

RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE SCIENTIFIQUE
UNIVERSITÉ MOHAMED BOUDIAF (USTO)
ORAN

FACULTÉ DES SCIENCES
DÉPARTEMENT INFORMATIQUE

MÉMOIRE DE MAGISTER
SPÉCIALITÉ : INFORMATIQUE

*Détection de contours et suivi d'objet dans une
séquence d'images par les réseaux de neurones
impulsionnels*

Présenté par : **Yedjour Hayat**

Devant le jury :

Mr. Benyettou Mohamed

Mr. Benyettou Abdelkader

Mr. Djebbar Bachir

Mme. Fizazi Hadria

Président de jury

Encadreur

Examineur

Examinatrice

Dédicace

A la mémoire de mon Père.

A ma mère, dont je souhaite une longue et heureuse vie,

A mon cher époux Djamel pour ses encouragements et ses conseils,

A mes trois mousquetaires Ahmed Fouad, Nabil et Ilies Abderraouf, qu'ils garderont ce travail comme beau souvenir de leur chère maman et qu'ils le prennent comme exemple dans leur carrière et leur avenir,

A mes affectueux frères et sœurs, neveux et nièces que je remercie pour leurs encouragements infiniment avec toute la grande famille,

A mes collègues du CHU d'Oran, pour leur soutien moral et logistique,

A mes collègues de la promotion 2007/2010, en particulier Hayet.

Yedjour Hayet

Remerciements

Remercions le Dieu, tout puissant clément et miséricordieux de m'avoir doté d'une patience et d'un courage inouï pour la réalisation de ce modeste travail.

Mes profonds remerciements sont adressés d'abord à mon respectueux encadreur Mr. A. Benyettou, qui a veillé à la qualité de cette recherche dont je souhaiterai être à la hauteur de l'ambition souhaitée.

Sage à celui auquel j'exprime toute l'honnêteté intellectuelle avec laquelle il m'a soutenu tant techniquement et méthodologiquement. Sage à celui auquel j'exprime toute ma sincère considération et ma reconnaissance pour l'élaboration de ce manuscrit dont le nom Mr B. Meftah reste gravé comme Co-encadreur.

Nos remerciements sont aussi adressés aux honorables membres du jury qui ont acceptés la lecture et l'évaluation de ce travail.

Enfin ! Sans oublier nos humbles enseignants qui méritent l'honneur et la reconnaissance de notre formation.

Résumé

Les réseaux de neurones impulsionnels (SNN) sont considérés comme la nouvelle génération des réseaux de neurones artificiels grâce à leurs propriétés très intéressantes comme l'intégration du temps au niveau du neurone qui leur permet de traiter des problèmes dynamiques où le temps est une composante essentielle.

La détection de contour est une étape préliminaire à de nombreuses applications de l'analyse d'images. Un contour peut être considéré comme une frontière entre deux régions différentes suivant un critère de similarité donné. L'obtention des contours est une étape importante dans le processus d'interprétation automatique d'une image puisque cela permet de matérialiser les contours des objets recherchés. En l'absence d'une méthode générale pour la détection de contours, différentes approches ont été proposées, s'appuyant soit sur la notion de région, soit sur celle de contour. Cependant le temps d'exécution, qui est l'une des principales difficultés de ce type d'application, n'est que rarement pris en compte.

Détecter un objet en mouvement est un problème souvent abordé dans le traitement des images. Parmi les méthodes les plus utilisées dans ce domaine, on distingue les techniques basées sur la différence d'images. Ces méthodes ont été développées dans le cadre d'application comme la vidéo-surveillance et le contrôle du trafic routier et sont robustes vis-à-vis des variations lentes et rapides rencontrées dans ces applications mais ne le sont pas vis-à-vis des variations frame à frame rencontrées dans les séquences d'images issues du milieu aquatique.

Mots-clés : Apprentissage, Détection de contour, Réseaux de neurones impulsionnels, Séquence d'images, Suivi d'objet.

Abstract

Spiking Neural Networks are considered as a new computation paradigm, representing the next generation of Artificial Neural Networks by offering more flexibility and degrees of freedom for modeling computational elements, taking into account the time of spike emission between neurons allows them to deal with dynamic problems where time is an essential component.

Detection of edges is a preliminary step with many applications of vision analysis. A contour can be regarded as boundarie between two different regions according to a given criterion of similarity. Obtaining contours is a significant stage in the process of interpretation automatic of an image since that makes it possible to materialize contours of

the required objects. In the presence of a general method of edge detection, different approaches have been proposed, based either on the concept of region or on the contour. However the execution time, which is one of the main difficulties of this type of application, is rarely taken into account.

Detecting moving object is a often discussed in image processing. Among the most widely used methods in this domain, there are techniques based on the difference image. These methods have been developed in the context of applications such as video surveillance and control of road traffic and are robust in the face of the fast and slow variations encountered in the applications but are not in the face of frame to frame variations encountered in image sequences from the aquatic environment.

Keywords : Edge detection, Learning, Object tracking, Sequence of images, Spiking Neural Networks.

Table des matières

Introduction générale	1
Introduction	1
1 Réseaux de neurones impulsionnels	6
1.1 Introduction	6
1.2 Vue biologique générale	7
1.2.1 Cerveau humain	7
1.2.2 Topologie des réseaux de neurones	7
1.2.3 Morphologie des neurones	7
1.2.4 Activité neuronale	8
1.3 Neurone biologique	10
1.3.1 Potentiel de la membrane	10
1.3.2 Potentiel de repos	10
1.3.3 Potentiel local	10
1.3.4 Potentiel d'action	11
1.3.4.1 Caractéristiques du potentiel d'action	12
1.3.5 Connexions synaptiques	13
1.3.6 Propriétés d'intégration	13
1.4 Générations des modèles connexionnistes	14
1.4.1 Les trois générations de modèles connexionnistes	14
1.4.2 Réseaux de 3ème génération	15
1.5 Modélisation des neurones impulsionnels	16
1.5.1 Modèle à base de conductance de Hodgkin-Huxley	16
1.5.2 Modèle impulsif intègre-et-tire et le modèle LIF	17
1.5.3 Neurone à réponse impulsif (SRM)	18
1.6 Réseaux de neurones impulsif	20
1.6.1 Topologie et dynamique d'un SNN	20

1.6.1.1	Connexion et délais	21
1.6.1.2	Topologie	21
1.6.1.3	Dynamique neuronale	22
1.6.2	Codage de l'information par les neurones	22
1.6.2.1	Codage par le taux de décharge	23
1.6.2.2	Codage par impulsions	23
1.7	Apprentissage dans les réseaux de neurones impulsionnels	25
1.7.1	Plasticité synaptique	25
1.7.1.1	Assemblées cellulaires	25
1.7.1.2	Plasticité synaptique par taux de décharge	26
1.8	Conclusion	26
2	Détection, segmentation d'images et opération morphologiques	28
2.1	Introduction	28
2.2	Approches de détection de contours	29
2.2.1	Approche par convolution	29
2.2.1.1	Détection des contours par dérivée première	30
2.2.1.2	Détection des contours par dérivée seconde	31
2.2.2	Approche par filtrage optimal	31
2.3	Segmentation et extraction de caractéristiques	32
2.3.1	Segmentation d'images	32
2.3.1.1	Segmentation par approche régions	33
2.3.1.2	Segmentation par approche frontière	34
2.3.1.3	Segmentation par classification ou seuillage	34
2.3.2	Extraction de primitives	35
2.3.2.1	Les indices visuels	35
2.3.2.2	Méthodes basées primitives	36
2.4	Opérations morphologiques	36
2.4.1	Erosion	37
2.4.2	Dilatation	38
2.5	Conclusion	39
3	Etat de l'art sur le suivi	40
3.1	Introduction	40
3.2	Modélisation d'un objet	41
3.2.1	Caractéristiques d'un objet	41

3.2.2	Représentation d'un objet	42
3.2.2.1	Modélisation par points	42
3.2.2.2	Modélisation par formes géométriques	42
3.2.2.3	Modélisation par silhouette et contour	43
3.2.2.4	Modélisation par des modèles de formes articulées	43
3.2.2.5	Modèles squelettiques	43
3.2.2.6	Densités de probabilités d'apparence d'objet	44
3.2.2.7	Les templates	44
3.2.2.8	Les modèles d'apparences multi-vues	44
3.3	Nature de la scène	45
3.3.1	Séquences à caméra fixe (Scène statique)	45
3.3.2	Séquences à caméra mobile (Scène dynamique)	46
3.3.3	Mono et multi-caméra	46
3.4	Détection d'objets	46
3.4.1	Catégorisation des méthodes de détection	46
3.4.1.1	Détections basées sur la différence inter-images	47
3.4.1.2	Détections basées sur la modélisation du fond	47
3.4.1.3	Les détections extrayant des couches de mouvement	48
3.4.1.4	Les détections utilisant la notion de cohérence	48
3.5	Suivi d'objets	49
3.5.1	But du suivi	49
3.5.2	Méthodes de suivi	50
3.5.2.1	Suivi par appariement de détections	50
3.5.2.2	Segmentation dynamique	51
3.5.2.3	Suivi déterministe de fenêtre englobante par détection sé- quentielle	53
3.5.3	Caractéristiques d'un bon algorithme de suivi	54
3.6	Conclusion	55

4 Conception et implémentation de la détection de contour par réseaux de neurones impulsionnels 56

4.1	Introduction	56
4.2	Modèle d'un réseau de neurones impulsionnels pour la détection de contours	56
4.2.1	Architecture du réseau de neurones impulsionnels	56
4.2.2	Description Fonctionnelle du réseau	57
4.3	Modèle du neurone impulsionnel et les champs récepteurs	58

4.4	Résultats de simulation	60
4.5	Conclusion	62
5	Conception et implémentation du suivi d'objet	64
5.1	Introduction	64
5.2	Détection d'objets en mouvement	65
5.2.1	Suppression du fond (Background substraction)	65
5.2.2	Suivi d'objets par la distance de Hausdorff	66
5.2.2.1	Comparaison d'images	66
5.2.2.2	Définition de la distance Hausdorff	67
5.2.2.3	Propriétés générales de la distance de Hausdorff	67
5.2.2.4	Distance Hausdorff totale	68
5.2.2.5	Distance Hausdorff partielle	68
5.3	Implémentation du suivi d'objet	69
5.3.1	De la video à l'image	69
5.3.2	détection de contour par réseaux de neurones impulsionnels	70
5.3.3	Soustraction du fond pour la détection d'un objet	71
5.3.4	Supression du bruit	71
5.3.5	Segmentation et extraction de caractéristiques	72
5.3.6	Suivi d'objet	74
5.3.7	Résultats sur une séquence d'images	76
5.4	Conclusion	76
	Conclusion et perspectives	78
	Conclusion	78
	Bibliographie	79

Table des figures

1.1	Morphologie d'arbres dendridiques	7
1.2	Neurone biologique	8
1.3	Potentiel d'action	9
1.4	Activation chimique et électrique d'un neurone	11
1.5	Une synapse	13
1.6	Intégration des signaux d'entrées en un seul signal de sortie	14
1.7	Les trois générations de réseaux de neurones	15
1.8	a.Interprétation Electrique du modèle Intègre et tire à fuite b.Interprétation Electrique du modèle Intègre et tire	18
1.9	Modèle d'un SNN	20
1.10	Topologie Feed-Forward	21
1.11	Topologie récurrente	22
1.12	Emission des Spikes	22
2.1	Profil de l'image et ses dérivées premières et secondes	30
2.2	Exemple d'un maillage carré et hexagonal	37
2.3	a.Erosion morphologique avec un élément structurant à 4 connexités, b.Erosion morphologique avec un élément structurant à 8 connexités	38
2.4	a. Dilatation morphologique avec un élément structurant à 4 connexités, b.Dilatation morphologique avec un élément structurant à 8 connexités	39
3.1	Points de contrôl	42
3.2	Boîtes englobantes	42
3.3	Modèle silhouette	43
3.4	Modèle d'apparence articulé	43
3.5	Modèle squelette	44
3.6	Modèle d'apparence multi-vues	45
3.7	Contour actif	52

3.8	Suivi de distribution	54
4.1	Modèle de réseau de neurones impulsionnels pour la détection de contours .	57
4.2	Image originale	61
4.3	Carte de fréquence de décharge de la couche de sortie	61
4.4	Comparaison de la carte de fréquence de décharge du neurone avec d'autres méthodes de détection de contours	62
5.1	Soustraction du fond	65
5.2	Processus du suivi d'objet	70
5.3	Carte de fréquence de décharge pour une image de la séquence	70
5.4	soustraction de fond	71
5.5	Influence du seuil	72
5.6	Soustraction de fond utilisant les cartes de contours SNN	73
5.7	Opérations morphologiques appliquée à l'image différence de contours . . .	73
5.8	Résultats de détection d'objet dans une séquence d'images	76

Liste des tableaux

Introduction générale

Les réseaux de neurones sont une classe de modèles de traitement d'information inspirée par la biologie du cerveau. Ce domaine interdisciplinaire a attiré beaucoup d'intérêt parmi des mathématiciens, physiciens, informaticiens et biologistes. Il couvre la modélisation d'un neurone isolé, les groupes de neurones ainsi que les phénomènes d'apprentissage et d'adaptation pour un but initial de mieux comprendre le fonctionnement du cerveau.

La première génération des modèles connexionniste fondée sur le neurone McCulloch et Pitts, était fondée sur des sorties digitales et synchrones (neurones binaires). La seconde génération de modèles a défini des sorties analogiques et toujours synchrones, permettant des tâches plus complexes de régression et de classification probabiliste, comme illustré par le célèbre perceptron multicouches. Actuellement nous voyons naître ou renaître (car là aussi les travaux fondateurs sont anciens) la troisième génération de modèles [Maa97], avec des neurones impulsionnels où le temps devient un élément explicite du modèle et l'activité est digitale et asynchrone.

Les neurones impulsionnels offrent la possibilité d'introduire naturellement une dimension temporelle dans les réseaux de neurones. Ils se présentent ainsi comme une forme particulière de système dynamique où la notion du temps joue un rôle important.

Quelques études se sont attachées à définir la puissance calculatoire des neurones impulsionnels [Maa97] et ont démontrées l'intérêt de ce type de modèle. Les travaux actuels sur des neurones impulsionnels emploient des modèles de neurones variés, souvent issus de travaux de modélisation physiologique à des niveaux de description variés. Les modèles utilisés dans des études en vue de réalisations technologiques ne s'éloignent pas beaucoup de ceux employés par les biologistes, bien que souvent choisis parmi les plus simples.

Avec la généralisation de l'utilisation d'images numériques, l'analyse du mouvement dans les vidéos s'est révélée être un outil indispensable pour des applications aussi diverses que la vidéo surveillance, la compression vidéo, l'imagerie médicale, la robotique, l'interaction homme-machine, l'analyse de séquences sportives...etc. En effet, les zones de mouvement d'une séquence d'images correspondent souvent à des événements sur lesquels un système de vision doit se focaliser.

L'analyse du mouvement est un vaste sujet qui englobe un certain nombre de probléma-

tiques. On peut notamment citer :

- la détection du mouvement, qui consiste à étiqueter chaque pixel d'une image suivant s'il correspond ou non à une région en mouvement dans la scène,
- la détection des objets en mouvement, c'est-à-dire la détection d'un ensemble de régions d'intérêt en mouvement dans la scène tridimensionnelle observée,
- la segmentation basée mouvement de la scène, pour laquelle chaque région de l'image ayant un mouvement distinct des autres est détectée et segmentée,
- l'estimation du mouvement, qui consiste à estimer, à partir d'une séquence d'images, le mouvement apparent des objets composants une scène tridimensionnelle,
- le suivi de primitives ou de régions, dont le but est de déterminer la position de chaque primitive ou région dans l'image à chaque instant,
- la reconnaissance et la modélisation d'activités ou de gestes.

Les trois premières problématiques (détection du mouvement, détection des objets en mouvement et segmentation basée mouvement), qui sont au cœur des travaux présentés, sont en général une première étape pour des outils automatiques de vision par ordinateur. Ces outils peuvent avoir pour vocation, soit uniquement de détecter, soit de détecter et reconnaître, soit de détecter et suivre des objets pour, par exemple, analyser le comportement ou la trajectoire de ces objets. Ainsi, des méthodes de détection fiables et automatiques sont indispensables pour de nombreuses applications de vision par ordinateur.

Les autres problématiques sont toutes aussi importantes et nécessitent la mise en place de méthodes simples et robustes. Tous ces sujets font l'objet d'un grand nombre de travaux, mais il n'existe pas, à l'heure actuelle, d'algorithmes aboutis s'adaptant à n'importe quelle situation.

L'analyse d'images numériques repose généralement sur un enchaînement de traitements permettant de transformer une matrice de pixels en une représentation structurée. L'un des principaux buts des traitements de bas-niveau est de détecter et de localiser avec précision les contours des objets. Cette étape s'avère cruciale dans de nombreuses applications. Un contour se matérialise par une rupture d'intensité dans l'image suivant une direction donnée. Plusieurs méthodes existent pour détecter cette rupture, les unes plus ou moins complexes, les autres plus ou moins gourmandes en calculs. L'application de détecteurs de contours sous la forme de filtres dérivateurs permet d'obtenir les contours des objets présents dans la scène. Mais ce genre de techniques est peu exploitable car elles donnent des contours non fermés, bruités ou des contours non détectés.

Contributions et description du mémoire

Le but de ce mémoire était finalement la détection et le suivi d'un objet en mouvement

dans une scène relativement statique avec un fond fixe et acquise avec une caméra immobile. L'objet est suivi par son contour, ce traitement a besoin d'une carte de contours pour chaque séquence de vidéo qui reflète la position de l'objet à chaque instant. La carte de contours doit donc être mise à jour le long de la séquence vidéo, et doit être calculée sur une durée assez courte pour ne pas influencer sur le temps d'exécution du module de suivi qui est considéré comme l'objectif principal de ce travail. L'intégration des contours de chaque image est réalisée par un modèle de réseau basé sur des neurones à impulsions qui s'inspirent du comportement des champs récepteurs biologiques.

Le processus de détection contours s'appuie donc sur la recherche dans une image les contours des objets se trouvant dans la scène à travers un modèle de réseaux de neurones impulsionnels. L'étape de suivi d'un objet mobile se fait par une approche basée sur l'emploi de la distance de Hausdorff pour comparer et déterminer la nouvelle position de l'objet localisé après l'étape de détection de contours.

Nous avons donc, dans ce contexte mis en place une application de détection de contours d'un objet en mouvement par un réseau de neurones impulsionnels et le suivi de sa trajectoire le long d'une séquence vidéo.

Les principales contributions de ce travail, qui seront présentées au cours des différents chapitres, sont les suivantes :

1. Chapitre 1 : ce chapitre est consacré à une présentation synthétique des modèles de réseaux de neurones impulsionnels, nous commençons par une description biologique et donnons en particulier une description des éléments qui font de ces neurones des unités de calcul particulières et pour lesquels la notion de temps joue un rôle important. Nous verrons donc comment les neurones répondent à la réception de trains d'impulsions ainsi que les mécanismes permettant la génération des impulsions, nous abordons ensuite les différents modèles mathématiques de Spiking Neurons qui rendent compte plus ou moins finement des aspects des neurones biologiques, nous verrons ensuite la mise en réseau des neurones impulsionnels " Spiking Neural Network " ou (SNN) qui tiennent compte explicitement de la synchronisation des entrées, et terminerons par présenter les différentes classes du code neural, la façon dont le cerveau code, transmet et traite l'information et expliquer le principe de la STDP qui correspond à la loi d'apprentissage pour les SNN.
2. Chapitre 2 : propose un état de l'art des méthodes de détection de contours et de segmentation dans une image visuelle. Quelques approches de bases ont été présentées dans ce chapitre, d'autres méthodes existent n'ont pas été présentées, il est clair que le sujet est trop vaste pour l'aborder dans son entière totalité ou même dans sa plus grande partie, le but étant de donner un éclaircissement ou une idée générale

sur les méthodes de détection de contours et de segmentation.

3. Chapitre 3 : ce chapitre présente un état de l'art des méthodes de détection et de suivi de mouvement. Nous différencions les méthodes de détection basées sur la différence inter-images, celles reposant sur une modélisation du fond, celles extrayant des couches de mouvement et finalement les approches de détection de mouvement basées sur la cohérence. Et dans le cadre du suivi, nous évoquons les méthodes de suivi par appariement de détection, le suivi par segmentation dynamique et le suivi déterministe par détection séquentielle. Toutes ces différentes techniques ont leurs propres avantages et leurs limites.
4. Les chapitres 4 et 5 : sont consacrés à la validation expérimentale des approches théoriques proposées dans les premiers chapitres de ce mémoire. Nous commencerons avec la description de l'organigramme générale de notre application et expliquer les différents modules de l'organigramme avec quelques détails. Enfin nous exposons quelques résultats de simulations appliquées à des scènes captées par un appareil photos numérique fixe.

Dans le chapitre 4 un modèle de réseau de neurones impulsionnel est utilisé pour détecter les contours dans une image visuelle. L'architecture du réseau est présentée avec quelques descriptions fonctionnelles du réseau ainsi que le modèle du neurone impulsionnel. L'efficacité de la méthode mise en œuvre a été testée sur divers types d'images, est comparable aux graphes de détection de contours de Sobel, Prewitt et Canny. Les résultats de simulation montrent que le réseau basé sur les neurones impulsionnels est capable d'effectuer la détection de contours dans un intervalle de temps de 100 ms. Cette durée de traitement est conforme à celle du système visuel humain.

Dans le chapitre 5 nous présentons le processus du suivi d'objet basé sur un modèle de contours pour la caractérisation de cible, les contours sont obtenus par un réseau de neurones impulsionnels déjà présenté dans le chapitre 4. Nous proposons de détecter et de suivre la cible par une méthode statique. La position de la cible est estimée en recherchant le modèle dans l'image courante en utilisant une distance Hausdorff. La cible est recherchée seulement dans une sous-fenêtre d'image contour courante. Le principe de l'approche mise en œuvre est détaillé dans cette partie. Il s'agit principalement de la description du principe général de l'application de suivi qui a été développé de façon à répondre aux différentes contraintes de notre contexte de l'étude menée dans les chapitres 1,2 et 3. Néanmoins nous nous intéressons aussi dans cette partie à détailler les expérimentations effectuées sur une séquence vidéo en discutant les résultats obtenus.

5. En conclusion nous résumerons les contributions et terminerons par évoquer quelques perspectives ouvertes par ces travaux, dans le contexte de détection de contours par les réseaux de neurones impulsionnels ainsi que le suivi d'objet dans une séquence d'images.

Chapitre 1

Réseaux de neurones impulsionnels

1.1 Introduction

L'intelligence artificielle est en continuelle progression depuis l'invention de l'ordinateur et l'utilisation de programmes informatiques. La biologie a apporté un grand nombre d'informations sur le fonctionnement du cerveau, des neurones...etc. Des mathématiciens ont alors tenté de reproduire le fonctionnement du cerveau en intégrant ces connaissances en biologie dans des programmes informatiques, et en leur donnant la possibilité d'apprendre. Le cerveau humain est un modèle dans le développement de l'intelligence artificielle. De simples enchevêtrements de neurones lui donnent toutes ses capacités, et c'est ce que tentent de reproduire les chercheurs grâce à l'électronique.

