacte et représentative de son contenu informationnel, plus exploitable que l'ensemble de ses points. Il s'agit de procéder à l'extraction d'indices visuels (primitives) pertinents, suffisamment corrélés avec les entités qui composent la scène d'où l'image est prise.

La diversité des travaux menés dans le but de la compréhension de la vision humaine montrent la complexité de la notion d'indice visuel et qu'il est bien difficile d'en donner une définition précise. Ceci étant, un indice visuel peut être défini comme une information perceptible directement à partir de la visualisation de l'image (contour, régions homogènes au sens des niveaux de gris et au sens de la texture, formes élémentaires). Ces indices visuels ou primitives correspondent à des phénomènes photométriques et/ou géométriques remarquables.

Dans ce Chapitre, nous allons nous efforcer d'esquisser un état de l'art de la segmentation d'images. Cette partie des différentes techniques et approches n'a pas pour but d'être exhaustive mais de montrer la grande diversité qui existe au sein de ce vaste domaine.

2 Définition de la segmentation

La segmentation est une tâche très importante et difficile, car elle dépend de la nature de l'image et de l'objectif de l'analyse. Plusieurs techniques ont été développées [38, 39, 40], mais la plupart d'entre elles ont besoin de plusieurs paramètres dont le réglage exige souvent l'expertise humaine.

La segmentation est un domaine de l'analyse d'image qui consiste à diviser une image en différentes régions homogènes connectées selon certaines caractéristiques probabilistes Figure 3.1 et Figure 3.2.

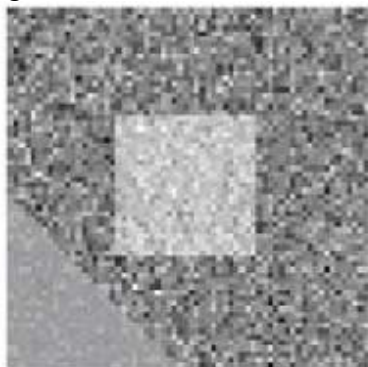


Fig.3.1 Image originale.



Fig.3.2 Image segmentée.

Il est difficile de trouver une seule définition de la segmentation car cette tâche est souvent confondue avec la classification ou l'étiquetage, on peut néanmoins essayer d'en trouver une afin de fixer le cadre de ce domaine et ainsi faire disparaître les ambiguïtés.

Segmenter une image signifie trouver ses régions homogènes et ses contours. Ces régions et contours sont supposés être pertinents, c'est-à-dire que les régions doivent correspondre aux parties significatives des objets du monde réel, et les contours leurs contours apparents. Une définition formelle d'un algorithme de segmentation a été donné par Horowitz et Pavlidis [41] en 1975.

Soit X le domaine de l'image et f la fonction qui associe à chaque pixel une valeur $f(x, y)$. Si nous définissons un prédicat P sur l'ensemble des parties de X , la segmentation de X est définie comme une partition de X en n sous-ensemble $\{R_1, \dots, R_n\}$ tels que :

1. $X = \bigcup_{i=1}^n R_i$
2. $\forall i \in \{1, \dots, n\} R_i$ est connexe
3. $\forall i \in \{1, \dots, n\} P(R_i) = \text{vraie}$
4. $\forall i, j \in \{1, \dots, n\}^2 R_i$ est adjacent à R_j et $i \neq j \Rightarrow P(R_i \cup R_j) = \text{faux}$

Le prédicat P est utilisé pour tester l'homogénéité des ensembles R_i . Ces sous-ensembles constituent les régions de l'image. Une segmentation de l'image est donc sa décomposition en un ensemble de régions homogènes, le critère d'homogénéité P restant à déterminer.

Zucker a résumé les conditions de la définition d'Horowitz comme suit : la première condition implique que tout pixel de l'image appartienne une région et une seule. Cela signifie que l'algorithme de segmentation ne doit pas se terminer avant d'avoir traité tous les points. La seconde condition implique que toute région doit être connexe. La connexité des régions étant induite par le voisinage défini sur l'image. La troisième condition implique que chaque région doit être homogène.

Enfin, la quatrième condition est une condition de maximalité indiquant que la fusion de deux régions ne doit pas être homogène. Il est important de remarquer que le nombre n de régions formant la partition de l'image reste indéterminé. Il peut donc exister plusieurs segmentations possibles pour un prédicat P donné.

Les psychophysiciens et les gestaltistes sont d'accord sur le fait qu'un processus de segmentation se produit aux toutes premières étapes du processus de perception visuelle. De plus, ils ont prouvé que ces premières étapes sont tout à fait indépendantes d'un quelconque apprentissage ou d'une connaissance a priori sur le monde (le domaine de l'image). Cela signifie qu'un algorithme programmé, traitant des images numériques, est capable de faire la même chose.

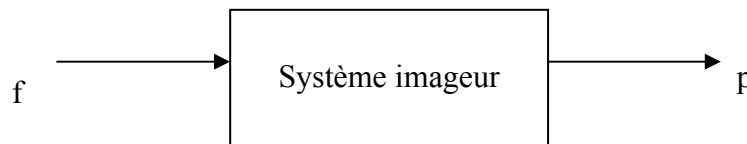


Fig.3.3 : Schéma d'un système imageur

L'image d'origine f est reliée à l'image observée p par une transformation qui fait intervenir deux types de phénomènes. Les premiers sont déterministes (ex : projection, défauts du système imageur, etc.) et les seconds sont aléatoires, connus uniquement en termes de statistiques (ex : bruit de mesure).

Le problème inverse consiste à inverser cette transformation : retrouver l'image f à partir de p , ce qui n'est pas toujours possible en pratique. En effet, la plupart des problèmes inverses en traitement d'image sont par nature mal posés. Au sens d'Hadamard, un problème est mal posé si une des conditions suivantes n'est pas respectée :

- Il existe une solution au problème,
- cette solution est unique,
- cette solution dépend continûment des données.

La seule connaissance des données observées est insuffisante pour assurer l'existence, l'unicité et la stabilité d'une solution. En général, on introduit un a priori sur la solution, que l'on appelle régularisation. La segmentation consiste donc à trouver une approximation de f suffisamment correcte pour effectuer les phases de travail en aval.

3 Les différentes approches de la segmentation d'image

La segmentation est une partie très importante de l'analyse d'image, c'est un des sujets qui a été le plus étudié dans ce domaine, car elle a des applications dans de nombreuses domaines pratiques, elle a des débouchés dans la reconnaissance de motifs, la détection d'objets, l'analyse d'image médicale [40, 42], la robotique [43], ou encore dans le domaine des images par satellites, pour les applications de recherche d'image par le contenu [44, 45] et bien d'autres encore. Plusieurs méthodes de segmentation sont connues; parmi elles nous avons la segmentation par région, la segmentation par frontière, et celle basée sur les seuils. Mais il existe cependant beaucoup d'autres méthodes..

Le domaine de la segmentation est un vaste sujet de recherche et fait partie des grands thèmes de l'imagerie numérique. A ce titre, de nombreuses publications font état de segmentations. Comment préférer l'une ou l'autre est un débat ouvert qui fait rage dans bien des laboratoires. En effet, pour valider correctement une segmentation d'objets naturels, comme en imagerie médicale, il faut disposer de la vérité terrain ; ce qui est bien difficile dans le cas de la segmentation, car comment définir de façon précise où commencent et où s'arrêtent les objets sur une image ? Il n'y a donc pas une mais des segmentations possibles sur une même image et elles sont bien souvent subjectives. De même, selon ce que nous voulons segmenter, certaines techniques seront plus à même d'y parvenir.

Nous allons donc présenter dans cette section, diverses techniques connues de segmentation en les organisant selon l'approche qui les régit.

3.1 Contour

Dans l'approche contour ou frontière, on considère que les primitives à extraire sont les lignes de contrastes séparant des régions de niveaux de gris différents et relativement homogènes, ou bien des régions de texture différentes. En pratique, il s'agit de reconnaître les zones de transition et de localiser au mieux la frontière entre les régions.

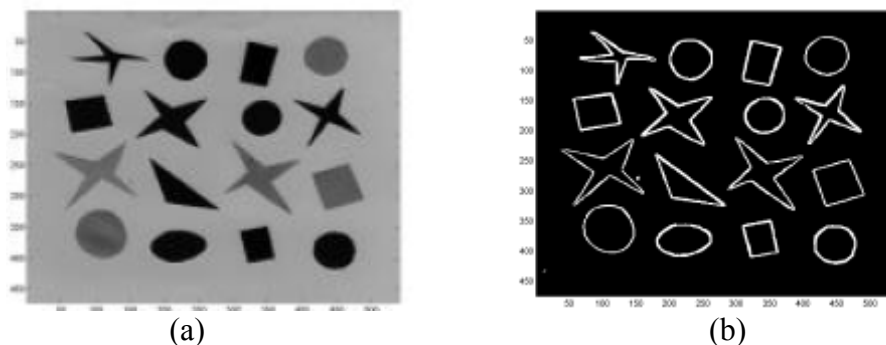


Fig.3.4 : Contours dans une image, a- Image originale, b- Contour des objets de l'image.

Les modèles dérivatifs consistent à modéliser les contours ou des zones d'images et supposent que l'image numérique provient de l'échantillonnage d'une fonction scalaire à support borné et dérivable en tout point. Ces variations d'intensité de l'image peuvent correspondre à des variations d'illuminations (ombres), des changements d'orientation ou de distance de l'observateur, des changements de réflectance de surface, des variations d'absorption des rayons, etc. Or, dans le traitement d'une image numérique, toutes ces grandeurs sont condensées en une seule variable bi ou tridimensionnelle; dans le cas monochrome, c'est l'intensité lumineuse.

- On distingue trois types de contours simples :
- Marche d'escalier : le contour est net (contour idéal).
 - Rampe : le contour est plus flou.
 - Toit : il s'agit d'une ligne sur un fond uniforme.

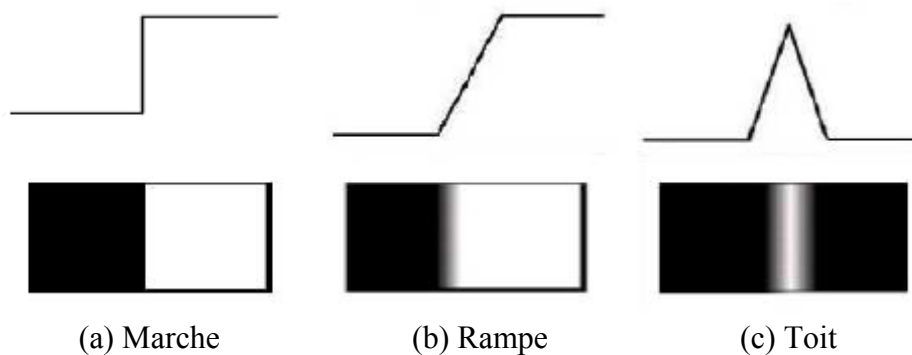


Fig.3.5 : Modèles de contours.

Ce sont des contours idéaux (Figure 3.5) qui ont permis une approche de détection par les opérateurs gradient et Laplacien. Les variations locales d'intensité constituent la source de ces opérateurs; ainsi, le gradient est une fonction vectorielle des pixels $[i, j]$:

$$\nabla f[i, j] = \left(\frac{\partial f}{\partial x}[i, j], \frac{\partial f}{\partial y}[i, j] \right) \quad (3.1)$$

Alors que le Laplacien est une fonction scalaire de $[i, j]$:

$$\Delta f[i, j] = \left(\frac{\partial^2 f}{\partial x^2}[i, j], \frac{\partial^2 f}{\partial y^2}[i, j] \right) \quad (3.2)$$

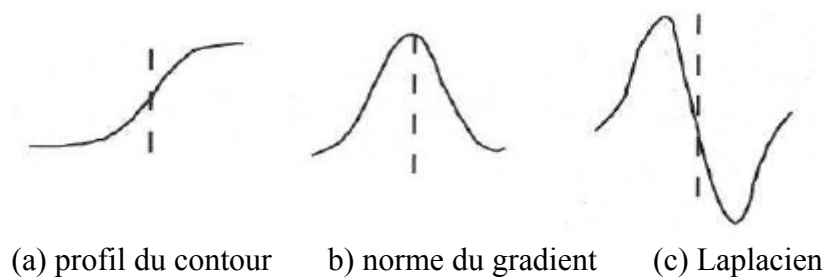


Fig.3.6 : Opérateurs dérivatifs.

Comme on le voit sur la figure 3.6, on peut trouver le point de contour par détermination du maximum de la norme du gradient ou bien en étudiant le passage par zéro

du Laplacien. Ainsi, les opérateurs de Roberts, Prewitt et Sobel donnent une estimation de la dérivée directionnelle de l'image par rapport un axe. Ces opérateurs sont très sensibles aux bruits mais ils ont le mérite d'avoir posé les bases de la détection de contours.

Les opérateurs Laplacien ne sont pas directionnels. En effet, si par hypothèse de départ, la fonction image est suppose continue, alors les propriétés de la dérivée seconde d'une fonction sont utilisées pour caractériser un contour par la présence d'un extremum local par le passage à zéro de la dérivée seconde.

Ainsi en utilisant également les propriétés de différentiation l'opérateur Laplacien peut s'écrire sous la forme suivante :

$$\Delta f(x, y) = f(x+1, y) + f(x-1, y) + f(x, y+1) + f(x, y-1) - 4f(x, y) \quad (3.3)$$

Généralement écrit par le masque de convolution suivant :

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Nous remarquons que l'opérateur Laplacien est symétrique quelque soit l'orientation choisie, il n'est donc pas directionnel. L'opérateur Laplacien est lui aussi fortement sensible aux bruits.

On peut aussi approximer le Laplacien par différence de deux lissages, l'un étant un lissage fort, l'autre un lissage faible. En employant un filtre gaussien, on a alors l'opérateur DOG (Difference Of Gaussians) de Marr et Hildreth [46].

Canny [50, 51] est l'instigateur d'une technique d'optimisation, il démontre qu'on ne peut pas obtenir la fois une bonne détection et une bonne localisation du contour. Pour palier à cela, Canny ajoute un critère supplémentaire pour définir un détecteur optimal : la non multiplicité des maxima locaux. Il développe un filtre à réponse impulsionnelle finie. L'approche de Deriche [47] a été de développer un filtre optimal réponse impulsionnelle infinie. Les différentes littératures présentent la technique de Deriche comme étant une référence dans les détections de contours. Deriche met en œuvre un opérateur particulier où les maxima du gradient sont recherchés dans la direction du gradient par l'opérateur suivant :

$$h(x) = \frac{c}{\omega} e^{-\alpha|x|} \sin(\omega x) \quad \text{avec } \alpha, \omega, c \in R \quad (3.4)$$

Les performances de ce filtre sont maximales quand ω est proche de 0. Selon la valeur de α , la qualité de segmentation diffère : quand α est petit, il y a peu de contours, une bonne qualité de détection mais les contours sont délocalisés ; quand α est grand, la localisation est meilleur, le nombre de contours est plus grand mais la qualité de détection est moins bonne.

Shen et Castan ont proposé un opérateur optimisant un critère incluant la détection et la localisation. Les critères qu'ils obtiennent correspondent aux critères de détection et de localisation de Canny et les filtres obtenus sont assez similaires dans la pratique. Le calcul du filtre de lissage optimal se fait par modélisation de la frontière par un échelon noyé dans un bruit blanc stationnaire additif de moyenne nulle. Le filtre de lissage obtenu par Shen et Castan s'écrit :

$$f(x) = ce^{-\alpha|x|} \quad (3.5)$$

Pour avoir un filtre normalisé :

$$c = \frac{1 - e^{-\alpha}}{1 + e^{-\alpha}} \quad (3.6)$$

et le filtre de dérivation correspondant :

$$h(x) = \begin{cases} d.e^{-\alpha x} & \text{si } x > 0 \\ d.e^{\alpha x} & \text{si } x \leq 0 \end{cases} \quad (3.7)$$

Avec :

$$d = 1 - e^{-\alpha} \quad (3.8)$$

Le paramètre α définit la largeur du filtre : plus α est petit, plus le lissage effectué par le filtre est important. Remarquons que la discontinuité d'ordre 1 au point 0 du filtre de Shen permet d'éviter une délocalisation importante des contours dans l'image lissée, même avec des valeurs faibles de α . Cependant, cette discontinuité peut entraîner la détection de contours multiples. Monga et coll. ont utilisé le paradigme des dérivées partielles dans le cadre de la segmentation d'images biomédicales, notamment appliqué aux IRM et aux scanners cardiaques. Ils ont adapté les filtres de Deriche et de Shen des images en 3 dimensions via une implémentation récursive [48] et ont fait le lien entre ces techniques de détection de contours et la modélisation de surface [49]. Ces techniques ont cependant des potentiels limités et ne sont plus guère utilisées de nos jours.

3.2 Région

L'approche duale de la détection des contours pour la décomposition d'une image en ses formes élémentaires est l'approche par régions. Elle utilise des techniques d'identification et de localisation d'ensembles connexes de pixels. Les méthodes par classification ont pour but de partitionner les images en plusieurs classes et constituent le plus souvent une étape dans la segmentation d'objet à proprement parler. Cependant, leur utilisation dans les méthodes de segmentation étant très répandue, il nous a paru judicieux d'en expliquer les ressorts.

Ces différentes classifications peuvent être séparées selon plusieurs critères : probabilistes ou déterministes, paramétriques ou non, supervisées ou non. Nous présenterons donc un panel de ces méthodes parmi lesquelles nous trouverons les réseaux de neurones (méthode déterministe supervisée), les approches K-Moyennes, Fuzzy C-Means, Adaptive Fuzzy C-Means et Mean Shift (déterministes non supervisées), les différentes mixtures de lois (probabilistes paramétriques) et les approches markoviennes et par Machine à Vecteurs de Support (probabilistes non-paramétriques).

3.2.1 Classification Déterministe Supervise : Réseaux de neurones

Un réseau de neurones artificiels [50, 51] est en général composé d'une succession de couches dont chacune prend ses entrées sur les sorties de la précédente. Chaque couche i est composée de N_i neurones (Fig.3.7) prenant leurs entrées sur les N_{i-1} neurones de la couche précédente.

A chaque synapse est associé un poids synaptique, de sorte que les N_{i-1} sont multipliés par ce poids, puis additionnés par les neurones de niveau i , ce qui est équivalent à multiplier le vecteur d'entrée par une matrice de transformation. Mettre l'une derrière l'autre, les différentes couches d'un réseau de neurones reviendrait à mettre en cascade plusieurs

matrices de transformation et pourrait se ramener à une seule matrice, produit des autres, s'il n'y avait à chaque couche, la fonction de sortie qui introduit une non-linéarité à chaque étape. Ceci montre l'importance du choix judicieux d'une bonne fonction de sortie : un réseau de neurones dont les sorties seraient linéaires, n'aurait aucun intérêt.

