

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université des Sciences et de la Technologie d'Oran
Mohamed Boudiaf

Faculté des Sciences
Département Informatique

Thèse de Doctorat en Sciences
Spécialité : Informatique

Contribution des Réseaux de Neurones Temporels dans le Traitement des Images Animées

Présentée par : **Boudjelal Meftah**

Membres de jury :

Mohamed Benyettou	Professeur	Université Mohamed Boudiaf	Président de jury
Abdelkader Benyettou	Professeur	Université Mohamed Boudiaf	Directeur de thèse
Olivier Lezoray	Professeur	Université de Caen-France	Co-directeur de thèse
Nacéra Benamrane	Maitre de conférences	Université Mohamed Boudiaf	Rapporteur
Okba Kazar	Maitre de conférences	Université de Biskra	Rapporteur
Khalid Benabdeslem	Maitre de conférences	Université de Lyon 1-France	Rapporteur

Dédicace

*A mes parents,
A ma femme,
Et à mes deux enfants.*

Meftah Boudjelal

Remerciements

Mes remerciements s'adressent dans un premier temps aux personnes qui m'ont fait l'honneur d'accepter de siéger dans le jury de cette thèse : le professeur Mohamed Benyettou, directeur du laboratoire LAMOSI, pour m'avoir fait l'honneur de présider le jury de cette thèse. Madame Nacéra Benamrane, maître de conférences à l'université Mohamed Boudiaf, Monsieur Okba Kazar, maître de conférences à l'université de Biskra et Monsieur Khalid Benabdeslem, maître de conférences à l'université de Lyon1-France, pour l'intérêt qu'ils ont porté à ce travail en acceptant de faire partie du jury.

Comment ne pas continuer ce paragraphe sans remercier mon directeur de thèse. Le professeur Abdelkader Benyettou, directeur du laboratoire SIMPA, a été un encadrant essentiel à la progression de mes travaux. J'ai pu bénéficier de ses conseils judicieux et de ses visions d'ensemble du domaine.

J'exprime toute ma reconnaissance à Monsieur Olivier Lezoray, mon co-directeur de thèse, professeur à l'université de Caen en France pour m'avoir accueilli dans l'équipe Image du laboratoire GREYC. J'ai ainsi pu apprécier ses qualités humaines, sa bienveillance et la pertinence de sa conduite scientifique.

Je continue ces remerciements pour exprimer ma gratitude aux personnes qui ont eu le courage de m'encourager ou de me supporter depuis ma naissance : mon père, ma mère, ma sœur, mon beau père et mon oncle Djilali.

Pour terminer, je voudrais tout particulièrement remercier celle qui partage ma vie qui a fait preuve d'une patience exemplaire et qui a toujours été là pour m'aider.

Résumé

Les réseaux de neurones classiques peuvent offrir des solutions très intéressantes dans des applications de reconnaissance des formes ou approximation de fonctions mais ne peuvent en aucun cas être appliqués sur des données où le temps joue un rôle déterminant dans la résolution du problème.

Comment ces réseaux de neurones statiques peuvent être adaptés pour être dynamiques ? Quelles sont les architectures de réseaux temporels qui existent en littérature ? Comment est mené l'apprentissage pour la prise en compte de la dimension temporelle des données ? Nous essaierons de répondre à ces questions à travers cette thèse consacrée à la représentation du temps dans les réseaux de neurones en correspondance avec le traitement d'images animées.

Nos intérêts de recherche concernent principalement deux domaines complémentaires, à savoir, d'une part, l'étude et la synthèse des réseaux de neurones temporels et d'autre part, la mise en application des réseaux de neurones impulsionnels en traitement d'images animées.

Mots clés :

Réseaux de neurones temporels, Réseaux de neurones impulsionnels, Apprentissage, Codage, Traitement d'images animées, Détection de contours, Segmentation, Suivi d'objets.

ملخص

الشبكات العصبية التقليدية يمكن أن توفر حلولا مثيرة جدا للاهتمام في تطبيقات التعرف على نمط أو وظيفة تقريبا ولكن لا يمكن بأي حال من الأحوال أن تطبق على البيانات حيث الوقت يلعب دورا حاسما في حل المشكلة.

كيف يمكن تكيف هذه الشبكات العصبية الثابتة لتكون ديناميكية؟ ما هي تصميمات الشبكات الزمنية الموجودة؟ كيف يتم التعلم الذي يؤدي إلى مراعاة البعد الزمني للبيانات؟ سنحاول الإجابة على هذه الأسئلة من خلال هذه المذكرة من خلال تمثيل الوقت في الشبكات العصبية مع تطبيقاتها في معالجة الصور.

اهتماماتنا البحثية تمحورت أساسا في مجالين متكاملين، وهما، أولا، دراسة وتحليل الشبكات العصبية الزمنية وثانيا، تنفيذ الشبكات العصبية الاندفاعية في معالجة الصور المتحركة.

كلمات مفتاحية :

الشبكات العصبية الزمنية، الشبكات العصبية الاندفاعية، التعلم، والترميز، معالجة الصور، الكشف عن الحافة، تجزئة الصور، تتبع الأشياء.

Abstract

Conventional neural networks can provide very interesting solutions in applications of pattern recognition or function approximation but can in no way be applied to data where time plays a decisive role in resolving the problem.

How these static neural networks can be adapted to be dynamic? What are the temporal network architectures that exist in literature? How is learning led to the consideration of the temporal dimension of data? We try to answer these questions through this thesis on the representation of time in neural networks in correspondence with image processing.

Our research interests mainly concern two complementary areas, namely, first, the study and synthesis of temporal neural networks and secondly, the implementation of spiking neural networks in image processing.

Keywords:

Temporal neural networks, spiking neural networks, Learning, Coding, Image processing, Edge detection, Segmentation, Object tracking.

Table des matières

Liste des figures

Liste des tableaux

Liste des abréviations

Introduction Générale 17

Chapitre 1. Qu'est-ce que le temps ?

1.1 Introduction..... 22

1.2 Perceptions culturelles du temps 22

1.2.1 Une multitude de rapport au temps.....23

1.2.2 Richesses descriptives : ce que le temps n'est pas24

1.3 Conceptualisation scientifique du temps 25

1.3.1 Le temps instantané de la mécanique classique26

1.3.2 Le temps dépendant de la relativité générale27

1.3.3 La flèche du temps28

1.4 Conclusion..... 29

Chapitre 2. Réseaux de neurones temporels : état de l'art

2.1 Introduction..... 32

2.2 Architectures neuronales temporelles..... 34

2.2.1 Représentation spatiale du temps34

2.2.2.1 Time Delay Neural Networks (TDNN)34

2.2.2.2 Time Delay Radial Basis Function (TDRBF)38

2.2.2 Représentation dynamique du temps41

2.2.2.1 Représentation implicite du temps : Réseaux de neurones récurrents.....41

2.2.2.2 Représentation explicite du temps dans les réseaux de neurones56

2.3. Conclusion et Analyse des Architectures Neuronales Temporelles.. 61

2.3.1. Limites de la Représentation spatiale du temps	62
2.3.2. Limites de la Représentation dynamique du temps.....	62
2.3.2.1. Modèle de Hopfield.....	63
2.3.2.2. Réseaux à longues mémoires court-terme (LSTM).....	63
2.3.2.3. Réseau de neurones récurrents bidirectionnel	64
2.3.2.4. Réseaux de neurones à délais temporels adaptatifs (ATNN).....	64
2.3.3. Représentation du temps à base de SOMs.....	64
2.3.4. Liste d'applications	65

Chapitre 3. Réseaux de neurones impulsionnels

3.1 Introduction..... 70

3.2 Les bases biologiques 71

3.2.1 Anatomie d'un neurone biologique	71
3.2.2 Physiologie d'un neurone biologique	75
3.2.2.1 Le potentiel d'action	75
3.2.2.2 Le potentiel de repos	77
3.2.2.3 Potentiel postsynaptique.....	77

3.3 Codage neuronal de l'information..... 78

3.3.1 Coder l'information avec un seul neurone	79
3.3.1.1 Qu'est-ce qu'un code neuronal ?.....	79
3.3.1.2 Le codage par taux de décharge.....	80
3.3.1.3 Le codage temporel.....	81
3.3.1.4 Synthèse	82
3.3.2 Coder l'information avec une population de neurones	83
3.3.2.1. Le taux de décharge dans une population.....	83
3.3.2.2. Codage temporel dans les populations	84
3.3.2.3. Alors, quel codage ?.....	86

3.4 Modélisation des neurones impulsionnels 86

3.4.1 Modèle de Hodgkin et Huxley	87
3.4.2 Modèle Intègre et Tire.....	91
3.4.3 Modèle d'Izhikevich.....	95
3.4.4 Modèle à réponse impulsionnelle	95

3.5 Architecture des réseaux de neurones impulsionnels 99

3.5.1 Réseaux feed-forward	100
----------------------------------	-----

3.5.2 Réseaux récurrents	100
3.5.3 Réseaux hybrides.....	101
3.5.3.1 Chaîne de décharge synchrone	101
3.5.3.2 Machine à états liquides.....	102
3.6 Quelques propriétés des réseaux de neurones impuls ionnels	103
3.6.1 Réponses synaptiques	103
3.6.2 Délais de transmission dans les réseaux de neurones impuls ionnels.....	106
3.7 Apprentissage des réseaux de neurones impuls ionnels.....	107
3.7.1 Apprentissage supervisé basé sur le gradient.....	108
3.7.2 Apprentissage non supervisé basé sur la STDP.....	110
3.7.2.1 Plasticité synaptique : base cellulaire de la mémoire et l'apprentissage	110
3.7.2.2 Modèles formels d'implémentations de la STDP	112
3.7.2.3 Mise en œuvre de la STDP	113
3.8 Conclusion.....	114

Chapitre 4. Segmentation et détection de contours par réseaux de neurones impuls ionnels

4.1 Introduction.....	117
4.2. Segmentation par classification des images couleurs.....	117
4.2.1 Définition d'un classifieur	118
4.2.2 Classification à partir d'histogrammes	119
4.2.3 Classification par clustering	120
4.2.3.1 Classification de pixels non supervisée.....	120
4.2.3.2 Classification de pixels supervisée	120
4.2.4 Précautions liées à l'utilisation de la classification pixellaire.....	120
4.2.5 Mesures de distorsion	121
4.3 Bases d'images utilisées	122
4.4 Architecture, codage et apprentissage du réseau de neurones impuls ionnels	124
4.4.1 Architecture du réseau des neurones impuls ionnels	124
4.4.2 Codage des entrées.....	126
4.4.3 Apprentissage du réseau des neurones impuls ionnels.....	127
4.5 Influence des différentes paramètres du réseau de neurones impuls ionnels sur la segmentation	129
4.5.1 Influence du nombre de classes sur le résultat de la segmentation.....	130