Le neurone, unité fonctionnelle du système nerveux, est une cellule spécialisée dans le traitement d'information. Au sein d'un réseau, il communique via des trains d'impulsions électriques [Bre03].

Nous allons présenter dans ce chapitre, l'étude sous l'angle de la biologie d'un neurone tout d'abord, ensuite nous allons présenter différentes modélisations d'un neurone impulsional ainsi que les représentations possibles des réseaux neuronaux impulsionnels (Spiking Neurons Networks) de point de vue dynamique, et terminerons par l'apprentissage de ces derniers.

1.2 Vue biologique générale

1.2.1 Cerveau humain

Le cerveau humain contient probablement plus de 100 milliards de cellules spécialisées dans le traitement de l'information : les neurones. On peut estimer à 11 à 14 milliards le nombre de cellules qui se trouvent dans le cortex, la couche supérieure du cerveau qu'on appelle aussi "matière grise" à cause de sa couleur. Chaque neurone peut être décrit comme un processeur très simple qui reçoit des signaux sous forme de potentiel électrique [Tch02], calcule la sommation des potentiels reçus par rapport à un seuil et émet un signal à son tour si le seuil est dépassé. Mais ces quelques milliards de processeurs élémentaires sont fortement interconnectés et travaillent en parallèle, ce qui fait du cerveau le système dynamique le plus complexe de l'univers.

1.2.2 Topologie des réseaux de neurones

La topologie des réseaux de neurones biologiques est très variable d'une région cérébrale à une autre (Figure 1.1). On y trouve toutes sortes de connectivités avec une très grande variété dans la densité des connections ainsi que dans la distribution des poids synaptiques [May02]. Tant structurellement que fonctionnellement, le cerveau est une structure profondément hétérogène.

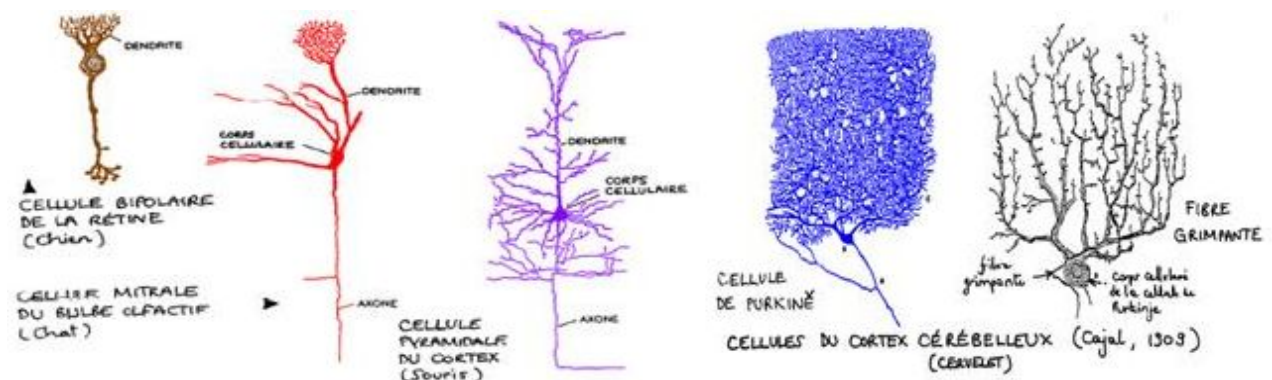


FIG. 1.1 – Morphologie d'arbres dendridiques

1.2.3 Morphologie des neurones

Les cellules neuronales mesurent entre 10 et 50 μm [Tch02]. Elles varient de forme et de propriétés anatomiques et on distingue plusieurs types de neurones. Les neurones pos-

sèdent un noyau et un corps cellulaire (soma) couvert d'une membrane comme toutes les cellules normales. La particularité du corps est qu'il se prolonge et forme des ramifications qu'on nomme "dendrites", (Figure 1.2). C'est par ces ramifications que les neurones s'interconnectent, reçoivent et transmettent l'information. La réception de l'information par la cellule se fait par de nombreuses connexions appelées synapses et l'émission par une des ramifications : l'axone, qui à son tour se ramifie en terminaisons axonales avec des synapses au bout. Les axones peuvent être enveloppés d'une matière spéciale : le myéline qui accélère plusieurs fois la vitesse de transition du potentiel électrique.

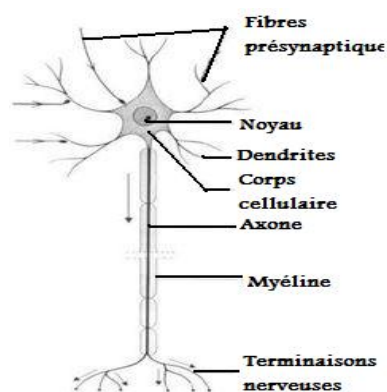


FIG. 1.2 – Neurone biologique

L'abondance des neurones est aussi un facteur de sécurité, elle permet la compensation fonctionnelle si une partie du cerveau est endommagée. Lors de lésions dans une zone corticale, il arrive que d'autres zones non atteintes récupèrent ses fonctions. C'est la fameuse plasticité neuronale qui se met en jeu pour reconstituer l'intégrité du système cognitif dans la mesure du possible.

1.2.4 Activité neuronale

Au premier plan se trouve la différence de potentiel qui existe entre l'intérieur et l'extérieur du neurone, régulée par des canaux ioniques situés sur la membrane du neurone, qui déterminent des différences de concentration en certaines espèces ioniques de chaque côté de cette membrane. Lorsque le neurone est au repos, la différence de potentiel est de -70mV à l'intérieur de la cellule [Tch02]. Lorsque le neurone est soumis à des excitations, ce potentiel membranaire va évoluer progressivement (de façon relativement linéaire par rapport à cette excitation), jusqu'à un seuil au-delà duquel on va observer une variation soudaine et importante de ce potentiel. Cette phase de comportement non-linéaire va se traduire par une excitation du neurone qui va devenir actif et transmettre un potentiel

d'action le long de son axone à destination des neurones efférents. La zone de transition entre le corps cellulaire et l'axone est d'un grand intérêt, en effet, c'est à cet endroit que l'essentiel du traitement non linéaire est effectué [Loi04].

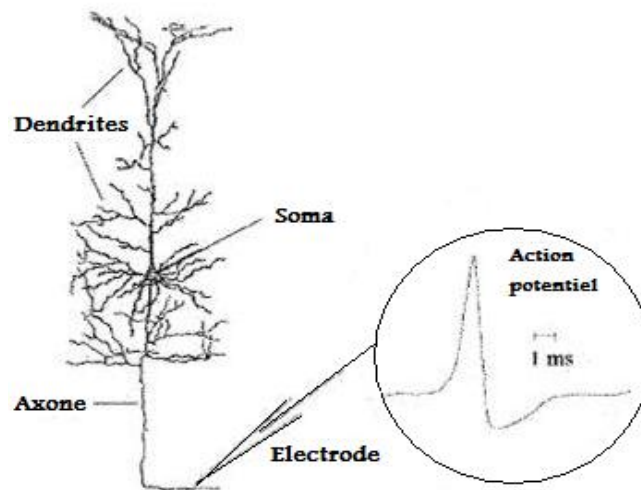


FIG. 1.3 – Potentiel d'action

Ce potentiel d'action ou spike est de courte durée (1ms) et de forme et d'amplitude invariable (Figure 1.3), même au cours de sa transmission le long de l'axone. Juste après cette émission, le potentiel retombe sous le niveau de repos (hyper-polarisation) et après une phase réfractaire où le neurone n'est plus excitable, il rejoint son niveau de repos et peut reprendre le cycle.

C'est l'arrivée d'un potentiel d'action dans les terminaisons axonales qui est à l'origine de la libération de neurotransmetteurs dans la fente synaptique, qui seront recueillis par le neurone post-synaptique et produiront un potentiel post-synaptique (PSP) au niveau de la membrane dendritique post-synaptique. Les variations du potentiel correspondent à des mouvements ioniques (principalement sodium Na^+ et potassium K^+ , mais aussi chlore Cl^- et calcium Ca^{++}) dans des canaux transmembranaires spécifiques à chaque type d'ions.

La nature inhibitrice ou excitatrice de la synapse permettra de distinguer entre potentiel post-synaptique inhibiteur ou excitateur (IPSP ou EPSP). Ce type de potentiel est de faible amplitude, relativement étendu dans le temps et retardé par rapport au temps de réception du spike. On détaillera par la suite les phénomènes cités dans ce paragraphe.

1.3 Neurone biologique

Afin d'imiter les réseaux de neurones biologiques comme le cerveau, il est important de comprendre le fonctionnement d'un neurone biologique ainsi que les modules des réseaux neurologiques. En outre, la majeure partie de la terminologie utilisée dans les réseaux de neurones artificiels provient de leurs contreparties biologiques [Boo04]. Dans cette section nous présenterons les neurones biologiques et expliquerons comment ils fonctionnent.

1.3.1 Potentiel de la membrane

On appelle potentiel membranaire, la différence de potentiel électrique entre l'intérieur du neurone et l'extérieur.

1.3.2 Potentiel de repos

Lorsqu'un neurone est au repos, on constate qu'il existe une différence de potentiel, typiquement d'environ -70 mV, entre l'intérieur et l'extérieur de la cellule. On dit alors que le neurone est polarisé. Cela est dû à une différence de concentration ionique entre les milieux intracellulaire et extracellulaire [Bre03], et à une perméabilité sélective de la membrane à certaines espèces d'ions. Les différences de concentration sont maintenues par des protéines de la membrane appelées pompes à ions, qui transportent des ions à l'encontre de leur gradient de concentration. D'autres protéines, les canaux ioniques, laissent passer spécifiquement certains ions. L'intérieur était donc chargé négativement et l'extérieur positivement, c'est-à-dire qu'il y a plus de charges positives à l'extérieur qu'à l'intérieur.

1.3.3 Potentiel local

Lorsqu'une synapse d'un neurone pré-synaptique, excite une dendrite d'un neurone post-synaptique, cela déclenche une petite modification du potentiel de membrane tout autour de cette dendrite. Il augmente en moyenne, de 1 mV. Cette petite modification s'appelle potentiel local. Les potentiels locaux créés s'additionnent (le résultat n'est pas une sommation arithmétique) et iront exciter le corps cellulaire de ce neurone, lui demandant d'envoyer un potentiel d'action si le potentiel membranaire du corps cellulaire a atteint le seuil critique (de l'ordre de -30 mV), C'est qu'aussitôt qu'on crée un déséquilibre dans la cellule, l'intensité de la réponse est beaucoup plus grande que celle de ce qui a provoqué ce déséquilibre.

Il existe aussi des potentiels locaux inhibiteurs (de -1mV d'intensité en moyenne) qui eux, contrairement aux potentiels locaux excitateurs, iront demander au corps cellulaire de ne pas envoyer de potentiel d'action. Le corps cellulaire vérifiera enfin si son potentiel membranaire a atteint le seuil de déclenchement, après avoir été inhibé et excité par les potentiels locaux reçus.

Lorsque ces potentiels locaux, déclenchés par une synapse connectée à une dendrite, sont rendus à la dendrite, ils commencent à perdre de leur intensité.

1.3.4 Potentiel d'action

En plus de messages chimiques communs à toutes les sortes de cellules, les neurones présentent la particularité de convoier des messages sous forme électrique. Cette propriété se traduit par l'émission de potentiels d'action (PA, ou spike) [Tch02]. au niveau du corps cellulaire (Figure 1.4).

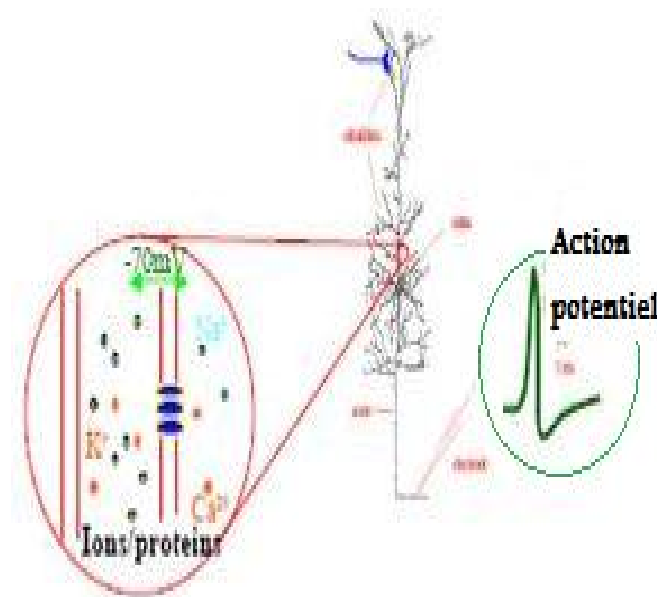


FIG. 1.4 – Activation chimique et électrique d'un neurone

Un potentiel d'action est un signal électrique, bref (durant une ou deux millisecondes), de forte amplitude (100 mV) et relativement stéréotypé (de forme et d'amplitude invariable) [Roc04].

Lorsque l'on injecte un courant assez faible dans le neurone, son potentiel se modifie en proportion du courant. Cependant, si le courant est suffisamment fort, on constate une montée brutale du potentiel (dépolariation), jusqu'à une valeur typique de $+30\text{ mV}$, on

dit que la membrane du corps cellulaire a atteint son seuil de déclenchement [Gil03], le corps cellulaire le réveille, en déclenchant un potentiel d'action de 100 mV d'intensité, et fait donc passer le potentiel membranaire de l'axone à 30 mV ($-70 \text{ mV} + 100 \text{ mV} = 30 \text{ mV}$).

Ce signal électrique est ensuite propagé le long de l'axone, et peut être acheminé sur de longues distances, jusqu'aux sites de jonction entre les neurones, les synapses. Une fois le PA émis, il est suivie d'une redescente rapide (hyperpolarisation) : le neurone entre en période réfractaire, c.à.d. qu'il ne peut pas émettre de nouveau PA pendant un certain temps, quelle que soit la stimulation qui lui parvient. Cela signifie qu'au-dessus d'un certain seuil, la perméabilité de la membrane change.

1.3.4.1 Caractéristiques du potentiel d'action

En plus des caractéristiques citées, un potentiel d'action possède d'autres caractéristiques telles que :

- Seuil de déclenchement : La première caractéristique du potentiel d'action est que, pour que celui-ci soit déclenché par le corps cellulaire, ce dernier doit avoir atteint un potentiel membranaire de -30 mV . Ceci est donc le seuil de déclenchement. Pour ce faire, les dendrites, connectées aux synapses des neurones pré-synaptiques, peuvent se faire soit exciter, soit inhiber par ces synapses. Lorsque le corps cellulaire reçoit plus de potentiels locaux excitateurs qu'inhibiteurs de ses dendrites, et qu'il atteint son seuil de déclenchement, il envoie un potentiel d'action sur l'axone. Le neurone est alors dépolarisé. Par contre, s'il reçoit plus de potentiels locaux inhibiteurs qu'excitateurs, et de ce fait, n'atteint pas son seuil, il n'envoie rien. Le neurone est alors hyperpolarisé.
- Loi du tout ou rien Le potentiel d'action est toujours de nature tout ou rien. Au moment où il est déclenché par le corps cellulaire sur l'axone, il garde toujours son intensité de 100 mV . Elle ne diminue donc jamais, qu'importe la distance avec laquelle le potentiel d'action doit parcourir.
- Période réfractaire Après qu'un potentiel d'action créé par le corps cellulaire, se soit propagé dans l'axone et dans ses terminaisons axonales, ces deux composants du neurone entrent en période réfractaire, d'une durée de 0.5 à 1 milliseconde. L'axone et ses terminaisons axonales se repolarisent donc, pour pouvoir accueillir et bien conduire un nouveau potentiel d'action

1.3.5 Connexions synaptiques

Les synapses correspondent aux jonctions entre les neurones. A ce niveau, le message électrique convoyé par l'intermédiaire d'un potentiel d'action est transformé en message chimique. Une synapse se compose de trois éléments : la membrane pré-synaptique, la fente synaptique, et la membrane post-synaptique.

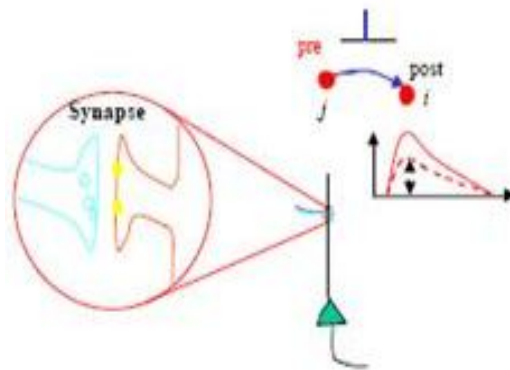


FIG. 1.5 – Une synapse

La membrane pré-synaptique est composée des vésicules, contenant des molécules chimiques, les neurotransmetteurs. L'arrivée d'un potentiel d'action a pour effet de libérer une certaine quantité de ces neurotransmetteurs dans la fente synaptique. Une fois cet espace traversé, les neurotransmetteurs vont se fixer, au niveau post-synaptique, sur des récepteurs, spécifiques à chaque type de neurotransmetteur. Cette association entre un neurotransmetteur et un récepteur induit l'entrée d'ions (K^+ , Ca^{2+}) dans la membrane. Ces ions sont responsables de la naissance d'un phénomène électrique au sein du neurone post-synaptique, le potentiel post-synaptique (PPS).

Le type de neurotransmetteur détermine le signe de la variation du potentiel dans le neurone post-synaptique et donc la nature excitatrice ou inhibitrice de la synapse.

1.3.6 Propriétés d'intégration

Les dendrites convoient ensuite cette information électrique jusqu'au corps cellulaire. Les signaux ont une amplitude (quelques mV) plus faible que les potentiels d'action, et le décours temporel de chaque signal dépend de la nature excitatrice ou inhibitrice du neurotransmetteur, de la morphologie de la synapse et de sa distance au corps cellulaire. Au niveau du corps cellulaire, les signaux provenant de l'ensemble des synapses sont intégrés en un signal unique [Meu07] (Figure 1.6), le potentiel de membrane, qui correspond à la sommation des différents signaux provenant des dendrites.

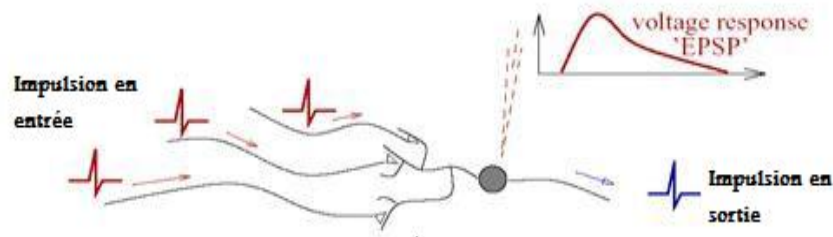


FIG. 1.6 – Intégration des signaux d’entrées en un seul signal de sortie

En l’absence de stimulation, le potentiel de membrane est à son potentiel de repos. Si le potentiel de membrane dépasse un certain seuil, alors le neurone émet un potentiel d’action, ce potentiel crée par les neurotransmetteurs s’appelle potentiel post-synaptique (PSP -Post-Synaptique Potential). Selon son signe il est excitateur (EPSP -Excitatory PSP-) ou inhibiteur (IPSP -Inhibitory PSP-) [May02].

1.4 Générations des modèles connexionnistes

1.4.1 Les trois générations de modèles connexionnistes

L’évolution des modèles connexionnistes est bien capturée dans la catégorisation proposée par Mass, qui classe les modèles neuromimétiques en trois générations [Roc04] selon les types de neurones qu’ils utilisent (Figure 1.7) :

1. Première génération (G1) : les réseaux basés sur des neurones de McCulloch et Pitts à sortie binaire. Ils peuvent calculer n’importe quelle fonction booléenne avec un seul niveau d’unités cachées. La simplicité des neurones de McCulloch et Pitts ne veut pas dire que tous les RNA qui les utilisent sont simples et faciles à analyser. Les réseaux de première génération incluent aussi les réseaux de Hopfield et la Machine de Boltzmann qui sont de type récurrent.
2. Deuxième génération (G2) : les réseaux basés sur des neurones à sortie analogique. Introduits par le groupe Parallel Distributed Processing, ces réseaux peuvent aussi calculer n’importe quelle fonction, mais sont aussi sensibles aux règles d’apprentissage par tracé de l’erreur.
3. Troisième génération (G3) : les réseaux utilisant des neurones qui pulsent et qui peuvent modéliser des phénomènes spatio-temporels. Ces réseaux sont neurobiologiquement les plus réalistes.

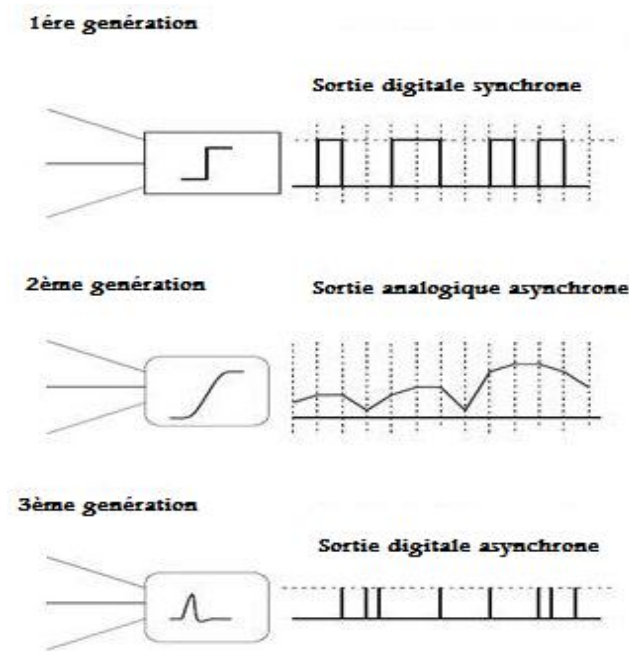


FIG. 1.7 – Les trois générations de réseaux de neurones

Des modèles hybrides, combinant les architectures multicouches des réseaux G2 avec des unités à pulsation des réseaux G3 existent aussi.

1.4.2 Réseaux de 3ème génération

La troisième génération de réseau de neurones artificiels selon Maass, utilise des neurones à pulsations qui miment le comportement des neurones réelles. Historiquement, les réseaux de troisième génération ont été proposés très peu de temps après les réseaux G2 et se positionnaient comme une alternative aux réseaux feed-forward [Roc04].

Les travaux actuels sur les réseaux de neurones impulsionnels emploient des modèles de neurones variés, souvent issus de travaux de modélisation physiologique à des niveaux de description variés et sont souvent choisis parmi les plus simples. Ce type de neurones sont considérés comme des unités de calculs particulière, pour lesquels la notion du temps joue un rôle important, les neurones impulsionnels offrent la possibilité d'introduire naturellement une dimension temporelle dans les réseaux de neurones. Ils se présentent ainsi comme une forme particulière de système dynamique.

Nous expliquerons par la suite les différents modèles de neurones impulsionnels existant ainsi les mécanismes leur permettant la génération d'impulsions et la manière dont ce type de neurones répondent à la réception de trains d'impulsions.