Au delà de cette structure simple, le réseau de neurones peut également contenir des boucles qui en changeant radicalement les possibilités mais aussi la complexité. De la même façon que des boucles peuvent transformer une logique combinatoire en logique séquentielle, les boucles dans un réseau de neurones transforment un simple dispositif de reconnaissance d'entrées, en une machine complexe capable de toutes sortes de comportements.

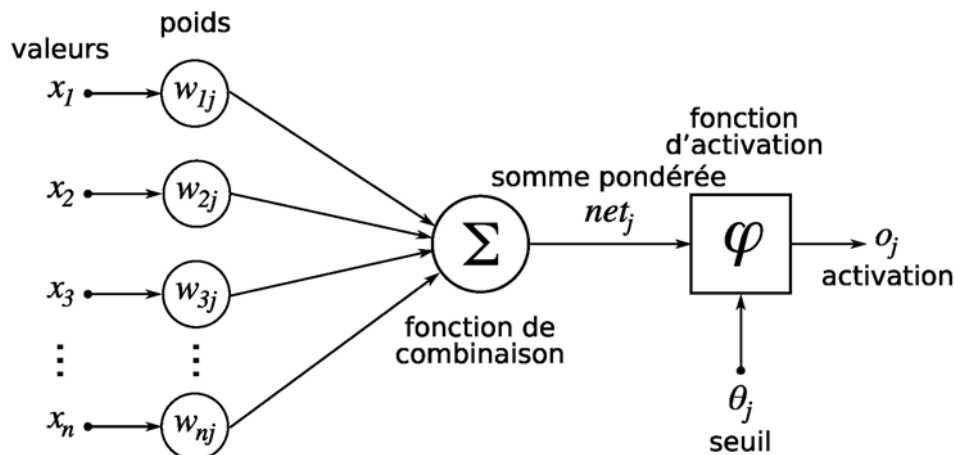


Fig.3.7 – Structure d'un neurone artificiel. Le neurone calcule la somme de ses entrées puis cette valeur passe à travers la fonction d'activation pour produire sa sortie.

Classiquement, en segmentation d'images médicales, certains réseaux de neurones sont utilisés comme classificateurs. Les poids synaptiques sont déterminés par apprentissage sur une base d'image dont le résultat de segmentation est connu, on parle alors de réseau de neurones supervisé. Souvent, les neurones d'entrées sont les différentes IRM disponibles et les neurones de sorties nous donnent alors les différentes classes recherchées [52]. Il est en outre possible d'introduire des informations a priori en plus des volumes et donc de donner plus de robustesse à cette classification.

Mais l'inconvénient majeur de cette méthode par réseaux de neurones est l'étape d'apprentissage qui demande une intervention manuelle pour donner cette vérité terrain dont le réseau a besoin pour calculer les poids synaptiques.

3.2.2 Classification Déterministe Non Supervise

a) K-Moyennes

L'algorithme des k-moyennes classe les objets selon leurs attributs en k parties (ou clusters) en supposant que les attributs des objets forment un espace vectoriel. L'objectif est de minimiser la variance intra-cluster :

$$v = \sum_{i=1}^k \sum_{x_j \in S_i} |x_j - \mu_i|^2 \quad (3.9)$$

Où S_i , $i = 1, 2, \dots, k$ sont les k clusters et μ_i est le centroïde ou point moyen des points $x_j \in S_i$. L'algorithme commence par partitionner les points en k ensembles initiaux, soit au hasard, soit en utilisant une heuristique. Il calcule ensuite le centroïde de chaque ensemble et construit une nouvelle partition en associant chaque point avec le centroïde le plus proche. S'ensuit une

alternance entre calcul des centroïdes des nouveaux clusters et appariement des points avec le centroïde le plus proche jusqu'à convergence. Celle-ci est obtenue quand plus aucun point ne change de groupe (ou bien quand les centroïdes ne changent plus).

Cet algorithme est très populaire car extrêmement rapide en pratique. En effet, le nombre d'itérations est typiquement inférieur au nombre de points. En termes de performance, cet algorithme ne garantit pas un optimum global. La qualité de la solution dépend grandement des ensembles initiaux et peut, en pratique, être bien au delà de l'optimum global. Comme l'algorithme est très rapide, une méthode courante est de le lancer plusieurs fois et de retourner la meilleure partition.

L'autre problème est qu'il est nécessaire de donner le nombre de clusters k à trouver; cela n'étant pas réellement handicapant dans le cas de la segmentation cérébrale puisque le nombre de classes est, le plus souvent, connu.

Cet algorithme a été utilisé pour segmenter le cerveau [53, 54] avec des résultats plutôt satisfaisants mais la qualité non constante de la solution en fait un algorithme à proscrire pour une automatisation du travail.

b) C-moyennes floues et c-moyennes floues adaptatives

l'algorithme des c -moyennes floues (ou fuzzy c -means - FCM) introduit par Dunn généralise l'algorithme des k -moyennes en permettant la classification floue basée sur la théorie des ensembles flous. Bezdek s'y est intéressé [55, 56] et a développé cet algorithme. Dans le cas des c -moyennes floues, la fonctionnelle L à minimiser est :

$$\mathfrak{J} = J_{\omega}(x, v : y) = \sum_{k=1}^c \sum_{i=1}^n x_{ik}^{\omega} \|y_k - v_i\|_A^2 \quad (3.10)$$

Où n est le nombre de points à traiter, c le nombre de classes désirées, $\omega \in [1, +\infty]$ est le poids de fuzzyfication, $v = (v_1, \dots, v_c)$ est le vecteur des centres de classes, $\|\cdot\|_A$ est un produit scalaire où A est une matrice définie positive et $x = [x_{ik}] \in \mathbb{R}^{c \times n}$, avec $x_{ik} \in [0, 1] \forall 1 \leq i \leq c$ et $1 \leq k \leq n$, est la c -partition floue de y et doit vérifier :

$$\sum_{i=1}^c x_{ik} = 1 \text{ pour } 1 \leq k \leq n, \quad \sum_{i=1}^n x_{ik} > 0. \text{ pour } 1 \leq i \leq c \quad (3.11)$$

3.2.3 Classification Probabiliste Paramétrique (Mélange de Lois)

Le problème classique de la classification automatique est de considérer qu'un échantillon de données provienne d'un nombre de groupes inconnus a priori qu'il faut retrouver. Lorsqu'on part du postulat que ces groupes suivent une loi de probabilité (quelconque), alors on se place nécessairement dans le cadre des modèles de mélanges. Si, en plus, on considère que les lois que suivent les individus sont normales, alors on se place dans le cadre des modèles de mélanges gaussiens.

Par la suite, on notera x , un échantillon composé de n individus (x_1, \dots, x_n) appartenant à \mathbb{R}^p (i.e. caractérisés par p variables continues). Dans le cadre des modèles de mélanges gaussiens, on considère que ces individus appartiennent chacun à un des g (g étant fixé a priori) G_1, \dots, G_g suivant chacun une loi normale de moyenne μ_k , ($k = 1, \dots, g$) et de matrice de variance-covariance Σ_k . D'autre part, en notant π_1, \dots, π_g les proportions des différents groupes, $\theta_k = (\mu_k, \Sigma_k)$ le paramètre de chaque loi normale et $\Phi = (\pi_1, \dots, \pi_g, \theta_1, \dots, \theta_g)$ le paramètre global du mélange, la loi mélange que suit l'échantillon peut s'écrire :

$$g(x, \Phi) = \sum_{k=1}^g \pi_k f(x, \theta_k) \quad (3.12)$$

Avec $f(x, \theta_k)$ la loi normale multidimensionnelle paramétrée par θ_k .

La principale difficulté de cette approche consiste à déterminer le meilleur paramètre Φ . Pour cela, on cherche habituellement le paramètre qui maximise la vraisemblance, donnée dans ce cas, par :

$$L(x; \Phi) = \sum_{i=1}^n \log(\sum_{k=1}^g \pi_k f(x_i, \theta_k)) \quad (3.13)$$

Une fois l'estimation effectuée, il s'agit d'attribuer à chaque individu la classe à laquelle il appartient le plus probablement. Pour cela, on utilise la règle d'inversion de Bayes. D'après celle-ci, on a :

$$P(G_k | x) = \frac{p(x|G_k) \cdot p(G_k)}{p(x)}, \quad (3.14)$$

Ce qui se traduit, dans notre cas, par :

$$P(G_k | x_i) = \frac{\pi_k f(x_i, \theta_k)}{\sum_{l=1}^g \pi_l f(x_i, \theta_l)}, \quad (3.15)$$

Il suffit alors d'attribuer chaque individu x_i à la classe pour laquelle la probabilité a posteriori $P(G_k | x_i)$ est la plus grande.

Bien que ce problème puisse sembler particulièrement ardu, l'algorithme EM (espérance-maximisation), en anglais Expectation-maximisation, permet de lever la difficulté de la détermination de Φ : Espérance-Maximisation alterne des étapes d'évaluation de l'espérance (E), où l'on calcule l'espérance de la vraisemblance en tenant compte des dernières variables observées, et une étape de maximisation (M), où l'on estime le maximum de vraisemblance des paramètres en maximisant la vraisemblance trouvée à l'étape E. On utilise ensuite les paramètres trouvés en M comme point de départ d'une nouvelle phase d'évaluation de l'espérance, et l'on itère ainsi.

3.2.4 Classification Probabiliste Non-Paramétrique (Champs Aléatoires de Markov)

La classification par mélange de lois ne tient pas compte de la répartition spatiale des voxels, pour ajouter cette information, on utilise les champs de Markov [57] qui modélisent les interactions entre un voxel et son voisinage. Notons x_s la valeur du descripteur au site s , $x^s = (x_s)_{t \neq s}$ la configuration de l'image excepté le point s , V_s le système de voisinage et $U(x) = \sum_{c \in C} U_c$ l'énergie globale de l'image (c'est la somme des potentiels de toutes les cliques). X est un champ de Markov si et seulement si :

$$P(x_s = x_s / x^s) = P(X_s = x_s / x_t, t \in V_s) \quad (3.16)$$

Autrement dit, le niveau de gris d'un site ne dépend que des niveaux de gris des pixels voisins. Les probabilités conditionnelles locales définies ci-dessus seront calculées par le théorème de Hammersley- Clifford grâce au champ de Gibbs donné par :

$$P(X = x) = \frac{1}{Z} \exp(-U(x)) \quad (3.17)$$

Où

$$Z = \sum_{x \in \Omega} \exp(-U(x)). \quad (3.18)$$

La définition du champ de Markov peut se réécrire alors :

$$P(X_s = x_s / x^s) = \frac{\exp(-U_s(x_s / V_s))}{\sum_{\xi \in E} \exp(-U_s(\xi / V_s))}, \quad (3.19)$$

Le formalisme des champs de Markov permet d'effectuer une segmentation de l'image en prenant en compte les interactions avec les pixels voisins. On considère que les k régions que l'on souhaite segmenter forment une partition de l'image. Chaque région est représentée par une fonction caractéristique et identifiée par une étiquette dans $\{1, \dots, k\}$. Le but de la segmentation est d'estimer le champ des étiquettes X à partir d'une réalisation bruitée de l'image Y . La démarche de la segmentation peut se formaliser comme un problème d'estimation bayésienne. On note S l'ensemble des sites de l'image, X le champ des étiquettes et Y le champ des observations (image de départ), on a donc dans le cas d'une image en 256 niveaux de gris :

$$X \begin{cases} S \rightarrow \{1, \dots, k\} \\ s \mapsto x_s \end{cases} \quad \text{et} \quad Y \begin{cases} S \rightarrow [0, \dots, 255] \\ s \mapsto y_s \end{cases} \quad (3.20)$$

En supposant que la réalisation de l'image est indépendante pour chaque pixel on a :

$$P(Y|X) = \prod_{s \in S} P(y_s | x_s) \quad (3.21)$$

Grâce au théorème de Bayes, il est possible d'exprimer la probabilité `a posteriori` d'un champ d'étiquettes étant donnée une observation Y . Connaissant Y , le champ des étiquettes est un champ aléatoire de distribution :

$$P(X|Y) = \frac{p(Y|X) \times p(X)}{p(Y)} \quad (3.22)$$

$P(Y)$ est un terme constant et ne nous intéresse pas pour l'estimation de X . $P(X|Y)$ mesure la similarité de la classification aux données que l'on observe, c'est le terme d'attache aux données. $P(X)$ est une probabilité a priori sur la distribution du champ d'étiquettes. On peut modéliser cet a priori grâce au modèle de Potts qui tend à favoriser des zones compactes et uniformes.

$$P(X) = \frac{1}{Z} \exp(-\beta \sum_{C=\{s,t\}} \phi(x_s - x_t)) \quad (3.23)$$

β joue le rôle de coefficient de régularisation. Plus β est grand, plus la transition entre régions est pénalisée et plus les régions obtenues sont grandes. Grâce au théorème de Hammersley-Clifford, on peut passer de la représentation probabiliste à une représentation en énergie. En prenant le logarithme des probabilités on obtient :

$$U(x) = \sum_{s \in S} \left[\frac{(y_s - \mu_{x_s})^2}{2\sigma_{x_s}} + \frac{1}{2} \log(2\pi)\sigma_{x_s} \right] + \beta \sum_{C=\{s,t\}} \phi(x_s - x_t) \quad (3.24)$$

Le choix de la fonction ϕ influe grandement sur le résultat (notamment en ce qui concerne la convexité de $U(x)$), ainsi, le choix d'une forme quadratique pure qui correspond la régularisation de Tikhonov pénalise les forts gradients, donnant alors une segmentation plus lisse.

Une estimation du champ des étiquettes peut se faire suivant le critère du Maximum a posteriori. Cela se fait en minimisant l'énergie grâce à l'algorithme des modes conditionnels itérés [58, 59] ou du recuit simulé. Une autre façon d'estimer le champ des étiquettes est d'utiliser une fonction de coût proportionnelle au nombre de pixels mal classifiés ce qui conduit à l'algorithme du maximum a posteriori de la marginale.

3.3 Approches Structurelles (Ligne de Partage des Eaux)

La morphologie mathématique consiste à comparer une structure inconnue (i.e. l'image que l'on étudie) à un ensemble de formes, les éléments structurants, dont on maîtrise toutes les caractéristiques. Ces comparaisons s'effectuent au moyen d'opérations booléennes telles que l'intersection et l'inclusion.

L'algorithme de la ligne de partage des eaux (LPE), proposé par Digabel et Lantuéjoul, utilise la description des images en termes géographiques. Une image peut en effet être perçue comme un relief si l'on associe le niveau de gris de chaque point à une altitude. Il est alors possible de définir la ligne de partage des eaux comme étant la crête formant la limite entre deux bassins versants.

Pour l'obtenir, il faut imaginer l'immersion d'un relief dans de l'eau, en précisant que l'eau ne peut pénétrer dans les vallées que par ses minima. La ligne de partage des eaux est représentée par les points où deux lacs disjoints se rejoignent au cours de l'immersion.

Une des difficultés à la mise en oeuvre de cette analogie intuitive est qu'elle laisse beaucoup de liberté quant à sa formalisation. Il existe deux classes principales d'implémentation : l'une est basée sur un algorithme récursif d'immersion [60] et une autre basée sur l'utilisation de fonctions de distances géodésiques [61]. Leur définition algorithmique ainsi que leur analyse critique est donnée dans l'article de Roerdink et Meijster [62].

Cette technique est souvent associée à une méthode de fusion de régions puisqu'elle donne une sur-segmentation de l'image. On trouve cette utilisation conjointe dans [63] pour des images multi-dimensionnelles et notamment des IRM de cerveaux.

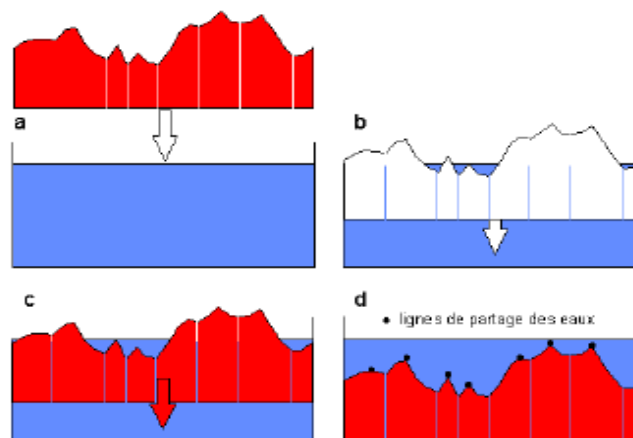


Fig.3.8: Principe de la Ligne de Partage des Eaux.

Peng et coll. proposent d'utiliser une immersion multi-degré pour améliorer l'algorithme LPE [64]. Betser et coll. utilisent la LPE dans un contexte de segmentation hiérarchique [65].

3.4 Théorie des Graphes et Hypergraphe

Dans les approches utilisant la théorie des graphes, l'idée directrice est de créer un graphe à partir de l'image selon des procédés assez simples et de travailler sur ces graphes, pouvant ainsi utiliser toute la panoplie d'outils développés dans le cadre de cette théorie. Un travail relativement important sera de valuer les arêtes puisque ce sont elles qui permettront de donner les caractéristiques de l'image à notre graphe.

Les hypergraphes, introduit par Claude Berge en 1969 [66] généralisent la notion de graphe dans le sens où les arêtes ne relient plus un ou deux sommets, mais un nombre quelconque de sommets (compris entre un et le nombre de sommets de l'hypergraphe).

Soient $V = \{v_1, v_2, \dots, v_n\}$ un ensemble, $E = \{E_1, E_2, \dots, E_m\}$ une famille de parties de V , avec $(m, n) \in \mathbb{N}^2$ deux entiers non nuls. Un hypergraphe H est un couple (V, E) tel que :

$$\forall i \in ([1, m] \cap \mathbb{N}) \quad E_i \neq \emptyset, \quad \bigcup_{i=1}^m E_i \subseteq V \quad (3.25)$$

À l'instar des graphes classiques, on dit que :

- Les éléments de V sont les sommets de H .
- Le nombre de sommets n est l'ordre de l'hypergraphe.
- Les éléments de E sont les arêtes ou hyper-arêtes de H .