4.5.2 Influence du nombre de sous-synapses sur le résultat de la segmentation	132
4.5.3 Influence du nombre de champs récepteurs sur le résultat de la segmentation.....	134
4.5.4 Influence de la taille du corpus sur le résultat de la segmentation	136
4.5.5 Influence du seuil de déclenchement téta sur le résultat de la segmentation.....	137
4.5.6 Influence du paramètre Tau sur le résultat de la segmentation	139
4.5.7 Influence des paramètres d'apprentissage b et c sur le résultat de la segmentation..	141
4.5.8 Meilleurs paramètres du réseau de neurones impulsionnels	144
4.6 Segmentation des cellules tumorales par réseaux de neurones impulsionnels	145
4.6.2 Résultats et évaluation de la segmentation cellulaire par les réseaux de neurones impulsionnels.....	147
4.7 Comparaison entre la segmentation par K-means, carte de Kohonen et les réseaux de neurones impulsionnels.....	149
4.8 Segmentation des images en contours	151
4.7 Conclusion.....	155

Chapitre 5. Suivi d'objets basé sur les réseaux de neurones impulsionnels

5.1 Introduction.....	157
5.2 Modèle d'un réseau de neurones impulsionnels pour la détection de contours	158
5.2.1 Architecture du réseau de neurones impulsionnels.....	158
5.2.2 Description Fonctionnelle du réseau.....	159
5.2.3 Modèle formel du réseau de neurones impulsionnels.....	160
5.2.4 Résultat de détection de contours.....	163
5.3 Extraction d'un modèle initial pour le suivi.....	164
5.4 Correspondance par la distance de Hausdorff.....	166
5.4.1 Distance de Hausdorff	166
5.4.1.1 Propriétés générales de la distance de Hausdorff.....	166
5.4.1.2 Distance de Hausdorff totale	166
5.4.1.3 Distance Hausdorff partielle.....	167
5.4.2 Mise en correspondance de blocs	167
5.5 Résultats sur une séquence d'images	168
5.5.1 Séquences d'images utilisées.....	168
5.5.2 Résultats de suivi.....	169

5.6 Conclusion.....	172
Conclusion Générale	174
Publications et citations de l'auteur	177
Bibliographie	180

Liste des figures

Chapitre 2. Réseaux de neurones temporels : état de l'art

Fig.2.1	Classification des réseaux de neurones temporels.....	32
Fig.2.2	Architecture du TDNN.....	36
Fig.2.3	Topologie du réseau TDRBF.....	39
Fig.2.4	Structure d'un réseau $x(y)$ -TDRBF pour une seule classe.....	40
Fig.2.5	Réseau de Hopfield complètement connecté.	43
Fig.2.6	Architecture de Jordan.	44
Fig.2.7	Architecture de Elman.	45
Fig.2.8	Architecture d'un réseau de neurones récurrents bidirectionnel.....	47
Fig.2.9	Propagation en avant de 2D RNN.	48
Fig.2.10	Propagation en arrière de 2D RNN.	49
Fig.2.11	Illustration de vanishing gradient problem.	50
Fig.2.12	Réseau de longue mémoire à court terme.....	51
Fig.2.13	Exemple d'un bloc de mémoire. Les cellules bleues représentent les unités multiplicatives.....	52
Fig.2.14	Schéma illustratif du fonctionnement du réseau SOM récursif.....	54
Fig.2.15	Schéma d'un réservoir de neurones.	55
Fig.2.16	Représentation d'un bloc de retards entre deux neurones.....	56
Fig.2.17	Exemple d'un ATNN à trois couches.....	57
Fig.2.18	Architecture du réseau PCNN.	59

Chapitre 3. Réseaux de neurones impulsifs

Fig.3.1	Morphologie d'arbres dendritiques.	72
Fig.3.2	Schéma d'un neurone biologique.	72
Fig.3.3	Microphotographie électronique d'une.	74
Fig.3.4	Propagation du potentiel d'action dans la synapse.	76
Fig.3.5	Description d'un potentiel d'action.....	77
Fig.3.6	Exemple de potentiels post-synaptiques.....	78
Fig.3.7	Codage par taux de décharge.	81
Fig.3.8	Codage temporel.....	81
Fig.3.9	Codage par modulation de l'activité de population.....	83

Fig.3.10 Codage temporel synchrone.	84
Fig.3.11 Codage temporel asynchrone.	85
Fig.3.12 Schéma électrique représentant les flux ioniques.	88
Fig.3.13 Schéma d'intégration de PPS par un neurone de type LIF.....	93
Fig.3.14 Forme du PA pour un neurone IF avec période.....	94
Fig.3.15 Schéma électrique du neurone Intègre et Tire à fuite.....	94
Fig.3.16 Exemple d'un PPSE.....	97
Fig.3.17 Comportement du modèle SRM.....	98
Fig.3.18 Architecture d'un réseau feed-forward avec multiples synapses.....	100
Fig.3.19 Architecture multi-couches de la chaîne de décharge synchrone.....	101
Fig.3.20 Schéma général d'une LSM à une entrée et une sortie.....	102
Fig.3.21 Trois formes de réponse synaptique les plus simples.....	105
Fig.3.22 Formes de réponses synaptiques.....	106
Fig.3.23 Variation d'amplitude du courant post-synaptique.....	111

Chapitre 4. Segmentation et détection de contours par réseaux de neurones impulsifonnels

Fig. 4.1 Quelques images de la base de Berkeley.....	123
Fig. 4.2 Quelques images microscopiques biomédicales.....	124
Fig. 4.3 Architecture du réseau de neurones impulsifonnels.....	125
Fig. 4.4 Codage d'une valeur réelle.....	127
Fig. 4.5 Fonction d'apprentissage gaussienne.....	129
Fig. 4.6 Résultat de segmentation par variation du nombre de classes.....	131
Fig. 4.7 Evaluation de la segmentation en variant le nombre de classes par MSE.	131
Fig. 4.8 Résultat de segmentation par variation du nombre de sous synapses...	133
Fig. 4.9 Evaluation de la segmentation en variant le nombre de sous synapses par MSE.....	135
Fig. 4.10 Résultat de segmentation par variation du nombre de champs récepteurs.....	135
Fig. 4.11 Evaluation de la segmentation en variant le nombre de champs récepteurs par MSE.....	135
Fig. 4.12 Résultat de segmentation par variation du corpus d'apprentissage.....	137
Fig. 4.13 Résultat de segmentation par variation du seuil téta.....	138
Fig. 4.14 Evaluation de la segmentation en variant téta par MSE.....	139
Fig. 4.15 Résultat de segmentation par variation du paramètre tau.....	140
Fig. 4.16 Evaluation de la segmentation en variant tau par MSE.....	141
Fig. 4.17 Résultat de segmentation par variation du paramètre b.....	142
Fig. 4.18 Evaluation de la segmentation en variant le paramètre b par MSE.....	143
Fig. 4.19 Résultat de segmentation par variation du paramètre c.....	144
Fig. 4.20 Evaluation de la segmentation en variant le paramètre c par MSE.....	144

Fig. 4.21 Topologie du réseau de neurones impulsionnels pour une sélection aléatoire.	146
Fig. 4.22 Topologie du réseau de neurones impulsionnels pour une sélection supervisée.....	146
Fig. 4.23 Segmentation des images microscopiques.....	148
Fig. 4.24 Résultat de segmentation par différentes approches pixellaires.....	150
Fig. 4.25 Architecture du réseau de neurones impulsionnels pour la détection de contours.....	151
Fig. 4.26 Zones d'activation des neurones de la couche de sortie.....	153
Fig. 4.27 Détection de contours par différentes approches.....	154

Chapitre 5. Suivi d'objet basé réseaux de neurones impulsionnels

Fig. 5.1 Processus de suivi.....	158
Fig.5.2 Modèle de réseau de neurones impulsionnels pour la détection de contours.....	159
Fig. 5.3 Différentes approches de détection de contours.....	164
Fig. 5.4 Soustraction de fond en utilisant les cartes de contours.....	165
Fig.5.5 Exemple de séquence sans occultations	169
Fig. 5.6 Exemple de séquence avec occultations	169
Fig. 5.7 Résultat de suivi d'une personne sans occultations.....	170
Fig. 5.8 Résultat de suivi d'une personne avec occultations.....	171

Liste des tableaux

Tableau 4.1 Meilleurs paramètres du réseau de neurones impulsionnels....	145
Tableau 4.2 Taux de classification.....	148
Tableau 4.3 Taux de segmentation en comparaison avec les travaux de Meurie.....	149
Tableau 4.4 Taux de segmentation par différentes approches pixellaires.....	151
Tableau 5.1 Taux suivi pour différentes séquences.....	172

Liste des abréviations

ATNN	Adaptable Time-Delay Neural Network
ATRBF	Adaptable Time Delay Radial Basis Function
BLSTM	Bidirectional Long Short Term Memory
BPTT	Back Propagation Through Time
EPSP	Excitatory Postsynaptic Potential
ESN	Echo State Network
H&H	Hodgkin and Huxley
IF	Integrate and Fire
LIF	Leaky Integrate and Fire
LSM	Liquid State Machine
LSTM	Long Short Term Memory
MDRNN	Multi-Dimensional Recurrent Neural Network
MLP	Multi Layer Perceptron network
M-SOM	Matricial Self-Organizing Map
MS-TDNN	Multistate Time Delay Neural Network
TDRNN	Time Delay Recurrent Neural Network
PCNN	Pulse Coupled neural networks
RecSom	Recursive Self-Organizing Map
RNN	Recurrent Neural Network
RSOM	Recurrent Self-Organizing Map
RTTL	Real Time Recurent Learning
SRM	Spike Response Model
STDP	Spike-Timing Dependent Plasticity
TDNN	Time Delay Neural Network
TDRBF	Time Delay Radial Basis Function
TDSOM	Time Delay Self-Organizing Map
TKM	Temporal Kohonnen Map
TOM	Time Organizing Map
TRBF	Time Radial Basis Function

Introduction générale

Introduction Générale

Comprendre les mécanismes à l'origine des fonctions supérieures du cerveau est l'objet de recherche au carrefour de la neurobiologie, de la psychologie, de l'informatique et de la physique. Dans ce vaste champ d'étude, le domaine des réseaux de neurones s'est tout particulièrement développée durant les années quatre-vingt. Les neurones sont des automates élémentaires dont le mode de fonctionnement s'inspire de celui des neurones biologiques. Il y deux motivations principales : d'une part la modélisation du cerveau, et d'autre part la réalisation d'algorithmes et de machines spécialisées dans des tâches de reconnaissance des formes par exemple, dont les performances pourraient être supérieures à celles des algorithmes et ordinateurs classiques.