1.5 Modélisation des neurones impulsifs

Les modèles récents de traitement neuronal de l'information portent le nom de modèles de " Spiking Neurons ", à l'encontre des modèles " classiques ", ceux-ci prennent en compte le fait que la sortie des neurones réels n'est pas caractérisée par une fonction continue, mais par une séquence discrète d'éléments indivisibles : les spikes.

Il existe plusieurs modèles " mathématiques " de neurones, qui rendent compte plus ou moins finement des aspects des neurones biologiques décrits dans la partie précédente, les plus utilisés sont : le modèle de Hodgkin-Huxley (HH), les neurones à pulses intègre-et-tire, les modèles de réponse synaptique.

1.5.1 Modèle à base de conductance de Hodgkin-Huxley

Une des plus grandes découvertes scientifiques en neurosciences a été la découverte par Hodgkin et Huxley (HH) [HH52] du processus cellulaire de génération d'un PA sur l'axone géant du calmar. Celle-ci, récompensée par un prix Nobel, est remarquable autant pour ses résultats que par ses méthodes, et reste aujourd'hui encore une référence. Ils ont pu modéliser l'évolution du potentiel de membrane par un jeu d'équations différentielles [Roc04].

$$C \frac{d\nu}{dt} = I - I_{Na} - I_k - I_L \quad (1.1)$$

Où, C est la capacité de la membrane, ν son potentiel et le courant I décrit l'ensemble des entrées du neurone. On pose :

$$I_{ion}(t) = G_{ion} \cdot (\nu(t) - E_{ion}) \quad (1.2)$$

Pour une modélisation plus précise, on remplace :

$$C \frac{d\nu}{dt} = I(t) - g_{Na} m^3 h (\nu - E_{Na}) - g_K n^4 (\nu - E_K) - g_L (\nu - E_L) \quad (1.3)$$

Les variables m, h et n décrivent l'ouverture et la fermeture des canaux et sont régis par les équations différentielles suivantes :

$$\frac{dm}{dt} = \alpha_m(\nu)(1 - m) - \beta_m(\nu)(m) \quad (1.4)$$

$$\frac{dn}{dt} = \alpha_n(\nu)(1 - n) - \beta_n(\nu)(n) \quad (1.5)$$

$$\frac{dh}{dt} = \alpha_h(\nu)(1 - h) - \beta_h(\nu)(h) \quad (1.6)$$

α_i et β_i pour $i=m,n,h$ définissent respectivement les vitesses d'ouverture et de fermeture des canaux ioniques et sont comprises entre 0 et 1.

Le modèle de Hodgkin et Huxley permet reproduire au mieux le fonctionnement du neurone biologique et d'expliquer de façon satisfaisante, le fonctionnement de la génération des impulsions, seulement ce type de modèle est trop complexe pour espérer s'en servir dans les réseaux de neurones à grande échelle. Toutefois, la complexité des mécanismes mis en jeu et la difficulté de manipuler ces systèmes d'équations ont conduit plusieurs auteurs à proposer des modèles plus simples tout en gardant le maximum de pertinence biologique.

1.5.2 Modèle impulsif intègre-et-tire et le modèle LIF

Un modèle impulsif est un modèle dont l'objet est de décrire la série d'impulsions générée par un neurone, une impulsion étant considérée comme instantanée et binaire. Autrement dit, la "sortie" du modèle est une suite d'instantanés, le temps des potentiels d'action.

Les modèles impulsifs sont souvent appelés " Intègre-et-tire " (Integrate-and-Fire), ce qui illustre la caractéristique fondamentale de ces neurones : le traitement de l'entrée et le potentiel d'action, plus rapide, sont traités séparément [Roc04]. L'objet du modèle n'est pas de comprendre comment le potentiel d'action est généré, mais de décrire les séries d'impulsions produites en fonction de l'entrée.

Historiquement, le premier modèle impulsif fut introduit par Louis Lapicque en 1907 (appelé aussi intégrateur à fuite) se présente sous la forme d'une équation différentielle linéaire [Roc04] :

$$C \frac{d\nu}{dt} = -g_L(\nu - \nu_0) - I(t) \quad (1.7)$$

Où ν est le potentiel membranaire, C est la capacité membranaire (due à la présence de charges électriques de part et d'autre de la membrane), ν_0 est le potentiel de repos (de l'ordre de -70 mV), g_L est la conductance de fuite et $I(t)$ est le courant injecté.

Cette équation modélise le comportement électrique du neurone sous le seuil (Figure 1.8). De plus le temps d'émission d'un PA est défini comme le temps où le potentiel de membrane croise la valeur du seuil $\nu(t^f) = \nu_S$ (seuil), avec la condition $\dot{\nu}(t^f) > 0$ juste après t^f , le potentiel est réinitialisé une valeur ν_{abs} (généralement proche de ν_0) durant une période d_{abs} afin de modéliser une période réfractaire absolue.

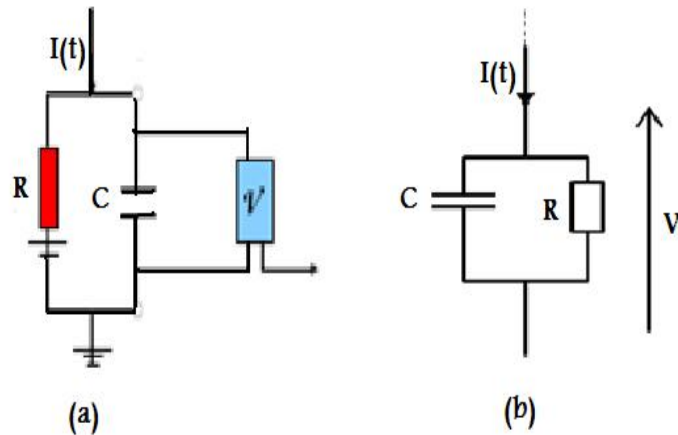


FIG. 1.8 – a.Interprétation Electrique du modèle Intègre et tire à fuite b.Interprétation Electrique du modèle Intègre et tire

Il existe un modèle impulsionnel plus simple encore que le modèle de Lapique : c'est l'intégrateur parfait. On omet le terme de fuite (le premier terme dans l'équation précédente), ce qui donne l'équation différentielle suivante :

$$C \frac{dv}{dt} = I(t) \quad (1.8)$$

Le modèle intègre-et-tire est un modèle simple utilisant le principe du seuil d'émission proposée précédemment, cette équation différentielle décrit l'évolution du potentiel (sous le seuil), son activité interne est mesurée simplement par son potentiel de membrane, et à l'image du neurone biologique, comprend une partie intégratrice (l'arbre dendritique), qui va correspondre à une somme linéaire des entrées pondérée par les poids synaptiques et une partie non-linéaire qui simule la génération d'un PA quand dépasse un seuil : on dit alors que le neurone "tire". Celui-ci est alors inhibé (pendant la période réfractaire) en fixant durant cette période son potentiel à une valeur dépolarisée.

1.5.3 Neurone à réponse impulsionnelle (SRM)

Le modèle SRM (Spike Response Model) proposé par W.Gerstner [Ger91] est une généralisation du LIF, qui intègre le passé du neurone, il permet une modélisation des réponses aux impulsions via l'utilisation des noyaux modélisant directement la forme des PPS.

Le principe de ce modèle est de mémoriser la réponse impulsionnelle, c'est-à-dire la réponse à un PA de la synapse (le PPS) depuis un modèle ou depuis des enregistrements neurophysiologiques.

Dans le modèle SRM le calcul effectué au niveau du neurone est une sommation de PPS ayant un décours temporel précis [HM07].

Le potentiel de membrane u au temps t est considéré comme la somme de deux termes : le premier terme correspond à la contribution des PPS arrivés au soma du neurone avant le temps t . Le second terme exprime la contribution des périodes réfractaires rencontrées par le neurone avant le temps t . On définit :

$F_j = \{t_j^{(f)}; 1 \leq f \leq n\} = \{t \mid u_j(t) = \nu \wedge \dot{u}_j(t) > 0\}$ comme l'ensemble des instants d'émission du neurone N_j .

$\mathcal{T}_j = \{i \mid N_i \text{ est amont de } N_j\}$ comme l'ensemble des neurones pré-synaptiques.

L'état du neurone N_j à t est alors défini par :

$$u_j(t) = \sum_{t_j^{(f)} \in F_j} w_{ij} \eta(t - t_j^{(f)} - d_{ij}^{ax}) + \sum_{i \in \mathcal{T}_j} \sum_{t_i^{(f)} \in F_i} w_{ij} \epsilon(t - t_i^{(f)} - d_{ij}^{ax}) \quad (1.9)$$

d_{ij}^{ax} est le délai de transmission axonale ; w_{ij} est le poids de la synapse, entre les neurones N_i et N_j , avec η et ϵ sont les fonctions noyaux, la première modélise la remise à zéro du potentiel après l'émission d'un PA, la seconde décrit la forme des PPS.

On a alors émission d'un PA lorsque : $t_j^{(n+1)} = t \iff u_j(t) = \vartheta$ la valeur du seuil

On prend habituellement, pour le neurone N_j , les expressions suivantes pour les fonctions noyaux :

$$\eta_j(s) = -\eta_0 \exp\left(-\frac{s - \delta^{abs}}{\tau}\right) H(s - \delta^{abs}) - KH(s)(\delta^{abs} - s) \quad (1.10)$$

Où : H et τ sont respectivement la fonction de Heaviside et une constante de temps.

δ^{abs} et η_0 sont respectivement la période réfractaire absolue, et l'amplitude de la période réfractaire.

$$\epsilon(s) = \left[\exp\left(-\frac{s}{\tau_m}\right) - \exp\left(-\frac{s}{\tau_s}\right)\right] H(s) \quad (1.11)$$

où τ_m et τ_s sont des constantes de temps.

Une variante plus simple du modèle, appelée le modèle SRM_0 , ne prend en compte que le dernier temps d'émission \hat{t}_j du neurone N_j pour la période réfractaire. On remplace alors la somme dans la formule (1.9) par une contribution unique :

$$u_j(t) = \sum_{i \in \mathcal{T}_j} w_{ij} \epsilon(t - \hat{t}_i - d_{ij}^{ax}) + \eta_j(t - \hat{t}_j) \quad (1.12)$$

avec émission d'un PA si $u_j(t) = \vartheta \implies \hat{t}_j = t$.

Bien que seulement descriptif, ce modèle peut simuler avec précision la réponse d'un neurone biologique. De plus, ce modèle est linéaire et, contrairement au modèle HH,

permet une implantation et une étude mathématique simple de la transformation du neurone. Il permet aussi par rapport au modèle LIF de décrire sans équations différentielles un comportement similaire en réduisant ainsi le nombre de variables en les remplaçant par des variables exogènes.

1.6 Réseaux de neurones impulsionnels

Les modèles de "Spiking Neurons" à l'encontre des modèles "classiques", prennent en compte le fait que la sortie des neurones réels n'est pas caractérisée par une fonction continue, mais par une séquence discrète d'éléments indivisibles appelés les spikes (Figure 1.9).

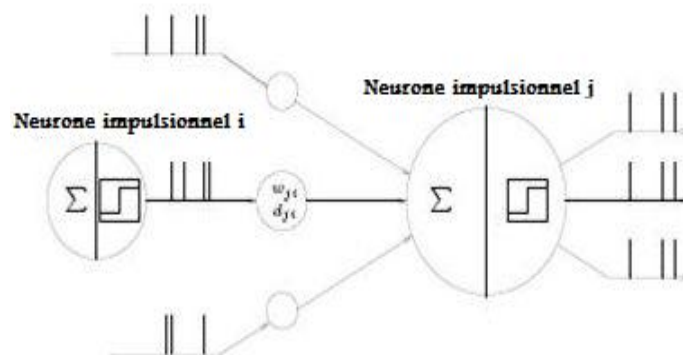


FIG. 1.9 – Modèle d'un SNN

Ce type de réseau de neurones appelé aussi " Spiking Neural Network " (SNN) sont des modèles qui tiennent compte explicitement de la synchronisation des entrées, l'entrée et la sortie du réseau sont habituellement représentés par des séries de temps de décharges appelées les trains d'impulsions ou " spike-trains " [Boo04].

1.6.1 Topologie et dynamique d'un SNN

Les mécanismes de réception d'impulsions déjà traités, nous permettent de modéliser la relation entre 2 neurones (pré-synaptique et post-synaptique), ces mécanismes peuvent être initiés de façon continue (tels que le modèle de HH), ou apparaître à un temps bien défini (tels que modèle SRM). Il est par ailleurs possible de cohabiter les 2 modèles ensemble dans un même réseau.

1.6.1.1 Connexion et délais

Dans un réseau de neurones impulsionnels, la connexion entre neurones se caractérise par l'ajout de délais, vue la notion du temps importante dans ce type de réseaux. Ce délai peut être lié à la synapse donc au neurone post-synaptique [Roc04]. De ce fait, il est possible de construire entre deux neurones plusieurs connexions, chacune avec un délai différent.

1.6.1.2 Topologie

Les topologies communes des réseaux de neurones classiques sont des réseaux de neurones de type feed-forward ou récurrents. Ces deux topologies peuvent aussi être utilisées pour les réseaux de neurones impulsionnels [Sch08] :

1. **Topologie Feed-Forward** : Cette architecture consiste en une topologie de réseaux classique de Feed-forward, cette topologie a été introduite par plusieurs auteurs. Dans ce type de topologie, pour chaque connexion entre 2 neurones consiste plusieurs connexions synaptiques, chacune est portée avec des poids et des délais différents le schéma suivant (Figure 1.10) illustre cette architecture.

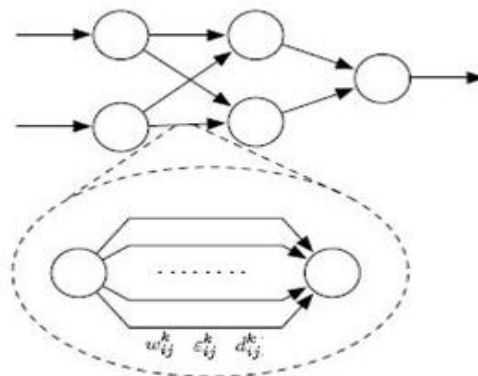


FIG. 1.10 – Topologie Feed-Forward

2. **Topologie récurrente** :

Dans les réseaux de neurones impulsionnels, les topologies récurrentes sont plus complexes (Figure 1.11), cependant elles sont bien adaptées à résoudre les problèmes temporels en raison de leurs qualités dynamiques temporelles intrinsèques. Mais il apparaît que sans l'étude d'une règle d'apprentissage appropriée, nous ne pouvons pas tirer profit de cette dynamique [Sch08].

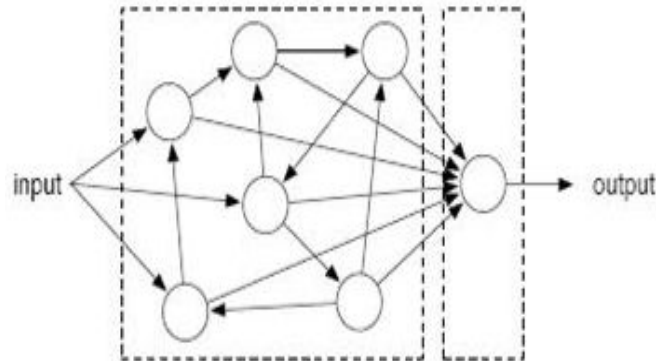


FIG. 1.11 – Topologie récurrente

1.6.1.3 Dynamique neuronale

Les différentes topologies citées en haut permettent de définir la dynamique neuronale. En fait, Les neurones encodent l'information dans les instants d'émission des impulsions [HM07]. Chaque spike pré-synaptique engendre un EPSP ; ceux-ci s'ajoutent (Figure 1.12) ; quand son potentiel atteint le seuil, le neurone post-synaptique N_j émet à son tour un spike.

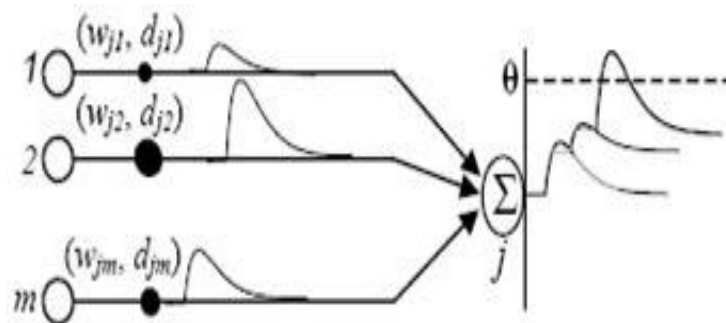


FIG. 1.12 – Emission des Spikes

1.6.2 Codage de l'information par les neurones

Le code neural, la façon dont le cerveau code, transmet et traite l'information, reste depuis toujours une question fondamentale dont les tentatives de réponses ne cessent d'être remises en cause [Ger91], [Nat99]. La recherche sur les différents modes de codage est particulièrement intense et plusieurs hypothèses sont en circulation. Une vision relativement ancienne supposait que l'information portée par le neurone était le taux de décharge du neurone (rate code). Plus tard, des travaux de recherche ont montré la nécessité de

prendre en compte l'instant précis des impulsions plutôt que le taux de décharge, Il est cependant important de noter qu'il n'existe probablement pas un unique code neural mais qu'il diffère vraisemblablement d'une région cérébrale à une autre.

De ce fait, on peut classer le codage neural en deux classes : le codage par le taux de décharge et un codage par les impulsions [Nat99].

1.6.2.1 Codage par le taux de décharge

Un bref aperçu sur la littérature expérimentale indique qu'il n'y a pas un concept bien défini ou un unique code du taux de décharge. En fait, il y a au moins trois notions différentes du taux qui sont souvent confuses et peuvent être utilisées simultanément [Nat99].

1. Cadence par le compte des impulsions

Le taux de décharge est habituellement défini par une moyenne temporelle et est essentiellement le calcul du nombre d'impulsions dans un intervalle de temps T divisé par la grandeur de cet intervalle [Loi04].

2. Cadence par la densité des impulsions

C'est une définition du taux de décharge qui est une moyenne sur plusieurs exécutions, elle fonctionne pour des stimulus stationnaires ou dépendants du temps [Loi04]. Dans ce cas, l'expérimentateur enregistre les impulsions générées par les neurones en simulant ceux-ci à l'aide d'une séquence d'entrées. Cette séquence est répétée plusieurs fois et le résultat est placé dans un histogramme. Pour chaque intervalle de temps court avant, pendant et après la simulation, l'expérimentateur compte le nombre de fois qu'une impulsion est produite et les additionne sur toutes les répétitions de l'expérience. Ensuite, le nombre d'occurrences des impulsions est divisé par le nombre de répétitions et par l'intervalle de l'histogramme pour obtenir la densité pour cet intervalle.

3. Cadence par l'activité d'une population

Il est possible de définir la cadence des impulsions à l'aide d'un ensemble (une population) de neurones. Celle-ci est justifiée, car souvent plusieurs neurones possèdent des propriétés et des réponses aux stimuli similaires [Loi04].

1.6.2.2 Codage par impulsions

Dans les années récentes, la mise en évidence de plusieurs expériences a montrée que le concept de taux de décharge basé sur la moyenne temporelle décrit de façon plus réelle

l'activité de cerveau.

Du point de vue du traitement de l'information nerveuse, un neurone peut être considéré comme une unité dynamique traitant des événements brefs. L'information nerveuse est ainsi codée dans les séquences temporelles de PA et elle est répartie sur l'ensemble du réseau [Boo04].

Nous allons présenter dans ce qui suit certaines possibilités d'un codage par des impulsions [Ger91],[Nat99].

1. Temps de la première impulsion

Est une méthode de codage par impulsions. Dans la version la plus pure de cette méthode, uniquement la première impulsion de chaque neurone compte. Toutes les impulsions subséquentes sont jugées inutiles [Loi04]. De la même façon, on peut assumer que chaque neurone n'émet qu'une impulsion par saccade et qu'elles sont par la suite inhibées. Il est évident que dans un tel cas, l'information est située dans la chronologie et non dans le nombre des impulsions.

2. Codage par ordre de rang

Est une stratégie inspirée du système visuel au niveau duquel de nombreux types de codes auraient de la difficulté à opérer à la vitesse impressionnante que l'on retrouve pour traiter certaines informations. Pour ce type de code, l'information est distribuée à travers une grande population de neurones et elle est représentée par le temps de décharge relatif des impulsions dans une seule vague de potentiels d'action [PK06].

3. Corrélacion et la synchronie

Peuvent aussi jouer un rôle dans le codage. Il est aussi possible d'utiliser d'autres neurones comme signal de référence pour un code impulsional. Par exemple, la synchronisation entre deux ou plusieurs neurones aurait une signification particulière [Loi04]. D'un autre côté, on pourrait utiliser non seulement la synchronisation, mais aussi des patrons spatio-temporels précis des impulsions. Il est à noter que la corrélation des neurones auditifs est dépendante du stimulus et pourrait fournir de l'information au-delà de la cadence des décharges.

1.7 Apprentissage dans les réseaux de neurones impuls ionnels

L'apprentissage joue un rôle important dans la dynamique des réseaux de neurones en général, ceci est obtenu en modifiant les interactions entre les neurones du réseau qui interviennent au niveau des paramètres libres du réseau essentiellement les poids synaptiques, et que ces paramètres influent sur la dynamique neuronale. Par conséquent une loi d'apprentissage doit intervenir sur la relation entre les neurones, Cette loi ou règle d'apprentissage doit jouer un rôle particulier qui consiste à les modifier pour obtenir un comportement désiré, ce phénomène s'appelle " la plasticité synaptique" [PK06].

1.7.1 Plasticité synaptique

Pour les réseaux de neurones impuls ionnels, Une loi d'apprentissage biologiquement plausible : la STDP (Spike-Time Dependent Plasticity). Cette règle de plasticité permet de renforcer les synapses dont les temps d'émissions des neurones pré et post-synaptiques suivent un enchaînement causal. Elle correspond à une forme hebbienne de plasticité [Meu07],[HM07].

Ainsi comme le présupposait Hebb, c'est l'activité même du réseau qui va déterminer les modifications conduisant à l'apprentissage. Les modifications d'une synapse ne suivent pas une règle globale contrôlée par une entité centrale, mais résultent d'une organisation spontanée due à un phénomène local se basant uniquement sur l'activité des neurones pré- et post-synaptiques [Cou06].

Plusieurs mécanismes biologiques participant à cette plasticité synaptique ont été mis en évidence, C'est l'un d'entre eux, la Spike Timing Dependent Plasticity (STDP), qui semble être la clé de l'apprentissage, les autres correspondant plutôt à des mécanismes d'ajustement permettant un bon fonctionnement du premier [Cou06].

1.7.1.1 Assemblées cellulaires

La modification postulée, au niveau synaptique, par la règle de Hebb permet d'associer des stimuli présentés de manière simultanée, et donc de créer des associations entre des caractéristiques de natures différentes au sein d'une même modalité, ou entre différentes modalités, par le fonctionnement en assemblées cellulaires (cell assemblies) [Meu07].