Un apport intéressant pour le traitement d'image est la notion d'hypergraphe de voisinage. Dans cet hypergraphe, les sommets sont les pixels de l'image et les hyper-arêtes relient les pixels voisins. En ajoutant un paramètre de sélection sur la distance colorimétrique (i.e. les hyper-arêtes ne contiennent que les pixels de couleurs "proches"), on obtient un hypergraphe de voisinage colorimétrique qui nous donne une notion de gradient d'intensité. D'autres critères peuvent être employés pour créer les hyper-arêtes, permettant ainsi la création d'hypergraphes de voisinage adaptatifs [67, 68].

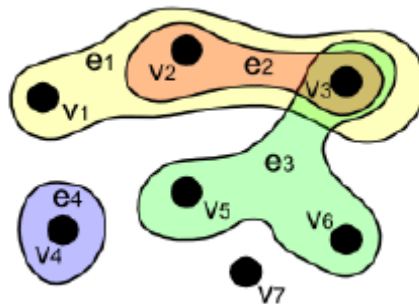


Fig.3.9 : Exemple d'hypergraphe par Claudio Rocchini.

Bretto et coll. proposent dans [69] une revue de l'apport des hypergraphes dans le traitement d'images et notamment dans la segmentation et la détection de contours par l'utilisation des familles d'arêtes intersectantes dites étoiles.

4 Conclusion

A travers ce chapitre, nous avons proposé un état de l'art des différentes techniques de segmentation d'images. Puisque la qualité du résultat est intrinsèquement subjective et que la finalité d'utilisation de la segmentation obtenue influe sur la perception de qualité de ce résultat, le choix d'une technique plutôt qu'une autre est assez ardu. On remarque néanmoins une certaine évolution des techniques au cours du temps. Ainsi, les plus anciennes utilisaient préférentiellement les observations brutes et les informations qu'elles contenaient alors que les techniques actuelles ont tendance introduire de plus en plus de connaissances a priori pour accroître la qualité de la segmentation.

Grâce à l'arrivée de machines de plus en plus puissantes, la capacité de calcul s'est accrue et il est désormais possible d'utiliser des techniques gourmandes en ressources mais d'une qualité indéniable. La problématique de la segmentation s'est donc peu à peu mue en un véritable travail d'optimisation des techniques.

Chapitre IV

Théorie des graphes et théorie de l'évidence

1 Théorie des graphes

La théorie des graphes ouvre un grand champ de modélisation conduisant à des solutions efficaces pour de nombreux problèmes. Il est difficile de dater très précisément le moment où l'idée de graphe a pu germer dans l'esprit des mathématiciens. Si les travaux de Leibniz ou d'Euler au XVIII^e siècle sont probablement précurseurs, il est intéressant de noter que la majeure partie des problèmes de combinatoire ou de recherche opérationnelle peut s'exprimer au moyen de la théorie des graphes. La théorie des graphes s'est alors développée dans diverses disciplines telles que la chimie, la biologie, les sciences sociales. Depuis le début du XX^e siècle, elle constitue une branche à part entière des mathématiques, grâce aux travaux de König, Menger, Cayley puis de Berge et d'Erdős.

1.1 Introduction

De manière générale, un graphe permet de représenter la structure, les connexions d'un ensemble complexe en exprimant les relations entre ses éléments : réseau de communication, réseaux routiers, interaction de diverses espèces animales, circuits électriques et dans plusieurs autres domaines. Les graphes constituent donc une méthode de pensée qui permet de modéliser une grande variété de problèmes en se ramenant à l'étude de sommets et d'arcs.

1.2 Définition d'un graphe

On appelle graphe G le couple (S, A) constitué d'un ensemble S de sommets ou nœuds et d'un ensemble A d'arêtes ou d'arcs. Chaque arc de A relie deux sommets de S . La cardinalité de l'ensemble S est appelée la taille du graphe. Si $e = \{x, y\}$ est une arête du graphe, les sommets x et y sont dénommés extrémités de e . Graphiquement, on représente souvent les nœuds par des points, et une arête $e = \{s, t\}$ par un arc de courbe joignant les deux points s et t (Fig.4.1). Un graphe peut être orienté ou non selon le sens que l'on donne aux arcs, les deux sommets qu'ils relient pouvant être considérées comme ordonné ou non.

1.3 Graphes non orientés

Un graphe simple G est un couple formé de deux ensembles, un ensemble S dont les éléments sont appelés sommets, et un ensemble A dont les éléments sont appelés arêtes. Le graphe est noté par $G(S, A)$ avec $a = \{x_i, x_j\}$ appelée l'arête de G d'extrémités x_i et x_j , ou que a joint x_i et x_j , ou que a passe par x_i et x_j . Les sommets x_i et x_j sont dits adjacents dans G (Fig.4.1).

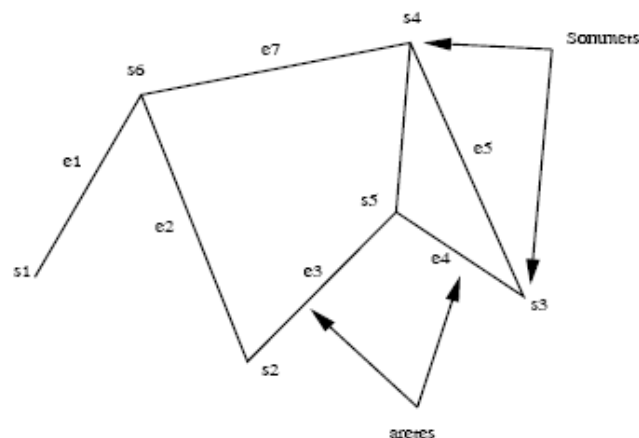


Fig.4.1: Exemple d'un graphe non orienté

Un multigraphe $G(S, A, f)$ est déterminé par :

- un ensemble S de sommets
- un ensemble A , cette fois abstrait
- une application $f : A \rightarrow P_2(S)$

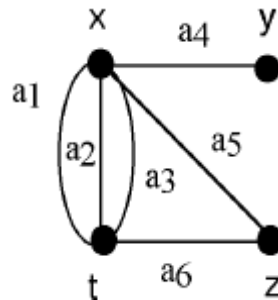


Fig.4.2: Liens entre sommets et arêtes

Dans cet exemple x, y, z, t sont les sommets du multigraphe et : $f(a_1) = f(a_2) = f(a_3) = \{x, t\}$; $f(a_4) = \{x, y\}$; $f(a_5) = \{x, z\}$; $f(a_6) = \{z, t\}$. Un multigraphe avec boucles est un triplet (S, A, f) où f est une application de A dans $p_2(S) \cup p_1(S)$; en d'autres termes, un multigraphe avec boucles peut comprendre des arêtes multiples entre deux sommets donnés ainsi que des boucles multiples en un sommet (Fig.4.2).

1.4 Graphes orientés

Un graphe orienté G est formé de deux ensembles, un ensemble X dont les éléments sont appelés sommets, et un ensemble A dont les éléments sont appelés arcs (Fig.4.3).

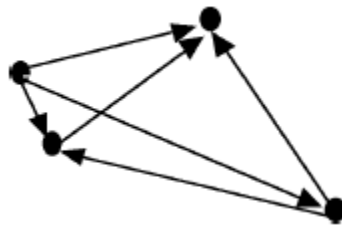


Fig.4.3: Graphe orienté

Si $a = (x, y)$ est un arc du graphe $G(S, A)$, x est l'extrémité initiale de a et y l'extrémité finale de a .

1.5 Définitions

- **Le degré d'un nœud** : Soit $G(S, A)$ un graphe simple, et x un sommet de ce graphe. Le degré de x , noté $d(x)$, est le nombre d'arêtes incidentes à x , c'est à dire contenant x . Lorsque $d(x) = 0$, on dit que le sommet est isolé, lorsque $d(x) = 1$, il est dit pendant.
- **Graphe régulier** : Un graphe simple est dit régulier de degré r , lorsque tous ses sommets sont de degré r .
- **Sous graphe** : $H = (Y, B)$ est un sous-graphe de $G = (X, A)$ si $Y \subseteq X$ et $B \subseteq A$.
- **Graphe partiel** : $H = (Y, B)$ est un graphe partiel de $G = (X, A)$ si $Y = X$ et $B \subseteq A$.
- **Ordre d'un graphe** : l'ordre d'un graphe est le nombre de sommets de ce graphe.
- **Chaîne** : suite finie de sommets reliés entre eux par une arête.
- **Chaîne simple** : chaîne qui n'utilise pas deux fois la même arête.
- **Chaîne eulérienne** : chaîne simple passant par toutes les arêtes d'un graphe.

- **Chaîne hamiltonienne** : chaîne simple passant par tous les sommets d'un graphe une et une seule fois.
- **Chemin** : suite de sommets reliés par des arcs dans un graphe orienté.
- **Cycle** : chaîne qui revient à son point de départ.
- **Cycle eulérien** : cycle simple passant par toutes les arêtes d'un graphe une et une seule fois.
- **Cycle hamiltonien** : cycle simple passant par tous les sommets d'un graphe une et une seule fois.
- **Graphe connexe** : un graphe G est dit connexe si pour toute paire de sommets $\{x, y\}$ de G , il existe une chaîne de premier terme x et de dernier terme y .
- **Arbre** : graphe connexe sans cycle simple et sans boucle.
- **Graphe eulérien** : graphe qui possède un cycle eulérien.
- **Graphe semi-eulérien** : graphe qui possède une chaîne eulérienne.
- **Graphe hamiltonien** : graphe qui possède un cycle hamiltonien.
- **Graphe semi-hamiltonien** : graphe qui possède une chaîne hamiltonienne.
- **Graphe valué** : graphe où des réels sont associés aux arêtes.
- **Longueur d'une chaîne** : nombre des arêtes qui composent la chaîne.
- **Valeur d'une chaîne** : somme des valeurs des arêtes (arcs) d'une chaîne d'un graphe valué.
- **Distance entre deux sommets** : longueur de la plus courte chaîne joignant ces deux sommets.
- **Diamètre d'un graphe** : maximum des distances entre les sommets d'un graphe.
- **Indice chromatique** : nombre minimal de couleurs permettant de colorier les arêtes d'un graphe, de telle sorte que deux arêtes adjacentes n'aient pas la même couleur.
- **Nombre chromatique d'un graphe** : nombre minimal de couleurs permettant de colorier les sommets d'un graphe, de telle sorte que deux sommets adjacents n'aient pas la même couleur.

1.6 Graphes eulériens (Théorème d'Euler 1766)

Un graphe simple connexe $G = (S, A)$ est eulérien si et seulement si pour tout sommet x de S , $d(x)$ est pair.

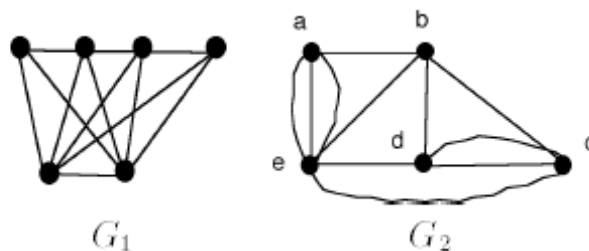


Fig.4.4 : G_1 n'est pas un graphe eulérien, G_2 est un graphe eulérien.

Un graphe simple connexe est semi-eulérien si et seulement si il admet aucun sommet ou exactement deux sommets de degré impair.

1.7 Graphes hamiltoniens

Contrairement aux graphes eulériens, il n'existe pas de caractérisation simple des graphes hamiltoniens ou semi-hamiltoniens. On peut cependant énoncer quelques propriétés et conditions suffisantes.

Soit $G = (S, A)$ un graphe simple d'ordre $n \geq 3$, Si pour toute paire $\{x,y\}$ de sommets non adjacents, on a $d(x) > d(y)$ alors G est hamiltonien.

Un graphe possédant un sommet de degré 1 ne peut être hamiltonien. Si un sommet dans un graphe est de degré 2, alors les deux arêtes incidentes à ce sommet doivent faire partie du cycle hamiltonien.

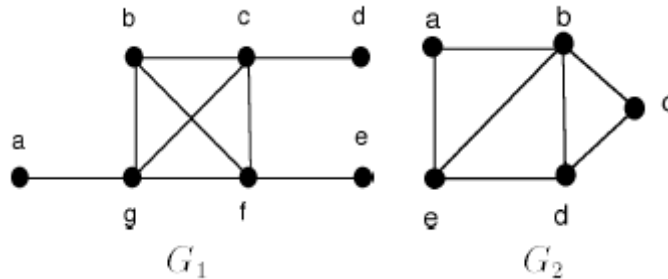


Fig.4.5: Le graphe G_1 n'est pas hamiltonien (sommet de degré 1), le graphe G_2 est hamiltonien.

1.8 Matrice d'adjacence

Soit $G = (S, A)$ un graphe orienté, avec $x = \{x_1, x_2, \dots, x_n\}$. La matrice d'adjacence du graphe G est la matrice $M(G)$ dont les coefficients $m_{i,j}$ sont définis par : $m_{i,j} = 1$ si $(x_i, x_j) \in A$ et $m_{i,j} = 0$ si $(x_i, x_j) \notin A$.

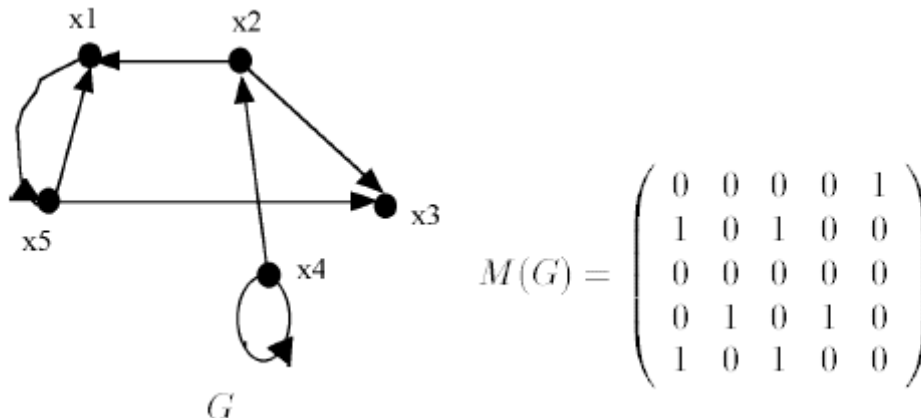


Fig.4.6: Graphe G et sa matrice d'adjacence.

La trace de $M(G)$ est égale au nombre de boucles du graphe G .

1.9 Distance

Soit $G = (S, A)$ un graphe orienté. Pour tout $(x, y) \in X^2$, la distance entre le sommet x et le sommet y est le nombre $d(x, y)$ défini par :

$D(x, y)$ est infinie s'il n'existe pas de chemin de x à y .

$D(x, y) = \min \{L(c) / c \text{ est un chemin de } x \text{ à } y\}$.

Où $L(c)$ désigne la longueur du chemin c . La matrice des distances (Fig.4.7) du graphe G est la matrice $D = d(i, j) = (d(x_i, x_j))$.

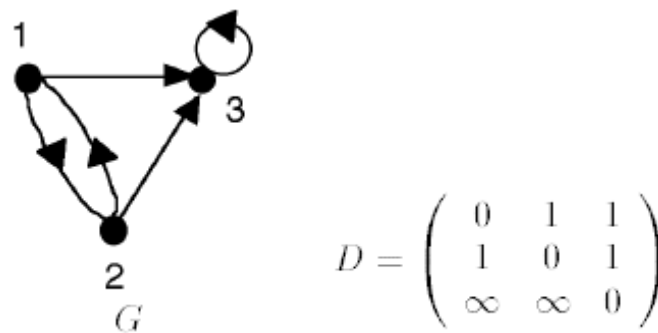


Fig.4.7: le graphe G et sa matrice des distances.

2 Théorie de l'évidence

Plusieurs sources d'information indépendantes sont souvent disponibles pour la classification de l'état d'un système. La combinaison de ces sources d'information par la théorie de Demspster-Shafer permet d'exprimer fidèlement chacune d'elle et de les combiner en exhibant la qualité de la classification résultante. La fusion d'informations consiste à combiner des informations hétérogènes issues de plusieurs sources afin d'améliorer la prise de décision.

L'établissement d'une décision à partir de la fusion d'informations a beaucoup évolué ces dernières années : dans différents domaines, et en particulier en vision et en robotique, les sources d'informations se sont multipliées, qu'il s'agisse de sources de capteurs, d'informations a priori, de connaissances génériques, etc. Chaque source fournissant en général une information bruitée ou incomplète, il est important d'en combiner plusieurs afin d'avoir une meilleure connaissance du monde à observer. La fusion d'informations peut alors se définir comme la combinaison d'informations, souvent imparfaites et hétérogènes, afin d'obtenir une information globale plus complète, de meilleure qualité, et permettant de mieux décider et agir. Parallèlement, les méthodes pour modéliser les connaissances et informations imparfaites et pour les combiner ont connu des développements théoriques importants et leurs champs d'applications se sont étendus. Ces méthodes, souvent issues des théories de la décision, de l'incertain et de l'intelligence artificielle, peuvent être soit numériques, soit symboliques. L'ampleur que prend la fusion d'informations suit celle que prennent les technologies et le traitement de l'information en général.

2.1 Introduction

Un des principaux challenges de la science réside dans l'accès à une information plus fiable, à sa représentation informatique ainsi qu'à son traitement. Ainsi, dans le domaine de l'intelligence artificielle, nombreuses sont les recherches qui ont permis d'étendre les formalismes mis en place pour le développement de systèmes intelligents au traitement de l'information imprécise et incertaine. Parmi les modèles étudiés, on retrouve les différentes théories des mesures de confiance qui inclut notamment un formalisme capable d'appréhender à la fois imprécision et incertitude, la théorie des fonctions de croyance. Cette description permet au lecteur néophyte de se familiariser avec les éléments mathématiques utiles à la représentation et à la manipulation de l'information. Celle-ci peut prendre plusieurs formes (mesures, données ou connaissances) selon le type de problèmes abordés (reconnaissance des formes, fusion d'informations). Ce modèle est suffisamment souple pour être appliqué dans de nombreux domaines des sciences et technologies de l'information et de la communication tels que l'analyse de données, le diagnostic, l'aide à la décision, la perception multicapteurs et le traitement d'images.

Dans les nombreux secteurs applicatifs, il s'agit de classer, décider, surveiller, commander à partir d'informations observées sur un système ou préalablement modélisées [70]. De manière générale, les finalités de ces applications consistent à aider un opérateur dans sa tâche, transmettre et enregistrer des savoir-faire, gérer des connaissances, améliorer les performances ou l'adaptabilité de systèmes artificiels. Dans [70], les auteurs dégagent quatre grandes classes de tâches de la problématique scientifique du traitement de l'information :

1. Mettre l'information sous une forme utile,
2. Stocker, retrouver, expliciter l'information,
3. Exploiter l'information pour décider et agir,
4. Communiquer l'information.