Les réseaux de neurones classiques peuvent offrir des solutions très intéressantes dans des applications de reconnaissance des formes ou approximation de fonctions mais ne peuvent en aucun cas être appliqués sur des données où le temps joue un rôle déterminant dans la résolution du problème. Comment ces réseaux de neurones statiques peuvent être adaptés pour être dynamiques ? Quelles sont les architectures de réseaux temporels qui existent en littérature ? Comment est mené l'apprentissage pour la prise en compte de la dimension temporelle des données ? Nous essaierons de répondre à ces questions à travers cette thèse consacrée à la représentation du temps dans les réseaux de neurones en correspondance avec le traitement d'images animées.

Afin de situer précisément le contexte dans lequel nos travaux s'inscrivent, nous effectuons dans cette thèse une revue et synthèse des principaux types de réseaux de neurones temporels. L'objectif visé par cette présentation est de mettre en exergue les capacités des réseaux de neurones temporels les plus couramment utilisés, principalement dans le domaine de traitement d'images animées. Par la suite, nous serons en mesure de mieux justifier nos choix quant à l'approche que nous avons privilégiée, les réseaux de neurones impulsionsnels. Une mise en œuvre de la segmentation d'images et la détection de contours ainsi que le suivi d'objets par réseaux de neurones impulsionsnels est exposée dans cette thèse.

Pourquoi les réseaux de neurones impulsionsnels ?

Malgré tous les succès des réseaux de neurones artificiels classiques, les méthodes mises en œuvre pour les améliorer sont de plus en plus souvent des méthodes déjà connues (une saturation), et surtout on s'est beaucoup éloigné de l'inspiration biologique : Les réseaux de neurones artificiels classiques se basaient sur le postulat que les temps d'émissions des potentiels d'action ne portaient pas d'information suffisamment précises pour qu'il soit intéressant de les prendre en

compte. L'apparition de données expérimentales sur le codage de l'information du neurone biologique a poussé les modélisateurs à s'intéresser à des modèles où l'on tient compte de certaines propriétés temporelles intrinsèques du neurone.

Donc, pour une modélisation plus réaliste, il faut prendre en compte les aspects temporels et utiliser de nouveaux modèles de neurones : “les neurones impulsionnels” (Paugam-Moisy, 2006).

Plusieurs points ont été avancés comme soutien à l'affirmation d'une génération nouvelle de modèle de neurones.

- Les neurones impulsionnels offrent la possibilité d'introduire explicitement une dimension temporelle dans des réseaux de neurones. Ils se présentent ainsi comme une forme particulière de système dynamique.
- Les neurones impulsionnels sont capables de reproduire avec exactitude le fonctionnement des neurones classiques (Maass, 1997).
- Du fait de leur construction plus proche des neurones biologiques, ils permettent d'étudier aussi bien les fonctionnements unitaires que les interactions entre populations

Ainsi la compréhension et la maîtrise du fonctionnement des neurones impulsionnels représentent aujourd'hui un enjeu majeur de l'évolution des neurosciences computationnelles.

Organisation du document

Le présent manuscrit s'organise en cinq chapitres.

Tout d'abord nous définissons le temps. Quoi de plus compliqué que de définir cette notion qui ne cesse de nous échapper? La définition du concept de temps de façon générale dépend en effet de nombreux domaines allant de la Philosophie à la Physique en passant par la Cognition. N'ayant aucunement la prétention d'être ou physicien ou philosophe, nous utiliserons donc ce mot dans son acception de sens commun sans en chercher une définition exacte. Nous pouvons cependant dégager un certain nombre de propriétés attendues de façon minimale de la notion de temps. Celles-ci nous permettront de décrire sur quelles bases sont fondées les architectures neuromimétiques temporelles présentées dans cette thèse.

Dans le deuxième chapitre, nous étudions la manière dont certains modèles neuromimétiques intègrent la contrainte temporelle au niveau de l'architecture. Elle présente une classification des diverses architectures neuronales temporelles existantes, illustrée par les modèles les plus connus, puis détaille certains travaux saillants sur le traitement d'images animées. Nous présentons dans ce chapitre, divers aspects de la relation entre les réseaux de neurones et le temps.

Le troisième chapitre pose le problème de la modélisation temporelle du neurone. Partant de considérations biologiques, il présente l'étude des réseaux de neurones impulsionnels.

Nous présentons dans le quatrième chapitre la méthode de segmentation par classification pixellaires et la détection de contours. Nous présenterons également les précautions à prendre dans le choix de certains paramètres de la classification pixellaire ainsi que dans la construction des différentes bases d'images. Les paramètres du réseau de neurones impulsionnels pour la segmentation et la détection du contour sont ainsi exposés dans ce chapitre. Enfin nous illustrerons certains résultats de segmentation et de détection de contours sur des images issues de la base Berkeley et sur des images microscopiques de cellules de tumeurs bronchiques. Le modèle impulsionnel implémenté est le SRM (Modèle à réponse implusionnel).

Le dernier chapitre sera consacré à la présentation d'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 de type intègre et tire. 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.

CHAPITRE 1

Qu'est-ce que le temps ?

Sommaire

1.1 Introduction	22
1.2 Perceptions culturelles du temps	22
1.2.1 Une multitude de rapport au temps.....	23
1.2.2 Richesses descriptives : ce que le temps n'est pas	24
1.3 Conceptualisation scientifique du temps	25
1.3.1 Le temps instantané de la mécanique classique	26
1.3.2 Le temps dépendant de la relativité générale	27
1.3.3 La flèche du temps	28
1.4 Conclusion	29

1.1 Introduction

Le temps est une notion si quotidienne qu'on croit savoir ce qu'elle cache. Pourtant, sa nature intime reste un mystère, même pour les physiciens. Est-ce une propriété fondamentale de notre univers, ou le temps est-il une forme d'illusion ? Pourquoi le temps est-il si difficile à définir ? "Qu'est-ce que donc le temps ? Si personne ne me le demande, je le sais ; mais si on me le demande et que je veuille l'expliquer, je ne sais plus". Cette célèbre phrase d'Augustin illustre le paradoxe qui se pose lorsque l'on parle du temps. Immergés dedans (impossible de nous en extraire), nous ne savons comment le définir.

Intuitivement, il nous apparaît incompressible, inextensible, irréversible. Le problème, c'est que toutes ces expressions présupposent ou contiennent déjà l'idée du temps. En effet, au moins quatre difficultés nous empêchent de saisir réellement ce qu'est le temps. D'abord la pauvreté du langage : le mot "temps" ne désigne rien de concret. Ensuite, on ne peut s'extraire du temps (même sans événement, il reste). De plus, on confond souvent le temps (qu'on ne se sait donc pas définir) et sa fonction (renouveler le présent). Et enfin, le temps n'est perceptible directement par aucun de nos sens.

Ils sont nombreux les penseurs, philosophes et scientifiques, à s'être posé ces questions. Y'a-t-il un ou des temps ? S'écoule-t-il toujours de la même façon ? Passe-t-il tout le temps ? Combien de temps dure l'instant ? Le temps est-il subjectif ou objectif ? Réversible ou pas ? A-t-il un début ? Est-il intérieur ou extérieur à l'homme ? S'écoule-t-il ou le traversons-nous ? Pourrions-nous le maîtriser ? Ou au contraire en sommes-nous les prisonniers ? N'est-il qu'une illusion ? A quelques unes de ces questions nous allons tenter d'apporter des éléments de réponses.

1.2 Perceptions culturelles du temps

Toutes les cultures ont apporté des réponses nombreuses au questionnement sur le temps, et la plupart d'entre elles tournent autour des mêmes thèmes, la permanence du cosmos et la vie bref de l'homme, sont autant de dimensions temporelles partagées par la plupart des peuples de la Terre. Elles s'expriment dans le langage, dans les arts... Pourtant, toutes ne portent pas la même vision intime du temps. Le partage le plus évident pour l'observateur des civilisations est

sans doute la séparation entre une vision linéaire du temps, prévalant en Occident, et une vision cyclique de l'ordre temporel, prévalant par exemple en Inde.

1.2.1 Une multitude de rapport au temps

Écrire un récit, prédire le retour d'une comète, lister une série de dates : chacune de ces actions est directement liée au temps. Pourtant, il y joue des rôles divers. Il peut être essentiellement un repère plus ou moins explicite, comme dans le récit ou la liste de dates. Mais il peut également être l'objet d'étude de la connaissance. Dans tous les cas, il est essentiel de le quantifier pour l'aborder dans le détail, que cette quantification soit figurée ou bien précise et effective (réalisée avec un instrument de mesure). Il semble que le temps s'offre à l'homme d'abord comme un objet ambigu, dont la mesure permet de créer des repères, mais pas de le définir complètement. Les cultures asiatiques ont cultivé le goût d'un temps mystique, à la fois momentané et infini, illustré par exemple par le haïku japonais : la notion de flux y est prépondérante.

Parallèlement, des peuples d'Amérique du sud tels les Incas, ont privilégié une dimension rituelle du temps, où la discontinuité prévaut. Pour autant, toutes ces approches reposent sur une même sensation intime : il est donc plus évident encore que ce que l'homme a connu du temps au fil de l'histoire n'a pas été le temps pour lui-même, mais quelque manifestation culturelle rendue possible par une singularité particulière du temps, qui se donne à nous par certains aspects seulement.

Toutes ces traditions "inconscientes" auront une influence non négligeable sur le développement du concept du temps, que ce soit en sciences ou en philosophie. Elles manifestent les croyances d'un peuple à une époque donnée, et la façon dont ces croyances traduisent par l'imaginaire le ressenti, l'expérience. Plus la confrontation au temps sera fine et consciente, plus la conceptualisation du temps sera complexe : en effet, une caractéristique forte du temps dans les premiers âges de réflexion était son lien direct et exclusif avec le divin. Au fil des siècles, ce lien deviendra plus distant et sera même rejeté par certains.

Les sociétés modernes et industrialisées modifient sensiblement le rapport culturel et traditionnel au temps. Même là où les mythes et la religion perdurent, le temps du quotidien subit les assauts de l'instantané : médias, nourriture, déplacement...etc. L'ensemble des actes quotidiens s'accélère, de sorte que les contraintes du temps se font moins sentir ou deviennent au contraire plus criantes quand les facilités s'estompent. Quels que soient les avantages ou les pertes occasionnés par cette mutation parfois brutale, le temps culturel n'a jamais été et

n'est pas le temps de l'économie. La lenteur est une caractéristique fondamentale du rythme des sociétés humaines : il s'agit peut-être de la force d'inertie qui assure leur cohésion. Temps de la réflexion et temps de l'action entre en concurrence et se distordent, jusqu'à parfois faire éclater les repères psychologiques. Ainsi constatons que les zones urbanisées, où le temps personnel est très souvent sacrifié sur l'autel des contraintes (aller plus vite, à un autre rythme, et tout ce que cela présuppose et entraîne) sont les noyaux durs de la consommation de médicaments du type psychotropes. Dans son développement accéléré, l'humanité prend le risque d'altérer durablement son rapport au temps. Mais, au juste, quel rapport entretient-elle avec le temps ?