Si l'on considère qu'un groupe de neurones est activé par la présentation d'un stimulus, alors les liens entre ces neurones vont être renforcés. Cette notion permet d'expliquer

comment une représentation distribuée dans le cortex peut être générée par l'activation spécifique d'un groupe de neurones.

La répétition des présentations du même percept consolide l'assemblée qui le représente, permettant ainsi de garder une trace de plus en plus forte de ce stimulus dans le système. Dès lors, la représentation d'un stimulus dans le système nerveux va pouvoir être plus facilement rappelée.

Ainsi, la formation puis la consolidation des assemblées cellulaires auraient un lien avec la plasticité du système nerveux : le renforcement des liens synaptiques serait responsable du stockage des engrammes, c.à.d. le support des souvenirs. Ce fonctionnement en assemblées permet d'avoir une forme explicite d'engrammes, c.à.d. de représentations des souvenirs au sein du système nerveux, correspondant à l'ensemble de ces liens renforcés [Meu07].

1.7.1.2 Plasticité synaptique par taux de décharge

Dans l'énonciation de la règle originale de Hebb, il n'est pas fait mention de diminution de la force des liens synaptiques. Un processus passif de régulation peut permettre de maintenir l'efficacité totale des synapses à une valeur stable, en diminuant l'efficacité des synapses n'ayant pas été potentiées par la règle de Hebb [Meu07].

Cependant, un mécanisme actif a été mis en évidence expérimentalement appelé dépression à long terme (DLT) : si le taux de décharge des deux neurones n'est pas suffisant (i.e. ne dépasse pas un certain seuil), alors l'efficacité de la synapse sera diminuée.

1.8 Conclusion

Ce chapitre fait l'étude des réseaux de neurones impulsionnels, ces réseaux qui ressemblent de près aux neurones biologiques par leurs comportements et leurs fonctionnements, sont actuellement très exploités, ils permettent de tenir compte du fonctionnement temporel du neurone de manière essentielle sans s'intéresser aux mécanismes sous-jacents. Cependant leurs modélisations imposent un compromis entre la complexité de calcul et la plausibilité biologique.

La réalisation de tels réseaux de neurones répond à deux besoins. D'une part, la réalisation de réseaux simples permet d'étudier certaines propriétés collectives de ces neurones. D'autre part, des réseaux de taille importante sont également développés pour réaliser des applications. Si on cherche à simplifier ces modèles, c'est aussi pour pouvoir les simuler plus simplement et également en plus grand nombre.

Il convient tout d'abord de mentionner que la famille de modèles que nous avons présentée ici est en fait une extension des modèles neuronaux classiques dans la mesure où, si l'on cherche l'état stable des équations différentielles mentionnées ici, on retrouve les équations habituelles des neurones artificiels classiques. Par leur aspect dynamique, ces modèles sont plus coûteux, mais aussi plus délicats à simuler. Cependant, ils offrent des caractéristiques inédites et très puissantes en intégrant des propriétés nouvelles du fonctionnement neuronal et en particulier des caractéristiques de calcul temporel asynchrone qui permettent de nouvelles perspectives pour le traitement et l'analyse des données.

Les réseaux de neurones impulsionnels sont donc plus puissants que leurs prédécesseurs non impulsionnels, ils peuvent coder l'information temporel dans leurs signaux, mais ont besoin biologiquement plus de règles pour leur plasticité synaptique.

Chapitre 2

Détection, segmentation d'images et opération morphologiques

2.1 Introduction

L'analyse d'images numériques repose généralement sur un enchaînement de traitements permettant de transformer une matrice de pixels en une représentation structurée [Elo94]. L'un des principaux buts des traitements de bas-niveau est de détecter et de localiser avec précision les contours des objets. Cette étape s'avère cruciale dans de nombreuses applications.

Les contours des objets constituent donc une information pertinente pour obtenir une description des objets. La mise en évidence des points représentant les contours d'objets dans une image peut servir à reconnaître des objets présents dans une scène, à différencier des zones de l'image, à faire de la segmentation d'images, à extraire une information réduite souvent pertinente pour caractériser l'image. Ces frontières sont généralement localisées par des variations locales significatives du niveau de gris. La détection de contours a comme objectif de mettre en évidence ces variations. De nombreuses méthodes de détection de contours ont été proposées.

Depuis quelques années, on assiste cependant à de nouvelles approches en extractions de contours qui procèdent par optimisation de critères prenant en compte un modèle prédéfini du contour à détecter [Elo94]. Ces approches qui permettent une meilleure formulation du problème se sont avérées nettement plus performantes que les toutes premières techniques proposées [Cou05] dont on peut citer les contours déformables.

L'extraction de contour est une technique très utilisée dans les domaines scientifiques et

techniques, le principe est donc essentiellement d'effacer tous les motifs à faible variation des niveaux de gris (ou de couleurs) pour ne conserver que les lignes de séparation entre régions homogènes [Cou05]. Les premiers modèles de segmentation s'appuient sur des détections de changement rapide d'intensité lumineuse ou de couleur, c'est-à-dire les contours de l'image.

L'application de détecteurs de contours sous la forme de filtres dérivateurs permet d'obtenir les contours des objets présents dans la scène. Il existe pour cela toute une gamme classique d'opérations de filtrage par convolution (notamment le filtre passe-haut). Il y a aussi les filtres différentiels (gradient, laplacien) qui donnent des valeurs élevées aux points où la variation des niveaux est rapide.

2.2 Approches de détection de contours

Un contour se matérialise par une rupture d'intensité dans l'image suivant une direction donnée. Plusieurs méthodes existent pour détecter cette rupture, les unes plus ou moins complexes, les autres plus ou moins gourmandes en calculs. Dans la plupart des cas, Elle s'applique en deux étapes : la première permet de localiser les contours à partir d'un calcul de Gradient ou de Laplacien dans des directions privilégiées. La seconde étape va permettre d'isoler les contours du reste de l'image à partir d'un seuillage judicieux.

Plusieurs méthodes permettent de déterminer le Gradient ou le Laplacien d'une image. Il en est de même des techniques de seuillage. L'application de détecteurs de contours sous la forme de filtres dérivateurs permet d'obtenir les contours des objets présents dans la scène. Nous pouvons citer les approches se basant sur les différences finies comme l'opérateur de gradient, l'opérateur Laplacien, les filtres de Sobel, Prewitt, Roberts ou bien des approches reposant sur des critères d'optimalité comme les filtres de Canny-Deriche. Mais ce genre de techniques est peu exploitable car elles donnent des contours non fermés, bruités ou des contours non détectés

2.2.1 Approche par convolution

Les contours correspondent à des maxima locaux de la dérivée première marquant les fortes variations d'intensité dans l'image. En 2D, la dérivée première correspond au gradient et on recherchera alors les points dont le module du gradient est maximal. Les contours correspondent également aux passages par zéro de la dérivée seconde (Figure 2.1) donc on peut aussi utiliser le laplacien afin de rechercher les points où il s'annule en changeant de signe [Cou05].

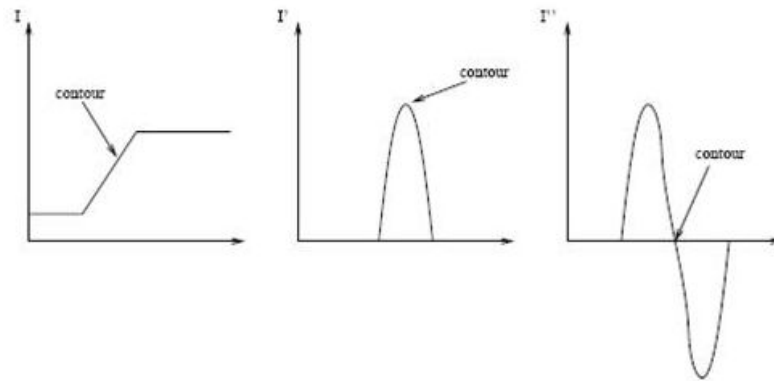


FIG. 2.1 – Profil de l'image et ses dérivées premières et secondes

Dans la plupart des techniques développées à ce jour, plusieurs opérateurs locaux (directionnels ou pas) de dérivation du premier ordre et du second ordre sont utilisés. Ils sont alors suivis respectivement d'une recherche de maxima locaux et de passage par zéro. Toutefois, ces opérateurs ne donnent souvent pas, à eux seuls, de bons résultats sur une image réelle où les changements d'intensité sont rarement nets et abrupts.

2.2.1.1 Détection des contours par dérivée première

Les filtres utilisés ici sont les filtres de dérivée première (appelés aussi filtres étroits) et l'on cherche alors le maximum de leur réponse. Leur prototype est le filtre de gradient mais la dérivation accentuant le bruit (pixels parasites de répartition aléatoire), des filtres dérivés plus robustes ont été proposés [Cou05]. Le gradient est une dérivation au premier ordre et est donné par la formule :

$$\nabla I(x, y) = \frac{\partial I}{\partial x} \hat{I}x + \frac{\partial I}{\partial y} \hat{I}y \quad (2.1)$$

$\hat{I}x$ (resp. $\hat{I}y$) est un vecteur unitaire suivant x (resp. suivant y).

Le gradient, en un pixel d'une image numérique, est un vecteur caractérisé par son amplitude et sa direction. L'amplitude est directement liée à la quantité de variation locale des niveaux de gris. La direction du gradient est orthogonale à la frontière qui passe au point considéré. La méthode la plus simple pour estimer un gradient est donc de faire un calcul de variation monodimensionnelle c'est-à-dire en ayant choisi une direction donnée. Le gradient étant un vecteur, l'approche la plus classique pour estimer le gradient consiste à choisir deux directions privilégiées (naturellement celles associées au maillage, ie : ligne et colonne) orthogonales sur lesquelles on projette le gradient. On peut donc obtenir une connaissance parfaite du gradient de l'image qui se calcule comme suit :

$$\nabla_x = \frac{\partial I(x, y)}{\partial x} \quad (2.2)$$

et

$$\nabla_y = \frac{\partial I(x, y)}{\partial y} \quad (2.3)$$

Ainsi, en chaque point (x, y) de l'image, on définit deux dérivées partielles, suivant x et suivant y . La direction du vecteur gradient maximise la dérivée directionnelle et sa norme est la valeur de cette dérivée. Le filtre le plus simple consiste à calculer les différences entre pixels voisins sur les horizontales puis sur les verticales. Chaque extremum correspond à un point d'un contour. On peut citer comme filtres utilisant la méthode du gradient pour une détection de contours : le filtre de Sobel, Prewitt, Roberts et Kirsh. Ces méthodes se caractérisent par ailleurs par une grande rapidité et un faible coût du fait de l'aspect local de ces filtres (application de petites matrices).

2.2.1.2 Détection des contours par dérivée seconde

Les filtres larges que nous allons maintenant étudier recherchent les zéros de la dérivée seconde (méthode permettant de bien mettre en évidence les maxima de la dérivée première) ou, plus précisément, du Laplacien qui est une dérivation au deuxième ordre. On décrit ici les opérateurs de type Laplacien définis comme étant :

$$\Delta I(x, y) = \frac{\partial^2 I(x, y)}{\partial x^2} + \frac{\partial^2 I(x, y)}{\partial y^2} \quad (2.4)$$

Les points de contours correspondent alors aux passages par zéro de l'image obtenue par convolution avec l'opérateur Laplacien. Une opération de seuillage sur la norme du gradient est généralement nécessaire afin d'éliminer les contours correspondants au bruit. Les considérations concernant le bruit dans la dérivée première sont encore plus importantes dans les calculs de dérivée seconde. On utilise donc couramment une combinaison de lissage et laplacien ce qui correspond au laplacien d'une gaussienne.

L'estimation de la dérivée seconde étant très sensible aux bruits, il convient de filtrer très fortement l'image avant d'en mesurer le laplacien. Ainsi, afin de limiter les réponses dues au bruit de l'image, le plus souvent, on fait appel à un filtrage gaussien dont le laplacien est plus connu sous le nom de "chapeau mexicain".

2.2.2 Approche par filtrage optimal

Cette approche a émergé dans les années quatre-vingt-cinq. Elle a permis une meilleure compréhension des conditions d'une bonne détection de contours et a ainsi conduit à des

détecteurs de très bonne qualité.

Les qualités attendues d'un filtre de détection de contours est un filtre de réponse impulsionnelle $h(x)$ qui permet une bonne détection des contours c'est-à-dire une réponse forte même à de faibles contours, une bonne localisation de ceux-ci et une faible multiplicité des maxima dus au bruit (assurer que pour un contour, il n'y aura qu'une seule détection). Ces trois critères proposés par Canny (A computational approach to edge detection, 1986) s'expriment par l'optimisation conjointe de trois fonctionnelles qui permettent ainsi de définir le filtre linéaire optimal [Cou05]. Ces trois critères d'optimalités sont :

- *Qualité 1 : bonne détection* : Plus le filtre lisse le bruit, plus la détection est bonne : on cherche à maximiser le rapport signal sur bruit, par conséquent, minimiser les fausses réponses,
- *Qualité 2 : bonne localisation* : Moins le filtre lisse l'image, meilleure est la localisation : il s'agit de minimiser la distance entre les points détectés et le vrai contour,
- *Qualité 3 : réponse unique* : On veut une réponse unique par contour. Il existe des cas où il est difficile de savoir si on est en présence de deux contours distincts ou un seul contour bruité : il s'agit de minimiser le nombre de réponse pour un seul contour.

En peut citer comme filtres utilisant : le filtre de Canny et le filtre de Deriche.

2.3 Segmentation et extraction de caractéristiques

L'analyse d'images a pour but l'extraction de l'information caractéristique contenue dans une image. Le résultat d'une telle analyse s'appelle très souvent la description structurelle. Essentiellement, l'analyse de l'image fait appel à la segmentation où l'on va tenter d'associer à chaque pixel de l'image un label en s'appuyant sur l'information portée (niveaux de gris ou couleur), sa distribution spatiale sur le support image, des modèles simples (le plus souvent des modèles géométriques) [Voi02].

2.3.1 Segmentation d'images

La segmentation est une étape essentielle en traitement d'images dans la mesure où elle conditionne l'interprétation qui va être faite sur ces images [Zha96]. De nombreux algorithmes ont ainsi été proposés durant les dernières décennies [JRdSV03],[SLV02],[HS85]. Ils sont basés sur différentes approches : contour, région, texture. Devant la multitude des méthodes proposées, le problème de l'évaluation de la qualité de la segmentation devient primordial. Habituellement, l'efficacité d'un nouvel algorithme est illustrée par la présentation de quelques résultats de segmentation, ce qui n'autorise que des conclusions

subjectives et qualitatives sur les performances de cet algorithme. Il n'existe pas à l'heure actuelle de méthode absolue et générique pour effectuer cette tâche d'évaluation.

La segmentation d'images ainsi définie est un domaine vaste où l'on retrouve de très nombreuses approches que l'on peut regrouper en quatre principales classes [Voi02] :

1. Segmentation fondée sur les régions. On y trouve par exemple : la croissance de région, décomposition/fusion
2. Segmentation fondée sur les contours
3. Segmentation fondée sur classification ou le seuillage des pixels en fonction de leur intensité
4. Segmentation fondée sur la coopération entre les trois premières segmentations

Toutes ces approches visent à l'extraction des indices visuels. Après de nombreuses années passées à rechercher la méthode optimale, les chercheurs ont compris que la segmentation idéale n'existait pas. On peut même montrer que le problème de la segmentation est le plus souvent un problème mal posé. Etant donnée une image, il existe toujours plusieurs segmentations possibles. Une bonne méthode de segmentation sera donc celle qui permettra d'arriver à une bonne interprétation. Elle devra donc avoir simplifié l'image sans pour autant en avoir trop réduit le contenu. Entre autres, elle devra éviter les choix irréversibles. L'avenir de la segmentation est dans le pilotage aval par l'application.

Globalement, la segmentation peut être définie comme une partition d'une image I en une ou plusieurs régions R_1, \dots, R_n telles que :

$$I = \bigcup_{i=1}^n R_i \text{ et } R_i \cap R_j = \emptyset \text{ pour tout } i \neq j \quad (2.5)$$

2.3.1.1 Segmentation par approche régions

La segmentation d'une image vis à vis d'un critère d'homogénéité H (par exemple "tous les pixels sont du même niveau de gris" ou "la variation de niveau de gris n'excède pas n niveaux") peut s'exprimer de la manière suivante :

La segmentation d'une image I en regard du critère H est une partition de l'image I en n régions homogènes X_1, \dots, X_n telles que :

- $\bigcup_{i=1}^n X_i = I$
- Pour tout i , X_i est connexe
- Pour tout i , $H[X_i]$ est connexe
- Pour tout couple (X_i, X_j) de régions voisines, $H[X_i, X_j]$ est faux

Cette définition conduit à deux remarques très importantes. Tout d'abord, une segmentation dépend du critère employé. Le choix du critère est donc primordial. Ensuite, la

décomposition obtenue n'est pas unique. Pour un critère donné, il existe plusieurs solutions.

2.3.1.2 Segmentation par approche frontière

Cette approche cherche à exploiter le fait qu'il existe une transition détectable entre deux régions connexes. Les méthodes les plus anciennes utilisent des opérateurs de traitement d'images, tels que le filtre de Canny, pour mettre en évidence les pixels qui semblent appartenir à un contour. La construction d'une partition est alors souvent difficile. On peut aussi faire intervenir des modèles déformables à l'aide de courbes paramétriques (courbe de Bézier, spline...etc.) ou de polygones (par exemple algorithme à bulle). Pour initier le processus, on recherche des points remarquables de l'image, tels que des points à l'intersection de trois segments au moins. De tels points sont appelés des graines .

L'intérêt principal des méthodes de segmentation selon l'approche frontières est de minimiser le nombre d'opérations nécessaires en cas d'itération du processus sur des séries d'images peu différentes les unes des autres (cas des images vidéo notamment). En effet, une fois que les contours des régions ont été trouvés dans la première image, l'application du modèle déformable à l'image suivante est plus efficace que de tout recalculer, si la différence entre les images est peu importante.

2.3.1.3 Segmentation par classification ou seuillage

Les techniques de classification ont très souvent été utilisées dans le domaine de la segmentation d'images. On cite souvent les travaux de Wacker en 1969 comme la première application des techniques de classification à la segmentation d'images. De nombreux algorithmes ont été développés depuis. On part ici d'un rapport qu'entretient chaque pixel individuellement avec des informations calculées sur toute l'image, comme par exemple la moyenne des niveaux de gris de l'ensemble des pixels, ou la médiane, permettant de construire n classes d'intensité. Lorsque les classes sont déterminées par le choix d'un seuil, on parle de seuillage. Les pixels appartenant à une même classe et étant connexes forment des régions.

Le seuillage a pour objectif de segmenter une image en plusieurs classes en n'utilisant que l'histogramme. On suppose donc que l'information associée à l'image permet à elle seule la segmentation, ie. qu'une classe est caractérisée par sa distribution de niveaux de gris. A chaque pic de l'histogramme est associée une classe.

Il existe de très nombreuses méthodes de seuillage d'un histogramme. La plupart de ces méthodes s'appliquent correctement si l'histogramme contient réellement des pics séparés.

De plus, ces méthodes ont très souvent été développées pour traiter le cas particulier de la segmentation en deux classes (ie. passage à une image binaire) et leur généralité face aux cas multi-classes n'est que très rarement garantie.

2.3.2 Extraction de primitives

Les méthodes d'extraction de primitives graphiques basées sur une approche structurelle pour appréhender une description d'une forme considèrent cette description comme une représentation du contenu d'une image. Il s'agit non plus d'extraire de l'image des mesures statistiques, mais des informations sémantiques relatives à la topologie de la forme. Le principe de ces méthodes est globalement de suivre la démarche de perception humaine pour décrire et reconnaître une forme. Il s'agit en général d'utiliser la structure des formes pour en obtenir une description unique.

2.3.2.1 Les indices visuels

Les indices visuels sont des objets extraits de l'image qui contiennent de manière concise une information importante pour son analyse. En général, ces indices sont soit des contours, soit des régions de l'image [Voi02]. Contours et régions sont complémentaires car on définit qualitativement les régions comme les zones de l'image homogènes au sens d'un certain critère, tandis que les contours sont les zones de transitions entre des régions homogènes. Les propriétés requises par les indices visuels dépendent de l'application et sont en général un sous-ensemble de la liste ci-dessous :

- *Compacts* : la représentation de l'image doit être aussi concise que possible pour réduire la complexité des algorithmes ultérieurs
- *Intrinsèques* : les indices visuels doivent correspondre à la projection dans l'image d'objets physiques ; en particulier, ils doivent être invariants par changement de point de vue
- *Robustes* : la représentation doit être peu sensible aux petites variations d'intensité dans l'image provoquées par des bruits divers tels que ceux liés à l'acquisition, la digitalisation, etc
- *Discriminants* : les indices visuels doivent posséder des propriétés qui permettent de les discriminer, afin de faciliter la mise en correspondance entre deux descriptions (pour la reconstruction tridimensionnelle ou pour l'appariement avec un modèle)
- *Précis* : la position des indices visuels doit pouvoir se calculer avec précision car la qualité de la localisation des objets physiques en dépend

- *Denses* : la densité des indices visuels doit être suffisante pour représenter tous les objets intéressants de la scène.

2.3.2.2 Méthodes basées primitives

Les caractéristiques de l'image à extraire en vue de la détection bas niveau des indices, pour le suivi d'objet sont nombreuses, il existe des descripteurs connus, partageant à la fois des caractéristiques des plusieurs approches basées primitives, ces dernières sont discutées en termes de propriétés invariantes, reconstruction, déformations attendues. Les performances des résultats obtenus prennent en compte les formats des caractéristiques extraites. Parmi les approches basées primitives, on peut citer :

1. Attributs géométriques :

L'objectif est de détecter des descripteurs liés à des points particuliers, et décrivant l'objet par un ensemble d'attributs géométriques (points, segments, courbes paramétriques, arêtes, contours), ou des régions de l'image. Ces méthodes ont l'avantage d'une bonne robustesse aux occultations car des associations qui n'ont pas pu se faire sur certains points de l'objet cachés dans l'image, ne mettent pas en échec le suivi sur l'ensemble des points. Gabriel et all [PGG03] détectent des véhicules par extraction des angles du toit et du capot. Ces primitives peuvent aussi être la taille, la position, la vitesse, le rapport des deux axes de l'ellipse. Elles sont obtenues par une extraction de contours et analyse de la norme du gradient de l'image [Der87],[SC92].

- #### 2. caractéristiques de l'image :
- Ces méthodes, parfois sans modèle, basées région ou suivi de " blobs ", basés sur la couleur, la texture, les primitives ponctuelles, les contours [Bre04],[ALP98], ce sont des procédés de mise en correspondance d'une image à la suivante. Les primitives des " blobs " sont extraites pour une mise en correspondance dans la séquence, par la distance euclidienne ou l'approche basée corrélation. La trajectoire peut alors être évaluée par le regroupement de ces primitives tout au long de la séquence. Cette approche n'identifie que quelques points sur l'objet suivi et non l'objet en entier, ce que font les approches contour et région.