Pour ce faire, les systèmes développés se basent sur un flot d'informations qui est soit préalablement collecté (apprentissage), soit recueilli en ligne (observation). Le niveau d'abstraction de cette information diffère selon son origine et peut se catégoriser en trois familles :

Une mesure : C'est le cas notamment pour les systèmes de perception multicapteurs. Dans ce contexte, il s'agit alors de fusionner les mesures provenant d'un ensemble de capteurs (signaux et/ou images) pour la commande d'un système ou la prise d'une décision. Le problème consiste généralement à fiabiliser les mesures (filtrer, débruiter la mesure, lui associer un degré d'incertitude) pour les rendre exploitables a posteriori.

Une donnée : Dans le cadre de la gestion des bases. Suivant le type de données (financières, commerciales, médicales,...), il peut s'agir d'extraire des variables discriminantes, d'identifier des structures sous-jacentes, d'extraire des informations utiles de manière générale, c'est le concept de fouille de données [71]. Une partie de la communauté scientifique se regroupe autour de cette problématique connue sous le vocable KDD pour Knowledge Discovery in Databases. Dans le cadre d'un système d'indexation, le problème peut consister à sélectionner les variables pertinentes de la base, catégoriser les données pour les indexer, fusionner les données avec la requête de l'utilisateur au niveau de la boucle de pertinence.

Une connaissance : On parle de modélisation de connaissances ou d'ingénierie des connaissances. Dans ce dernier cas, les difficultés proviennent de la capture des connaissances, de leur explicitation et de leur interprétation sans introduction d'un biais. Un exemple peut provenir de la structuration de connaissances en vue de leur formalisation logique.

La mesure d'une grandeur physique est généralement entachée d'une imprécision liée au choix et à la technologie du capteur employé. Cette notion d'imprécision est généralement bien maîtrisée mais est souvent liée à un second concept qui se différencie par le fait qu'il ne fait pas référence au contenu de l'information mais à sa qualité. Un des nombreux problèmes couramment rencontré dans de nombreuses situations concerne le manque d'informations. C'est le cas lors du dysfonctionnement d'un capteur ; on parle alors d'incomplétude. Modéliser (représenter, accéder et traiter) des informations aussi imparfaites devient alors rapidement un problème crucial et constitue l'un des principaux challenges du traitement de

l'information. En réalité, nombreuses sont les causes de l'imperfection des informations. Dans [70] les auteurs parlent d'information ambiguë, bruitée, biaisée, incomplète, imprécise, incertaine, incohérente, redondante.

Nombreuses sont les recherches qui se sont attachées à développer des outils théoriques pour la gestion de ces informations imparfaites [72, 73]. Ainsi, la théorie des sous-ensembles flous (Fuzzy Sets Theory) développée par L. Zadeh [74] a été un cadre précurseur. Un autre cadre théorique très séduisant est certainement la théorie des fonctions de croyance.

Elle permet dans un même formalisme d'encoder l'imprécision et l'incertitude contenue dans une information. Un des principaux ambassadeurs de ce formalisme est sans doute Ph. Smets qui, via le modèle des croyances transférables [75], continue à populariser la théorie développée par A. Dempster et G. Shafer[76].

2.2 Eléments mathématiques

Les travaux d'A. Dempster sur les bornes inférieures et supérieure d'une famille de distributions de probabilités [76] ont permis à G. Shafer d'asseoir les bases de la théorie des fonctions de croyance. Ce formalisme théorique peut revêtir plusieurs interprétations et appellations (théorie des fonctions de croyance ou théorie de Dempster-Shafer). G. Shafer a montré l'intérêt des fonctions de croyance pour la modélisation de connaissances incertaines. L'utilité de ces fonctions, comme une alternative aux probabilités subjectives, a été démontrée plus tard de manière axiomatique par Ph. Smets [77] au travers du modèle des croyances transférables. Ce modèle fournit une interprétation non probabiliste cohérente et permet de clarifier le concept sous-jacent à la théorie. Dans cette partie, nous abordons quelques éléments mathématiques liés aux fonctions de croyance.

2.2.1 Représentation de l'information

Soit $\Omega = \{\omega_1, \dots, \omega_k, \dots, \omega_K\}$ un ensemble fini, généralement appelé cadre de discernement. Une fonction de croyance bel est une fonction de 2^Ω vers $[0,1]$ définie par :

$$bel(\emptyset) = 0$$

$$\forall n \geq 1, \forall i = 1, \dots, n, A_i \subseteq \Omega$$

$$bel\left(\bigcup_{i=1, \dots, n} A_i\right) \geq \sum_{I \subseteq \{1, \dots, n\}, I \neq \emptyset} (-1)^{|I|+1} bel\left(\bigcap_{i=1, \dots, n} A_i\right) \quad (4.1)$$

Le cas particulier où $n = 2$ et $A_1 \cap A_2 = \emptyset$ illustre le fait que la croyance attribuée à l'union de deux sous-ensembles disjoints de Ω est supérieure ou égale à la somme des croyances attribuées à chacun des sous-ensembles.

$$bel(A_1 \cup A_2) \geq bel(A_1) + bel(A_2) \forall A_1, A_2 \subseteq \Omega \text{ si } A_1 \cap A_2 = \emptyset \quad (4.2)$$

Si l'inégalité est transformée en égalité dans l'équation (4.1), la fonction résultante bel, appelée généralement fonction de croyance bayésienne, devient une fonction de probabilité classique. Une fonction de croyance peut également être définie mathématiquement par une fonction (ou allocation) de masse, notée m définie de 2^Ω dans $[0,1]$, qui vérifie :

$$\sum_{A \subseteq \Omega} m(A) = 1. \quad (4.3)$$

Chaque sous-ensemble $A \subseteq \Omega$ tel que $m(A) > 0$ est appelé élément focal de m . Ainsi, la masse $m(A)$ représente le degré de croyance attribué à la proposition A et qui n'a pas pu, compte tenu de l'état de la connaissance, être affectée à un sous-ensemble plus spécifique que A . Une fonction telle que $m(\emptyset) = 0$ est dite normale. Dans le modèle des croyances transférables, la condition :

$$\sum_{\phi \neq A \subseteq \Omega} = 1. \quad (4.4)$$

n'est pas supposée et $m(\emptyset) > 0$ est acceptée. Cette spécificité permet d'introduire la notion de monde ouvert en supposant que la croyance ne peut être placée sur un sous-ensemble de Ω . Dans ce contexte, l'ensemble vide \emptyset peut être interprété comme une hypothèse non clairement définie dans l'ensemble de discernement par opposition au monde fermé où l'ensemble des hypothèses est exhaustif. Etant donnée une fonction de masse m , on peut définir bel comme :

$$bel(A) = \sum_{B \subseteq A, B \neq \emptyset} m(B) \quad \forall A \subseteq \Omega. \quad (4.5)$$

Les fonctions m et bel représentent deux facettes de la même information. De plus, on peut retrouver m à partir de bel en utilisant la transformée de Möbius [78]. L'implémentation de cette transformation peut être obtenue par de simples calculs matriciels [79]. Issues des travaux de G. Shafer, les fonctions de croyance sont de nos jours reconnues pour la modélisation d'informations incertaines.

Une des principales difficultés consiste à modéliser la connaissance disponible en initialisant de manière adéquate les fonctions de croyance. L'ignorance partielle peut évidemment être codée elle aussi sous la forme d'une allocation de masse [80]. Modéliser le plus fidèlement l'information disponible confère aux fonctions de croyance l'avantage de ne pas introduire d'information a priori contrairement aux fonctions de probabilités. En effet, le concept d'équiprobabilité n'est pas identique au concept d'ignorance totale mais constitue déjà une forme précise de connaissance.

Dans le modèle des croyances transférables deux niveaux peuvent être distingués : le niveau crédal où les croyances sont modélisées et révisées, et le niveau pignistique dans lequel les fonctions de croyance sont transformées en fonctions de probabilité pour la prise de décision [81].

Lorsque la source d'information dont la fonction de croyance est extraite n'est pas totalement fiable, il est possible d'introduire une opération d'affaiblissement. Dans ce cas, un coefficient α , qui représente une sorte de métaconnaissance concernant la fiabilité de la source, permet de transférer une partie de la croyance vers l'ensemble Ω . Ainsi, une fonction de masse affaiblie, notée m_α , peut se déduire de m par :

$$m_\alpha(A) = \alpha m(A) \quad \forall A \subset \Omega, A \neq \Omega. \quad (4.6)$$

$$m_\alpha(\Omega) = 1 - \alpha + \alpha m(\Omega). \quad (4.7)$$

Ainsi, une fonction de croyance est capable de représenter plusieurs types de connaissance et constitue par ce biais un cadre riche et flexible pour la représentation

d'informations incertaines. Dans la section suivante, nous introduisons quelques-uns des outils qui permettent de manipuler l'information dans une chaîne de traitement.

2.2.2 Outils associés

Un des premiers outils que nous choisissons de décrire concerne la combinaison de deux fonctions de croyance. Supposons que nous disposons de deux allocations de masse m_1 et m_2 définies sur le même référentiel Ω . Ces deux fonctions peuvent être agrégées par un opérateur de combinaison conjonctif noté \cap . Le résultat de cette opération conduit à une fonction de croyance unique à laquelle correspond une fonction de masse, notée m_{\cap} , qui peut être définie par :

$$m_{\cap}(A) = (m_1 \cap m_2)(A) = \sum_{B \cap C = A} m_1(B)m_2(C) \quad \forall A \subseteq \Omega. \quad (4.8)$$

Cette règle conjonctive est parfois dénommée règle de combinaison de Dempster non normalisée. Si nécessaire, l'hypothèse de normalisation $m_{\cap}(\emptyset) = 0$ peut être retrouvée en divisant chaque masse par un coefficient adéquat. L'opérateur résultant, qui est connu sous le nom de règle de Dempster et dénoté \oplus , est défini par :

$$(m_1 \oplus m_2)(A) = \frac{(m_1 \cap m_2)(A)}{1 - m_{\cap}(\emptyset)} \quad \forall \emptyset \neq A \subseteq \Omega. \quad (4.9)$$

La quantité $m_{\cap}(\emptyset)$ est le degré de conflit entre les fonctions m_1 et m_2 et peut être calculée par:

$$m_{\cap}(\emptyset) = (m_1 \cap m_2)(\emptyset) = \sum_{B \cap C = \emptyset} m_1(B)m_2(C). \quad (4.10)$$

L'utilisation de la règle de Dempster est possible si les fonctions de masse m_1 et m_2 ne sont pas en conflit total c'est-à-dire s'il existe deux éléments focaux B et C respectivement de m_1 et m_2 tels que $B \cap C \neq \emptyset$. Par ce fait, cette règle, qui permet de combiner des informations incertaines extraites sous forme de fonctions de croyance, a été très discutée [82]. Elle a cependant été justifiée en introduisant des axiomes particuliers bien que certains auteurs ont envisagé d'autres solutions [83, 84] comme par exemple des règles disjonctives [85].

Dans le modèle des croyances transférables, la fonction de masse a une interprétation naturelle puisque la masse $m(A)$ représente le degré de croyance attribué à la proposition $A \subseteq \Omega$ qui n'a pas pu être assignée à un sous ensemble plus spécifique. Si maintenant, on apprend que la croyance allouée au sous-ensemble $B \subseteq \Omega$ est totale alors la masse $m(A)$ initialement allouée à A est transférée au sous-ensemble $A \cap B$ d'où le nom du modèle introduit par Ph. Smets. La fonction de masse résultante est celle obtenue par l'application de la règle dite de conditionnement qui n'est qu'un cas particulier de la règle de combinaison de Dempster. Ainsi, le résultat de la combinaison d'une fonction de masse m conditionnellement à B produit une fonction de croyance, notée $bel(.|B)$ qui peut se calculer par :

$$\forall A \subseteq \Omega, \quad bel(A|B) = bel(A \cap B) - bel(\overline{B}). \quad (4.11)$$

Au niveau pignistique, une fonction de croyance unique, sorte de résumé exhaustif de l'information disponible au niveau crédal, est utilisée pour la prise de décision. En basant son

raisonnement sur des arguments de rationalité développés dans le modèle des croyances transférables, Ph. Smets propose de transformer une fonction de masse m en une fonction de probabilité $BetP_m$ définie sur Ω (appelée fonction de probabilité pignistique) qui se formalise pour tout ω_k appartenant à Ω par :

$$BetP_m(\omega_k) = \frac{1}{1 - m(\emptyset)} \sum_{A \ni \omega_k} \frac{m(A)}{|A|}. \quad (4.12)$$

Où $|A|$ représente la cardinalité du sous-ensemble $A \subseteq \Omega$. Dans cette transformation, la masse de croyance $m(A)$ est uniformément distribuée parmi les éléments de A . A partir de cette distribution de probabilité, il est alors possible d'utiliser les outils classiques de la théorie de la décision statistique pour la prise de décision [81]. On suppose ainsi qu'il existe une fonction de coût $\lambda : A \times \Omega \rightarrow \mathbb{R}$ où A désigne un ensemble fini d'actions défini telle que $\lambda(a, \omega)$ quantifie le coût de choisir l'action a quand la vérité est ω . A partir de la probabilité $BetP_m$, on peut associer à chaque action $a \subseteq A$ un risque $BetR$ défini comme le coût espéré si l'on choisit l'action a lorsque la vérité est ω :

$$BetR(a) = \sum_{\omega \in \Omega} BetP_m(\omega) \lambda(a, \omega). \quad (4.13)$$

Il suffit alors de choisir l'action qui minimise ce risque, aussi appelé risque pignistique. D'autres règles de décision basées sur les fonctions de croyance elles-mêmes sont néanmoins envisageables [86, 87].

5 Conclusion

Dans ce chapitre, une introduction à la théorie des graphes a été présentée avec les différentes définitions nécessaires pour leur utilisation. Ensuite, un état de l'art sur l'application des fonctions de croyance à la gestion d'informations imparfaites a été présenté. Les principaux outils associés à cette théorie ont été brièvement présentés. En brochant un état de l'art, il a été montré que ce cadre théorique est suffisamment souple pour être appliqué dans de nombreux domaines applicatifs des sciences et technologies de l'information et de la communication tels que l'analyse de données, le diagnostic, l'aide à la décision, la perception multicapteurs et le traitement d'images. Enfin, en ce qui concerne les applications de la théorie des fonctions de croyance, il semble que les perspectives d'utilisation peuvent être nombreuses.

Chapitre V

Analyse d'une scène par graphe probabiliste évolutif avec raffinement par la théorie de l'évidence

1 Introduction

La segmentation d'image est un prétraitement qui améliore l'état des informations contenues dans l'image avant le traitement souhaité pour l'application désirée. L'objectif est de séparer le plus fidèlement possible les objets et le fond qui composent l'image [88, 89].

La segmentation de l'image possède des applications dans de nombreux domaines pratiques, elle a des débouchés dans la reconnaissance de motifs, la détection d'objets, l'analyse d'image médicale, la robotique, ou encore dans le domaine des images par satellites et bien d'autres encore.

Plusieurs techniques de segmentation développées se basent sur un nombre de classes prédéfini dans l'étape initiale de l'algorithme, qui s'occupe de la classification de chaque pixel de l'image dans sa classe la plus probable. Segmentation par région, la segmentation par contour, segmentation par seuils, et celle basée sur la méthode des k-means. Classification par la théorie de l'évidence, appelée aussi théorie de Dempster-Shafer ou théorie des fonctions de croyances, elle permet de traiter d'une part des informations incertaines et d'autre part de combiner des informations provenant de plusieurs sources, avant d'appliquer des règles de décision pour la sélection de la classe d'affectation [90, 91]. Classification par les chaînes de Markov cachées [92, 93]. Classification bayésienne qui se base sur la détermination des probabilités conditionnelles pour estimer l'appartenance d'un individu à une classe, classification floue [94].

Au cours de ce chapitre on présentera une nouvelle méthode de segmentation non supervisée d'image basée sur l'évolution d'un graphe qui traduit les regroupements des occupants d'une image [100, 101, 102]. Cette méthode est originale, elle se base sur deux phases. Dans un premier lieu, la méthodologie, qui s'articule autour de deux phases principales, est représentée. L'étape de création du graphe des classes est tout d'abord réalisée, l'astuce dans cette phase est de repousser la prise de décision à la fin de la traversée d'une région. Au terme de cette étape, la frontière de chaque région est imprécise. L'étape suivante consiste à tracer la carte des zones de conflit qui décèlent la présence de contour ou de bruit. Le rôle de cette phase est de bien déceler les frontières de chaque région. L'originalité dans cette phase est la procédure de raffinement des zones détectées dans l'image. La dernière étape traite la phase finale de la segmentation. Dans cette partie, l'image des classes est établie à partir du graphe et de la carte des zones de conflit. En fin, des résultats de segmentation sur quelques images seront présentés.

2 Méthodologie générale et support technique

Le principe de cette analyse se divise en deux étapes principales. La première phase consiste à déterminer le nombre de classes contenues dans l'image et de localiser spatialement les régions pures. La méthode repose sur le parcours de l'image de la scène et d'établir un graphe probabiliste qui évolue le long du parcours selon les informations récoltées. A la fin du parcours, le graphe résultant représente un descripteur des composantes de la scène à analyser. Chaque nœud du graphe représente une classe. Une arête décrit le voisinage entre deux classes. La deuxième phase consiste à utiliser la méthode des fonctions de croyance pour classer les pixels des régions à forte variance et qui constituent les zones de bruit et des transitions. Ces zones peuvent altérer significativement le résultat du traitement. Le travail est développé sur PC de vitesse 3GHz.

3 Etablissement du graphe

La méthode consiste à détecter les régions pures de l'image qui représentent une faible perturbation et de générer un graphe qui évolue selon la présence de ces régions. Les zones de forte variance sont omises et laissées pour la deuxième phase du traitement. La première phase de la segmentation regroupe les étapes suivantes :

3.1 Parcours de l'image

Afin d'éviter le saut dans le balayage qui peut engendrer un changement de région sans détection de transition, cette situation indésirable peut provoquer un déroutement de l'algorithme, dans ce cas deux régions distinctes seront considérées appartenir à la même classe. Afin d'éviter ce genre de situations, un chemin qui serpente l'image est choisi.