1.2.2 Richesses descriptives : ce que le temps n'est pas

Pour l'homme, le temps est essentiellement un support, et à ce titre il est orienté: il "coule" du passé au futur. Grâce au profond sentiment de durée, l'homme peut agir, se souvenir, imaginer, mettre en perspective... si bien que le temps lui est essentiel. Le niveau de complexité du rapport au temps est assez bien traduit par le langage, quoique imparfaitement : certaines cultures primitives ont peu de mots porteurs d'un sens temporel, et se situent essentiellement dans le présent et le passé. Pour les peuples anciens de Mésopotamie¹, par exemple, le futur est "derrière" et le passé, connu, est placé "devant". Dès lors qu'un peuple s'intéresse à l'avenir, toutefois, cet ordre intuitif s'inverse : on attend du temps qu'il nous apporte le moment suivant. Ce qui constitue une première confusion entre temps et mouvement. La simplicité de ce rapport s'estompe rapidement : bientôt, l'homme essaye de se jouer du temps. "Perdre son temps" ou "prendre son temps", ou toutes autres expressions de quelque langue que se soit, traduisent la volonté séculaire de gagner un contrôle sur ce temps subit.

Somme toute, c'est encore d'une conception faussement spatiale qu'il s'agit : pouvoir agir sur notre flèche du temps intime, la tendre, la distordre, l'infléchir. Mais le temps reste fidèle à lui même, et sa dimension rigide est également exploitée avec ténacité, par la quête de la juste et précise mesure. Quantifier, voilà une autre façon de décrire le temps qui fut engagée très tôt. Bien que privilégiée des sciences, elle n'en est pas moins source d'amalgames et de tromperie toujours renouvelées. Ainsi, compter le temps n'est pas le saisir en soi, car l'action de compter le temps, présuppose du temps. Quel est donc ce "vrai" temps qui mesure le temps, celui invoqué par le vocable "laisser le temps au temps" ? Cette question a laissé muettes des générations entières de penseurs; les disciplines modernes

¹ Désigne le pays entre deux fleuves est une région du Moyen-Orient. Elle correspond pour sa plus grande part à l'Irak actuel.

tentent d'y répondre en exhibant un temps pluriel, physique, biologique, psychologique, mais le temps de la vérité évidente ne semble pas encore venu.

Pour réfléchir au concept du temps, l'homme s'appuie sur son langage ; mais les mots sont trompeurs et ne nous disent pas ce qu'est le temps, ils viennent souvent nous dicter notre pensée et l'encombrer de préjugés sémantiques. La dimension paradoxale du langage temporel n'est pas très complexe : il suffit de s'attarder sur une simple expression courante comme "le temps qui passe trop vite" pour s'en rendre compte. Cette expression désigne un temps qui s'accélérerait. Mais l'accélération, c'est bien encore une position (spatiale) dérivée par rapport au temps : voilà que ressurgit le "temps cadre" immuable ! Le temps n'est ni la durée, ni le mouvement : en clair, il n'est pas le phénomène temporel. Ce n'est pas parce que des événements se répètent que le temps est nécessairement cyclique. Cette prise de recul, distinction entre temps et phénomène, sera relativement effective au cours de l'histoire en sciences et peut-être moins en philosophie, parfois victime des apparences sémantiques.

Toutefois, en distinguant ainsi le temps et les événements portés par lui surgit une dualité embarrassante : dans quelle réalité placer ces phénomènes qui adviennent, si ce n'est dans le temps lui-même ? Le sage dira, dans le "cours du temps". Cette scène animée des phénomènes est séduisante et juste, mais il faut prendre garde au piège sémantique. Le cours du temps illustre la sensation de chronologie imposée, qui est une propriété du temps pour lui-même. Il faut noter que rien ici n'indique encore l'idée de changement ou de variation. Il s'agit véritablement d'un cadre du Chronos. L'homme, pour sa part, devient, et les phénomènes, eux, adviennent. C'est là l'affaire de la flèche du temps, qui modélise les transformations au cours du temps, ou plutôt, "au cours du cours du temps". Elle est une propriété des phénomènes. Ces deux notions sont importantes et non intuitives, elles sont mélangées et brouillées par le langage en un seul et même tout, une fausse idée première du temps. La science, notamment, s'est appuyée sur elles pour édifier plusieurs visions successives du temps au fil de ses progrès.

1.3 Conceptualisation scientifique du temps

On distingue traditionnellement deux dimensions conceptuelles vis-à-vis du temps. La première peut être qualifiée d'objective : mathématisée, elle correspond au temps étudié par la science physique et à ce que les philosophes désignent usuellement par le terme de "temps". Coexiste avec cette première voie de conceptualisation une dimension plus subjective du temps, soulevant principalement la problématique de sa perception par l'homme. Il ne s'agit pas

seulement de la perception des durées, mais plus généralement de la compréhension de l'écoulement du temps.

La première de ces deux dimensions, le temps objectif, est affaire de mesure et de grandeur, lesquelles sont définies en rapport à une horloge invariable. L'unité légale du temps dans le système international est la seconde (et ses multiples), mais la quantification de cette unité a varié au cours de l'histoire. Le temps subjectif peut, en première approche, être défini comme l'altération psychologique du temps objectif. En "temps subjectif" les secondes peuvent paraître des heures et les heures, des secondes. La durée (i.e. l'impression subjective du temps) dépend particulièrement des émotions ressenties par la personne qui l'évalue. Autrement dit, il existe une horloge subjective et variable, qui bat la mesure en raison inverse de la concentration du sujet. Pour le sujet très pris par son activité (par exemple quand il s'amuse), l'horloge bat très lentement, il ne s'écoule que quelques battements et le sujet "ne voit pas le temps passer". Inversement, quand on s'ennuie ou qu'une situation est dramatique (lors d'un accident, par exemple), les phénomènes semblent se ralentir (l'attention cherche un appui), les battements s'accélèrent et on "trouve le temps long".

Toutefois, temps objectif et subjectifs ne suffisent pas à donner une vue complète de la complexité posée par la conceptualisation du temps. Premièrement, ce qui semblait objectif pour les premiers astronomes - pionniers de l'estimation du temps - s'est révélé bien subjectif, les premières mesures venues. Par la suite, plusieurs révolutions conceptuelles ont pris le contre-pied d'un modèle du temps objectif absolu et indépendant des phénomènes physiques. La relativité de la mesure du temps n'est ainsi pas une particularité de la condition humaine.

Deux domaines des sciences ont des implications profondes avec le temps : la physique et la biologie. Pour les physiciens, le temps a une importance cruciale dans la formalisation des systèmes et des lois de la Nature. D'abord considéré comme un paramètre, le temps a acquis une stature nouvelle avec la Théorie de la relativité. Les biologistes ne peuvent pas non plus ignorer le temps, puisque l'essentiel de l'étude du vivant a affaire avec les interactions dynamiques et les ajustements au cours du temps, quelle que soit l'échelle considérée : de la théorie de l'évolution à la multiplication des microbes, le temps impose son cours.

1.3.1 Le temps instantané de la mécanique classique

Le temps "scientifique", qui hérite directement des idées aristotéliennes, apparaît pour ainsi dire soudainement avec la révolution de la mécanique en Europe, au début du XVIIe siècle. Avec la formalisation mathématique du système Galileo-newtonien, le temps est défini comme un cadre absolu, un repère

inaltérable. Temps et espace constituent dans la mécanique dite "classique" les éléments indépendants d'un référentiel au pouvoir qu'on pourrait qualifier de régalien : les phénomènes physiques y trouvent leur place en tant que produits des lois, lesquelles sont totalement subordonnées au couple d'absolus formé par l'espace et le temps ce qui constitue une formalisation simple et élégante de la causalité, tout à fait intuitive pour le non scientifique.

Cette vision du monde physique vient en fait directement d'Aristote ; elle n'en a pas moins constitué une grande avancée dans la conceptualisation du temps car elle acquière, par les mathématiques, un statut nouveau : démystifié, débarrassé de ses attributs divins, le temps n'est plus considéré comme une qualité mais déjà comme une quantité. Cela est notamment lié aux progrès effectués dans la mesure du temps : le moyen-âge a vu l'horlogerie se développer considérablement, et l'idée d'un temps universel, sur la base duquel tous ces instruments de mesure pourraient être synchronisés, transparait finalement dans la physique et la pensée scientifique.

Ainsi, le temps prend le contre-pied de la simplicité mais aussi de l'histoire. C'est la naissance du premier temps physique construit, grâce au calcul différentiel. Exprimé sous la forme d'un rapport entre deux moments successifs, d'une "transformation invariante", le flux continu qu'est le temps en mécanique classique perd dès lors de sa substance : les lois absolues ne considèrent en effet que les conditions initiales et leur devenir au cours du temps désormais mathématisé.

Un tel temps, sans consistance propre, si éloigné de l'ontologie, fera dire à Kant que le temps est, non pas du tout une catégorie a priori de l'entendement, mais la forme a priori de la subjectivité humaine. Pour le physicien en mécanique classique, il s'agit très pragmatiquement de considérer que le temps n'est plus une qualité, comme chez la majorité des penseurs grecs, mais un mode de l'être et une condition de son devenir. Ainsi, un corps change, se déplace, mais le temps vient avec lui, dans son absolue invariabilité, comme une "persévérance de l'être". Il doit donc être supposé régulier - toujours régulier : le calcul différentiel construit et donne naissance au temps instantané, qui fixe une fois pour toutes les relations de la causalité et de la logique (les lois de la Nature). Cette conceptualisation du temps fut d'abord géométrique ; la mécanique analytique de Lagrange aidera à affiner la physique du mouvement sur la base de ce temps newtonien.

1.3.2 Le temps dépendant de la relativité générale

À l'époque du modèle newtonien, tous les physiciens ne sont pas en accord avec le modèle de temps et d'espace absolus qui prévaut désormais. Ainsi, Ernst

Mach considérait-il le temps comme un moyen d'exprimer facilement certaines relations entre les phénomènes, à l'instar de Leibniz. Cette façon de voir les choses peut sembler absurde à la lumière de la physique classique. Pourtant, les développements théoriques de Lorentz jusqu'à nos jours lui donneront, en partie tout du moins, raison.