2.4 Opérations morphologiques

La morphologie mathématique est un ensemble de méthodes d'analyse d'images mis au point à l'Ecole des Mines de Fontainebleau par l'équipe de Jean Serra dans les années 60,

Elle est constituée d'un ensemble de techniques basées sur une étude des objets en fonction de leur forme, de leur taille, de leur texture. Elle offre un grand nombre d'outils très puissants de traitement et d'analyse d'images, que l'on retrouve sous différentes rubriques dans certains logiciels d'analyse d'images et même de retouche d'images. Les outils proposés ont été développés au départ pour traiter des images binaires : on fait alors de la morphologie mathématique ensembliste. Leur utilisation a été ensuite étendue aux images en niveaux de gris : on parle, à ce moment là, de morphologie mathématique fonctionnelle. Certains opérateurs fonctionnent désormais sur des images couleur. Les opérations de base sont l'érosion et la dilatation ; les autres outils de morphologie mathématique en sont en fait dérivés. La morphologie mathématique, ensembliste ou fonctionnelle, s'appuie sur un élément structurant dont on choisit la forme et la taille en fonction de ce que l'on souhaite faire. D'autre part, on peut considérer une configuration de voisinage, pour "travailler", qui soit en maille carrée ou en maille hexagonale. Mais, c'est la configuration en maille carrée qui est la plus utilisée.

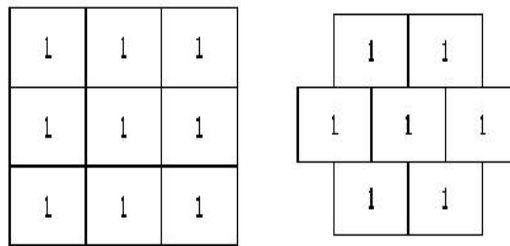


FIG. 2.2 – Exemple d'un maillage carré et hexagonal

2.4.1 Erosion

L'érosion d'une forme X par un élément structurant B est notée " $X \otimes B$ ". Elle est définie par : $X \otimes B = \{p \in S, \text{tel que } Bp \subseteq X\}$. Il s'agit donc de l'ensemble des pixels P d'affixe p du support S de l'image, qui vérifient $Bp \subseteq X$, lorsqu'ils sont pris comme centre de l'élément structurant B (ie. translation de B par p).

La figure 2.3 présente deux cas d'érosion. Les deux érosions sont réalisées sur la même image de départ, mais avec deux éléments structurant différents :

Dans les deux cas, on observe qu'une érosion élimine les pixels isolés sur le fond et érode le contour des objets.

Après une érosion :

- les objets de taille inférieure à celle de l'élément structurant vont disparaître,
- les autres seront amputés d'une partie correspondant à la taille de l'élément structurant,

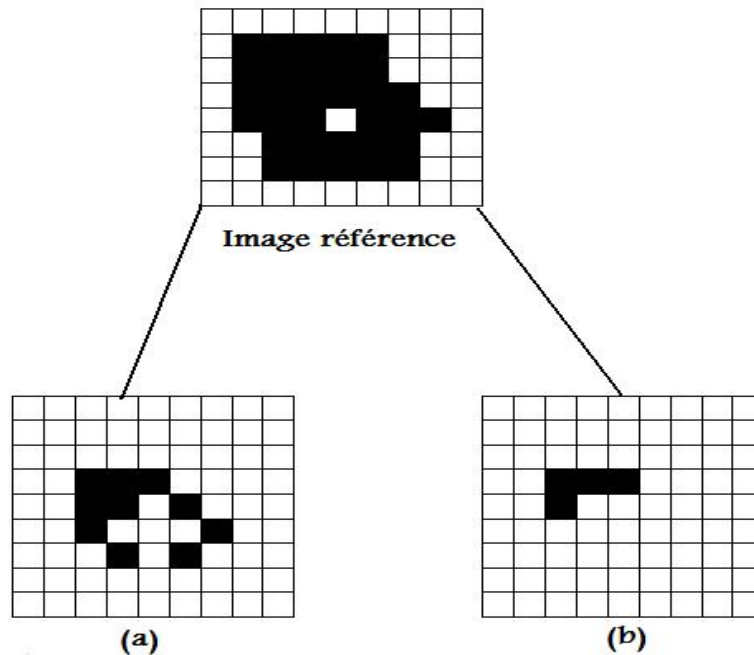


FIG. 2.3 – a.Erosion morphologique avec un élément structurant à 4 connexités, b.Erosion morphologique avec un élément structurant à 8 connexités

- s'il existe des trous dans les objets, c'est 'a dire des "morceaux" de fond 'a l'intérieur des objets, ils seront accentués,
- les objets reliés entre eux par un élément de taille inférieur à celle de l'élément structurant vont être séparés.

2.4.2 Dilatation

La dilatation d'une forme X par un élément structurant B est notée " $X \oplus B$ ". Elle est définie par : $X \oplus B = \{p \in S, \text{telque } \bar{B}p \cap X \neq \emptyset\}$. Il s'agit donc de l'ensemble des pixels P d'affixe p , tel que le translaté $\bar{B}p$, de l'élément structurant symétrique \bar{B} , ait une intersection non vide avec X . La figure 2.4 présente deux cas de dilatation.

Dans les deux cas, on observe qu'une dilatation élimine les trous isolés dans les objets et dilate le contour des objets en tenant compte de l'élément structurant.

Après une dilatation :

- Tous les objets vont "grossir" d'une partie correspondant à la taille de l'élément structurant,
- s'il existe des trous dans les objets, ils seront partiellement ou intégralement comblés ,
- Si des sont situés à une distance moins grande que la taille de l'élément structurant, ils vont fusionner.

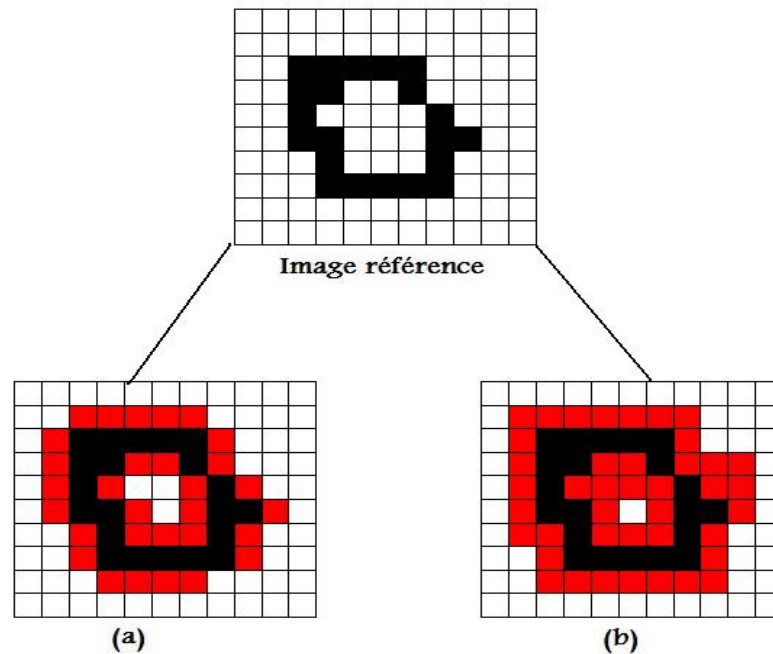


FIG. 2.4 – a. Dilatation morphologique avec un élément structurant à 4 connexités,
b. Dilatation morphologique avec un élément structurant à 8 connexités

2.5 Conclusion

Quelques approches de bases ont été présentées dans ce chapitre, d'autres méthodes de détection de contours et de segmentation d'images existent n'ont pas été présentées, il est clair que le sujet est trop vaste pour l'aborder dans son entière totalité ou même dans sa plus grande partie, le but étant de donner un éclaircissement ou une idée générale sur les méthodes.

Chapitre 3

Etat de l'art sur le suivi

3.1 Introduction

Le suivi visuel d'objets dans des séquences d'images est devenu une des tâches les plus utilisées, ce domaine connaît un développement important dans la vision par ordinateur [GC04]. Avec la généralisation de l'utilisation d'images numériques, l'analyse du mouvement dans les vidéos s'est révélée être un outil indispensable pour des applications aussi diverses que la vidéoconférence (suivi des interlocuteurs), la vidéosurveillance (détection, suivi, reconnaissance du comportement de personnes, d'intrus), la compression vidéo, l'imagerie médicale [Noc08], la segmentation et l'édition de séquences vidéo, le langage des signes pour les mal-entendants, l'analyse de séquences sportives (tennis, golf), sans oublier les applications plus traditionnelles de la robotique [Her04] : suivi d'obstacles pendant une phase d'évitement, suivi d'indices visuels dans une tâche asservie sur la vision, suivi d'un opérateur (corps, visage, main...etc.) pour définir des modes d'interaction évolués entre l'Homme et la Machine...etc. En effet, les zones de mouvement d'une séquence d'images correspondent souvent à des événements sur lesquels un système de vision doit se focaliser.

L'analyse du mouvement est un vaste sujet qui englobe en général trois problématiques [GC04],[Lei04] :

- La détection,
- Le suivi,
- La reconnaissance

La détection consiste à segmenter l'image d'une séquence en objets mobiles et stationnaires, permettant de voir le changement d'environnement observé et d'en déduire une

statistique sur le mouvement de l'objet de la séquence. Le suivi consiste à suivre un objet le long de sa vie dans la scène dans le but de déterminer sa position dans l'image à chaque instant t . La reconnaissance est la modélisation d'activités ou de gestes dans le but d'identifier les objets et d'interpréter leurs mouvements.

De nombreuses applications en vision par ordinateur nécessitent la détection et le suivi des objets en mouvement dans une séquence d'images, la réussite du suivi dépend de la justesse de l'étape de détection [Lei04].

Le but de ce rapport est de donner un état de l'art des différentes techniques de détection et de suivi dans une séquence d'images. L'idée est plus de donner un bref aperçu des méthodes les plus couramment utilisées et des différents types d'approches existantes pour la détection de mouvement en vision par ordinateur. La première étape concerne la détection, la deuxième celle du suivi, la reconnaissance ne fera pas l'objet de ce rapport. Avant d'entamer les parties concernant la détection et le suivi, nous allons présenter d'une manière générale la représentation d'un objet dans un scénario de suivi, puisqu'il constitue l'élément principal dans ce type d'application.

3.2 Modélisation d'un objet

3.2.1 Caractéristiques d'un objet

Dans le domaine du mouvement d'un objet dans une séquence, un objet peut avoir différentes caractéristiques [Lei04], il peut être :

- Constant ou variable : La forme, le mouvement, les couleurs ou les textures d'un objet varient ou non au cours du temps. Ces considérations jouent en faveur de l'adoption ou non d'un modèle de l'objet pour en améliorer le suivi.
- Rigide ou non rigide : pour la forme nous parlerons d'objet rigide ou non-rigide c.à.d. la distance entre les points de l'objet est constante ou variable au cours du temps.
- Unique ou multiple : dans le cadre d'un suivi labellisé où chaque objet porte une étiquette et n'en change pas durant l'opération de son suivi, une distinction entre chaque objet est nécessaire c.à.d que chaque objet doit être unique. L'unicité repose, là aussi, sur un modèle. Cependant celui-ci n'est pas nécessairement connu au préalable. Cependant une connaissance des critères potentiellement discriminants permet une amélioration de la charge de calcul. La description de chaque objet selon des critères pertinents assure une labellisation fiable.

3.2.2 Représentation d'un objet

Un objet, dans un scénario de suivi, est une entité indépendante pourvue de son identité spatiale (sa forme, son contour,...etc.) dans un environnement particulier. Cependant, on ne dispose pas toujours de toutes les informations qui caractérisent l'objet à suivre, mais on utilise un modèle de représentation de son état à surveiller. Une modélisation d'objets par catégories peut être présentée [SKS01].

3.2.2.1 Modélisation par points

L'objet à suivre est représenté soit par un point qui représente le centre de gravité de l'objet concerné [HNvdB01], soit par un ensemble de points [PGG03](Figure 3.1). En général cette dernière est efficace pour le suivi d'objets de petites tailles ou régions.

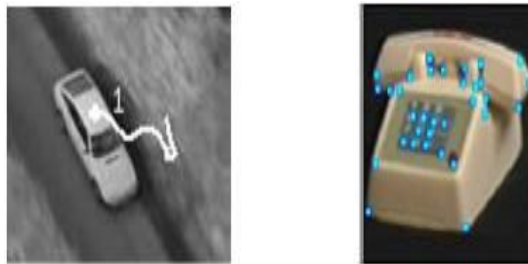


FIG. 3.1 – Points de contrôl

3.2.2.2 Modélisation par formes géométriques

La forme d'un objet est représentée par un rectangle ou (boîtes englobantes) (Figure 3.2) ou une ellipse, le mouvement d'objet défini par cette représentation est souvent modélisé par une translation, projection et autre transformation, cependant ces formes géométriques sont plus appropriées à représenter les objets rigides simples, elles sont également employées pour suivre les objets souples.



FIG. 3.2 – Boîtes englobantes

3.2.2.3 Modélisation par silhouette et contour

Par définition, un contour est la frontière qui sépare deux objets dans une image. La région à l'intérieur du contour s'appelle la silhouette de l'objet (Figure 3.3). La représentation par le contour et la silhouette est appropriée pour représenter des objets non rigides et les modèles déformables complexes.



FIG. 3.3 – Modèle silhouette

3.2.2.4 Modélisation par des modèles de formes articulées

Les objets articulés sont composés de parties du corps qui sont liés avec des joints. Par exemple le corps humain est un objet articulé avec torse, jambes, mains, tête et pieds reliés par des joints (Figure 3.4). On peut modéliser ces parties en utilisant des cylindres ou des ellipses, les rapports entre eux sont régis par des modèles de cinématique de mouvement, par exemple l'angle commun etc.

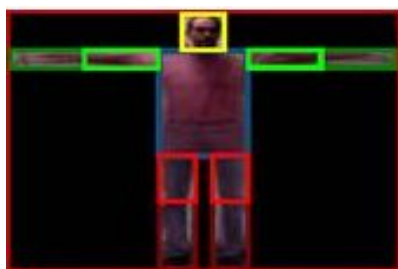


FIG. 3.4 – Modèle d'apparence articulé

3.2.2.5 Modèles squelettiques

Le squelette d'objet peut être extrait en appliquant la transformation de l'axe médiane à la silhouette d'objet (Figure 3.5). Ce modèle est généralement utilisé pour la reconnaissance de l'objet et du suivi des modèles cinématiques complexes et la reconnaissance de mouvement.

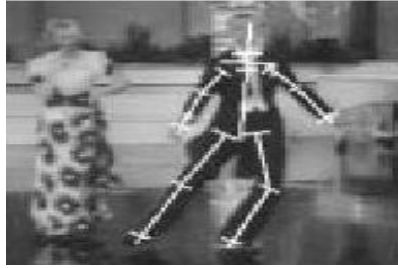


FIG. 3.5 – Modèle squelette

Il y a un certain nombre de voies pour représenter les caractéristiques d'apparence d'objets, ces représentations de cette forme d'objet peuvent être également combinées avec les représentations d'apparence pour le suivi.

3.2.2.6 Densités de probabilités d'apparence d'objet

La densité de probabilité (Pdf : Probability Density Function) décrit l'observation de couleur, de texture, elle fournit une estimation raisonnable pour l'objet et décrit les co-occurrences des multiples attributs d'objets, les évaluations de densité de probabilité d'apparence d'objet peuvent être paramétrique, comme gaussien ou un mélange de gaussiens ou non paramétriques comme l'histogramme. La densité de probabilité des caractéristiques d'apparence d'objet (couleur, texture) peut être calculée en spécifiant des régions dans l'image à l'aide d'une représentation de forme connue (rectangle, ellipse).

3.2.2.7 Les templates

Les templates sont formés d'une forme géométrique simple ou une silhouette, l'avantage des templates est qu'elles portent toutes des informations spatiales de l'objet, elles diffèrent des densités de probabilités par l'utilisation des informations spatiales.

3.2.2.8 Les modèles d'apparences multi-vues

Cette représentation utilise les différentes vues d'apparences d'objets pour le représenter [MOP97] (Figure 3.6), on représente les vues comme un sous espace en utilisant les méthodes d'analyse telle que l'analyse de composantes principales (ACP).

Une fois le choix de la représentation de l'objet est effectué, il faut pouvoir extraire des éléments des images pour comparer ces éléments au modèle qui a été choisi. Les informations que l'on peut extraire d'une image sont nombreuses, et le choix est souvent guidé

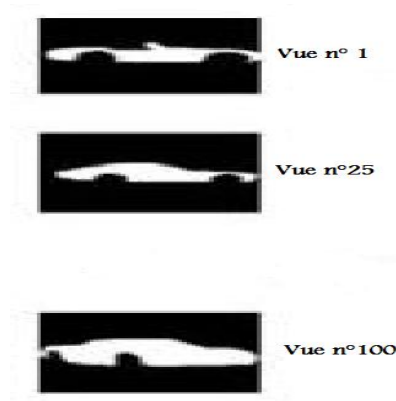


FIG. 3.6 – Modèle d'apparence multi-vues

par le modèle choisi précédemment, cependant plusieurs types d'informations peuvent être utilisés, voire combinés pour rendre la comparaison plus fiable et le suivi plus robuste.

3.3 Nature de la scène

Selon l'objet présenté et la position de la caméra, le contenu de la séquence vidéo peut changer fortement.

Dans les méthodes de détection de mouvement, le mouvement des objets n'a pas besoin d'être calculé, les objets sont caractérisés par le fait que ce sont les seules parties de la vidéo qui ont un mouvement important. Il s'agit d'abord de distinguer les séquences à caméra fixe et les séquences à caméra mobile.

3.3.1 Séquences à caméra fixe (Scène statique)

La grande majorité des systèmes telle que les systèmes de vidéosurveillance utilisent des caméras fixes par rapport à la scène filmée. L'image issue des caméras est alors relativement stable, dans le sens où les changements temporels visibles dans l'image correspondent généralement à des objets d'intérêt. Lorsque la scène n'est pas complètement statique, les variations sont souvent modélisables. Par exemple un changement global de luminosité entre la nuit et le jour est suffisamment lent pour être dissocié facilement des changements intéressants comme le mouvement d'un objet, ou bien comme par exemple des nuages dans le ciel ou bien des branches d'arbre qui se déplacent avec le vent mais ce ne sont pas les mouvements dominants dans la scène et la plupart des pixels du fond restent statiques.

3.3.2 Séquences à caméra mobile (Scène dynamique)

Dans le cas des systèmes tel que le cas d'un système embarqué dans un véhicule [Der87]. La mobilité de ce dernier fait perdre à la scène sa stabilité, même lorsque la caméra est fixée. Nous ne pouvons donc pas faire l'hypothèse d'une scène statique et les variations présentes sont très difficiles à modéliser.

Les variations temporelles se produisant dans une séquence vidéo typique sont de plusieurs sortes. On peut les classer en variations dues au mouvement d'un élément de la scène, et en variations dues à des modifications d'illumination. Il est nécessaire d'estimer le mouvement de la caméra afin de le compenser pour pouvoir segmenter les objets en mouvement. Le mouvement dominant de la scène appliqué à tous les pixels est considéré comme étant le mouvement de la caméra et les pixels ne suivant pas ce mouvement global forment les objets en mouvement. Il faut donc compenser le mouvement de la caméra afin de se replacer dans un cadre de caméra fixe. L'idée est d'utiliser une estimation de caméra la plus robuste possible.

3.3.3 Mono et multi-caméra

Des méthodes de suivi existent dans le cas d'une seule caméra [JS04] ou de plusieurs caméras. La différence essentielle entre ces deux méthodes provient du fait que les images prises de plusieurs caméras doivent se projeter dans le même espace, ce qui implique la calibration des caméras. Cela n'est pas nécessaire dans le cas du suivi mono caméra. La plupart des méthodes de suivi du corps en mouvement sont basées sur des séquences mono caméras. En ce qui concerne l'avantage à utiliser plusieurs caméras est de réduire les ambiguïtés causées par les problèmes d'occultations.

3.4 Détection d'objets

Dans le domaine de la vision par ordinateur, la détection consiste à percevoir une scène statique ou dynamique, et de savoir le changement de mouvement approprié en déduisant ainsi la cinématique de tous les objets mobiles par rapport à la caméra. La détection est une tâche difficile, elle dépend généralement de la nature de la scène et de l'application correspondante.

3.4.1 Catégorisation des méthodes de détection

Les méthodes de détection peuvent être classées en général en quatre catégories [Bug07] :

1. Les détections basées sur la différence inter-images,
2. Les détections basées sur la modélisation du fond,
3. Les détections extrayant des couches de mouvement,
4. Les détections utilisant la notion de cohérence.

3.4.1.1 Détections basées sur la différence inter-images

La méthode de détection de mouvement la plus simple consiste à faire la soustraction entre deux images successives [TR04],[PGG03]. Elle porte aussi le nom de méthode de différence temporelle dans d'autres littératures. Les pixels dont l'intensité résultante est proche de zéro sont assimilés comme étant les pixels du fond.

Les méthodes basées sur la différence inter-images s'adaptent très bien aux environnements dynamiques mais laisse des "trous" dans les objets en mouvements. Par exemple, la différence temporelle seuillée permet de détecter des objets en mouvement dans une séquence vidéo. Une amélioration de cette méthode consiste à extraire trois images de différence au lieu de deux permettant de décider si un pixel a bougé.

Ces méthodes ne peuvent permettre à elles seules de détecter les objets mobiles dans des séquences d'images. Néanmoins, ces techniques peuvent donner une bonne indication des pixels en mouvement et peuvent donc être utilisées comme prétraitement pour améliorer et accélérer les autres techniques de détection ou de segmentation de mouvement.

3.4.1.2 Détections basées sur la modélisation du fond

Une autre catégorie de méthodes, très populaire en détection de mouvement, repose sur la modélisation du fond. Plusieurs conditions sont nécessaires à l'utilisation de ces techniques. Tout d'abord, la caméra doit être maintenue fixe et les occultations de parties du fond par des objets en mouvement doivent rester temporellement minoritaires. De plus, il est préférable d'avoir, au début de la séquence, un certain nombre d'images sans objets en mouvement afin de pouvoir apprendre correctement le fond.

Ces techniques peuvent être prédictives ou non prédictives, ces dernières ayant fait l'objet d'un nombre plus important de recherche. La modélisation du fond, tout comme les méthodes basées sur la détection inter-images, permet de décider si un pixel est en mouvement ou non.

- **Méthodes non prédictives** : Les techniques non prédictives de modélisation du fond se décomposent en deux étapes principales : l'apprentissage du fond [DJG00], le plus souvent en chaque pixel de l'image, et l'étiquetage binaire des pixels comme appartenant

ou non au fond dans l'image courante. L'apprentissage du fond doit tenir compte de la fluctuation des intensités observées.

- **Méthodes prédictives** : Les méthodes prédictives utilisent un modèle dynamique pour prédire la valeur d'un pixel à l'aide des observations précédentes. Elles emploient généralement un filtre de Kalman pour modéliser la dynamique de l'état d'un pixel [HNvdB01].

Des méthodes plus récentes utilisent des modèles plus complexes basés sur les textures dynamiques, ces derniers permettent de représenter une image à l'aide d'une combinaison linéaire d'images de base. Chaque paramètre de l'image est projeté comme un point dans un espace de caractéristiques.