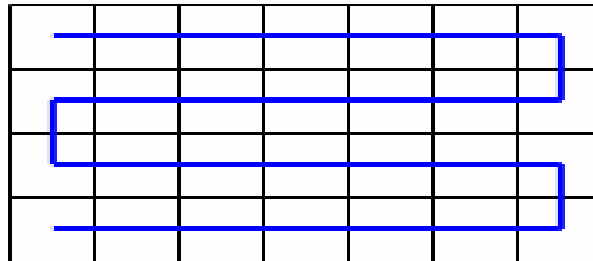


Fig.5.1 : Balayage de l'image

3.2 L'état initial

La région de départ est représentée par le nœud unique qui forme le graphe dans son état initial. A chaque pas de parcours, et afin de mesurer l'état du chemin, un vecteur d'attributs caractérisant le site et son voisinage est calculé. Il est composé du niveau de gris du pixel centre, la moyenne et l'écart type de la fenêtre de voisinage.

3.3 Détection d'une transition

En présence d'une fluctuation relativement importante de l'attribut écart type, l'appartenance du voisinage du site actuel à un contour indiquant le franchissement d'une frontière d'une nouvelle région devient très probable. Un seuil est prédéfini pour indiquer l'approche d'une transition.

3.4 Première transition

Afin de minimiser la probabilité d'erreur de décision, la première transition d'une frontière sera mémorisée comme un contour limite amont. La prise de décision concernant l'affectation de la présente portion de région incluse entre deux frontières sera faite en se basant sur les informations fournies par l'ensemble des pixels visités le long de la portion, et cela après la détection du contour limite aval.

3.5 Deuxième transition

Après la détection du deuxième contour, la portion de région parcourue est identifiée et une décision est prise soit pour affecter la portion à une classe déjà existante, soit pour la création d'un nouveau nœud dans le graphe qui traduit la présence d'une nouvelle classe.

3.6 Création d'un nouveau nœud

La classification de la région \mathbf{R} parcourue et située entre deux transitions se fait après le calcul des probabilités d'appartenance. L'univers de discrimination \mathbf{H} est constitué des hypothèses \mathbf{H}_i d'appartenance à la classe i déjà identifiée à ce stade du parcours et de l'hypothèse \mathbf{H}_0 de la non appartenance. La probabilité d'appartenance de la région à classifier \mathbf{R} à la classe i est calculée par :

$$pr(H_i) = \alpha_i e^{-\beta_i d_i} \quad i = 1, \dots, k. \quad (5.1)$$

Où k est le nombre de classes décelé à ce niveau de traitement, α_i est une constante et β_i est l'inverse de la distance maximale entre le centre de la région \mathbf{R} et les centres des classes existantes. d_i est la distance entre le centre de \mathbf{R} et le centre de la classe \mathbf{C}_i . La probabilité attribuée à l'hypothèse \mathbf{H}_0 qui stipule l'éloignement de la région \mathbf{R} des classes existantes est donnée par :

$$pr(H_0) = 1 - \max_i \{pr(H_i)\} \quad i = 1, \dots, k. \quad (5.2)$$

Avec $i=1, 2, \dots, k$, k étant le nombre de classes décelées avant la prise de décision de fusion ou de création d'une nouvelle classe. Le verdict d'affectation de la région \mathbf{R} obéit au critère de la probabilité maximale :

$$H^* = \max_i \{pr(H_i)\} \quad i = 0, 1, \dots, k. \quad (5.3)$$

La région est fusionnée avec l'ensemble des pixels de la classe de l'hypothèse élue. Dans le cas où l'hypothèse \mathbf{H}_0 possède la plus grande probabilité, une nouvelle classe est créée et un nouveau nœud dans le graphe apparaît.

3.7 Matrice des classes

La matrice des classes $\mathbf{M}(m \times n \times l)$, présentée par l'image de la figure 5.2, est constituée des l matrices $\mathbf{M}_k(m \times n)$ des cartes des classes, m et n étant les dimensions de l'image et l le nombre de classes. Chaque carte représente un nœud du graphe. De cette manière la classe peut être décomposée en sous classes ou régions qui la constituent. La carte $\mathbf{M}_i(m \times n)$ de la classe i est formée par l'agrégation des pixels dans le plan i , de la matrice des classes \mathbf{M} , après chaque identification d'une appartenance à la classe i . Les régions \mathbf{R}_i^j de chaque classe i sont déterminées au niveau de chaque plan de la matrice par la localisation des regroupements des pixels qui sont spatialement séparés.

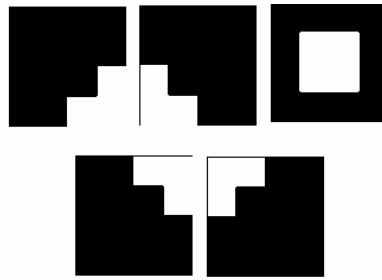


Fig.5.2 : Matrice des régions d'une image synthétique.

3.8 Carte des contours

La carte des contours figure 5.3 est un plan sur lequel sont tracées les courbes de séparation des différentes régions qui constituent l'image. Les contours sont les lieux de variations significatives de l'information. De plus, la transition étant stricte, un contour doit être une chaîne de pixels d'épaisseur 1. Cette restriction sur la nature du contour est imposée dans le but de bien séparer les régions tout en préservant les formes qui sont probablement les plus proches de la scène prise en image.

La transition d'une région à une autre est décelée par une variation relativement importante. La probabilité de présence d'un contour croît si le gradient est localement maximum ou si la dérivée seconde présente un passage par zéro. En premier lieu, deux plans sont construits, le plan de la norme du gradient qui exprime le taux de variation dans l'image. Cette amplitude est donnée par :

$$G(s) = \max_{d_i} [abs(s^+ - s^-) + (abs(s^+ - s))] \quad (5.4)$$

Cette formule de calcul attribue au site frontière le maximum de variation et élimine la situation des deux raies au bord d'une transition issue d'un calcul de gradient normal.

Le deuxième plan mémorise la direction du maximum local de la variation. Cette direction est déterminée par la localisation du minimum de variation de G qui représente la direction de passage du contour. La direction recherchée est perpendiculaire à la direction du contour.

$$d_i = i \pm 2 \quad / G \text{ est minimale} \quad (5.5)$$

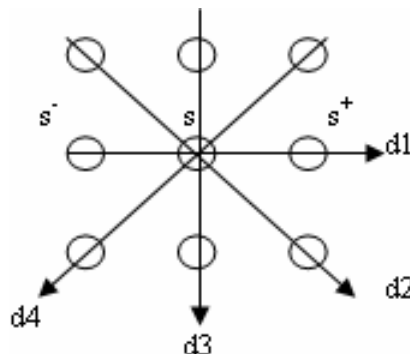


Fig.5.3 : Choix des directions

En deuxième lieu, la carte des contours est établie à partir de l'image des variations et de la matrice des directions. Le site observé est pris pour un point de contour dans le cas où il présente un maximum local dans la direction de la variation maximale.

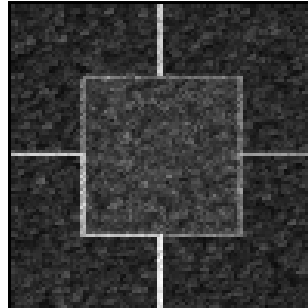


Fig.5.4 : Carte des contours d'une image synthétique.

3.9 L'image segmentée

L'image segmentée est éditée à partir du graphe des régions et de la carte des zones de transition où la région est peuplée d'individus hétérogènes qui présentent des attributs de classes différentes.

4 Raffinement des extrémités (Classification dans les régions de conflit)

L'image segmentée, issue des différentes étapes citées précédemment, représente les régions de l'image qui sont peuplées par des individus présentant un certain degré d'homogénéité satisfaisant un critère sévère. L'image ainsi établie trace les zones pures, les sites de conflit sont laissés sans classification. De cette manière la première phase de la segmentation et la détermination du nombre de classes sont basées sur un traitement qui ne touche que la population localement pure de l'image. Pour la classification des sites impurs, une méthode probabiliste appelée théorie de l'évidence est utilisée, cette méthode est basée sur un traitement des sources d'information locales du site à classifier.

4.1 Théorie de l'évidence

Avant d'appliquer cette procédure, nous rappelons d'une façon succincte quelques concepts de base de la théorie des fonctions de croyance et de la partition crédale. Introduite par Dempster [95], la théorie de l'évidence fut reprise par Shafer [96], il a montré l'intérêt des fonctions de croyance pour la modélisation de connaissances incertaines. L'utilité des fonctions de croyance, comme alternative aux probabilités subjectives, a été démontrée plus tard de manière axiomatique par Smets [98, 99] au travers du Modèle de Croyances Transférables. Cette théorie peut être considérée comme une généralisation de la théorie des probabilités. Deux niveaux de traitement de l'information sont utilisés, le niveau crédal où les croyances sont manipulées et le niveau pignistique pour la prise de décision.

4.2 Niveau crédal (Modélisation de la connaissance)

Dans le cadre de la théorie de l'évidence, l'ensemble $\Omega = \{\omega_1, \dots, \omega_k, \dots, \omega_K\}$ est un ensemble fini, généralement appelé cadre de discernement. On suppose que les hypothèses dans Ω sont exclusives et que le cadre de discernement est exhaustif. On définit une masse de probabilité

élémentaire, appelée masse de croyance, qui caractérise la véracité d'une proposition A pour une source d'information S . La masse m associée à cette source est alors définie par :

$$m : 2^\Omega \rightarrow [0,1] \quad (5.6)$$

et vérifie les propriétés suivantes :

$$m(\emptyset) = 0 \quad (5.7)$$

$$\sum_{A \subseteq \Omega} m(A) = 1 \quad (5.8)$$

Cette fonction se différencie d'une probabilité par le fait que la totalité de la masse de croyance est répartie non seulement sur les hypothèses singletons ω_q , mais aussi sur les hypothèses combinées. La modélisation issue de la fonction m est appelée jeu de masses. A partir de la fonction m , on définit respectivement les fonctions de crédibilité Cr et de plausibilité Pl par :

$$Cr(A) = \sum_{B \subseteq A} m(B) \quad (5.9)$$

$$pl(A) = \sum_{(B \cap C) \neq \emptyset} m(B) = 1 - Cr(\bar{A}) \quad (5.10)$$

Où \bar{A} représente l'événement contraire de la proposition A . La crédibilité $Cr(A)$

mesure la force avec laquelle on croit en la véracité de la proposition. La plausibilité $Pl(A)$, fonction duale de la crédibilité, mesure l'intensité avec laquelle on ne doute pas de A . Deux fonctions de masse initiale m_1 et m_2 représentant les informations respectives de deux sources distinctes, peuvent être fusionnées en utilisant la loi de combinaison de Dempster. Celle-ci, qui s'avère commutative et associative, est définie par :

$$(m_1 \oplus m_2)(A) = \frac{1}{1-k} \sum_{B \cap C \neq \emptyset} m_1(B)m_2(C) \quad (5.11)$$

Où k est défini par :

$$k = \sum_{B \cap C = \emptyset} m_1(B)m_2(C) \quad (5.12)$$

Le coefficient k reflète le conflit existant entre les deux sources S_1 et S_2 et le quotient

$\frac{1}{1-k}$ est un terme de normalisation.

4.3 Niveau pignistique (Prise de décision)

Pour la prise de décision, l'étape d'agrégation précédemment définie permet d'obtenir un résumé exhaustif de l'information sous forme d'une fonction de croyance unique m qui est utilisée pour la prise de décision. Après combinaison, il reste à prendre une décision quant à l'hypothèse la plus probable de Ω . La règle de décision parmi les plus utilisées reste celle du maximum de probabilité pignistique. Cette règle de décision présentée par Smets [99] utilise la transformation pignistique, qui permet de transformer m en une fonction de probabilité P_{Bet} sur Ω , appelée fonction de probabilité pignistique, définie pour tout $\omega \in \Omega$ comme :

$$BetP(\omega) = \sum_{\omega \in A \subseteq \Omega} \frac{m(A)}{|A|}, \quad \forall \omega \in \Omega \quad (5.13)$$

Où $|A|$ est le cardinal de $A \subseteq \Omega$. Dans cette transformation la masse de croyance $m(A)$ est uniformément distribuée parmi les éléments de A. La décision penche alors vers l'élément de Ω dont cette valeur est la plus grande :

$$\omega^* = Arg \left\{ \max_{\omega \in \Omega} [BetP(\omega)] \right\} \quad (5.14)$$

5 Application à la classification dans les zones de conflit

L'objectif, dans cette phase, est de raffiner les frontières des composants de l'image en classifiant les pixels des régions limites des classes qui sont maculés par le bruit de la transition, et cela en appliquant la théorie des fonctions de croyance.

La décision de la classification du pixel P est établie en combinant n fonctions m_i . Le voisinage $V(p)$ du pixel P est pris comme sources d'information, chaque masse m_i est relative à un pixel $P_i \in V(P)$ et définie sur $\Omega = \{\omega_1, \dots, \omega_k, \dots, \omega_K\}$. Ce cadre de décernement reflète les classes des pixels qui composent le voisinage $V(P)$. A chaque sous-ensemble ω_k de Ω correspond une classe C_k . On définit la fonction de croyance relative au pixel P_i de classe C_k comme suit :

$$m_i(\{\omega_k\}) = \alpha \quad (5.15)$$

$$m_i(\Omega) = 1 - \alpha \quad (5.16)$$

Avec le paramètre $\alpha \in [0,1]$. La fonction de croyance du voisinage est obtenue par la combinaison conjonctive du jeu de masse m_i du voisinage.

$$m = \bigcap_i^n m_i \quad (5.17)$$

Etant donné que chacune des fonctions m_i ne possède que deux éléments focaux qui sont le singleton ω_k et Ω , la fonction de croyance m est déterminée par :

$$m(\{\omega_k\}) = \alpha(1 - \alpha)^{n-n_k} - (1 - \alpha)^n \quad (5.18)$$

$$m(\Omega) = (1 - \alpha)^n \quad (5.19)$$

Avec n^k est le nombre de pixels appartenant à la classe C_k dans le voisinage du pixel P considéré.

La masse attribuée à l'ensemble vide $m(\emptyset)$ n'est pas négligeable, il est nécessaire d'analyser cette répartition du conflit.

6 Application de la méthode à des images à niveaux de gris

Une image à niveau de gris est une projection d'une scène tridimensionnelle dans un plan bidimensionnel, prise par un capteur à niveaux de gris (une bande). Dans une image à niveaux de gris, des phénomènes physiques, géométriques et photométriques dans la scène engendrent des variations dans l'image. Ces variations sont appelées contours et forment les frontières des objets qui composent la scène à analyser dans le but d'en tirer les informations nécessaires à l'accomplissement de la tâche prédéfinie et pour laquelle l'image a été prise.

L'image à niveau de gris réelle est une représentation de la scène tridimensionnelle captée par une caméra et transformée par la chaîne d'acquisition à une matrice bidimensionnelle composée d'éléments représentatifs des pixels à niveau de gris.

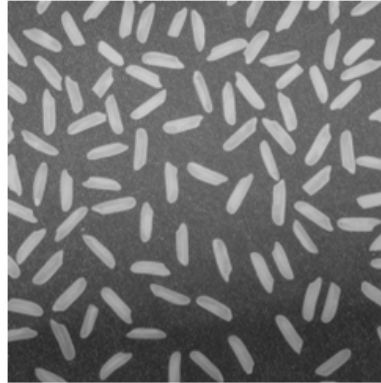


Fig.5.5 : Image réelle à segmenter

La figure 5.5 présente une image réelle à niveaux de gris, elle est utilisée pour tester l'efficacité de la méthode vis-à-vis à des scènes réelles, et cela par une appréciation visuelle du résultat de la décomposition de l'image.

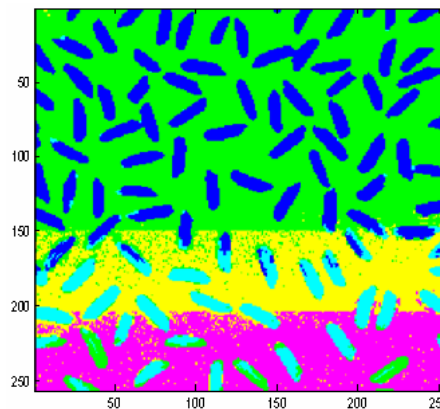


Fig.5.6 : L'image segmentée

La figure 5.6 présente le résultat de la segmentation, on remarque la dégradation de l'illumination de la scène traduite par la détection des trois régions de fond de l'image segmentée. Les grains sont bien décelés et reproduits dans l'image segmentée.

7 Application de l'algorithme à des images couleurs

On a appliqué notre algorithme de segmentation sur des images en couleurs synthétiques pour vérifier l'efficacité de la méthode. On a testé notre algorithme sur des images réelles en couleurs dans le but de voir le comportement de la méthode en face de scène en couleurs réelles.

7.1 Segmentation d'une image couleur RVB synthétique bruitée avec un bruit Gaussien $N(0,0.01)$

Pour cette application on a utilisé un seuil de création d'un nouveau nœud $ds = 0.7$, qui correspond à la probabilité de la véracité de l'hypothèse la dernière région parcourue n'appartenir à aucune classe déjà décelée.

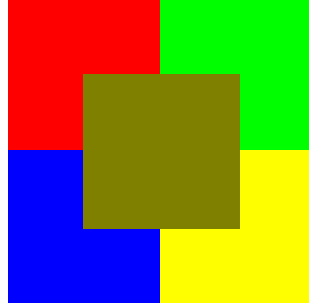


Fig.5.7 : L'image synthétique pure à segmenter.

La figure 5.7 représente une image en couleurs synthétique réalisée dans le but de tester notre méthode sur une image connue. Elle est composée de cinq régions dont les caractéristiques sont résumées dans le tableau.

Région	Région 1	Région 2	Région 3	Région 4	Région 5
Composantes RVB	(1, 0, 0)	(0, 1, 0)	(0.5, 0.5, 0)	(0, 0, 1)	(1, 1, 0)
Nbr pixels sans raff	1543	1632	2303	1563	1697
Nbr pixels avec raff	1680	1751	2597	1751	1825

Table1 : Caractéristiques de l'image synthétique

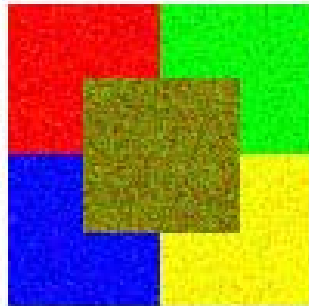


Fig.5.8 : L'image synthétique bruitée.

La figure 5.8 représente l'image synthétique de la figure.1 bruitée avec un bruit gaussien, ce bruit suit la loi normale $N(0, 0.01)$.

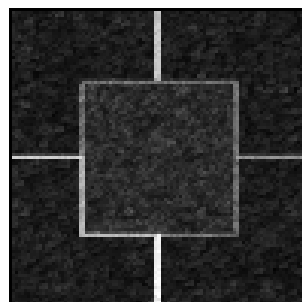


Fig.5.9: L'image des zones de conflit.