Ce n'est pas tout : la relativité (générale) inverse une autre notion. Le temps et l'espace ne sont plus le cadre des phénomènes mettant en jeu la matière, mais au contraire, ce sont les corps qui influent sur le temps et l'espace. De plus, l'ordre des événements, du passé vers le futur, devient inséparable des propriétés de la lumière. Voilà donc un temps qui se dilate (pour un corps se déplaçant à grande vitesse par rapport à un autre) ou se raccourcit (pour un corps qui se déplace à faible vitesse par rapport à un autre), et qui semble tout sauf intuitif. Il est pourtant celui qui permet d'expliquer des phénomènes comme les lentilles gravitationnelles ou les trous noirs (pont d'Einstein-Rosen).

En 1936, Einstein et son collaborateur Nathan Rosen ont imaginé un raccourci dans l'espace-temps, appelé "pont d'Einstein-Rosen", pour relier deux points très éloignés dans l'univers par une incursion dans une quatrième dimension spatiale. Dans les années cinquante, le physicien John Archibald Wheeler leur donnera le nom de "Trous de Ver". Les trous de ver apparaissent dans des régions de l'univers où l'espace-temps est très courbé. Ces trous de ver constituent l'hyperespace, la quatrième dimension, l'au-delà, l'univers parallèle dans lequel se déplacerait celui qui voyage à une vitesse supérieure à celle de la lumière et le voyageur du temps.

Autre conséquence troublante : la notion de simultanéité perd de son absolu car tout dépend de l'observateur. Ce phénomène n'est toutefois flagrant que si les observateurs se déplacent l'un par rapport à l'autre à des vitesses proches de la vitesse de la lumière. La mesure du temps est donc différente selon la vitesse du référentiel dans lequel on se place. La théorie de la relativité indique donc que le temps soi disant objectif du physicien n'en est pas moins variable.

1.3.3 La flèche du temps

Les équations de la physique sont en effet bien souvent symétriques par rapport à une "inversion temps". C'est le cas de toutes les équations qui décrivent les phénomènes à une échelle microscopique. Ainsi, si on passe l'enregistrement d'une interaction physique se produisant à échelle microscopique, il est impossible de dire si l'enregistrement est passé à l'endroit ou à l'envers. Pourtant, à l'échelle macroscopique, certains phénomènes ne peuvent évidemment pas se passer à

l'envers. Ainsi, un œuf qui tombe par terre et qui se casse, ne rebondira jamais sur la table dans le même état qu'avant sa chute fatale.

Ludwig Boltzmann a tenté d'expliquer comment des phénomènes réversibles par rapport au temps à l'échelle microscopique peuvent conduire à une flèche du temps évidente à l'échelle macroscopique. Pour cela, il a développé la physique statistique, où les probabilités jouent un rôle très important. L'échelle microscopique, les constituants que sont les atomes ont un comportement individuel discontinu qu'on peut modéliser fidèlement par voie statistique - jusqu'à approcher la complète certitude pour certaines propriétés.

L'irréversibilité du temps, qui est exactement pour nous la direction de son écoulement, passe donc par la définition d'une orientation générale. Examiner le monde tel qu'il nous est donné, que ce soit par la physique ou par nos sens, flèche du temps et direction semblent cohérents. L'idée de remonter le temps, notamment, au sens d'en annuler son déroulement et d'en inverser le cours, est fondamentalement contre intuitive.

Pour l'heure, c'est donc la flèche du temps qui porte le concept moderne de temps, aussi bien en sciences qu'ailleurs. Ainsi, la physique n'impose pas une flèche du temps universelle : la biologie introduit son propre cours constant des choses, et on peut réfléchir à une flèche du temps psychologique. Et la physique elle-même a mis à jour plusieurs flèches du temps, adaptées à l'un ou l'autre de ses champs disciplinaires : on trouve ainsi pêle-mêle la flèche gravitationnelle, qui voit son expression la plus spectaculaire dans l'effondrement stellaire ; la flèche radiative, qui révèle que toute source radiative est vouée à s'éteindre. Nous nous créons intuitivement notre propre flèche du temps, du passé vers le futur, mais il est vrai que nous sommes plus sensibles aux durées qu'à la direction du temps - aussi la science conserve telle un "monopole" de l'étude des flèches temporelles. La diversité de ces modélisations est peut-être un indice de leur imperfection.

1.4 Conclusion

Nous pouvons cependant dégager un certain nombre de propriétés attendues de façon minimale de la notion de temps. Celles-ci nous permettront de décrire sur quelles bases sont fondées les architectures neuromimétiques temporelles qui seront décrit au chapitre suivant.

La première notion qui vient à l'esprit à propos du temps est celle d'**ordre** : tel évènement s'est produit avant tel autre ; il existe un passé et un futur. De plus, cet ordre est parfois total : lorsque deux évènements peuvent toujours être classés,

suivant la connaissance ou la conceptualisation que l'on en a. La notion d'ordre évoque naturellement celle de plus petit élément ("la nuit des temps") et de plus grand élément ("la fin du monde"). Ce sont là des questions qui dépassent notre propos car tous les problèmes temporels traités informatiquement (voire même par l'humain) sont à support fini. Ne serait-ce que pour des raisons pratiques, ils commencent et finissent à des moments parfaitement définis.

Au delà de l'ordre, une autre propriété attendue du temps est celle de sa **mesure**: "j'ai tapé cette phrase 1 seconde avant la prochaine" : la durée a une mesure. La métrique est donc une propriété intuitivement associée au temps.

Elle nous renvoie alors à celle de **densité** : existe-t-il toujours un instant entre deux autres ? Très souvent dans les problèmes d'ordre temporel, surtout avec les ordinateurs, le temps est échantillonné en une suite d'instant. Il perd sa notion (intuitive ?) de "continuité" ou "densité".

Lorsque cela n'est pas le cas, le temps est souvent (et surtout par les personnes ayant une culture scientifique) pensé comme la droite réelle \mathbb{R} (comme en Physique par exemple). Il faut remarquer que le modèle réel du temps incorpore bien entendu les notions précédentes (ordre, mesure, densité) mais y joint de plus toutes les propriétés mathématiques d'un paramètre par rapport auquel on peut par exemple dériver.

Un dernier aspect associé au temps mais malheureusement pas présent dans le modèle réel est celui de la flèche du temps : jusqu'à présent, et en dépit des nombreux fantasmes que cela a engendré, on ne peut pas retourner dans le passé. Il faut cependant noter que cette notion, même si elle n'est pas présente dans tel ou tel modèle de temps employé, est pratiquement toujours présente dans l'esprit du concepteur : on déplacera tel buffer de gauche à droite en non le contraire, on présentera toujours tel exemple après tel autre, jamais avant. L'ordre du temps ne peut être parcouru que dans le sens croissant.

Et pour terminer ce chapitre, je tiens à orienter le lecteur de ce manuscrit vers un ensemble de livres rédigé récemment sur le temps. On peut citer « Au commencement du temps » ([Bogdanov, 2009](#)), « Une belle histoire du temps », ([Hawking, 2009](#)), Trous noirs et distorsions du temps : L'héritage sulfureux d'Einstein ([Thorne, 2009](#)), « Voyage à travers le temps, un guide pour les débutants » ([Brennan, 2008](#)), « Qu'est-ce que la flèche du temps? » ([Chardin, 2007](#)), « Qu'est-ce que le temps ? Qu'est-ce que l'espace ? » ([Rovelli, 2006](#)) et « La Nature de l'espace et du temps » ([Penrose et al., 2003](#)).

CHAPITRE 2

Réseaux de neurones temporels : état de l'art

Sommaire

2.1 Introduction	32
2.2 Architectures neuronales temporelles	34
2.2.1 Représentation spatiale du temps.....	34
2.2.2.1 Time Delay Neural Networks (TDNN)	34
2.2.2.2 Time Delay Radial Basis Function (TDRBF).....	38
2.2.2 Représentation dynamique du temps	41
2.2.2.1 Représentation implicite du temps : Réseaux de neurones récurrents	41
2.2.2.2 Représentation explicite du temps dans les réseaux de neurones.....	56
2.3. Conclusion et Analyse des Architectures Neuronales Temporelles	61
2.3.1. Limites de la Représentation spatiale du temps.....	62
2.3.2. Limites de la Représentation dynamique du temps	62
2.3.2.1. Modèle de Hopfield.....	63
2.3.2.2. Réseaux à longues mémoires court-terme (LSTM).....	63
2.3.2.3. Réseau de neurones récurrents bidirectionnel	64
2.3.2.4. Réseaux de neurones à délais temporels adaptatifs (ATNN).....	64
2.3.3. Représentation du temps à base de SOMs.....	64
2.3.4. Liste d'applications.....	65

2.1 Introduction

Il faut commencer par distinguer les modèles qui utilisent le temps des modèles qui le traitent. Un modèle utilise le temps si d'une façon ou d'une autre son calcul se déroule dans le temps alors que la notion de temporalité est totalement absente du problème traité par l'architecture.

On dira qu'un modèle traite le temps s'il se préoccupe de problèmes d'ordre temporel (reconnaissance de parole, reconnaissance de mouvement,...etc.) en gardant d'une façon ou d'une autre la trace du temps (c'est-à-dire sans pré-codage extérieur au modèle qui transformerait le problème temporel en un problème purement statique). Dans ce dernier cas, on peut distinguer deux sous catégories : les modèles qui traitent le temps de façon externe, c'est-à-dire sans qu'aucun fonctionnement interne ne soit, à proprement parler, temporel, et les modèles qui traitent le temps de façon réellement interne.

L'hierarchie des réseaux de neurones temporels manipulant les images animées est illustrée dans la figure (Figure 2.1). D'autres approches temporelles existent et elles sont décrites dans la conclusion de ce chapitre.