3.4.1.3 Les détections extrayant des couches de mouvement

L'objectif de ces méthodes est de fournir une description d'une scène en terme d'objets et de leur mouvement dans le plan image ou, autrement dit, en terme de couches de mouvement. La segmentation de mouvement peut donc être vue comme une décomposition de la scène en différentes couches [MOP97], chacune étant associée à un modèle de mouvement. En effet, la plupart des méthodes d'extraction de couches de mouvement repose sur le calcul d'un certain nombre de modèles paramétriques et supposent que le mouvement 2D d'un objet 3D projeté dans l'image suit un modèle paramétrique.

L'extraction de couches de mouvement permet de détecter directement les masques de chaque objet. L'utilisation de ce type de méthodes suppose que la scène peut être décomposée en plusieurs régions bien distinctes toutes caractérisées par un mouvement bien défini. L'avantage de ces méthodes est qu'elles n'ont ni besoin d'une caméra fixe ni des résultats aux instants précédents. Elles se basent sur une estimation préalable du mouvement dans l'image.

3.4.1.4 Les détections utilisant la notion de cohérence

Une dernière approche consiste à définir un objet mobile comme une région ayant un mouvement cohérent. Dans ce contexte, un mouvement cohérent peut être défini comme un mouvement ayant de grande chance de provenir d'une "cible" classique (personne, véhicule).

Cette définition a été utilisée pour la segmentation de mouvement en ajoutant l'hypothèse suivante : un objet avec un mouvement cohérent se déplace dans une direction approximativement constante pendant une courte période (en pratique quelques images). Les modèles de mouvement dérivés du flot optique sont utilisés comme primitives pour

la détection ou le suivi. Le flot optique a pour rôle de décrire le mouvement cohérent des points entre des images successives, des mesures de flot optique cohérentes en direction sont accumulées pendant quelques pas de temps. Une estimation du déplacement de chaque pixel à travers une séquence d'images est ainsi obtenue [JB03]. Il est alors possible de distinguer les objets mobiles, qui se déplacent avec une direction constante, des mouvements parasites.

Une région cohérente peut être directement vue comme une couche. La différence et l'avantage ici est que seuls les objets cohérents sont détectés. Cela permet l'application directe d'autres traitements tels que l'identification ou le suivi.

3.5 Suivi d'objets

Le suivi est un problème plus complexe que la détection, car il faut prendre en compte d'autres contraintes fortes, en particulier, des objets peuvent disparaître du champ de la caméra pendant un instant et réapparaître ensuite lorsqu'elles sont cachées par une autre personne.

Le but central du suivi visuel consiste à déterminer la configuration d'une cible ou les caractéristiques d'un objet dans une séquence d'images, tant que cette cible est en mouvement apparent dans les images.

Les méthodes de suivi permettent de construire les trajectoires d'objets mobiles en s'accommodant du bruit, des déformations subies par les objets et d'éventuelles occultations. Certaines méthodes de suivi nécessitent une initialisation manuelle tandis que d'autres utilisent directement des initialisations obtenues préalablement par une méthode de détection d'objets. Généralement, le suivi d'un objet mobile nécessite l'utilisation d'un filtre de poursuite et d'un filtre d'association.

3.5.1 But du suivi

Le suivi de plusieurs objets dans une séquence d'images peut s'avérer difficile selon les scènes complexes qu'il traite, cette complexité réside là où les objets peuvent avoir des tailles différentes. Ces objets peuvent être rigides (par exemple une voiture) ou non-rigides (par exemple une personne) et s'occultent les uns les autres [Lei04].

De ce fait, des méthodes de suivi sont apparues, le but de ces méthodes est d'estimer au fil du temps des paramètres d'une cible (ou plusieurs) présente dans le champ de vision de la caméra et initialement détectée par un moyen quelconque. Les paramètres peuvent être divers : position dans l'image, à laquelle peuvent s'ajouter la taille et l'orientation

apparente, l'attitude, l'apparence, etc. Pour qu'une méthode de suivi soit complètement automatique, il faut non seulement suivre la ou les cibles mais aussi les initialiser automatiquement par une méthode de détection et gérer leurs éventuels arrêts et/ou sorties du champ de la caméra.

La plupart des systèmes de suivi actuels ne marchent pas bien avec ces conditions difficiles. En fait, la majorité d'entre eux est basée soit sur l'extraction de régions en utilisant un modèle du fond, soit sur l'extraction de contours avec un filtre tel que le filtre à particules [TR04],[HNvdB01].

3.5.2 Méthodes de suivi

Les méthodes existantes pour le suivi simultané de plusieurs objets dans une séquence d'images impliquent généralement que des caractéristiques telles que des histogrammes couleurs, des mesures de corrélation, des estimations de vitesse, ou des mesures de distance entre régions puissent être utilisées pour suivre chaque objet. Selon les méthodes, ces objets peuvent être suivis indépendamment les uns des autres, ou de manière unifiée pour une meilleure robustesse aux occultations et aux erreurs d'association.

Plusieurs classifications des méthodes de suivi d'objets ont été proposées dans la littérature; elles dépendent autant des auteurs, que du but pour lequel ces méthodes ont été conçues. Nous considérons une classification des méthodes de suivi, qui peuvent être divisées en trois principales catégories [Bug07],[AYS06] :

1. Suivi par appariement de détection ("detect-before-track"),
2. Suivi par segmentation dynamique,
3. Suivi déterministe par détection séquentielle.

Chacune de ces catégories a ses propres avantages et inconvénients

3.5.2.1 Suivi par appariement de détections

Dans le cas où, à chaque instant, toutes les cibles visibles peuvent être détectées de façon fiable et rapide, le problème de suivi peut être défini comme un problème d'appariement de détections entre images successives [PGG03]. Les cibles candidates à chaque instant sont appelées observations. Elles sont généralement obtenues par soustraction de fond mais peuvent aussi être le résultat de méthodes de détection plus complexes (telle que les méthodes traitées dans la partie détection). Le suivi se fait alors en assignant les observations aux pistes en cours d'estimation. Ces méthodes, dites "detect-before-track" sont très populaires en pistage sonar et radar. Elles peuvent être déterministes ou probabilistes.

1. **Méthodes déterministes** : Le principe des méthodes déterministes est d'associer les observations aux pistes en cours de suivi en minimisant une distance calculée sur certaines caractéristiques de l'objet. Les caractéristiques des objets couramment utilisées sont la proximité (hypothèse de déplacement limité) et l'apparence (similarité de forme et/ou de contenu photométrique et/ou de mouvement) [Noc08]. Les modèles de l'objet basés sur l'apparence peuvent être des densités (histogrammes de couleur ou de contour), une carte de contours (contour ouvert ou fermé de l'objet) [Her04] ou une combinaison de ces modèles.
2. **Méthodes probabilistes** Les observations obtenues par un algorithme de détection sont très souvent corrompues par du bruit. De plus, le mouvement ou l'apparence d'un objet peut légèrement varier entre deux images consécutives. Les méthodes probabilistes permettent de gérer ces variations en ajoutant une incertitude au modèle de l'objet et aux modèles des observations. Le suivi d'une seule cible est alors obtenu par des méthodes de filtrage (filtres de Kalman, filtrage particulaire) [HNvdB01]. Le suivi de plusieurs objets peut lui aussi se faire avec ces méthodes de filtrage mais une étape préalable d'association de l'objet avec l'observation la plus probable doit être ajoutée [HNvdB01].

Il existe de nombreuses méthodes statistiques d'association. Les deux techniques les plus utilisées sont le MHT ("Multiple Hypothesis Tracking") et le PDAF ("Probability Data Association Filtering").

3.5.2.2 Segmentation dynamique

Les méthodes de suivi par segmentation dynamique sont utilisées lorsque l'on souhaite extraire la silhouette de la cible à chaque instant, et ce sans connaissance a priori sur sa forme.

Ces approches reposent sur une succession de segmentations. Elles font généralement évoluer le contour de l'objet à l'instant précédent jusqu'à sa nouvelle position à l'instant courant. Ce type de suivi peut se faire avec des méthodes utilisant des modèles d'état ou avec des méthodes minimisant des fonctions d'énergie sur le contour.

1. **Suivi avec une représentation explicite du contour** Un certain nombre de méthodes de suivi par segmentation dynamique cherche à représenter le contour par un ensemble de paramètres avant de le suivre par une méthode de filtrage. Il s'agit en fait de définir un modèle du contour de l'objet (généralement défini par la forme de l'objet) qui servira de modèle d'état à l'algorithme de filtrage.

Dans [HNvdB01] des points de contrôle sont placés le long du contour à suivre. Le

modèle d'état est alors représenté par leur dynamique. Cette dynamique donne une prédiction de ces points de contrôle (filtre de Kalman) qui est ensuite corrigée en utilisant des mesures de gradient de l'image.

D'autres méthodes consistent à définir l'état comme un ensemble de splines et de paramètres de mouvement. Les mesures utilisées sont les contours de l'image calculés dans la direction normale à la silhouette. Les auteurs utilisent un algorithme de filtrage particulaire pour mettre à jour l'état. Ces méthodes ont été étendues au suivi de plusieurs objets et à la gestion des occultations dans qui sont mis à jour à chaque instant en maximisant une loi a posteriori.

2. **Minimisation d'une fonction d'énergie sur contour** Les méthodes décrites ci-dessus utilisent une représentation explicite du contour, et ne peuvent donc pas suivre les changements topologiques, la division ou la fusion de régions.

Des méthodes basées sur la minimisation d'une fonctionnelle d'énergie permettent de suivre un contour ou une région en tenant compte des changements de topologie [Pre04],[SLV02]. La région occupée par la cible peut être définie par un masque binaire ou par la courbe de niveau zéro d'une fonction à support continue [GC04]. La fonction d'énergie est définie en utilisant de l'information sous la forme de gradient temporel (flot optique) [JB03] ou en se basant sur l'apparence de l'objet et du fond (modélisation des intensités à l'intérieur et/ou à l'extérieur de l'objet). Généralement cette fonction encourage les formes régulières s'appuyant sur des contours forts de l'image (Figure 4.3).



FIG. 3.7 – Contour actif

L'avantage des méthodes utilisant les gradients temporels d'intensité de l'image est qu'elles permettent, de part le calcul du flot optique, de prédire la position du contour avant de procéder à la minimisation de la fonctionnelle, c'est-à-dire à la segmentation. Au contraire, les méthodes se basant sur des modélisations de l'apparence de l'image utilisent comme initialisation le contour de l'objet à l'instant précédent. Elles ne gèrent donc pas les grands mouvements des objets.

3.5.2.3 Suivi déterministe de fenêtre englobante par détection séquentielle

La dernière catégorie de méthodes évoquée dans cette partie regroupe les méthodes de suivi d'images [Lei04]. Une image est une boîte (en général un rectangle mais parfois aussi une ellipse) entourant ou à l'intérieur de l'objet à suivre. Il s'agit en fait d'une petite portion de l'image.

Les techniques de suivi considérées dans cette section sont basées sur la conservation de l'apparence (généralement couleur et/ou luminosité) de l'objet pendant au moins deux instants consécutifs.

On distingue les algorithmes faisant l'hypothèse de conservation de l'apparence localement (en chaque point de l'objet) et ceux utilisant une caractérisation globale de la cible (hypothèse globale de conservation de l'apparence).

1. **Suivi différentiel d'images** Cette approche consiste à chercher la portion d'image la plus proche dans l'image courante de celle à l'instant précédent. La recherche se fait souvent autour de la position de l'objet à l'instant précédent. Ces approches sont appelées "block matching" ou "template matching". Elles font appel à une mesure de similarité telle que la corrélation ou la SSD ("sum of square differences"). Les intensités et les couleurs étant sensibles aux changements d'illumination, il est aussi possible d'exploiter les gradients de l'image. Le coût de calcul de ce type de méthodes peut être élevé. Il dépend de la taille du voisinage dans lequel l'image est recherchée.

Pour suivre de plus grands objets, on peut appliquer la méthode précédente à un ensemble d'images contenues dans la cible. Une autre approche pour suivre une région de forme prédéfinie est de calculer son déplacement (champ dense) par une méthode de type flot optique.

2. **Suivi de distributions** L'utilisation de distributions permet de rendre moins critique la phase de remise à jour de l'image de référence. Elles sont moins cruciales lorsque l'on se donne une caractérisation globale de l'apparence de la cible.

La recherche de la cible à l'instant courant se base alors sur des distributions (histogrammes ou mélanges de gaussiennes) des couleurs dans une région de géométrie simple. La méthode la plus utilisée dans cette catégorie est le suivi par mean shift (Figure 2.8). La distribution est un histogramme de couleur. L'algorithme consiste alors à déplacer une fenêtre d'analyse (noyau) de manière à trouver l'histogramme contenu dans la fenêtre qui coïncide le mieux avec l'histogramme de référence.

A partir d'une image de la séquence et d'un modèle de joueur, une image représentant la répartition des pixels les plus probables est calculée. Cette image est

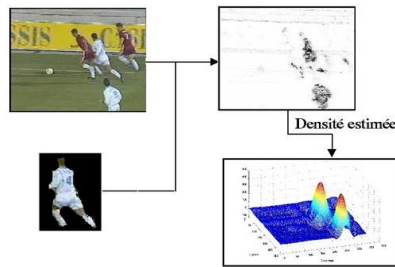


FIG. 3.8 – Suivi de distribution

considérée comme un nuage de points, puis d'utiliser la procédure Mean Shift pour trouver tous les modes (maxima de la fonction de densité), qui correspondent aux positions de tous les joueurs

D'autres méthodes permettant le suivi de distributions sans calculer explicitement les distributions. Elle consiste à ajouter, en plus de la couleur, une information sur la géométrie des objets.

3.5.3 Caractéristiques d'un bon algorithme de suivi

Il ressort des techniques de suivi présentées précédemment des points importants permettant de définir la qualité d'un algorithme de suivi.

Les caractéristiques d'un bon algorithme de suivi sont les suivantes [Bug07] :

- La méthode doit être capable d'initialiser automatiquement les cibles et doit gérer les arrêts et les sorties du champ de la caméra.
- Elle doit de plus être robuste aux changements d'illumination et aux éventuels changements de topologie. La topologie n'est importante que dans le cas où une segmentation de la cible est exigée.
- Enfin, la méthode doit permettre de continuer à suivre la cible même en cas d'occultations partielles ou totales par un autre objet ou par le fond.

La première caractéristique ne peut être obtenue qu'avec l'ajout d'une méthode de détection des objets, des observations étant apportées à chaque instant ou à des instants espacés par un faible pas de temps.

Les changements d'illumination ou de photométrie ne sont bien pris en compte que par les méthodes utilisant des histogrammes ou des mélanges de gaussiennes sur l'intensité ou la couleur. Ces distributions doivent régulièrement être mises à jour.

Enfin, la gestion des occultations n'est bien considérée qu'avec l'utilisation d'observations ou en ajoutant une loi dynamique au processus de suivi, la difficulté étant alors de déterminer cette loi. La plupart du temps les objets peuvent avoir un mouvement quelconque

qu'aucune loi ne peut caractériser.

3.6 Conclusion

De nombreuses applications en vision par ordinateur nécessitent la détection et le suivi des objets en mouvement dans une séquence d'images. La plupart des méthodes existantes ne donnent de bons résultats que pour des séquences avec des fonds peu changeants, ou si le fond et les objets sont rigides.

La mise en œuvre de l'évaluation d'un système de suivi d'objets est un problème difficile du fait de la complexité du système lui-même, qui est constitué d'au moins trois composants [MMB07] de base qui sont :

- Le modèle d'apparence qui décrit ce à quoi un objet doit ressembler dans une image,
- L'algorithme d'optimisation, qui tente d'estimer la position de l'objet en optimisant la correspondance entre l'apparence actuelle et le modèle d'apparence,
- Les contraintes spatio-temporelles, qui donnent un a priori sur la position de l'objet en fonction du suivi passé.

Ainsi, Le degré de précision d'une méthode de suivi est plus ou moins grand, selon le type d'application utilisée. Un modèle de représentation des objets bien choisi permet souvent d'améliorer la qualité d'un suivi afin d'obtenir la précision souhaitée en rapport avec l'application.

Chapitre 4

Conception et implémentation de la détection de contour par réseaux de neurones impulsionnels

4.1 Introduction

Le système visuel humain effectue la détection de contours d'une manière très efficace. Les neuroscientifiques ont constaté qu'il y a une variété de champs récepteurs à partir des cellules simples dans le cortex strié jusqu'à celles de la rétine et du noyau géniculé latéral [Kan81] et les neurones peuvent être simulés par le modèle de Hodgkin et de Huxley. Basé sur ces champs récepteurs et sur ce modèle de neurone, un modèle de réseau est proposé pour détecter les contours dans une image visuelle.

4.2 Modèle d'un réseau de neurones impulsionnels pour la détection de contours

4.2.1 Architecture du réseau de neurones impulsionnels

Le réseau de neurones impulsionnels sur lequel on a travaillé [QWB06] se présente comme le montre la Figure 4.1 Il possède 3 couches : une couche entrée, une couche intermédiaire et une couche de sortie.

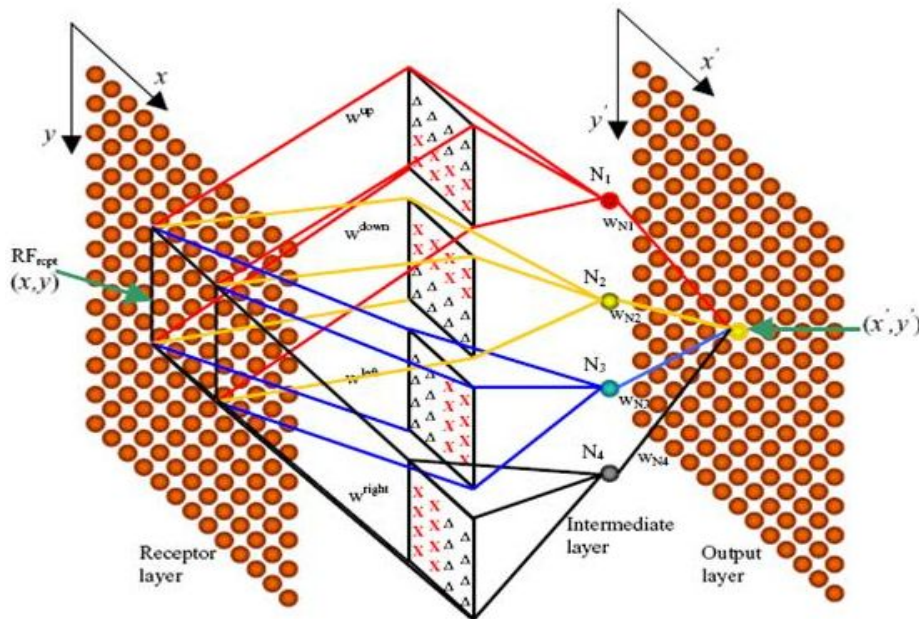


FIG. 4.1 – Modèle de réseau de neurones impulsionnels pour la détection de contours

La première couche représente les photorécepteurs. Chaque pixel correspond à un récepteur. La couche intermédiaire est composée de quatre types de neurones correspondants respectivement à quatre champs récepteurs différents. Dans les connexions synaptiques 'X' représente une synapse excitatrice. 'Δ' représente une synapse inhibitrice. Chaque neurone dans la couche de sortie intègre les quatre sorties correspondantes aux neurones de la couche intermédiaire.

La carte de la fréquence de décharge de la couche de sortie donne un graphe de contours correspondant à l'image d'entrée.

4.2.2 Description Fonctionnelle du réseau

Il y a quatre matrices parallèles de neurones dans la couche intermédiaire chacune de la même dimension que la couche d'entrée. Ces matrices sont marquées N1, N2, N3 et N4 et seulement un neurone dans chaque matrice est montré pour simplifier le schéma. Chacune de ces matrices dans cette couche effectue le traitement respectif des contours hauts, bas, gauches et droits et sont reliées à la couche d'entrée par différentes matrices de poids. Ces matrices de poids peuvent être de tailles variables pour représenter la largeur du champ récepteur en considération. Par exemple dans la Figure 4.1, le neurone N1 est relié au champ récepteur RFrept de la couche d'entrée à travers une matrice wup de distribution des poids synaptiques et répond à un contour haut dans ce champ. Si une image uniforme

dans RFrept donne une sortie uniforme, les sorties par les synapses wup atteindront le neurone N1. Les connexions par la moitié supérieure de la matrice de poids représentent les synapses inhibitrices qui déprécient le potentiel de membrane du neurone N1 tandis que les connexions par les synapses excitatrices de la moitié inférieure renforcent le potentiel membranaire du neurone N1. Par conséquent le potentiel de membrane du neurone N1 n'est pas modifié, et aucune impulsion n'est générée par le Neurone N1. Cependant, si une image contour à l'intérieur du RFrept apparait sur les récepteurs de moitié inférieure avec un signal fort et les récepteurs de moitié supérieure avec un signal très faible, alors le signal fort renforcera (en raison des synapses excitatrices) le neurone N1, mais le signal faible ne dépréciera pas le potentiel de membrane d'une manière significative. Le potentiel de membrane du neurone N1 augmente de façon très rapide et produit des impulsions fréquentes pour répondre à un contour haut dans son champ récepteur. La matrice wup de distribution de synapse joue le rôle d'un filtre pour le contour haut dans le champ récepteur. Par analogie, le neurone N2 avec la distribution de poids synaptiques wdown peut mieux répondre à un contour bas dans le champ récepteur, le neurone N3 avec la distribution de poids synaptiques wleft peut mieux répondre à un contour gauche, et le neurone N4 avec la distribution de poids synaptiques wright peut mieux répondre à un contour droit. Le neurone (x', y') dans la couche de sortie intègre les sorties de ces quatre neurones de toutes les matrices de neurones dans la couche intermédiaire, et peut répondre à n'importe quelle direction de contour dans le champ récepteur RFrept.

4.3 Modèle du neurone impulsionnel et les champs récepteurs

Les résultats de simulation ont montrés que la conductance basée sur le modèle intègre-et-tire ressemble à celle du modèle de neurone de Hodgkin et Huxley [Koc99],[QWB06]. La conductance basée sur le modèle intègre-et-tire est appliquée au modèle du réseau présenté ci-dessus.