La figure 5.9 illustre l'image des zones de conflit, c'est des régions qui révèlent l'existence des écarts appréciables entre les attributs des occupants de ces zones. On remarque que ces zones tracent les frontières des régions.

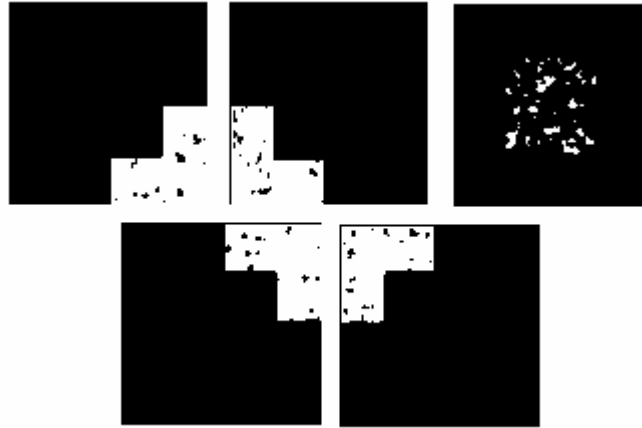


Fig.5.10 : l'image des classes avant raffinement

Dans la figure 5.10 l'image des classes identifiées est présentée, on aperçoit les régions construits dans l'image synthétique d'origine avec l'absence des pixels qui présentent des écarts satisfaisant la contrainte d'omission. Ces zones sont révélées par la présence des taches noires dans les zones blanches.

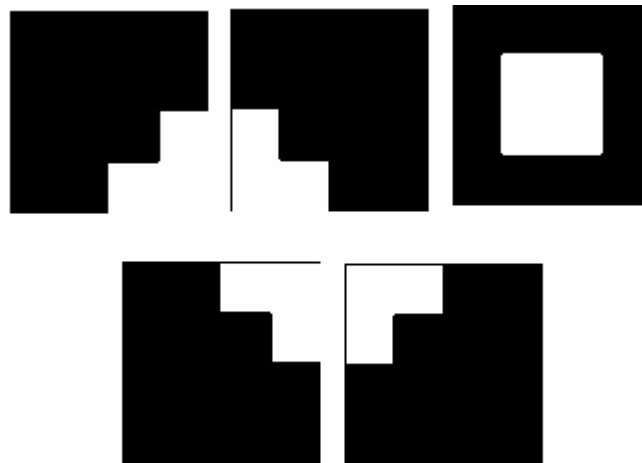


Fig.5.11 : L'image des classes après raffinement.

La figure 5.11 illustre l'image des classes après l'application de la procédure de raffinement, on peut observer la netteté des régions et l'absence des taches noires.

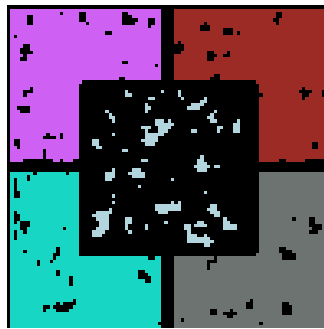


Fig.5.12 : L'image segmentée sans raffinement.

L'image de la figure 5.12 présente le résultat de la segmentation sans raffinement, on peut voir les cinq classes avec la présence des zones noires qui révèlent les sites des pixels non classifiés.

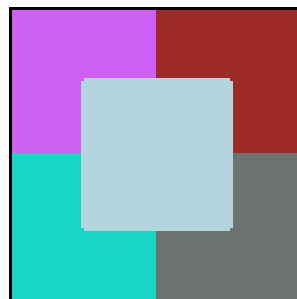


Fig.5.13 : L'image segmentée avec raffinement.

L'image de la figure 5.13 présente le résultat de la segmentation après raffinement, on peut voir clairement la forme et l'endroit de chaque région des cinq classes de l'image synthétique d'origine.

Région	Région 1	Région 2	Région 3	Région 4	Région 5	Région 6
Composantes RVB	(1, 0, 0)	(0, 1, 0)	(0.5, 0.5, 0)	(0, 0, 1)	(1, 1, 0)	(0.6, 0.4, 0)
Nbr pixels sans raff	1543	1619	2294	1563	1691	2
Nbr pixels avec raff	1687	1761	2558	1751	1825	22

Table 2 : Les caractéristiques des classes après la segmentation avec $ds = 0.8$

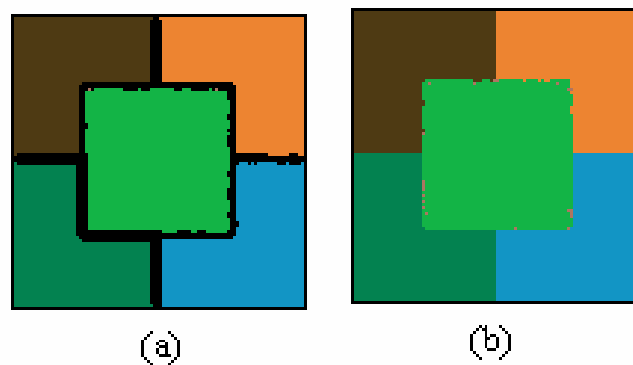


Fig.5.14 : Image segmentée avec $ds = 0.8$, $k=6$: (a) sans raffinement, (b) avec raffinement

La figure 5.14 illustre l'image segmentée sans raffinement (a) et l'image segmentée avec raffinement avec un seuil de création d'un nouveau nœud, donc une classe, de 0.8. Le seuil étant plus élevé provoque l'augmentation du nombre de classes qui atteint 6 classes.

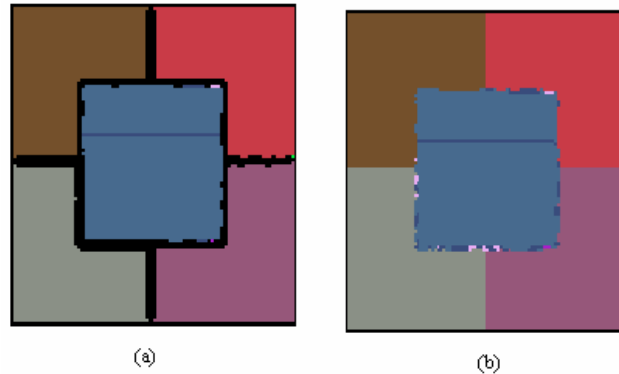


Fig.5.15 : Image segmentée avec $ds = 0.9$, $k = 9$: (a) sans raffinement, (b) avec raffinement

La figure 5.15 montre l'image segmentée sans raffinement (a) et l'image segmentée avec raffinement avec un seuil de création d'une nouvelle classe de 0.9. Le choix d'un seuil proche de 1 augmente considérablement le nombre de classes qui atteint 31 classes.

7.2 Segmentation d'une image couleur synthétique bruitée avec un bruit Gaussien $N(0,0.02)$

Pour cette application le seuil de création d'un nouveau nœud est $ds = 0.7$.

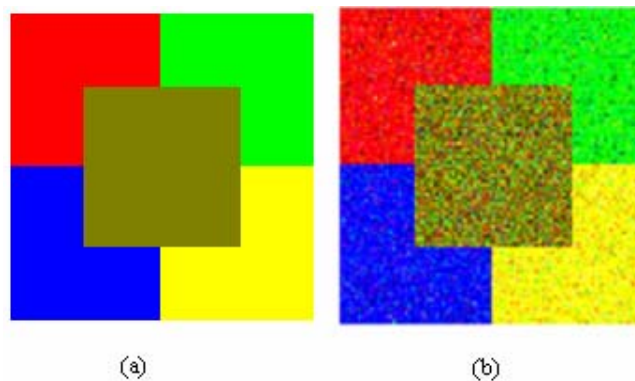


Fig.5.16 : L'image synthétique à segmenter avec $ds = 0.7$: (a) L'image synthétique pure.
(b) L'image bruitée avec un bruit Gaussien $N(0,0.02)$.

La figure 5.16 montre l'image synthétique pure en (a) et l'image synthétique bruitée avec un bruit Gaussien qui suit la loi $N(0,0.02)$ en (b). Dans cette image le bruit est plus important que le bruit de l'image précédente.

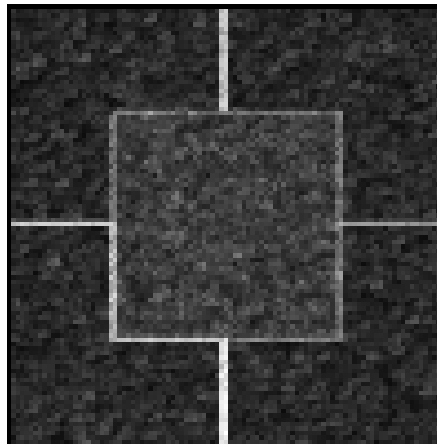


Fig.5.17 : L'image des zones de conflit.

Cette figure 5.17 illustre les zones de conflit de l'image bruitée, on remarque une surface des zones non noires plus importante que celle de la figure 3.

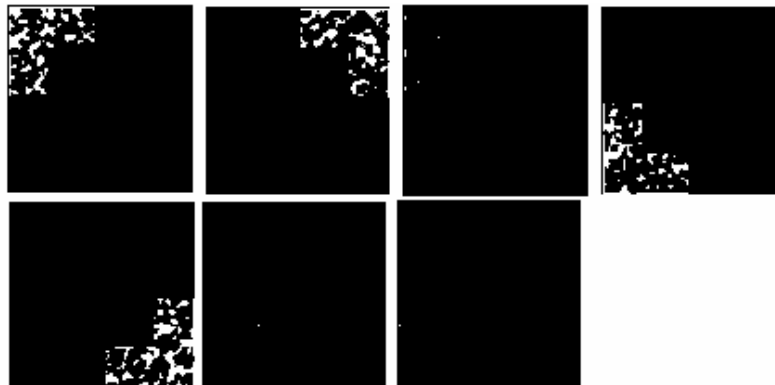


Fig.5.18 : l'image des classes avant raffinement.

Cette figure montre les images des différentes classes décelées avant l'application de la procédure de raffinement. Le nombre de classes est 7. Le nombre de classe a augmenté à cause du bruit.

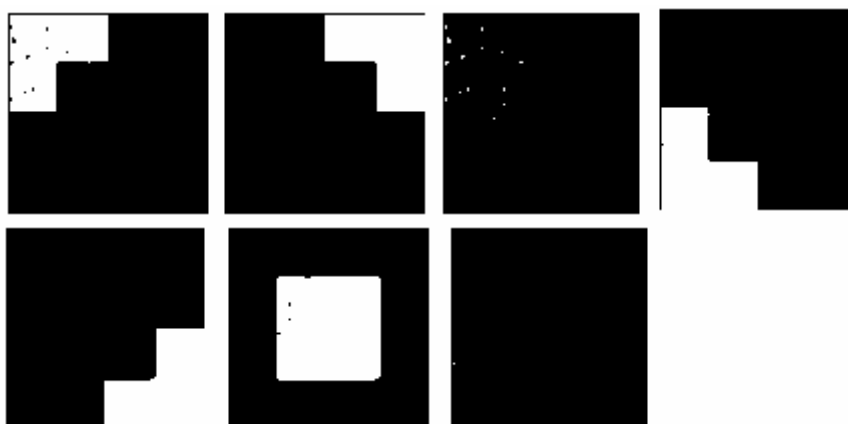


Fig.5.19 : L'image des classes après raffinement.

La figure 5.19 illustre les images des différentes classes décelées après l'application de la procédure de raffinement. Le nombre de classes à nombre de pixels important est 5. Le nombre de classe total est 7.

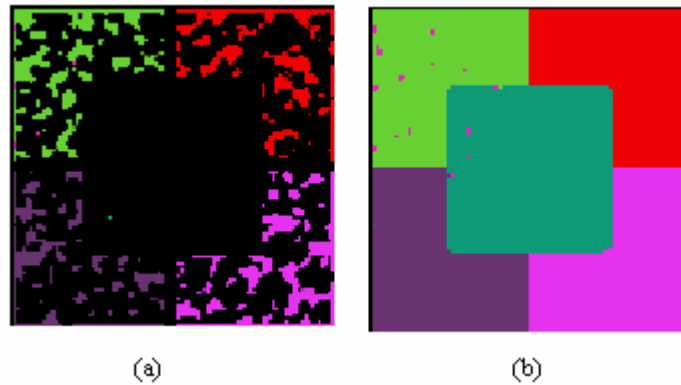


Fig.5.20 : L'image segmentée : (a) sans raffinement. (b) avec raffinement.
 La figure 5.20 montre l'image segmentée avec un seuil de création d'un nouveau nœud $ds = 0.7$. La figure (a) illustre l'image segmentée avant la procédure de raffinement. La figure (b) montre l'image segmentée avec application de la procédure de raffinement.

Nombre de classe $k=7$:

Classe	Composante R	Composante V	Composante B	Nombre de pixels
1	0.9631	0.0448	0.0431	475
2	0.0425	0.9589	0.0445	444
3	0.7829	0.1216	0.0370	8
4	0.0437	0.0530	0.0370	390
5	0.9590	0.9585	0.0474	519
6	0.4452	0.6212	0	1
7	0.0909	0.4485	1	1

Table 3 : Caractéristiques des classes de l'image segmentée

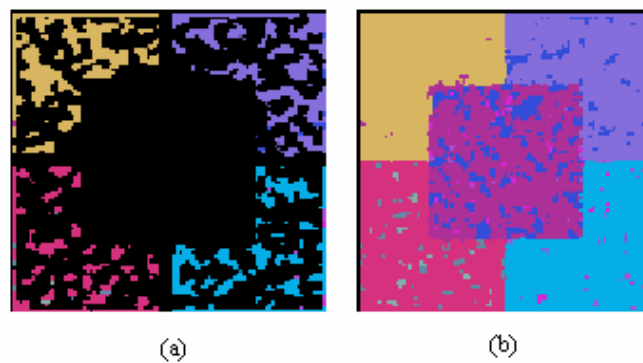


Fig.5.21 : Image segmentée avec $ds = 0.8$, $k=9$: (a) sans raffinement. (b) avec raffinement.

La figure 5.21 représente l'image synthétique segmentée avec un seuil de création d'un nouveau nœud $ds = 0.8$. Le nombre de classes détectées est $k = 9$, il a augmenté de 2. L'image segmentée sans raffinement (a) et après l'utilisation de la procédure de raffinement (b).

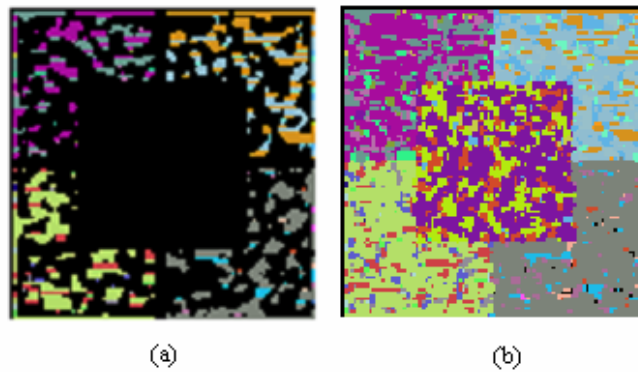


Fig.5.22 : Image segmentée avec $ds = 0.9$, $k=30$: (a) segmentation sans raffinement.
(b) segmentation avec raffinement.

La figure 5.22 représente l'image synthétique segmentée avec un seuil de création d'un nouveau nœud $ds = 0.9$. Le nombre de classes détectées est $k = 30$, le nombre de classe a augmenté considérablement à cause de la sévérité dans l'hypothèse de création d'un nouveau nœud dans le graphe. La figure illustre l'image segmentée sans raffinement (a) et après l'utilisation de la procédure de raffinement (b).

7.3 Segmentation d'image en couleurs réelle avec $ds=0.7$

L'image réelle est souvent très riche en informations visuelles, c'est une source d'informations très importante. Segmenter une image réelle pour en tirer les informations désirées est une tâche délicate et dépend énormément de l'objectif désiré dans l'acquisition de l'image de la scène filmée.



Fig.5.23 : Image réelle d'un puzzle à segmenter

Pour tester notre algorithme, on choisie une image réelle en couleurs d'un puzzle figure. C'est une image riche en couleurs.

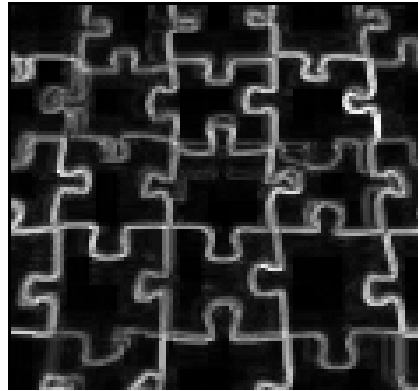


Fig.5.24 : L'image des zones de conflit.

L'image de la figure 5.24 révèle les sites de conflit de l'image à segmenter. Elle trace les courbes des frontières des objets contenus dans l'image.

Classe	Composante R	Composante V	Composante B	Nombre de pixels
1	0.9	0.4	0.2	1415
2	1	0.9	0.4	1249
3	0.5	0.8	0.9	566
4	1	0.7	0.7	391
5	0.9	0.9	0	565
6	1	0.3	0.4	690
7	0.3	0.7	0.2	870
8	0.1	0.5	0.8	1849
9	0	0	0	535
10	0.4	0.8	0.8	302
11	0.1	0.2	0.3	5

Table 4 : Caractéristiques de classes de l'image segmentée

Après la segmentation, l'image résultante est construite par la composition de 11 classes dont les caractéristiques sont données dans le tableau.

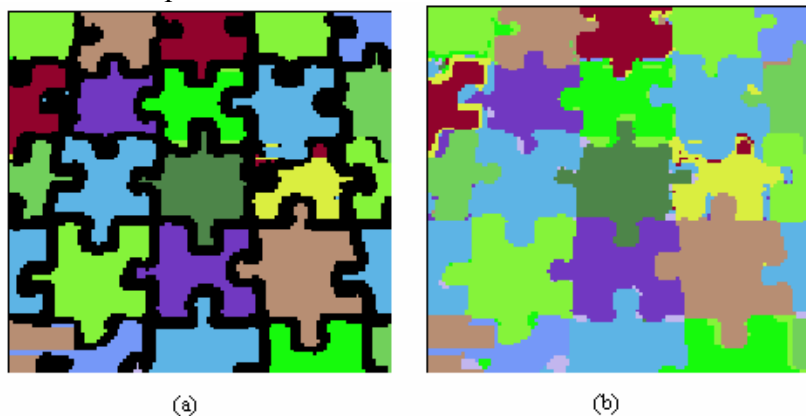


Fig.5.25 : L'image segmentée avec $ds = 0.7$: (a) Segmentation sans raffinement. (b) Segmentation avec raffinement.