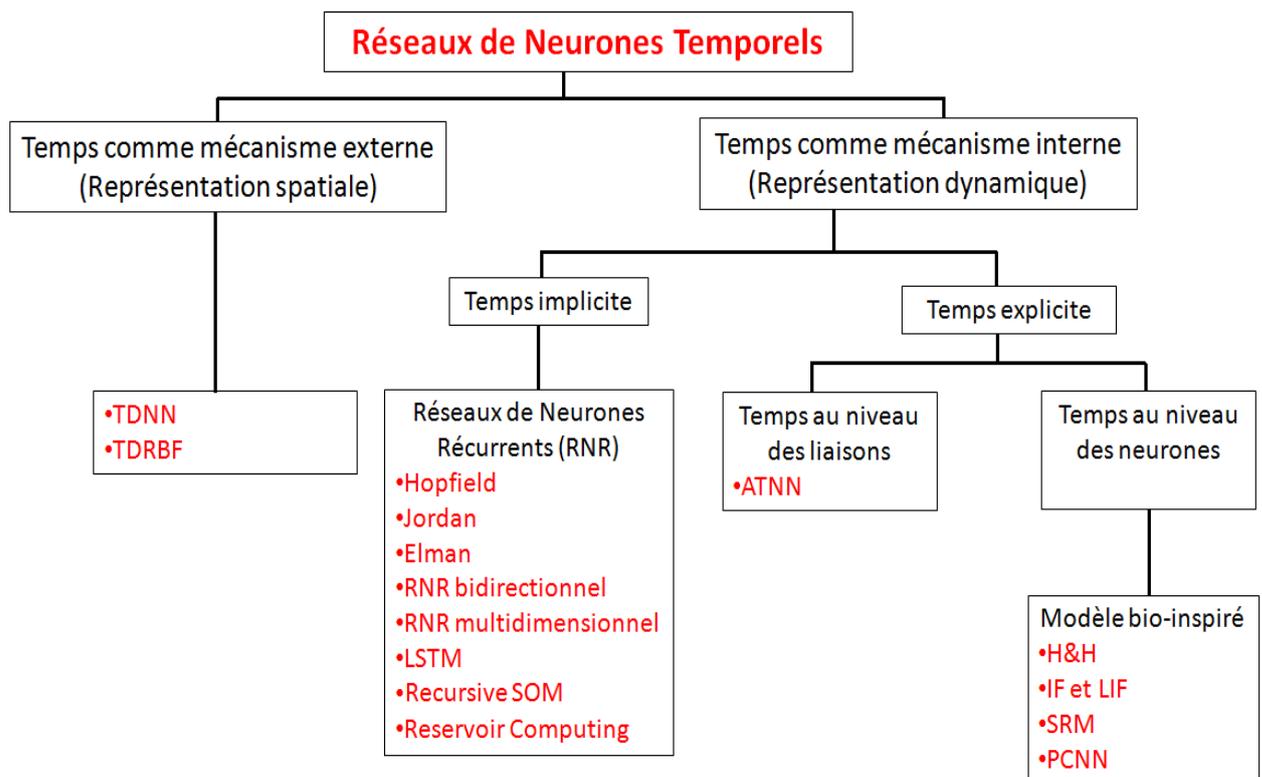


Fig. 2.1 Classification des réseaux de neurones temporels en fonction du degré d'intégration du temps dans le modèle.

La façon la plus primaire de traiter le temps est de le faire de façon externe de telle sorte qu'une architecture statique classique puisse effectuer en aval un traitement spatial statique. C'est par exemple le cas des TDNN (Time Delay Neural Networks) le temps est uniquement représenté par une transformation en espace. L'ordre temporel est conservé par des familles de poids liés (fenêtre glissante). Le réseau en aval est un MLP (Perceptron multicouches) qui accède à une information purement spatiale dont une dimension contient la sémantique du temps.

Une représentation toujours superficielle du temps, fondée sur des architectures classiques, est réalisée en réintroduisant en entrée du réseau l'état précédent du réseau (ou une partie de celui-ci). C'est le cas des réseaux récurrents. Le temps n'est utilisé ici que comme un index permettant la réintroduction successive du contexte (partie de l'état du réseau à l'instant précédent). Un tel modèle utilise donc le temps comme un mécanisme mais pas nécessairement pour traiter une dimension temporelle propre au problème considéré. On parlera alors de représentation implicite du temps lorsque seul son aspect ordre est préservé.

Représenter explicitement le temps dans le modèle neuromimétique employé constitue l'étape suivante de l'intégration. Elle peut se faire au seul niveau des liaisons du réseau ou au niveau du neurone lui-même. Introduire le temps au niveau des liaisons ou au niveau du neurone n'est cependant pas équivalent.

Introduire le temps au niveau des liaisons ou au niveau du neurone n'est cependant pas équivalent : dans le premier cas il résulte d'un processus global au niveau du réseau. La sortie de celui-ci dépendant alors d'une "tranche de temps" des entrées du réseau dont la durée dépend directement de la conception de l'architecture. Dans le cas d'un modèle temporel de neurone, il y a un traitement local du temps.

Dans ce chapitre, nous essaierons de répondre aux deux questions principales suivantes à travers un état d'art de la représentation du temps dans les réseaux de neurones ainsi que leurs applications en traitement d'images animées.

- Comment les réseaux de neurones statiques peuvent être adaptés pour être dynamiques ?
- Quelles sont les architectures de réseaux de neurones temporels qui existent en littérature ?

2.2 Architectures neuronales temporelles

2.2.1 Représentation spatiale du temps

La façon la plus simple et immédiate de représenter le temps dans les réseaux de neurones est d'utiliser une représentation spatiale du temps. L'information temporelle contenue dans les données est alors transformée en une information spatiale, c'est à dire une forme qu'il s'agit de reconnaître. Dès lors, les techniques de classification par réseaux de neurones habituellement employées deviennent applicables. Cette transformation du temporel en spatial s'obtient par l'utilisation classique de ligne à retard. Au lieu de présenter au réseau chaque événement, dès son apparition, il convient d'attendre un certain temps avant de procéder à la classification de la forme obtenue. Chaque retard temporel représente une dimension de la représentation spatiale.

Ce type de représentation du temps fait donc appel à un mécanisme externe qui est chargé de retarder ou de retenir un certain temps les données, ce qui conduit à l'appeler également représentation externe du temps. Nous présentons dans cette partie deux architectures neuronales utilisant ce principe : le TDNN (Time Delay Neural Networks) et le TDRBF (Time Delay Radial Basis Function).

2.2.2.1 Time Delay Neural Networks (TDNN)

Principe du fonctionnement

Les TDNN ont été développés à l'origine dans le cadre de la reconnaissance de phonème (Waibel et Lang, 1989) mais ont été depuis adaptés à d'autres domaines tels que la reconnaissance de l'écriture manuscrite (Guyon et al., 1991). Une approche liée montre qu'il a aussi des propriétés intéressantes en matière d'analyse de mouvement sur des séquences d'images (Ambellouis et Cabestaing, 1996).

Ce réseau de neurones devait satisfaire certaines conditions parmi lesquelles :

- le réseau doit contenir suffisamment de couches et d'interactions entre ces couches,
- pour pouvoir apprendre des données complexes non linéaires, le réseau doit être capable de représenter les relations temporelles entre les données d'entrée,
- le réseau doit être insensible aux positions temporelles absolues des données,

- l'algorithme d'apprentissage ne doit également pas prendre en compte les positions absolues des données,
- le nombre de poids dans tout le réseau doit être considérablement petit par rapport au nombre des données d'apprentissage. Cette condition permet au réseau d'encoder certaines régularités des données.

Les TDNN sont constitués comme les MLP (Perceptron multicouches) d'une couche d'entrée, de couches cachées et d'une couche de sortie mais ils se différencient de part l'organisation des liaisons inter-couches. Les TDNN introduisent des contraintes qui leurs permettent d'avoir un certain degré d'invariance par décalage temporel et déformation. Celles-ci utilisent trois idées : poids partagés, fenêtre temporelle et délai.

- Les poids partagés permettent de réduire le nombre de paramètres du réseau neuronal et induisent ainsi une capacité de généralisation plus importante. Les poids sont partagés suivant la direction temporelle, c'est à dire que pour une caractéristique donnée, la fenêtre associée à celle-ci aura les mêmes poids selon la direction temporelle. De plus, cette contrainte entraîne une capacité d'extraire les différences au fur et à mesure du balayage du signal. Ce concept de poids partagés est le comportement présumé du cerveau humain ou plusieurs neurones calculent la même fonction sur des entrées différentes.
- Le concept de fenêtre temporelle implique que chaque neurone de la couche $I+1$ n'est connecté qu'à un sous ensemble de la couche I (connectivité partielle). La taille de cette fenêtre est la même entre deux couches données. Cette fenêtre temporelle permet que chaque neurone n'ait qu'une vision locale du signal, il peut être vu comme une unité de détection d'une caractéristique locale du signal.
- En plus des deux contraintes précédentes, des délais sont introduits entre deux fenêtres successives pour une couche donnée.

De plus chaque couche a deux directions : une direction temporelle et une direction caractéristique.

L'architecture du réseau ([Figure 2.2](#)) comporte deux parties principales. La première, correspondant aux couches basses, implémente les convolutions successives permettant de transformer progressivement les caractéristiques en grandeurs de plus en plus significantes vis-à-vis du problème. La seconde correspond à un MLP classique, il reçoit en entrée l'ensemble des sorties de la partie TDNN. Ces deux blocs sont complètement paramétrables.