On pose, $G_{x,y}$ représente le niveau de gris du point $(x, y) \in RFrept$, $q_{x,y}^{ex}$ représente la conductance maximale causée par le courant excitateur du récepteur au point (x,y) , et $q_{x,y}^{ih}$ représente la conductance maximale causée par un courant inhibiteur d'un récepteur au point (x,y) . Pour simplifier, supposons que chaque récepteur peut transformer une valeur en niveau de gris à une conductance maximale par les expressions suivantes.

$$q_{x,y}^{ex} = \alpha G_{x,y}; q_{x,y}^{ih} = \beta G_{x,y} \quad (4.1)$$

Où α et β sont des constantes. Selon le modèle à base de conductance intègre et tire [QWB06][QWG05], le neurone N1 est régi par les équations suivantes.

$$\frac{g_{x,y}^{ex}}{dt} = -\frac{1}{\tau_{ex}} g_{x,y}^{ex}(t) + \alpha G_{x,y} \quad (4.2)$$

$$\frac{g_{x,y}^{ih}}{dt} = -\frac{1}{\tau_{ih}} g_{x,y}^{ih}(t) + \alpha G_{x,y} \quad (4.3)$$

$$c_m \frac{d\nu_{N1}}{dt} = g_l(E - \nu_{N1}(t)) + \sum_{(x,y) \in RF_{recp}} \frac{W_{x,y}^{upex} g_{x,y}^{ex}(t)}{A_{ex}} (E_{ex} - \nu_{N1}(t)) + \sum_{(x,y) \in RF_{recp}} \frac{W_{x,y}^{upih} g_{x,y}^{ih}(t)}{A_{ih}} (E_{ih} - \nu_{N1}(t)) \quad (4.4)$$

$g_{x,y}^{ex}(t)$ et $g_{x,y}^{ih}(t)$ sont respectivement les conductances des synapses excitatrices et inhibitrices,

τ_{ex} et τ_{ih} sont respectivement les constantes de temps pour les synapses excitatrices et inhibitrices,

$\nu_{N1}(t)$ est le potentiel de la membrane du neurone N1,

E_{ex} et E_{ih} sont les potentiels de réversion respectivement des synapses excitatrices et inhibitrices,

c_m représente la capacité de la membrane,

g_l représente la conductance de la membrane, ex désigne excitatrice et ih pour l'inhibitrice, $W_{x,y}^{upex}$ représente le poids des synapses excitatrices et $W_{x,y}^{upih}$ représente le poids des synapses inhibitrices,

A_{ex} est la région de la surface de la membrane reliée aux synapses excitatrices, et A_{ih} est la région de la surface de la membrane reliée aux synapses inhibitrices.

Selon la description biologiques des champs récepteurs [Kan81], les valeurs de $W_{x,y}^{upex}$ et de $W_{x,y}^{upih}$ sont données par les expressions suivantes.

$$w_{x,y}^{upex} = \begin{cases} 0 & si (y - y_c) \leq 0 \\ w_{emax} \exp\left(\frac{(x-x_c)^2}{\delta_x^2} - \frac{(y-y_c)^2}{\delta_y^2}\right) & si (y - y_c) > 0 \end{cases} \quad (4.5)$$

$$w_{x,y}^{upih} = \begin{cases} 0 & si (y - y_c) \leq 0 \\ w_{imax} \exp\left(\frac{(x-x_c)^2}{\delta_x^2} - \frac{(y-y_c)^2}{\delta_y^2}\right) & si (y - y_c) > 0 \end{cases} \quad (4.6)$$

Où $(x - x_c)$ est le centre du champ récepteur $RF_{rept}(x, y) \in RF_{rept}$, δ_x et δ_y sont des constantes, w_{emax} et w_{imax} sont les poids maximaux respectivement des synapses excitatrices et des synapses inhibitrices.

Par analogie, les neurones N2, N3 et N4 sont régis par un ensemble d'équations semblables à celle pour le neurone N1.

Quand le potentiel de la membrane atteint un seuil ϑ_{th} , le neurone génère une impulsion, ensuite il rentre dans une période réfractaire. Après la période réfractaire τ_{ref} le neurone peut intégrer des entrées pour produire d'autres impulsions.

Soit $S_{N1}(t)$ qui représente un train d'impulsions généré par le neurone N1.

$$S_{N1}(t) = \begin{cases} 1 & \text{si le neurone N1 dcharge l'instant } t \\ 0 & \text{si le neurone N1 ne dcharge pas l'instant } t \end{cases} \quad (4.7)$$

Par analogie, posons $S_{N2}(t)$, $S_{N3}(t)$ et $S_{N4}(t)$ représentant respectivement les trains d'impulsions pour les neurones N2, N3 et N4.

Le neurone $N_{\hat{x},\hat{y}}$ dans la couche de sortie est régi par les équations suivantes :

$$\frac{g_{\hat{x},\hat{y}}^{ex}(t)}{dt} = -\frac{1}{\tau_{ex}} g_{\hat{x},\hat{y}}^{ex}(t) + (w_{N1}S_{N1}(t) + w_{N2}S_{N2}(t) + w_{N3}S_{N3}(t) + w_{N4}S_{N4}(t)) \quad (4.8)$$

$$c_m \frac{d\nu_{\hat{x},\hat{y}}}{dt} = g_l(E_l - \nu_{\hat{x},\hat{y}}(t)) + \frac{g_{\hat{x},\hat{y}}^{ex}(t)}{A_{ex}}(E_{ex} - \nu_{\hat{x},\hat{y}}(t)) \quad (4.9)$$

Notons que le neurone $N_{\hat{x},\hat{y}}$ est connecté aux neurones intermédiaires seulement par les synapses excitatrices.

Posons $s_{\hat{x},\hat{y}}(t)$ représentant un train d'impulsions généré par le neurone $N_{\hat{x},\hat{y}}$ dans la couche de sortie. La fréquence de décharge pour le neurone $N_{\hat{x},\hat{y}}$ est calculée par l'expression suivante.

$$r_{\hat{x},\hat{y}} = \frac{1}{T} \sum_t^{t+T} S_{\hat{x},\hat{y}}(t) \quad (4.10)$$

En traçant cette fréquence de décharge comme une image avec une barre de couleurs, un graphe de contours est obtenu pour l'image d'entrée.

4.4 Résultats de simulation

L'image qui est montrée dans la Figure 4.2 est présentée dans le réseau, la carte de la fréquence de décharge dans la couche de sortie est obtenue comme on la voit dans la Figure 4.3 reflétant les contours de l'image présentée en entrée.

Les lignes de couleurs claires montrent que les neurones correspondants déchargent avec une haute fréquence et indiquent des contours avec un fort contraste. Les lignes de couleurs foncées montrent que les neurones correspondants déchargent avec une faible fréquence et indiquent des contours avec un faible contraste. En utilisant ces fréquences de décharges, des contours de différents contrastes peuvent être séparés.



FIG. 4.2 – Image originale

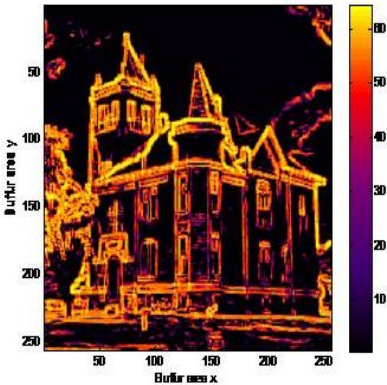


FIG. 4.3 – Carte de fréquence de décharge de la couche de sortie

Une comparaison a été faite avec les méthodes de détection de contours Sobel, Canny et Prewitt, les résultats de détection sont montrés dans la Figure 4.4.



(a) Contours Sobel



(b) Contours Prewitt



(c) Contours Canny



(d) Carte de fréquence de décharge du neurone

FIG. 4.4 – Comparaison de la carte de fréquence de décharge du neurone avec d'autres méthodes de détection de contours

4.5 Conclusion

Les réseaux de neurones impulsionnels sont dotés d'une structure hiérarchique qui est composée de neurones impulsionnels avec divers champs récepteurs et plasticités synaptiques. Les modèles de neurones impulsionnels fournissent une fonctionnalité très puissante pour l'intégration des entrées et la génération des impulsions.

Les synapses sont capables d'effectuer différents traitements, de filtres, d'adaptation ainsi que de propriétés dynamiques [LW04]. Divers champs récepteurs et des structures hiérarchiques des neurones impulsionnels permettent au réseau de neurones impulsionnels d'exécuter des calculs très compliqués, des tâches d'apprentissage et des comportements intelligents du cerveau humain.

Ce chapitre démontre comment un réseau de neurones impulsionnels peut détecter des contours dans une image. Bien que le fonctionnement du neurone dans le cerveau pour la détection de contours ne soit pas très clair, le modèle du réseau proposé est une solution

possible basée sur les neurones impulsionnels. Dans notre simulation, la carte de la fréquence de décharge du neurone peut être obtenue dans un intervalle de temps de 100 ms. Cet intervalle de temps est conforme à celui du système visuel biologique. Par conséquent, ce modèle peut être appliqué aux systèmes artificiels intelligents.

Chapitre 5

Conception et implémentation du suivi d'objet

5.1 Introduction

Le problème de suivi d'objet dans une séquence d'images peut être divisé en deux sous-problèmes, extraction de l'objet cible et son suivi le long de la séquence. En général ces deux tâches ne sont pas faciles en raison de présence des milieux encombrés et des occlusions fréquentes. Cependant, nous pouvons aisément trouver un grand nombre d'études sur des méthodes diverses employant une variété de descripteurs comme par exemple la texture, optical flow, le contour...etc.

Nous présentons dans cette partie une application du suivi d'objet basé sur un modèle de contours pour la caractérisation de cible, les contours sont obtenus par un réseau de neurones impulsionnels déjà présenté dans la partie précédente. Nous proposons de détecter et de suivre la cible par une méthode statistique. La position de cible est estimée en recherchant le modèle dans l'image courante en utilisant une distance Hausdorff. La cible est recherchée seulement dans une sous-fenêtre d'image contour courante.

Le modèle de cible est mis à jour dans chaque itération en employant la distance Hausdorff entre ce dernier et celui de l'image précédente de la séquence. Les paramètres utilisés dans notre programme sont ajustés de manière efficace le long de la trajectoire de la cible pour assurer un bon suivi. L'application a montré sa robustesse face aux changements d'illumination et aux variations de pose.

5.2 Détection d'objets en mouvement

Dans le cadre du suivi d'objet dans une séquence d'images, il est nécessaire dans une première étape de procéder à un prétraitement de l'image.

La première étape de prétraitement consiste à extraire le fond de l'image. La modélisation du fond présente l'avantage de permettre une détection plus précise et rapide des objets. La seconde étape détecte les pixels en mouvement, ceci en utilisant une technique de différence d'images entre chaque nouvelle image de la séquence et celle du fond pour détecter les objets mobiles dans la séquence d'images.

5.2.1 Suppression du fond (Background subtraction)

La méthode de soustraction de fond permet dans le cadre d'une scène statique d'extraire facilement les informations de premier plan de l'image. Parmi les méthodes les plus utilisées, on distingue les méthodes dites à différence temporelle seuillée. Cette technique rapide permet de soustraire le fond de l'image reçue et donne comme résultat une image binaire (Figure 5.1). La modélisation du fond de l'image se fait en premier lieu par une caméra fixe avec un fond peu dynamique. Les pixels appartenant au fond sont mis à 0, Les pixels appartenant à un objet mouvant sont mis à 1, le seuil de différence est déterminé par l'utilisateur.

Le principe est le suivant :

$$X_t(x, y) = \text{Seuil}(I_t(x, y) - b_t(x, y)) \quad (5.1)$$

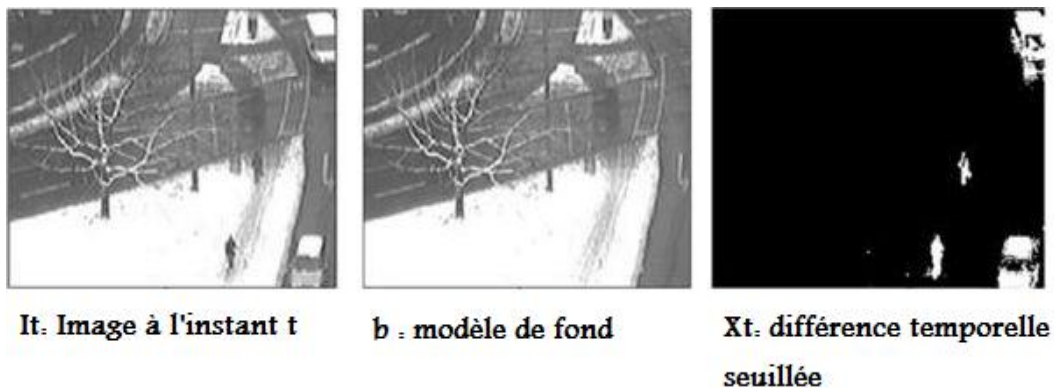


FIG. 5.1 – Soustraction du fond

Plusieurs critères peuvent être distingués pour cette technique, on peut avoir une différence des niveaux de gris des pixels, ou une différence des couleurs comme on peut avoir un ou plusieurs seuils de différence.

Comme toute autre technique utilisée, plusieurs problèmes peuvent apparaître durant cette étape dont on peut citer : camouflage, bruit, texture animée ou alors camera instable. Pour remédier à ces problèmes, et avoir un résultat efficace, il est nécessaire d'appliquer après l'étape de soustraction de fond, des opérations morphologiques.

5.2.2 Suivi d'objets par la distance de Hausdorff

Dans la vision par ordinateur, le suivi d'objet reste un problème difficile vu la complexité de la scène elle-même due à un certain nombre de transformations géométriques telles que la rotation, changement d'échelle, la translation et à l'illumination etc. Nous présentons dans cette partie une application d'un suivi qui permet d'atteindre le cheminement stable et précis de la cible.

La méthode utilisée dans cette partie permet après une phase d'extraction de repérer un objet non rigide se déplaçant tout en maintenant stable le mouvement d'appareil-photo. Pour l'extraction d'objet, nous détectons d'abord l'objet en utilisant la technique de background subtraction déjà définie et extrayons ensuite un modèle référence de l'objet dans chaque image de la séquence. La position de l'objet est localisée en utilisant une mesure de Hausdorff entre l'image contours référence et l'image contours cible, ceci revient à faire une mise en correspondance de celles-ci.

Le but de l'extraction d'objet dans une image est de mesurer la similitude entre deux objets extraits, le premier définissant le modèle le deuxième la cible. Cette similitude est définie par une mesure basée sur la distance de Hausdorff entre deux ensembles de points. Cette distance peut être employée pour comparer deux ensembles de points de contours extraits à partir de deux objets quelconques [DJ94].

L'approche de comparaison d'images utilisant la distance de Hausdorff permet de donner une mesure définie entre deux ensembles de points, ici représentant un modèle référence et un modèle cible. La distance de Hausdorff est fiable même lorsque l'image contient des objets multiples, du bruit, des occlusions, en outre, elle permet de résoudre les problèmes de rotation, scaling et de translation entre deux images [Ten02] en appliquant le processus de réduire au minimum la distance de Hausdorff sur les deux ensembles de points.

5.2.2.1 Comparaison d'images

La comparaison d'images intervient dans plusieurs domaines, elle permet de déterminer la similarité entre deux images en un sens prédéfini et peut prendre plusieurs formes.

Une méthode générale consiste à effectuer une comparaison directe des images, le cas des images binaires se ramène à cette situation. Dans le cas où les images binaires ne sont pas

de simples formes mais des images plus complexes composées éventuellement de plusieurs parties, une possibilité est de développer une méthode de comparaison locale qui puisse rendre compte des dissimilarités locales.

C'est cette étude que nous avons menée et qui est présentée dans cette partie. Elle a conduit au développement d'une méthode de comparaison directe d'images binaires qui donne une mesure locale des dissimilarités entre les images comparées, nommée distance Hausdorff.

5.2.2.2 Définition de la distance Hausdorff

La distance Hausdorff est une mesure scalaire d'une distance entre deux ensembles de points. En pratique, ces deux ensembles de points peuvent être obtenus par une détection de contour d'une image référence et une image cible [VL02] pour déterminer la position courante de l'objet sélectionné à l'intérieur de l'image.

Considérons l'interprétation de cette distance d'un simple point X d'un ensemble de points A . Quand nous disons que X est à une distance d de A , d est souvent considéré la distance euclidienne de X au point le plus proche de A . La distance Hausdorff élargie ce concept entre deux ensembles de points notés A et B .

Si nous déterminons la distance pour chaque point de B d'un ensemble de points comme ci-dessus, nous aurons alors N mesures de distances euclidiennes, où N est le nombre d'éléments de B . Puisque nous cherchons une mesure scalaire de la distance, nous choisissons la valeur maximum de ces mesures de distance qui est connu en tant que distance de Hausdorff directe.

5.2.2.3 Propriétés générales de la distance de Hausdorff

L'intérêt de cette mesure vient tout d'abord de ses propriétés métriques : positivité, identité, symétrie, et inégalité triangulaire. Ces propriétés correspondent généralement à notre intuition pour la comparaison d'images. En effet, une forme image est identique à elle-même, et l'ordre de comparaison n'a pas d'importance en général, dans le cas où l'ordre de comparaison est important, la propriété d'asymétrie de la distance de Hausdorff directe peut être exploitée. Enfin, l'inégalité triangulaire évite qu'une image soit similaire à deux images complètement dissimilaires en même temps.

De plus, la distance de Hausdorff ne nécessite pas l'appariement des points pour effectuer la mesure, elle est donc robuste aux déformations locales non rigides. Une autre qualité de la DH provient de la propriété de la translation. Cela implique que pour de petites translations, la DH sera petite, ce qui correspond à notre attente pour une mesure de

dissimilarité.

5.2.2.4 Distance Hausdorff totale

Etant donné deux ensembles finis de points $A = a_1, \dots, a_p$ et $B = b_1, \dots, b_q$, la distance Hausdorff totale est définie par :

$$H(A, B) = \max(h(A, B), h(B, A)) \quad (5.2)$$

où

$$h(A, B) = \max_{a \in A} \min_{b \in B} \|A - B\| \quad (5.3)$$

et $\|\cdot\|$ est une certaine norme fondamentale sur les points de A et B (par exemple la norme euclidienne).

La fonction $h(A, B)$ s'appelle la distance Hausdorff directe de A à B. elle identifie le point $a \in A$ qui est le plus loin de tout point de B et mesure la distance de a à son voisin plus proche de B, de même pour la distance Hausdorff inverse $h(B, A)$.

Les distances Hausdorff directe et inverse ne sont pas forcément égales, Pour créer une mesure de distance symétrique, on définit la distance Hausdorff totale H comme étant le maximum des distances directe et inverse. Plus la distance H est minime, mieux est la ressemblance entre les ensembles A et B.

Cette distance est théoriquement intéressante et pratiquement répandue. Cependant, elle est très sensible aux points isolés. De nombreuses versions modifiées ont été proposées pour améliorer sa robustesse, avec des applications qui vont de la recherche de forme dans une image [HR93],[Fel], pour l'extraction d'objet en mouvement et son suivi dans une séquence vidéo en temps réel [PL03],[HR93] à la reconnaissance de visages [TZL04].

5.2.2.5 Distance Hausdorff partielle

La distance Hausdorff directe $h(A, B)$ peut avoir une valeur petite lorsque chaque point de A est proche de certains points de B, par conséquent cette distance est une mesure très fragile parce qu'elle dépend d'un seul pixel isolé dans l'image entière causant une valeur de $h(A, B)$ grande. Cette particularité est commune lorsque l'objet est partiellement occulté, ou présente un bruit dans l'image dû à un détecteur de contours non fiable ou un bruit d'acquisition de caméra.

La définition de l'équation 5.3 est ainsi remplacée par une généralisation plus robuste de la distance de Hausdorff d'ordre statistique, parfois appelée "distance Hausdorff partielle"

en prenant la distance de rang k^{th} plutôt que de prendre la distance maximale. Ainsi la distance Hausdorff généralisée ou partielle est définie par :

$$h_K(A, B) = K_{a \in A}^{th} \min_{b \in B} \|a - b\| \quad (5.4)$$

Où k^{th} dénote la k^{th} valeur par ordre croissant de $d(a, B)$.

La distance Hausdorff partielle a été utilisée pour la mise en correspondance d'objet avec occlusion éventuelle. Elle donne de bons résultats en présence de bruit impulsionnel.

5.3 Implémentation du suivi d'objet

Nous proposons dans cette partie une méthode pour l'extraction et le suivi d'un objet non rigide se déplaçant tout en maintenant stable le mouvement d'un appareil-photo numérique. Pour l'extraction d'objet nous détectons d'abord un objet en utilisant la technique de soustraction de fond et extrayons ensuite les pixels en mouvement sous forme de blobs ou régions.

Le mouvement est obtenu par le suivi d'un blob ou région la plus significative de l'objet, c.à.d. la région qui contient le maximum de pixels en mouvement. La position de l'objet dans l'image courante de la séquence est déterminée en employant une mesure de Hausdorff entre l'image précédente considérée comme image référence et celle courante considérée comme image cible.

L'application consiste à réaliser les différents blocs de l'organigramme dressé dans la Figure 5.2, ce dernier explique le processus du principe général de l'application. Les tests de détection de contours et leurs suivis sont portés sur une séquence vidéo qui montre une scène filmée par un appareil photo numérique fixé, l'objet visé en mouvement avec 30 frames par seconde.

5.3.1 De la video à l'image

Pour le suivi du mouvement d'un objet dans une séquence d'images, il est nécessaire dans un premier temps d'extraire les images à partir d'un fichier vidéo, pour ce faire on a utilisé un logiciel de traitement d'images : ImageGrab 4.2.0.

ImageGrab est un petit logiciel qui ouvre toutes sortes de fichiers vidéo : avi, mpeg, wmv, des VCD ou des SVCD, des flux vidéo live et permet d'en extraire des images soit au format bmp, soit en jpeg avec une qualité ajustable. Il permet aussi de les copier dans le presse-papier de manière à les utiliser dans n'importe quelle application favorite.

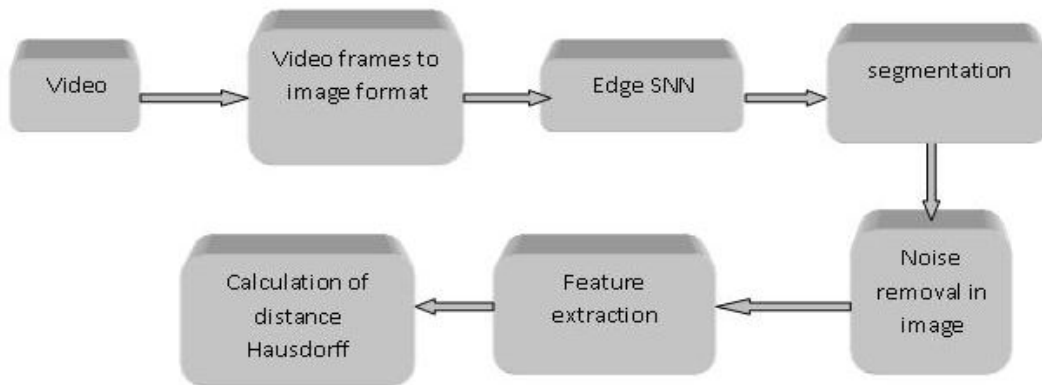


FIG. 5.2 – Processus du suivi d'objet

ImageGrab repose intensément sur directshow (les fonctions vidéos de DirectX), il est donc un logiciel entièrement consacré à l'extraction d'images de vidéo. Son ambition est de rendre cette opération facile, précise et efficace.

5.3.2 détection de contour par réseaux de neurones impulsionnels

Il s'agit de faire un suivi de contours avec une distance HD " tracking with hausdorff distance matching ", donc tout d'abord il faut détecter les contours dans chaque image de la séquence, ces derniers sont obtenus par un réseau de neurones impulsionnels déjà présenté dans le chapitre 4, chaque image de la séquence est présentée au réseau, une carte de la fréquence de décharge dans la couche de sortie est obtenue pour chaque image reflétant les contours de l'image présentée en entrée.