La figure 5.25 illustre les images du résultat de la segmentation de l'image du puzzle. L'image (a) représente le résultat de la segmentation avant raffinement. L'image (b) montre le résultat de l'image segmentée avec raffinement.

7.4 Segmentation d'image réelle avec $ds=0.8$

Dans le but de rendre le critère de création d'un nouveau nœud dans le graphe plus sévère, on augmente le seuil à 0.8. Le résultat de la segmentation révèle un nombre de classe $k=19$, c'est un nombre plus élevé qui se traduit par l'augmentation du nombre de détails dans l'image segmentée.

Classe	Composante R	Composante V	Composante B	Nombre de pixels
1	1	0.3	0.3	640
2	1	0.9	0.4	1249
3	0.6	0.9	1	303
4	0.9	0.4	0.2	1138
5	1	0.8	0.8	155
6	0.4	0.8	0.8	559
7	0.9	0.9	0	563
8	0.9	0.3	0.4	306
9	0.3	0.7	0.2	870
10	0	0.5	0.8	1133
11	0.1	0.7	0.7	7
12	0	0	0	535
13	0.2	0.5	0.8	713
14	0.8	0.3	0.2	14
15	0.4	0.8	0.6	2
16	1	0.6	0.6	236
17	0.2	0.1	0.2	2
18	0.8	0.6	0.1	2
19	0	0.2	0.3	3

Table 5 : Les caractéristiques des classes de la segmentation avec $ds=0.8$

On remarque dans le tableau 5 un certain nombre de classes qui possèdent une population très faible comme la classe 11, 17, 18 et 19.

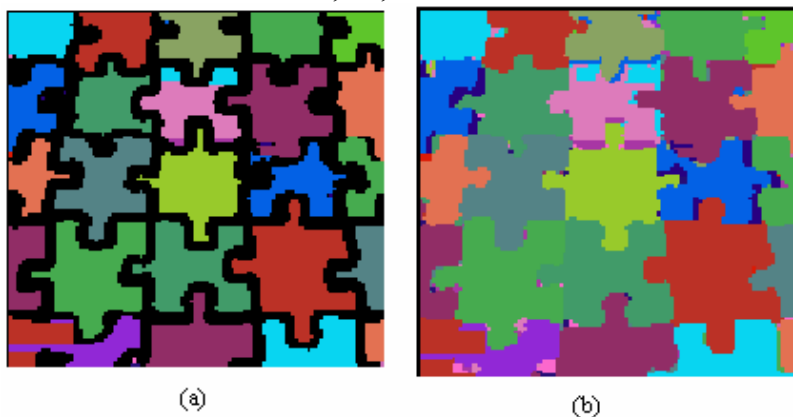


Fig.5.26 : L'image segmentée avec $ds = 0.8$: (a) Segmentation sans raffinement.
(b) Segmentation avec raffinement.

7.5 Segmentation d'image réelle avec $ds=0.9$

Pour un seuil de création d'un nouveau nœud dans le graphe très sévère, au voisinage de 1, le nombre de classe de vient important. Pour une segmentation avec $ds = 0.9$ on obtient un nombre de classes $k = 38$.

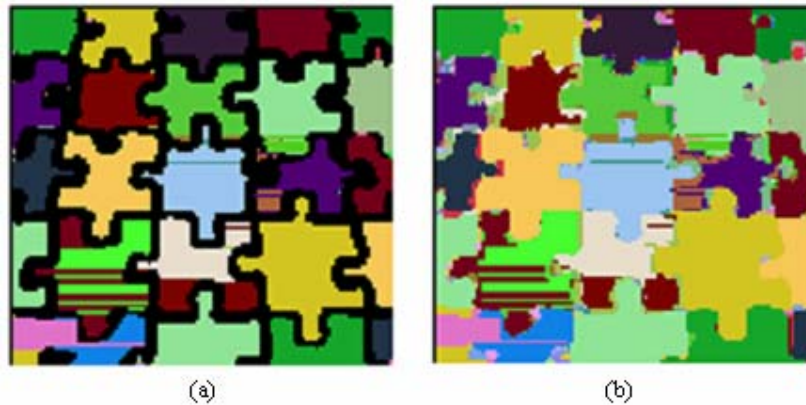


Fig.5.27 : L'image segmentée avec $ds = 0.9$: (a) Segmentation sans raffinement.
(b) Segmentation avec raffinement

La figure 5.27 illustre le résultat de segmentation avec $ds = 0.9$, on remarque l'apparition d'un nombre important de détails dans l'image.

8 Segmentation d'image synthétique riche en couleurs

La figure 5.28 présente une image riche en couleurs, elle représente un damier multicolore.

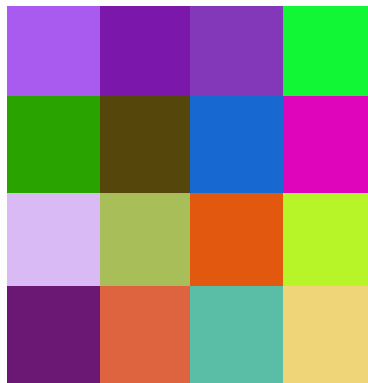


Fig.5.28 : Image synthétique à segmenter

La figure 5.29 illustre l'image des zones de conflit.

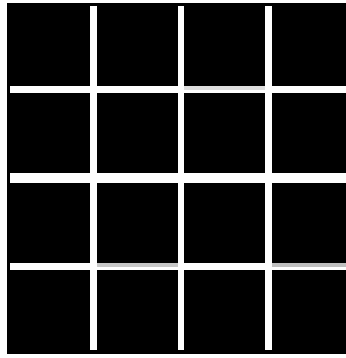


Fig.5.29 : Image des zones de conflit



Fig.5.30 : Image segmentée sans raffinement

L'image de la figure 5.30 montre le résultat de la segmentation avant le raffinement.

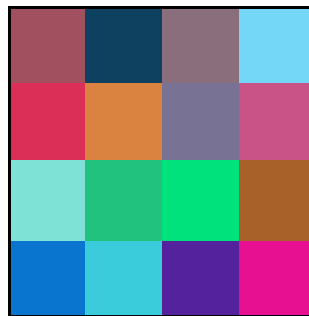


Fig.5.31 : Image segmentée avec raffinement

La figure 5.31 présente l'image segmentée avec raffinement. Le nombre de classe détecté est $k=16$, c'est le nombre de zones qui composent l'image synthétique. Le tableau 6 illustre les caractéristiques des classes détectées dans l'image de la figure 5.28.

Classe	Composante R	Composante V	Composante B	Nombre de pixels
1	168.5377	91.2944	239.2409	484.0000
2	124.3551	23.2025	171.8277	462.0000
3	131.2945	56.5027	184.8774	462.0000
4	17.4029	245.8517	52.9524	484.0000
5	41.0851	162.7465	0.0582	462.0000
6	85.5864	70.1505	11.3545	441.0000
7	23.9436	104.5493	208.3075	441.0000
8	221.9819	5.7516	185.4301	462.0000
9	183.1957	244.3446	39.9929	462.0000
10	225.4256	88.5472	15.1666	441.0000
11	167.3695	189.2878	87.9663	441.0000
12	216.2424	185.7930	243.5502	462.0000
13	106.1700	23.9789	114.7363	484.0000
14	221.6339	99.8624	64.4598	462.0000
15	90.3674	189.4594	165.9622	462.0000
16	239.6472	212.3638	119.8444	484.0000

Table 6 : Les caractéristiques des classes détectées dans l'image
9 Segmentation d'image réelle

L'image à segmenter est une image réelle de faible qualité Figure 5.32

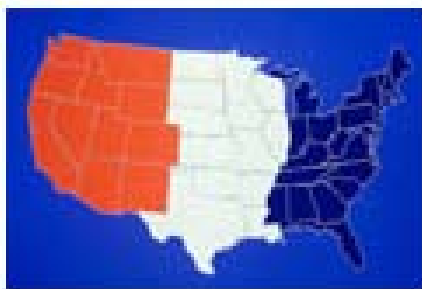


Fig.5.32 : Image réelle à segmenter

La figure 5.33 illustre l'image des zones de conflit. On remarque que cette image riche en zones de conflit.

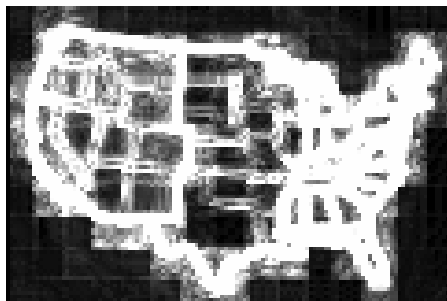


Fig.5.33 : Image des zones de conflit

La figure 5.34 présente l'image segmentée avant raffinement.

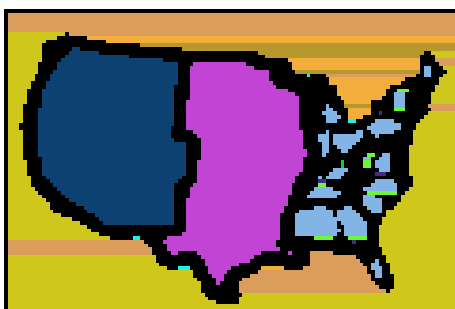


Fig.5.34 : Image segmentée sans raffinement

La figure 5.35 donne l'image résultat de la segmentation avec raffinement.

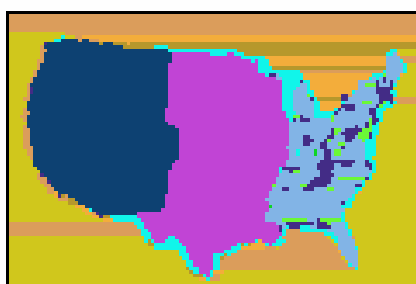


Fig.5.35 : Image segmentée avec raffinement

10 Apport de la nouvelle méthode comparativement aux autres méthodes

Toutes les méthodes qui existent actuellement donnent le même intérêt à chaque pixel de l'image à segmenter, par conséquent le même traitement avec sa complexité est appliqué à chaque pixel, ce qui nécessite un temps considérable.

La technique de segmentation développée dans cet ouvrage est une nouvelle méthode, elle englobe les deux phases de la segmentation. En premier lieu un balayage de l'image est effectué pour établir le nombre de classes qui composent cette image et la carte des zones de conflit qui présente le lieu géométrique des éléments difficiles à classifier. Enfin, un raffinement de l'image pré segmentée est réalisé, il consiste à appliquer une technique complexe qui nécessite un temps de calcul pour classifier les éléments de la zone de conflit. En général le nombre des éléments de la zone de conflit est faible par rapport au nombre des pixels de l'image, ce qui rend la méthode rapide et efficace. La phase de pré segmentation rend cette technique utilisable dans plusieurs domaines d'application.

11 Conclusion

Dans ce chapitre, on a décrit une nouvelle approche de segmentation non supervisée d'image. On a utilisé une stratégie originale qui consiste à faire évoluer un graphe à progression probabiliste. Ce graphe représente dans sa structure les différentes composantes de l'image à segmenter. Le nombre de nœuds du graphe établit à la fin de sa progression donne le nombre de classes et leurs centres. Une matrice des classes est générée localisant ainsi chaque classe dans l'espace de l'image. Cette image de la classe présente l'isolation de l'objet des différents constituants de l'image.

La dernière phase de traitement consiste à classifier les sites omis lors de la première phase de la segmentation. Ces sites possèdent des attributs qui peuvent dérouter le système de segmentation, et présentent ainsi des sources d'erreurs. Pour leur classification on a opté pour l'utilisation d'un système de classification puissant basé sur la théorie de l'évidence.

Ce travail fait partie de la conception et de la réalisation d'un robot mobile possédant un système sensoriel équipé entre autre d'une caméra destinée à la navigation dans un environnement inconnu. L'objectif de cette navigation est de localiser et éviter les obstacles et de planifier une trajectoire d'évolution du robot dans le but d'atteindre une cible dans un environnement totalement inconnu. On a utilisé des images à niveaux de gris et des images en couleurs pour tester notre algorithme. Les résultats obtenus sont satisfaisants.

**Conclusion
Générale
Et
Perspectives**

Nous avons présenté dans cette thèse les travaux qui concernent l'étude des fonctionnalités visuelles sur des scènes qui composent l'environnement d'évolution et ses applications à la robotique mobile. Ces Travaux gravitent autour de la phase qui s'intéresse à l'extraction des informations issues d'un capteur visuel qu'il convient d'embarquer sur un robot mobile ou non, afin qu'il puisse se déplacer dans son environnement. Plus précisément ils ont trait aux méthodes d'identification visuelle d'objets ou d'entités extraites d'une image, pour la navigation d'un robot mobile autonome.

Si l'homme sait naturellement séparer des objets dans une image c'est grâce à des connaissances de haut niveau (compréhension des objets et de la scène). Mettre au point des algorithmes de segmentation de haut niveau (chaque région est un objet sémantique) est encore un des thèmes de recherche les plus courants en traitement d'images. Bien séparer les différentes entités qui composent une image, pour les applications souhaitées, est un problème très important en traitement et analyse d'image. Cette opération appelée segmentation est quasi obligatoire dans tous les systèmes de vision artificielle. Elle influence fortement la qualité de l'interprétation, ce qui justifie les nombreux travaux qui lui ont été consacrés.

L'objectif de ce travail a été de concevoir un nouveau système rapide de segmentation d'image, qui présente un intérêt très important dans les applications à temps réel, qui entre dans le cadre des travaux développés sur les applications de la vision artificielle destinés à la navigation d'un robot mobile. Un intérêt tout particulier a été accordé à l'aspect automatique et non supervisé du dispositif.

Nous nous sommes tout d'abord intéressés à la robotique mobile, qui présente le domaine d'application vers lequel nos travaux sont orientés. Nous nous sommes penchés, au cours de ce chapitre, sur une introduction à la robotique mobile ainsi que sur les concepts et outils permettant à un robot mobile, évoluant dans un environnement dynamique incertain, de se déplacer d'une façon autonome vers un objectif. Afin d'exécuter la tâche désirée, le robot doit généralement exécuter deux tâches, le calcul préalable d'une trajectoire de référence à l'aide de méthodes de planification, puis la poursuite de cette trajectoire à l'aide d'une méthode réactive qui permet de prendre en compte les petits imprévus et de rattraper les erreurs de suivi.

Posséder une trajectoire de référence n'assure pas la sécurité du robot, il doit la suivre fidèlement. L'absence de connaissance complète a priori sur l'environnement rend la trajectoire calculée incertaine sur un temps indéfini, elle doit être recalculée régulièrement. L'aspect dynamique de l'environnement contraint le robot à rester réactif et à adapter son évolution à la dynamique de son environnement.

Nantir un robot mobile d'un système de vision artificielle allège considérablement le problème de la navigation et ouvre une fenêtre sur la planification de tâches plus complexes, qui présentées de grandes difficultés pour la navigation d'un système aveugle.

Ensuite, nous avons scruté quelques concepts de fonctionnement du système visuel humain, nous avons étalé le long de cette partie, les différentes composantes qui assurent le fonctionnement du cerveau visuel biologique. Plusieurs opérations sont réalisées dans le cerveau pour passer de la sensation, stimulation des organes des sens, à la perception, impression obtenue à partir des sens et appréhendé par l'esprit. La perception fait ainsi suite à la sensation, en incluant des traitements des stimuli perçus consciemment ou inconsciemment et en prenant en considération des informations contextuelles et des expériences passées, stockées dans une bibliothèque de taille impressionnante. Cette bibliothèque inclut des images, des sons, des odeurs.

Utiliser l'image d'une scène issue d'un capteur vidéo pour recueillir et utiliser les informations de l'espace environnant est une tâche particulièrement difficile. Si de nombreuses applications permettent d'automatiser des tâches liées à la perception humaine, comme le contrôle qualité sur des objets manufacturés ou la conduite autonome de véhicules, ceci ne peut être réalisé que dans des cas bien particuliers, avec des contraintes très fortes quant à l'environnement et la nature des objets eux-mêmes.

Puis, nous avons effectué un état de l'art des différents travaux réalisés dans le cadre de la segmentation d'image. Il existe évidemment dans la littérature de nombreuses méthodes de segmentation. Les différentes techniques ont leurs avantages et leurs limites et leur utilisation dépend généralement du problème à résoudre. L'enjeu est donc de trouver la méthode la plus adaptée à une famille d'application donnée. Le but de nos travaux est de rechercher une stratégie de

segmentation rapide et robuste, permettant ainsi de traiter le plus grand nombre de situations possibles.

L'approche de segmentation d'image décrite dans cette oeuvre, possède une architecture basée sur deux étapes. La première étape consiste à établir un graphe représentatif de l'image à segmenter, fournissant ainsi, le nombre de classes qui composent l'image avec les caractéristiques de chaque classe et cela à partir des individus relativement homogènes. La seconde phase est consacrée à l'amélioration de la segmentation de l'image pré-segmentée, et cela en classifiant les individus qui présentent un certain degré d'hétérogénéité par rapport aux éléments des classes identifiées. Une procédure probabiliste fondée sur la théorie de l'évidence est utilisée à cette fin. L'évaluation des performances obtenues sur un ensemble de collections d'images très différentes a permis de tester la robustesse du processus.

Enfin, nous souhaitons terminer en évoquant des travaux que nous pourrions mener dans le futur sur la thématique présentée dans cette thèse. Tout d'abord, de nombreuses améliorations peuvent être apportées aux modules logiciels développés, augmentation du nombre de nœuds du graphe en augmentant les contraintes des seuils, application de techniques puissantes pour une fusion des nœuds jugés homogènes. Nous considérons comme une perspective importante l'implémentation de cette méthode en l'introduisant dans le système de traitement d'un robot mobile équipé d'une caméra évoluant dans un environnement dynamique.