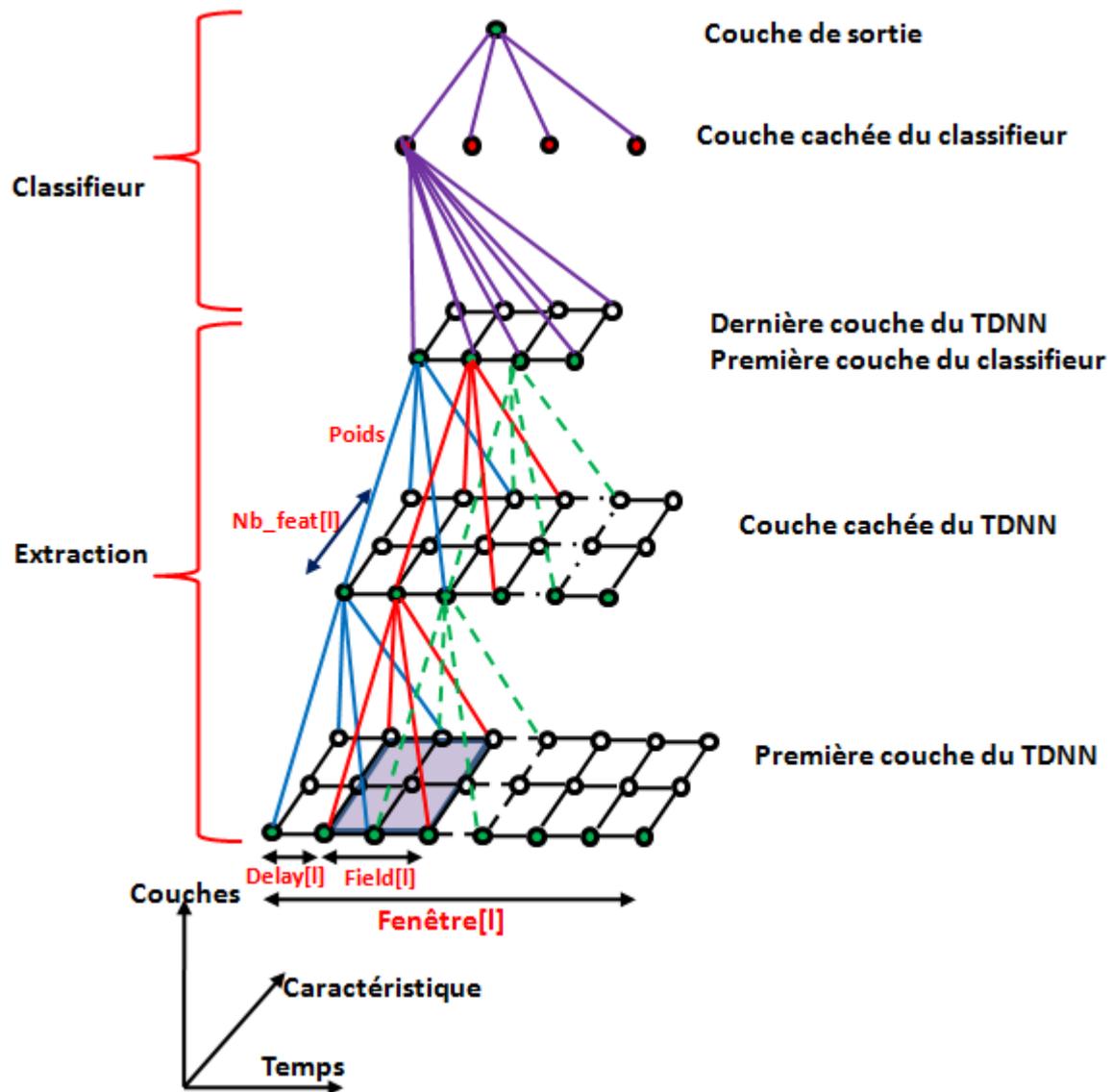


Fig. 2.2 Architecture du TDNN.

La partie extraction se caractérise par le nombre de couches, le nombre de neurones de chaque couche selon la direction temporelle (fenêtre), le nombre de neurones de chaque couche selon la direction caractéristique (nb_feat), la taille de la fenêtre temporelle vue par chaque couche (sauf celle d'entrée) soit le nombre de neurones de la couche i vu par un neurone de la couche $i+1$ ($field$) et le délai temporel (nombre de neurones) entre chaque fenêtre ($delay$). La partie classifieur se caractérise par le nombre de couches et le nombre de neurones de chaque couche.

Un neurone est identifié par sa couche i , sa caractéristique f , et son emplacement temporel t .

Le but du TDNN est non pas d'apprendre basiquement le signal mais d'extraire les caractéristiques de celui-ci. La première couche acquière le signal puis une ou plusieurs couches cachées transforment le signal en des vecteurs de caractéristiques. Un neurone donné détecte une caractéristique locale de la variation de la courbe. Le champ de vision du neurone est restreint à une fenêtre temporelle limitée. Avec la contrainte des poids partagés, le même neurone est dupliqué dans la direction temps (la même matrice de poids dupliquée) pour détecter la présence ou l'absence de la même caractéristique à différentes places le long du signal. En utilisant plusieurs neurones à chaque position temporelle, le réseau de neurone effectue la détection de caractéristiques différentes : les sorties des différents neurones produisent un nouveau vecteur caractéristique pour la couche supérieure.

La composante temporelle du signal d'origine est peu à peu éliminée au fur et à mesure de sa transformation en caractéristique par les couches supérieures, pour compenser cette perte d'information on augmente le nombre de neurones dans la direction caractéristique.

Applications en traitement d'images animées

Dans leur article, Le Callet et al. ([Le Callet et al., 2006](#)) décrivent une application de réseaux de neurones TDNN pour une mesure objective conçue pour évaluer automatiquement la qualité perçue des vidéos numériques. L'objectif de leur travail vise à automatiser le jugement humain en qualité subjective afin de remplacer son évaluation complexe et minimiser le temps consommé.

Une corrélation audiovisuelle est d'abord apprise par un TDNN, qui effectue ensuite une recherche spatio-temporelle au cours des séquences pour détecter les personnes en mouvement ([Zou et Bhanu, 2005](#)).

Après avoir extrait les de mouvement trajectoires, un TDNN est utilisé pour reconnaître les gestes de la main ([Ming et al., 2002](#)).

Cutler et Davis ([Cutler et Davis, 2000](#)) ont utilisé un TDNN pour apprendre la corrélation audiovisuelle afin de rechercher la personne qui parle dans une scène. Cependant, il est supposé que les objets dans la scène (par exemple, l'expression des visages) ne se déplacent pas de façon spectaculaire. Par conséquent, ils ne peuvent répondre aux changements dynamiques des objets.

Wohler présente un concept pour l'estimation simultanée des formes et analyse du mouvement par les TDNN ([Wohler et Anlauf, 1999a](#)) avec adaptation de la fenêtre temporelle. L'application de cette méthode a été conçue pour la

reconnaissance d'objets et l'analyse du mouvement des scènes du trafic urbain en temps réel.

Dans leur article, Yang et al. (Yang et al., 1997) abordent le problème de l'amélioration de la sécurité de conduite sur autoroute. Le champ de vision du conducteur dans le rétroviseur, qui se trouve limité par la présence d'un angle mort, ne lui permet pas toujours de détecter à temps un obstacle potentiel. Ils utilisent les images issues d'une caméra placée au niveau du rétroviseur latéral gauche, et effectuent une détection d'obstacles en utilisant un reéchantillonnage suivi d'une analyse de mouvement. L'estimation du mouvement est réalisée par un réseau de neurones à temps de propagation (TDNN).

Ambellouis et al. ont proposé une méthode d'analyse du mouvement local basée sur l'utilisation de réseaux de neurones à temps de propagation sur les connexions TDNN (Ambellouis et Cabestaing, 1996). Leur méthode s'avère très efficace pour estimer de façon précise les paramètres du mouvement lorsque la vitesse de l'objet varie très peu dans la séquence d'image.

2.2.2.2 Time Delay Radial Basis Function (TDRBF)

Principe du fonctionnement

Le TDRBF a été introduit par Berthold en 1994 (Berthold, 1994a) pour la reconnaissance de phonèmes. Ce réseau combine les caractéristiques du TDNN et des RBF. En d'autres termes, le TDRBF profite du traitement spatial des données d'entrée du TDNN et de la souplesse de l'apprentissage des RBF. Nous montrons sur la Figure 2.3 l'architecture du réseau. Une fenêtre temporelle constituée de plusieurs vecteurs d'entrée à différents instants ($t, t-1, \dots, t-\Delta t$) peut être vue comme un seul vecteur de grande taille. Chaque neurone caché calcule alors la distance par rapport à son prototype et donne en sortie la réponse de sa fonction d'activation.

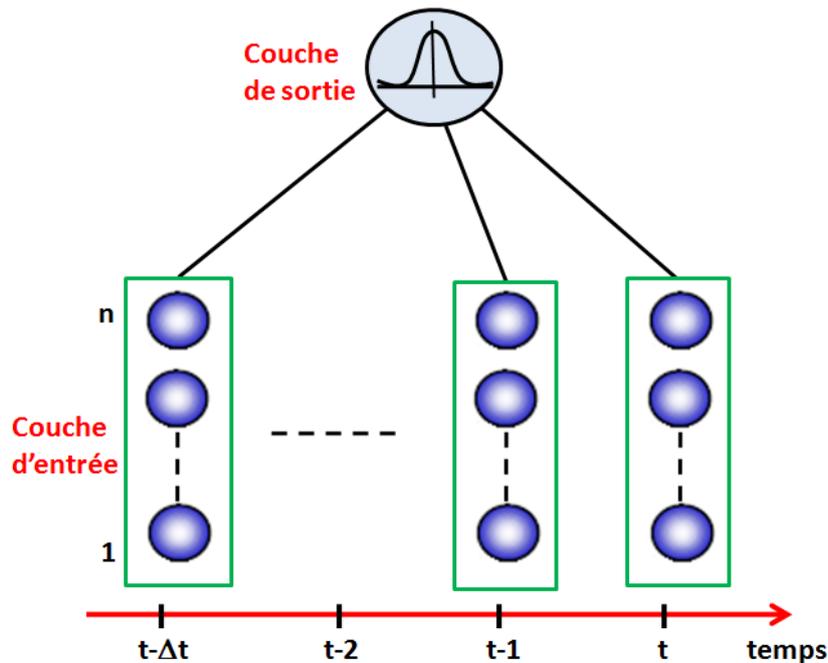


Fig. 2.3 Topologie du réseau TDRBF.

L'inconvénient de cette architecture est sa forte sensibilité aux positions temporelles absolues. En d'autres termes, si on présente au réseau un vecteur déjà mémorisé mais décalé dans l'espace temps, le réseau est incapable de le reconnaître. Pour contourner cet obstacle, les auteurs (Berthold, 1994b) ont proposé une couche supplémentaire appelée couche d'intégration qui permet d'intégrer dans le temps les sorties de chaque neurone à fonction de base radiale (Figure 2.4).

Les auteurs ont donc appelé ce réseau : $x(y)$ -TDRBF ; où x représente la taille de la fenêtre temporelle que chaque neurone RBF doit traiter et y le nombre de fenêtres que la couche d'intégration doit intégrer. Cette structure peut être moins sensible aux positions absolues des données d'entrée. Les auteurs ont utilisé la sigmoïde comme fonction d'activation de la couche d'intégration.