FIG. 5.3 – Carte de fréquence de décharge pour une image de la séquence

Toutes les images de la séquence sont donc présentées au réseau SNN, les cartes de fréquences de décharges obtenues sont sauvegardées dans une matrice de la taille de la

séquence en vue d'un traitement ultérieur.

5.3.3 Soustraction du fond pour la détection d'un objet

La première étape de tout système de traitement de séquence d'images pour le suivi de personnes consiste à détecter le mouvement des régions mobiles dans l'image.

Dans notre cas on a retenu la méthode par soustraction du fond procède par soustraction de l'image courante avec une image du fond pour segmenter l'avant-plan.

La soustraction du fond permet d'extraire les objets en mouvements mais le fond doit être bien modélisé, et cette méthode est très sensible aux changements de lumière ou aux mouvements des objets dans le fond mais elle est plus rapide que les autres méthodes.

Cette méthode calcule une carte binaire de mouvement, et les pixels de mouvement sont regroupés en " blobs ", régions de pixels connexes.



FIG. 5.4 – soustraction de fond

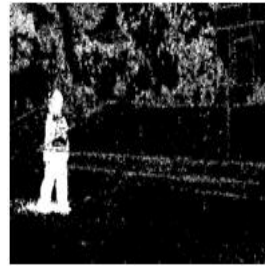
Le fond étant modélisé par une méthode statistique sans objet en mouvement. L'image résultante de la différence est ensuite seuillée afin d'obtenir un masque binaire.

En passant notre séquence par le notre réseau SNN, on obtient les résultats montrées dans la figure (Figure 5.6).

5.3.4 Supression du bruit

La technique de différence d'images seuillée présentée dans la Figure (5.5) montre qu'il existe toujours du bruit malgré que le seuil soit augmenté, ce qui pourra fausser les résultats attendus dans notre projet qu'est le suivi.

Pour cela, on a procédé a appliquer des opérations morphologiques : érosion suivi d'une dilatation, ce qui nous a conduit à de bons résultats. Voir Figure (5.7).



(a)Image 40 pour un seuil=10 (b)Image 85 pour un seuil=10



(c)Image 40 pour un seuil=30 (d)Image 85 pour un seuil=30

FIG. 5.5 – Influence du seuil

5.3.5 Segmentation et extraction de caractéristiques

La segmentation est un processus pour analyser plusieurs régions dans une image. L'algorithme de différence image est employé dans notre cas pour séparer les objets mobiles dans une image et donne de bons résultats quand un seul objet mobile doit être suivi.

L'extraction de caractéristiques nous permet dans le cadre de notre application de visualiser l'objet mobile dans la séquence d'images en dessinant un rectangle ou une boîte englobante dans la région considérée par le mouvement tout en utilisant ses dimensions.

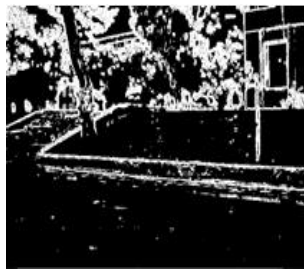
La caractéristique région se caractérise par l'extraction dans l'image courante des régions dénommées " blobs ", un ensemble de pixels connexes, regroupés en fonction d'un critère donné, ces régions homogènes seront considérées comme des éléments de base pour le suivi au cours de la séquence.

Quand l'algorithme est mis en application pour l'image produite à partir de la segmentation, des régions dans l'image binaire sont produites, ces dernières sont étiquetées et sont reliés aux régions correspondantes dans l'image binaire, une frontière rectangulaire de la surface est tracée pour la région considérée pour le suivi. Les régions correspondantes à l'image résultante après la phase d'extraction des caractéristiques seront étiquetées par des valeurs de nombre entier supérieur ou égal à 0. Les Pixel marqués par 0 correspondent au fond. Les Pixel marqués 1 correspondent à la région1, les Pixel ont marqués 2 correspondent à la deuxième région, et ainsi de suite.

Dans ce cas, le suivi temporel des objets revient à un problème de mise en correspondance



(a) Image référence



(b) Carte de contours par SNN de l'image référence



(c) Image 85



(d) Carte de contours par SNN de l'image 85



(e) Dif. de contours pour seuil=6

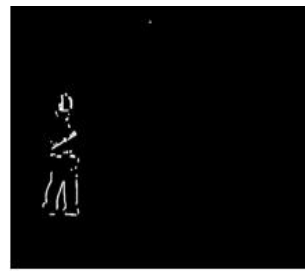


(f) Dif. de contours pour seuil=10

FIG. 5.6 – Soustraction de fond utilisant les cartes de contours SNN



(a) Dif. de contours pour seuil=10



(b) Erosion suivi d'une dilatation

FIG. 5.7 – Opérations morphologiques appliquée à l'image différence de contours

des régions segmentées entre deux images successives.

5.3.6 Suivi d'objet

Une segmentation au sens du mouvement par différence de l'image courante avec un modèle du fond est effectuée, elle sera suivie d'une mise en correspondance dynamique de régions.

L'utilisation de la méthode par soustraction du fond est faite pour de la détection des objets en mouvement pour un suivi dans un environnement complexe avec un fond connu. Une méthode de mise en correspondance des primitives des blobs par distance Hausdorff dans une séquence d'images est proposée. Le fond étant modélisé par une méthode statique.

Dans notre travail, nous présentons une méthode par comparaison d'images : après avoir détecté les blobs, les primitives sont extraites et comparées avec les primitives de blobs dans les images précédentes via une distance Hausdorff. Le suivi est désigné par un rectangle autour du blob contenant l'information de mouvement.

Pour les méthodes de mise en correspondance, on suppose que l'image est divisée en régions, chacune correspondant à un mouvement particulier et donc à un objet, le principe et le suivant : Trouver dans ces deux images $I(t-1)$ et $I(t)$ les régions qui correspondent au même objet physique, personnes ou autre.

– Mise en correspondance de blobs

La mise en correspondance de blobs peut être considérée comme un problème de comparaison de formes. Dans notre cas, il s'agit de faire un suivi basé sur la comparaison de modèles : un modèle de référence et un modèle de cible, dans lequel le modèle associé à l'objet est recherché dans l'image courante. La recherche est réalisée en calculant la similitude entre le modèle cible tirée de l'image courante avec le modèle référence tirée de l'image précédente.

Dans cette approche, Le modèle d'objet, qui est habituellement sous forme de carte de contours détecté par notre réseau de neurones impulsionnels, est mis à jour continuellement dans chaque image de la séquence après que l'objet soit localisé, c.à.d que le modèle cible devient le modèle référence. Cette mise à jour est conseillée afin de ne considérer que les petits mouvements d'un objet non rigide entre deux images consécutives.

La mise en correspondance de bloc appelée " block matching " exploite les redondances temporelles entre les images consécutives. Pour simplifier, on considère le mouvement entre deux images successives, l'image courante et l'image précédente appelée " image

de référence ". L'image courante est divisée en blocs de tailles variables. On suppose le mouvement uniforme dans chacun des blocs et pour chaque bloc d'une zone déterminée appelée "fenêtre de recherche" dans l'image précédente, on calcule un critère de comparaison entre les deux blocs. Le critère de comparaison entre deux blocs est une mesure de ressemblance entre les valeurs des pixels contenus dans chaque bloc.

Les objets extraits sous la forme de "blobs" à l'instant $t-1$ sont mis en correspondance avec ceux détectés à l'instant t . tout d'abord, on a pris la plus grande région au sens mouvement (c.à.d) celle qui contient le plus grand nombre de pixels en faisant un tri sur la taille des régions, cette région marquée comme X_i devient alors l'image référence. Ensuite, pour chaque image de la séquence ou image cible, après une phase de segmentation et d'extraction de régions I_k , il faut trouver la région $R_i \in I_k$ possédant la plus grande ressemblance avec une région $X_i \in I_{ref}$, une fois trouvée, la région R_i devient alors la référence. Autrement dit la position de l'objet à l'instant t est déterminé en fonction de sa position à l'instant précédent $t-1$, à partir d'un modèle de référence. Donc Chaque région est reliée à la région de l'image suivante par une meilleure similarité.

– **Choix de la distance Hausdorff minimale :**

Une fois les régions extraites, le but de cette extraction est de mesurer la similitude entre deux objets extraits, le premier définissant le modèle ou référence, le deuxième la cible. Cette similitude est définie par une de mesure basée sur la distance de Hausdorff entre deux ensembles de points.

Hausdorff métrique est une mesure mathématique pour comparer deux ensembles de points, Il existe plusieurs techniques pour calculer cette distance, la méthode que nous avons choisi est la mieux adapté à notre étude pour la comparaison :

Soit $X_i \in I_{ref}$, (I_{ref} : image reference) Et $R_i \in I_k$, (I_k image cible)

$$H(X_i, R_i) = \max(h(X_i, R_i), (R_i, X_i)) \quad (5.5)$$

$$h(X_i, R_i) = \max_{x \in X_i} \min_{r \in R_i} \|x - r\| \quad (5.6)$$

La comparaison de régions dénommées "blobs" est réalisée en utilisant une représentation basée sur les contours, c.à.d. que les deux régions sont deux ensembles de points de contours extraits à partir du même objet dans deux images de la séquence. On a employé la distance de Hausdorff totale pour construire une surface de corrélation à partir duquel le minimum est choisi comme nouvelle position de l'objet.

5.3.7 Résultats sur une séquence d'images

Le suivi d'un objet mobile est évalué sur un PC Pentium 4, avec une vitesse CPU de 3 Ghz et une RAM de 1 Go, les images extraites de la vidéo sont de type BMP de taille (480x640), la durée d'exécution est d'environ de 30mns sur une séquence de 150 images. L'exécution donne un suivi efficace, après la phase de séparation des objets de fond, l'objet est visualisé avec un rectangle tracé autour de la région considérée par le suivi. La distance hausdorff de déplacement entre 2 frames consécutives est calculée et comparée par rapport à un seuil, ce seuil est déterminée par l'utilisateur, s'il s'agit d'une même région cette distance devient nulle. Notre test porte sur une scène vidéo prise par un appareil photos numérique, le fond étant modéliser d'abord sans objet, ensuite le suivi est assuré par un rectangle tracé autour de celle-ci. L'application fonctionne très bien aussi longtemps que l'objet se déplace le long de la séquence. Quand l'objet disparaît de la scène, la boîte ou le rectangle autour d'elle disparaît aussi. le résultat d'exécution de quelques images de la séquence vidéo est monré dans la figure suivante :



FIG. 5.8 – Résultats de détection d'objet dans une séquence d'images

5.4 Conclusion

Dans ce chapitre, nous avons présenté une application de suivi d'objet en se basant les contours détectés lors de la phase de détection de contours avec un réseau de neurones impulsionnels. L'application de suivi avec mise en correspondance utilisant comme mesure de similarité la distance Hausdorff a montrée sa résistance au bruit.

Cette application peut être implémentée sur d'autres machines plus sophistiquées pour réduire le temps d'exécution. Dans l'avenir on peut l'améliorer pour un suivi robuste de plusieurs cibles puisque la mesure Hausdorff a prouvée sa robustesse face au bruit ainsi

qu'à l'occlusion. On peut aussi étendre l'idée de ce travail pour utiliser des applications en temps réel et à la classification d'objets.

Conclusion et perspectives

Ce travail a été l'occasion d'explorer les diverses possibilités d'application de méthodes d'analyse d'images au contexte de détection et suivi d'objet dans une séquence d'images par un réseau de neurones impulsionnels. Les spécificités des séquences vidéo traitées qui sont fixes et classiques ainsi que le choix de la méthode de suivi (tracking with hausdorff distance matching). Ce problème de détection d'objet dans une scène est plus difficile pour un objet non rigide que celui de tout autre objet rigide, du fait de la nature déformable des objets non rigides d'une part, et de la gestion des occultations, des conditions de lumière, des changements brusques dans le mouvement, ...etc d'autre part.

Dans notre mémoire nous avons démontré comment un réseau de neurones impulsionnels peut détecter des contours dans une image. Bien que le fonctionnement du neurone dans le cerveau pour la détection de contours ne soit pas très clair, le modèle du réseau utilisé est une solution possible basée sur les neurones impulsionnels. Dans notre simulation, la carte de la fréquence de décharge du neurone peut être obtenue dans un intervalle de temps de 100 ms. Cet intervalle de temps est conforme à celui du système visuel biologique. Par conséquent, ce modèle peut être appliqué aux systèmes artificiels intelligents.

Dans notre cas, un soin particulier a été donné aux aspects robustesse de notre méthode retenue dans laquelle nous avons intégré au modèle des caractéristiques calculées sur les régions pour assurer une correspondance des objets ainsi qu'à leur implémentation efficace en termes de temps de calcul via l'utilisation d'un réseau de neurone impulsionnels pour la détection de contours.

Bibliographie

- [ALP98] H. Fujiyoshi A.J. Lipton and R.S. Patil. Moving target classification and tracking from real-time video. *Proceedings of the DARPA Image Understanding Workshop(IUW'98), Monterey, USA*, pages 129–136, 1998.
- [AYS06] O. Javed A. Yilmaz and M. Shah. Object tracking : a survey. *Acm Computing Surveys*, 38(4) :9–11, 2006.
- [Boo04] O. Booi. Temporal pattern classification using spiking neural networks. Master's thesis, Université d'Amsterdam, 2004.
- [Bre03] R. Brette. *Modèles Impulsionnels de Réseaux de Neurones Biologiques*. PhD thesis, Université de Paris, Décembre 2003.
- [Bre04] Bremond. *Environnement de résolution de problèmes pour l'interprétation de séquences d'images*. PhD thesis, INRIA-Université de Nice Sophia-Antipolis, 2004.
- [Bug07] A. Bugeau. *Détection Et Suivi D'objets En Mouvement Dans Des Scènes Complexes, Application A La Surveillance Des Conducteurs*. PhD thesis, Université De Rennes 1, 2007.
- [Cou05] A. Coutant. La méthode des contours actifs en traitement des images. Master's thesis, Conservatoire National des Arts et Métiers Paris, Février 2005.
- [Cou06] A. Coulon. Dynamique et plasticité des réseaux de neurones à impulsions, etude du codage en population et des propriétés de traitement du signal. Master's thesis, Institut National des Sciences Appliquées, Lyon, Juin 2006.
- [Der87] R. Deriche. Using canny's criteria to derive a recursively implemented optimal edge detector. *International Journal of Computer Vision*, 2 :167–187, 1987.
- [DJ94] M.P. Dubuisson and A. K. Jain. A modified hausdorff distance for object matching. *Proceedings of the 12th IAPR International Conference on Pattern Recognition*, 1 :566–568, 1994.
- [DJG00] D. Weinshall D. Jacobs and Y. Gdalyahu. Classification with nonmetric distances : Image retrieval and class representation. *PAMI*, 6(22) :583–600, 2000.

- [Elo94] Y. Elomary. *Modèles déformables et multirésolution pour la détection de contours en traitement d'images*. PhD thesis, Université Joseph Fourier, Grenoble, 1994.
- [Fel] F.P. Felzenszwalb. Learning models for object recognition. Technical report, Artificial Intelligence Laboratory Massachusetts Institute of Technology Cambridge, MA 02139.
- [GC04] A. Ganoun and R. Canals. A new fast level set method. In *6th NORDIC Signal processing Symposium, ESPOO, Finland*, pages 9–11, 2004.
- [Ger91] W. Gerstner. Associative memory in a network of biological neurons. *Advances in Neural Information Processing Systems, Morgan Kaufmann Publishers, San Mateo*, 3 :84–90, 1991.
- [Gil03] M. Gilson. Entraînement de réseaux de neurones récurrents à pulses appliqué à la modélisation d'un tissu neuronal biologique. Master's thesis, Université de Montréal, Novembre 2003.
- [Her04] A. Marin Hernandez. *Vision dynamique pour la navigation d'un robot mobile*. PhD thesis, Institut National Polytechnique de Toulouse, 2004.
- [HH52] A. Hodgkin and A. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of physiology*, 117(4) :500–544, August 1952.
- [HM07] R. Martinez et S. Bengio H. Paugam-Moisy. A supervised learning approach based on stdp and polychronization in spiking neuron networks. *Verleysen, M., editor, ESANN'2007*, pages 427–432, 2007.
- [HNvdB01] M. Worring H.T. Nguyen and R. van den Boomgaard. Occlusion robust adaptative template tracking. *International Conference on Computer Vision, Vancouver, Canada*, 1 :678–683, 2001.
- [HR93] D.P. Huttenlocher and W.J. Rucklidge. A multi-resolution technique for comparing images using the hausdorff distance. *Proceedings of the IEEE, Conference on Computer Vision and Pattern Recognition*, pages 705–706, 1993.
- [HS85] R. H. Haralick and L.G. Shapiro. Image segmentation techniques. *Computer Vision Graphics Image Processing*, 29 :100–132, 1985.
- [JB03] S. Jehan-Besson. *Modèles de contours actifs basés régions pour la segmentation d'images et de vidéos*. PhD thesis, Université de Nice Sophia Antipolis, France, 2003.

- [JRdSV03] A. Shats J. Ruiz-del Solar and R. Verschae. Real-time tracking of multiple persons. *IEEE Computer Society Washington, DC, USA*, pages 109–114, 2003. 12th International Conference on Image Analysis and Processing.
- [JS04] B. Jung and Gaurav S. Sukhatme. Detecting moving objects using a single camera on a mobile robot in an outdoor environment. *8th Conference on Intelligent Autonomous Systems, Amsterdam, The Netherlands*, pages 980–987, 2004.
- [Kan81] Shwartz-J.H. Kandel, E.R. *Principles of Neural Science*. Edward Arnold, 1981.
- [Koc99] C. Koch. Information processing in single neurons. *Biophysics of Computation, Oxford University Press*, 1999.
- [Lei04] C. Leignel. *Une architecture à tableau noir pour la détection et le suivi d'une personne*. PhD thesis, Rfia, Toulouse, 2004.
- [Loi04] S. Loisel. *Exploration de réseaux de neurones à décharges dans un contexte de reconnaissance de parole*. PhD thesis, Université du Québec à Chicoutimi, 2004.
- [LW04] L.F.Abbott and W.G.Regehr. Synaptic computation. *Nature.*, 431 :796–803, 2004.
- [Maa97] W. Maas. Fast sigmoidal networks via spiking neurons. *Neural Computation*, (9) :279–304, 1997.
- [May02] R. Mayrhofer. A new approach to a fast simulation of spiking neural networks. Master's thesis, Université Linz, Juillet 2002.
- [Meu07] D. Meunier. *Une modélisation évolutionniste du liage temporel*. PhD thesis, Université Lumière Lyon2, Octobre 2007.
- [MMB07] R. Megret M. Mikram and Y. Berthoumieu. Evaluation des performances de descripteurs pour le suivi d'objets. *Colloque GRETSI, Troyes*, pages 11–14, Septembre 2007.
- [MOP97] P. Sinha E. Osuna M. Oren, C. Papageorgiou and T. Poggio. Pedestrian detection using wavelet templates. *In IEEE Proceedings Conf. on Computer Vision and Pattern Recognition*, pages 193–199, June 1997.
- [Nat99] T. Natschlager. *Efficient Computation in networks of spiking neurons, Simulations and Theory*. PhD thesis, Institute for theoretical computer Science, Graz, Aout 1999.

- [Noc08] A. Noce. *Vision et Reconstruction 3d Basée sur la Texture : Application à la Robotique Médicale*. PhD thesis, Université Montpellier II, Novembre 2008.
- [PGG03] J. Piater P. Gabriel, J. Verly and A. Genon. The state of the art in multiple object tracking under occlusion in video sequences. *Advanced Concepts for Intelligent Vision Systems*, pages 166–173, 2003.
- [PK06] F. Ponulak and A. Kasinski. Resume learning method for spiking neural networks dedicated to neuroprostheses control. *Dynamical principles for neuroscience and intelligent biomimetic devices In Proc. of EPFL LATSIS Symposium*, pages 119–120, 2006.
- [PL03] S.C. Park and S.W. Lee. Object tracking with probabilistic hausdorff distance matching. *Intelligent Robotics Development Program*, pages 136–173, 2003.
- [Pre04] F. Precioso. *Contours actifs paramétriques pour la segmentation d'images et vidéos*. PhD thesis, Université de Nice - Sophia Antipolis, Septembre 2004.
- [QWB06] L.P. Maguire B. Glackin Q.X. Wu, T.M. McGinnity and A. Belatreche. Learning mechanism in networks of spiking neurons. *Studies in Computational Intelligence, Springer-Verlag*, 35 :171–197, 2006.
- [QWG05] L.P. Maguire A. Belatreche Q.X. Wu, T.M. McGinnity and B. Glackin. Adaptive co-ordinate transformation based on spike timing-dependent plasticity learning paradigm. *LNCS, Springer*, 3610 :420–429, 2005.
- [Roc04] O. Rochel. *Une approche événementielle pour la modélisation et la simulation de réseaux de neurones impulsifs*. PhD thesis, Université Henri Poincaré-Nancy 1, Mars 2004.
- [SC92] J. Shen and S. Castan. An optimal linear operator for step edge detection. *Computer Vision, Graphics and Image Processing*, 54(2) :13–17, Mars 1992.
- [Sch08] B. Schrauwen. *Towards Applicable Spiking Neural Networks*. PhD thesis, Université de Gent, 2008.
- [SKS01] Z. Rasheed S. Khan, O. Javed and M. Shah. Human tracking in multiple cameras. In *Proceedings of the 8th IEEE International Conference on Computer Vision (ICCV 2001)*, pages 331–336, July 2001.
- [SLV02] V. Tiberghien S. Lefèvre, L. Mercier and N. Vincent. Multiresolution color image segmentation applied to background extraction in outdoor images. *IST European Conference on Color in Graphics, Image and Vision, Poitiers, France*, pages 363–367, 2002.

- [Tch02] A. Tchobanov. *Représentations et apprentissage des primitives phonologiques : approche neuromimétique*. PhD thesis, Université de Paris X-Nanterre, Février 2002.
- [Ten02] Y.C. Teng. *Remote-sensing Image Processing and Recognition Using Wavelet Transform and Hausdorff Distance*. PhD thesis, Institute of Computer Science and Information Engineering National Central University Chung-li, Taiwan 320, June 2002.
- [TR04] F.Jurie T.Chateau and R.Marc. Reconnaissance de gestes par vision monoculaire en temps réel : application la formation des chargés de manœuvres pour la conduite des ponts polaires. In *Workshop Acquisition du geste humain par vision artificielle, RFIA04, Toulouse, 2004*.
- [TZL04] R. Nevatia T. Zhao and F. Lv. Segmentation and tracking of multiple humans in complex situations. *IEEE PAMI*, 9, 2004.
- [VL02] D. Vignon and B. Lovell. Real-time hausdorff-based tracking. *DICTA2002 : Digital Image Computing Techniques and Applications, Melbourne, Australia*, pages 21–22, January 2002.
- [Voi02] N. Voisine. *Approche adaptative de coopération hiérarchique de méthodes de segmentation : application aux images multicomposantes*. PhD thesis, Université de Rennes I, 2002.
- [Zha96] Y. J. Zhang. A survey on evaluation methods for image segmentation. *Pattern Recognition*, 29(8) :1335–1346, 1996.