Références Bibliographiques

- [1] T.R. REED and J.M. HAS DU BUF. "A review of recent texture segmentation and feature extracting techniques". *CVGIP: IU*, 57 pp. 359-372, Mai 1993.
- [2] J.P. COCQUEREZ et S. PHILIPP. "Analyse d'images : filtrage et segmentation". Collection Enseignement de la physique. Masson, Paris, 1995.
- [3] R. Arkin. "Towards the unification of navigational planning and reactive control". In *Proceedings of the AAAI Spring Symposium on Robot Navigation*, pages 1-5, 1989.
- [4] O. Trullier and J. A. Meyer. "Biomimetic navigation models and strategies in animats". *AI Communications*, 10 :79-92, 1997.
- [5] O. Trullier, S. Wiener, A. Berthoz, and J. A. Meyer. "Biologically-based artificial navigation systems : Review and prospects". *Progress in Neurobiology*, 51 :483-544, 1997.
- [6] P. Gaussier, C. Joulain, J.P. Banquet, S. Lepretre, and A. Revel. "The visual homing problem: an example of robotics/biology cross-fertilisation". *Robotics and autonomous systems*, 30(1-2) :155-180, 2000.
- [7] D. Lambrinos, R. Möller, T. Labhart, R. Pfeifer, and R. Wehner. "A mobile robot employing insect strategies for navigation". *Robotics and Autonomous Systems, special issue : Biomimetic Robots*, 30 :39-64, 2000.
- [8] S. Gourichon and J.-A. Meyer. "Using colored snapshots for short-range guidance in mobile robots". *International Journal of Robotics and Automation*, submitted for publication, 2001.
- [9] T. S. Levitt and D. T. Lawton. "Qualitative navigation for mobile robots". *Artificial Intelligence*, 44 :305-360, 1990.
- [10] B. J. Kuipers and Y. T. Byun. "A robot exploration and mapping strategy based on a semantic hierarchy of spatial representations". *Robotics and Autonomous Systems*, 8 :47- 63, 1991.
- [11] S. Thrun. "Learning metric-topological maps for indoor mobile robot navigation". *Artificial Intelligence*, 99(1): 21-71, 1999.
- [12] C. Kunz, T. Willeke, and I. Nourbakhsh. "Automatic mapping of dynamic office environments". In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA-97)*, volume 2, pages 1681-1687, 1997.
- [13] U. Nehmzow and C. Owen. "Robot navigation in the real world : Experiments with manchester's fortytwo in unmodified, large environments". *Robotics and Autonomous Systems*, 33(4) :223-242, 2000.
- [14] R. Simmons and S. Koenig. "Probabilistic navigation in partially observable environments". In S. Mellish, editor, *Proceedings of IJCAI'95*. Morgan Kaufman Publishing, 1995.
- [15] G. Von Wichert. "Mobile robot localization using a self-organised visual environment Representation". *Robotics and Autonomous Systems*, 25 :185-194, 1998.
- [16] D. Filliat and J.-A. Meyer. "Map-based navigation in mobile robots - i. a review of localisation strategies". *Journal of Cognitive Systems Research*, submitted for publication, 2001.
- [17] M. Franz, B. Scholkopf, P. Georg, H. Mallot, and H. Bulthoff. "Learning view graphs for robot navigation". *Autonomous Robots*, 5 :111-125, 1998.
- [18] O. Trullier and J. A. Meyer. "Animat navigation using a cognitive graph". *Biological Cybernetics*, 83(3) :271-285, 2000.
- [19] M. Betke and K. Gurvits. "Mobile robot localization using landmarks". In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA-94)*, volume 2, pages 135-142, 1994.

- [20] C. Madsen, C. Andersen, and J. rensen. "A robustness analysis of triangulation-based robot self-positioning". In *Proceedings of the 5th Symposium for Intelligent Robotics Systems*, 1997.
- [21] A. Arsenio and M. I. Ribeiro. "Absolute localization of mobile robots using natural landmarks". In *Proceedings of the International Conference on Electronics, Circuits and Systems*, 1998.
- [22] A. Saffiotti and L. P. Wesley. "Perception-based self-localization using fuzzy locations". In *Reasoning with Uncertainty in Robotics*, volume 1093 of *Lecture Notes in Computer Science*. Springer-Verlag, 1995.
- [23] P. Moutarlier and R. Chatila. "An experimental system for incremental environment modelling by an autonomous mobile robot". In *Experimental Robotics 1*, pages 327–346. Springer-Verlag, 1990.
- [24] C. F. Olson. "Probabilistic self-localization for mobile robots". *IEEE Transactions on Robotics and Automation*, 16(1), 2000.
- [25] J.-C. Latombe. "Robot Motion Planning". Boston : *Kluwer Academic Publishers*, 1991.
- [26] D. Kortenkamp, M. Huber, F. Koss, W. Belding, J. Lee, A. Wu, C. Bidlack, and S. Rogers. "Mobile robot exploration and navigation of indoor spaces using sonar and vision". In *Proceedings of the AIAA/NASA Conference on Intelligent Robots in Field, Factory, Service, and Space (CIRFFSS 94)*, pages 509–519, 1994.
- [27] Chauvin A, "Perception des scènes naturelles : étude et simulation du rôle de l'amplitude, de la phase et de la saillance dans la catégorisation et l'exploration de scènes naturelles" *Thèse de doctorat, Université Joseph Fourier, Grenoble*, 2003.
- [28] J.P. HATON, "Les systèmes à base de connaissances en reconnaissance et interprétation de formes.", COGNITIVA, Paris, La vilette, pp.73-80, 18-22 May 1987.
- [29] J. ALOIMONOS, "Purposive and qualitative active vision". In *ICPR*, pp. 346-360, Atlantic City, New Jersey, 1990.
- [30] R. BAJCSY, "Active Perception". *Proceedings of The IEEE*, 76(8) pp. 996-1005, August 1988.
- [31] Haralick R.M., Shapiro L.G., "Computer and robot vision". Vol. 2, *Addison-Wesley Publishing Company*, 1993.
- [32] Haralick R.M., Shapiro L.G., "Computer and robot vision". Vol. 1, *Addison-Wesley Publishing Company*, 1992.
- [33] Marion A., "Introduction aux techniques de traitement d'images". *Eyrolles*, 1987.
- [34] Kunt M., Granlund G., Kocher M., "Traitement numérique des images". *Collection Electricité, Traitement de l'Information: Vol. 2, Presses Polytechniques et Universitaires Romandes et CNET-ENST*, 1993.
- [35] Horaud R., Monga O., "Vision par ordinateur. Outils fondamentaux". *Traité des Nouvelles Technologies, Serie Informatique, Hermès*, 1993.
- [36] Rzevski G. (Ed.), prepared by an Open University Course Team, "Perception, cognition and execution". *Mechatronics, Designing Intelligent Machines*, Vol. 1, *Butterworth-Heinemann* in association with *the Open University*, 1994.
- [37] Jain A.K., "Fundamentals of digital image processing". *Prentice Hall Information and System Sciences Series*, T. Kailath (Ed.), *Prentice Hall*, 1989.
- [38] A. N. Zahid. O. Abouelala, M. Limouri, "Unsupervised fuzzy clustering", *Pattern Recognition Letters*, 20 (1999), pp. 123-129.

- [39] T. Uchiyama & M.A. Arbib, "Color Image Segmentation using competitive learning", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. PAMI-16, n° 12, 1994, pp. 1197-1206.
- [40] N. Malpica, C. Ortiz de Solorzano, J.J. Vaquero, A. Santos, I. Vallcorba, J.M. Garcia-Sagredo, F. del Pozo, "Applying Watershed Algorithms to the Segmentation of Clustered Nuclei", *Wiley-liss, Inc., Cytometry* 28:289-297, 1997.
- [41] S. Horowitz et T. Pavlidis: "Picture segmentation by a directed split-and-merge procedure". *Rapport technique, Departement of Electrical Engineering, Princeton University*, 1975.
- [42] Z. Ying, T. Jie, L. NingNing, H. ZhiGang, D. RuWei, "An Interactive Image Segmentation Method Based on Dynamic Programming and its Application in Medical Image Analysis", *The 14th International Conference on Pattern Recognition (ICPR-14)*, August 1998.
- [43] J. Bruce, T. Balch, M. Veloso, "Fast and Inexpensive Color Image Segmentation for Interactive Robots", *Proceedings of the 2000 IEEE/RSS International Conference on Intelligent Robotical Systems (IROS'00) Vol.3*, pp.2061-2065, Takamatsu, Japan, October 2000.
- [44] O.C. Ozcanli, F.T. Yarman Vural, "A Content Based Image Retrieval System Based on Localization of the Query", *Madison, Wisconsin, USA, submitted to CVPR June 16-22 2003*.
- [45] Y. Xu, P. Duygulu, E. Saber, A.M. Tekalp, F.T. Yarman-Vural, "Object Based Image Retrieval Based on Multi-Level Segmentation", *SPIE Electronic Imaging 2001, Storage and Retrieval for Media Databases*, San Jose, CA, USA, January 21-26, 2001.
- [46] J. Canny: "A computational approach to edge detection". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8:679-714, 1986.
- [47] R. Deriche: "Using Canny's criteria to derive a recursively implemented optimal edge detector". *The International Journal of Computer Vision*, 1(2):167-187, mai 1987.
- [48] O. Monga, R. Deriche, G. Malandain et J.P. Cocquerez: "Recursive filtering and edge closing: two primary tools for 3-d edge detection". *Dans European Conference on Computer Vision*, pages 56-65. Springer-Verlag, 1990.
- [49] O. Monga, N. Ayache et P. Sander: "From voxels to curvature features". *Dans International workshop on Visual form: analysis and recognition*, pages 399-407. Plenum Press, 1992.
- [50] P. Simpson: "Artificial neural systems: foundations, paradigms, applications, and implementations". *Pergamon Press*, 1990.
- [51] S. Haykin: "Neural networks: a comprehensive foundation". *2nd Edition. Prentice Hall*, 1998.
- [52] R. Hult: "Grey-level morphology combined with an artificial neural networks approach for multimodal segmentation of the hippocampus". *Dans International Conference on Image Analysis and Processing*, pages 277-282, 2003.

- [53] B. Vemuri, S. Rahman et J. Li: "Multiresolution adaptive K-means algorithm for segmentation of brain MRI". *Dans International Computer Science Conference on Image Analysis Applications and Computer Graphics*, pages 347–354, 1995.
- [54] M. Yan et J. Karp: "Segmentation of 3D brain MR using an adaptive K-means clustering algorithm". *Dans Nuclear Science Symposium and Medical Imaging Conference*, pages 1529–1533, 1995.
- [55] J. Bezdek: "Pattern Recognition with Fuzzy Objective Function Algorithms". *Kluwer Academic Publishers*, 1981.
- [56] J. Bezdek, J. Keller, N. Pal et R. Krisnapuram: "Fuzzy Models And Algorithms For Pattern Recognition And Image Processing". *Kluwer Academic Publishers*, 1995.
- [57] S. Geman et D. Geman: "Stochastic relaxation, Gibbs distribution, and the Bayesian restoration of images". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6:721–741, 1984.
- [58] C. Tan et J. Rajapakse: "Tissue segmentation of multi-channel brain images with inhomogeneity correction". *Dans International Symposium on Image and Signal Processing and Analysis*, volume 1, pages 571–576, septembre 2003.
- [59] A. Capelle, O. Alata, C. Fernandez, S. Lefevre et J.C. Ferrie: "Unsupervised segmentation for automatic detection of brain tumors in MRI". *Dans International Conference on Image Processing*, volume 1, pages 613–616, septembre 2000.
- [60] L. Vincent et P. Soille: "Watershed in digital spaces, an efficient algorithm based on immersion simulation". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(6):583–598, juin 1991.
- [61] F. Meyer: "Skeletons in digital spaces. Image Analysis and Mathematical Morphology". 2:257–296, 1988.
- [62] J. Roerdink et A. Meijster: "The watershed transform: Definitions, algorithms and parallelization strategies". *Fundamenta Informaticae*, 41(1-2):187–228, 2000.
- [63] J. Thiran, V. Warscotte et B. Macq: "A queue-based region growing algorithm for accurate segmentation of multi-dimensional digital images". *Signal Processing*, 60(1):1–10, 1997.
- [64] S. Peng et L. Gu: "A novel implementation of watershed transform using multi-degree immersion simulation". *Dans International Conference of the Engineering in Medicine and Biology Society*, pages 1754–1757, septembre 2005.
- [65] J. Betser, S. Delest et R. Bon: "Unbiased watershed hierarchical 3D segmentation". *Dans Visualization, Imaging, and Image Processing*, 2005.
- [66] C. Berge: "Graphes et hypergraphes". *Dunod*, 1969.
- [67] A. Bretto et L. Gillibert: "Hypergraph-based image representation". *Dans Graph-based Representations in Pattern Recognition*. pages 1–11, 2005.
- [68] S. Chastel, P. Colantoni et A. Bretto: "Displaying image neighborhood hypergraphs line-graphs". *Dans International Conference on Discrete Geometry for Computer Imagery*, pages 124–135. Springer-Verlag, 2002.

- [69] A. Bretto, H. Cherifi et D. Aboutajdine : "Hypergraph imaging : an overview". *Pattern Recognition*, 35(3):651–658, 2002.
- [70] D. Dubois et H. Prade. "La problématique scientifique du traitement de l'information". *Information-Interaction-Intelligence*, 1(2), 2001.
- [71] D. Heckerman, D. Geiger et D. Chickering. "Learning bayesian networks: The combination of knowledge and statistical data". *Machine Learning*, 20:197–243, 1995.
- [72] S. Parsons. "Current approaches to handling imperfect information in data and knowledge bases". *IEEE Transactions on Knowledge and Data Engineering*, 8(3):353–372, 1996.
- [73] R.R. Yager. "Uncertainty representation using fuzzy measures". *IEEE Transactions on Systems, Man and Cybernetics - Part B*, 32(1):13–20, February 2002.
- [74] L. Zadeh. "Fuzzy sets as a basis for a theory of possibility". *Fuzzy Sets and Systems*, 1:3–28, 1978.
- [75] Ph. Smets et R. Kennes. "The Transferable Belief Model". *Artificial Intelligence*, 66(2):191–234, 1994.
- [76] A. Dempster. "Upper and lower probabilities induced by multivalued mapping". *Annals of Mathematical Statistics*, AMS-38:325–339, 1967.
- [77] Ph. Smets. "The normative representation of quantified beliefs by belief functions". *Artificial Intelligence*, 92(1-2):229–242, 1997.
- [78] R. Kennes et Ph. Smets. "Computational aspects of the Möbius transform". In P.P. Bonissone, M. Henrion, L.N. Kanal et J.F. Lemmer, éditeurs, *Uncertainty in Artificial Intelligence 6*, pages 401–416. North Holland, Amsterdam, 1991.
- [79] Ph. Smets. "The application of the matrix calculus to belief functions". *International Journal of Approximate Reasoning*, 31(1-2):1–30, October 2002.
- [80] D. Dubois, H. Prade et Ph. Smets. "Representing partial ignorance". *IEEE Systems, Machine and Cybernetics*, 26:361–377, 1996.
- [81] Ph. Smets. "Decision making in a context where uncertainty is represented by belief functions". In R. P. Srivastava et T.J. Mock, éditeurs, *Belief Functions in Business Decisions*, pages 17–61. Physica-Verlag, Heidelberg, Germany, 2002.
- [82] Ph. Smets. "The combination of evidence in the transferable belief model". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(5):447–458, 1990.
- [83] R.R. Yager. "Hedging in the combination of evidence". *Journal of Information and Optimization Sciences*, 4(1):73–81, 1983.
- [84] C. K. Murphy. "Combining belief functions when evidence conflicts". *Decision Support Systems*, 29:1–9, 2000.
- [85] Ph. Smets. "Belief functions: The disjunctive rule of combination and the generalized bayesian theorem". *International Journal of Approximate Reasoning*, 9:1–35, 1993.
- [86] T. Denoeux. "Analysis of evidence-theoretic decision rules for pattern classification". *Pattern Recognition*, 30(7):1095–1107, 1997.
- [87] T. Strat. "Decision analysis using belief functions". *International Journal of Approximate Reasoning*, 4:391–417, 1990.
- [88] S. K. Pal et al, "A Review on Image Segmentation Techniques", *Pattern Recognition* 29 1277-1294, 1993.
- [89] H. D. Cheng et al, "Color Image Segmentation : Advences and Prospects", *Pattern Recognition* 34, 2259-2281, 2001.
- [90] Christophe Charrier et Anne-Laure Jousselme, "Quantification vectorielle d'images couleur par la théorie de l'évidence", *Université de Caen*, 2003.

- [91] Said Kharbouche et Patrick Vannoorenberghe, “Un filtre crédal pour la segmentation d’images”, *Université de Rouen*, 2005.
- [92] Serradura, L., Brouard, T. , Slimane, M., Venturini, G., “Image segmentation with hidden Markov models”, *7th Conference of the International Federation of Classification Societies. Data Analysis, Classification, and Related Methods*. July 11-14, Namur, Belgium 2000.
- [93] E. Monfrini, T. Ledru, E. Vaie, W. Pieczynski, “Segmentation non supervisée d’images par arbres de Markov cachés”, *Actes du 17ème Colloque GRETSI, GRETSI’99*, Vannes, France, pp. 43-46, 13- 17 septembre 1999.
- [94] R. L. Cannon, J.V. Dave, J.C. Bezdek, “ Efficient Implementation of the Fuzzy C-means Clustering Algorithms”, *IEEE Trans. Pattern Anal. Mach. Intell.* PAMI-8 (2), pp: 249-255, 1986.
- [95] A. Dempster, “A generalization of Bayesian inference”. *Journal of the royal statistical society*, Vol30, pages 205-245, 1968.
- [96] G. Shafer, “A mathematical Theory of Evidence”, *Vo.12702, Princeton*, 1976.
- [97] PH. Smets, “Data Fusion in the transferable Belief Model”, *Proc. 3rd Intern. Conf. Information Fusion*, Paris, France, Pages 21-33, 2000.
- [98] P. Smets. “What is Dempster-Shafer’s model ?”. *In R.R. Yager, M. Fedrizzi, and J. Kacprzyk, editors, Advances in the Dempster-Shafer Theory of Evidence*, pages 5–34. Wiley, 1994.
- [99] P. Smets and R. Kennes. “The transferable belief model”. *Artificial Intelligence*, 66(2):191–234, 1994.
- [100] N. AMMOUR, A.GUESSOUM, D. BERKANI, ” Scene analysis using both a probabilistic evolutionary graph and the evidence theory”, *WSEAS TRANSACTIONS on BIOLOGY and BIOMEDICINE, Issue 2, Volume 4*, February 2007, ISSN: 1109-9518, pages: 22-28.
- [101] N. AMMOUR, A.GUESSOUM, D. BERKANI, ” Using Both a Probabilistic Evolutionary Graph and the Evidence Theory for Color Scene Analysis”, *American Journal of Applied Sciences* 5 (12), ISSN 1546-9239: 1635-1641, 2008.
- [102] N. AMMOUR, A.GUESSOUM, D. BERKANI, ” Using evolutionary graph for image segmentation”, *Proceedings of the 11th WSEAS International Conference on SYSTEMS*, Agios Nikolaos, Crete Island, Greece, July 23-25, 2007 ,pp 131-135.