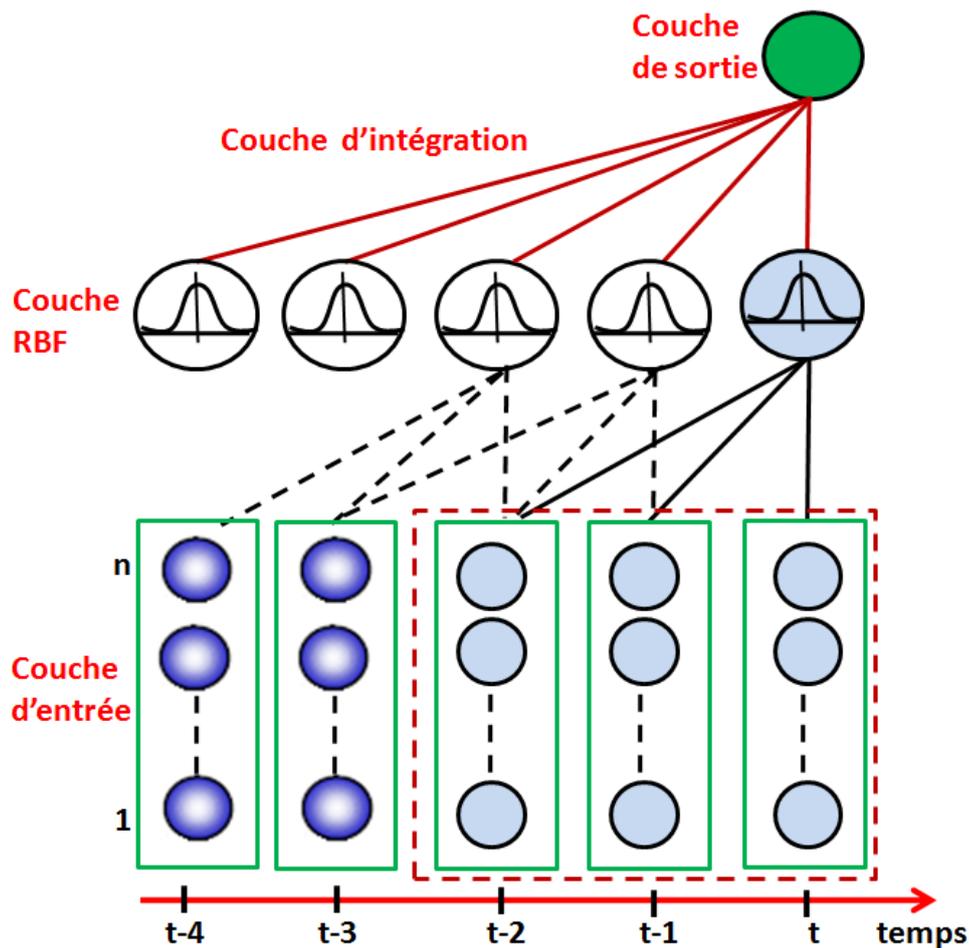


Fig.2.4 Structure d'un réseau x(y)-TDRBF pour une seule classe avec une fenêtre temporelle de taille 3 et une intégration de 5 fenêtres.

Le seul réglage nécessaire pour ce type de réseau correspond à trouver la relation entre l'espace temps de la couche RBF (noté x précédemment) et la taille de la fenêtre de la couche d'intégration (noté y précédemment). Ce dernier a été ajusté par tâtonnement. L'apprentissage du réseau consiste à déterminer les prototypes ainsi que les rayons d'influence des neurones RBF.

Applications en traitement d'images animées

Howell et Buxton (Howell et Buxton, 2002 ; 1999) proposent l'utilisation des réseaux TDRBF prenant en compte la dynamique temporelle des gestes. Un mécanisme de délais temporel est ajouté afin de manipuler le contexte temporel. Les expérimentations réalisées par Howell et Buxton concernant la reconnaissance de gestes dans le contexte d'interaction visuelle médiatisée. Les gestes réalisés face à une caméra permettant de la contrôler dans des applications de conférence vidéo. Quatre gestes sont considérés : deux gestes déictiques constituant à pointer

à gauche ou à droite et deux gestes sémiotiques de mouvement rapide de la main au dessus ou en dessous de la tête. Ils obtiennent, dans un premier temps, entre 69% et 75% pour un apprentissage effectué par une seule personne et une reconnaissance avec deux personnes différentes.

On peut également citer d'autres applications du TDRBF pour la reconnaissance du mouvement (Howell et Buxton, 2001) et des expressions du visage (Howell et Buxton, 1998).

2.2.2 Représentation dynamique du temps

Contrairement aux réseaux de neurones qui interprètent le temps comme un mécanisme spatial ou externe, les réseaux de neurones dits dynamiques selon Elman (Elman, 1990) ou « représentation interne » selon Chappelier (Chappelier, 1996a) traitent le temps d'une façon totalement interne au réseau.

Nous trouvons à l'intérieur de cette représentation dynamique du temps deux sous catégories : la représentation implicite du temps et la représentation explicite du temps.

Dans les deux cas, le réseau possède la capacité de mémoriser des informations soit implicitement par la récurrence des connexions, soit explicitement par des retards au niveau des connexions. Nous présentons ces deux techniques internes de traitement du temps.

2.2.2.1 Représentation implicite du temps : Réseaux de neurones récurrents

Les réseaux de neurones récurrents sont des réseaux de neurones comportant une ou plusieurs boucles de rétroaction. Dans ce type d'application, les cycles dans le graphe du réseau permettent de maintenir une représentation du contexte temporel associé aux entrées. Par conséquent, tout cycle du graphe des connexions d'un réseau de neurones récurrent doit comprendre au moins une connexion de retard unitaire (Dreyfus et al., 2010). Ceci donne lieu à des comportements dynamiques fort intéressants mais qui peuvent être très complexes (Haykin, 1999).

Ainsi, ces réseaux possèdent une représentation implicite du temps : l'ordre temporel n'est pas représenté explicitement en tant qu'ordre des composantes du vecteur d'entrée, mais de manière implicite, par les effets qu'il a sur les traitements.

Indépendamment du grand nombre d'architectures récurrentes spécifiques, il existe plusieurs algorithmes d'apprentissage s'inspirant de l'algorithme de rétro-propagation des réseaux « feedforward » et l'adaptant aux réseaux récurrents. Il s'agit des algorithmes de "Back-Propagation Through Time" (BPTT) (Werbos, 1990) et du "Real Time Recurrent Learning" (RTRL) (Williams et Zipser, 1989).

Nous référons au chapitre 15 du livre de Simon Haykin (Haykin, 1999) pour une présentation détaillée de la théorie associée aux réseaux récurrents et au chapitre intitulé « Recurrent Networks: Learning Algorithms » pour une étude plus fine des algorithmes d'apprentissage de réseaux de neurones récurrents (Arbide, 2003).

Une liste non exhaustive d'architectures de réseaux de neurones récurrents est présentée dans l'article de Chung (Chung, 1998).

a. Le réseau de neurones de Hopfield

Principe du fonctionnement

D'abord constitué de neurones binaires (Hopfield, 1982), puis de neurones à valeur réelles (Hopfield, 1984), le réseau de John Hopfield est basé sur le modèle de William Little (Little, 1974).

Une différence importante du réseau de Hopfield par rapport à celui de Little concerne l'asynchronisme du réseau. En effet, Hopfield considère qu'il n'existe aucune preuve de synchronisme dans le cerveau et par conséquent chaque neurone du réseau est responsable de la mise à jour de son état.

Il faut noter qu'un réseau de Hopfield est dépourvu d'entrées externes: son comportement est autonome, dicté uniquement par sa dynamique propre. Pour assurer qu'un tel réseau est stable ², et pour calculer facilement ces états d'équilibre, Hopfield introduit une règle qui n'a rien de biologique : la symétrie des connexions. Les connexions sont symétriques : le poids w_{ij} de la connexion reliant le neurone i au neurone j est égal au poids w_{ji} de la connexion qui relie le neurone j au neurone i ; de plus, à chaque connexion est associé un retard égal à une unité de temps.

La Figure 2.5 présente le modèle général de Hopfield. Il n'y a pas de distinction entre neurone d'entrée et neurone de sortie ; seule importe la matrice des poids. Les neurones sont en effet reliés deux à deux par des connexions

² C'est-à-dire que, quel que soit son état initial, il évolue jusqu'à ce qu'il ait atteint un état d'équilibre, indépendant du temps.

symétriques. Ces connexions supportent les motifs de connaissance que le réseau apprend.

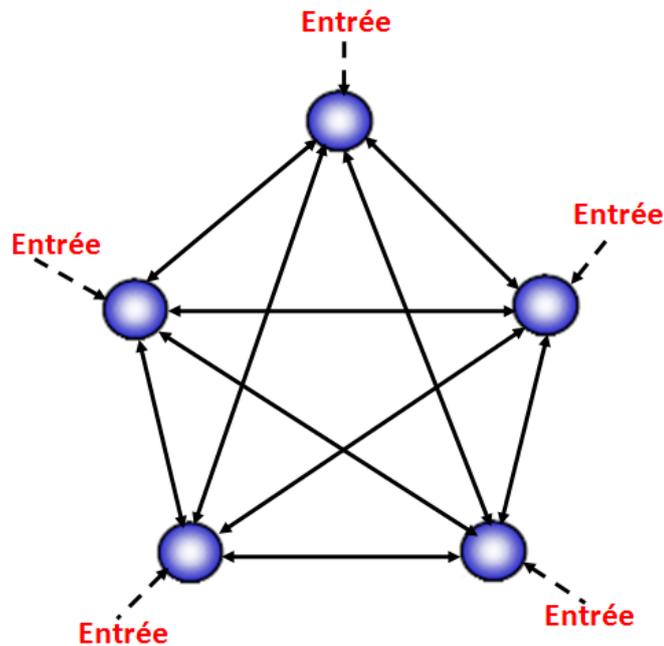


Fig. 2.5 Réseau de Hopfield complètement connecté à connexions symétriques.

L'apprentissage d'un tel réseau consiste à présenter des vecteurs d'entrée prédéfinis. Le but du réseau est de pouvoir reconnaître dans un autre vecteur d'entrée une forme apprise relativement proche. L'apprentissage du réseau consiste alors à calculer les paramètres du réseau de telle manière que les codes des informations que l'on souhaite mémoriser soient des états stables du réseau. L'application d'une règle de type Hebb ou bien delta agit sur les connexions afin de stocker le motif correspondant.

Le modèle de Hopfield présente tout de même quelques inconvénients. En premier lieu, un tel réseau n'est capable d'activer qu'un seul motif simultanément. Si des caractéristiques appartenant à deux motifs ne possédant aucun neurone commun sont activées, le réseau n'en reconnaîtra qu'un seul, selon un choix basé aléatoirement sur le temps de réponse et de mise à jour des neurones. Ensuite, les motifs doivent être appris simultanément. Un apprentissage incrémental donne des résultats chaotiques durant la phase de rappel des connaissances.

Applications en traitement d'images animées

Sulehria et al (Sulehria et Zhang, 2008 ; Sulehria, 2007) ont fourni un rapport détaillé sur l'application du réseau de neurones Hopfield dans plusieurs

domaines et particulièrement en traitement d'images (Sun et al., 2000) et en reconnaissance de formes (Pervej et al., 2008 ; Young et al., 1997).

Pajares (Pajares, 2006) propose une approche pour la détection de changements dans une suite d'images en utilisant le réseau de Hopfield. Pour valider son modèle, l'auteur utilise quatre bases de séquences d'images. Les résultats trouvés sont performants.

b. Architecture de Jordan et de Elman

Principe du fonctionnement

Dans l'architecture (Figure 2.6) proposée par Jordan (Jordan, 1986), les unités de la couche de sortie sont dupliquées sur une couche appelée couche de contexte. Les unités de cette couche tiennent également compte de leur propre état à l'instant précédent. Cette connexion récurrente d'une unité de contexte à elle-même lui donne une dynamique ou une mémoire individuelle.

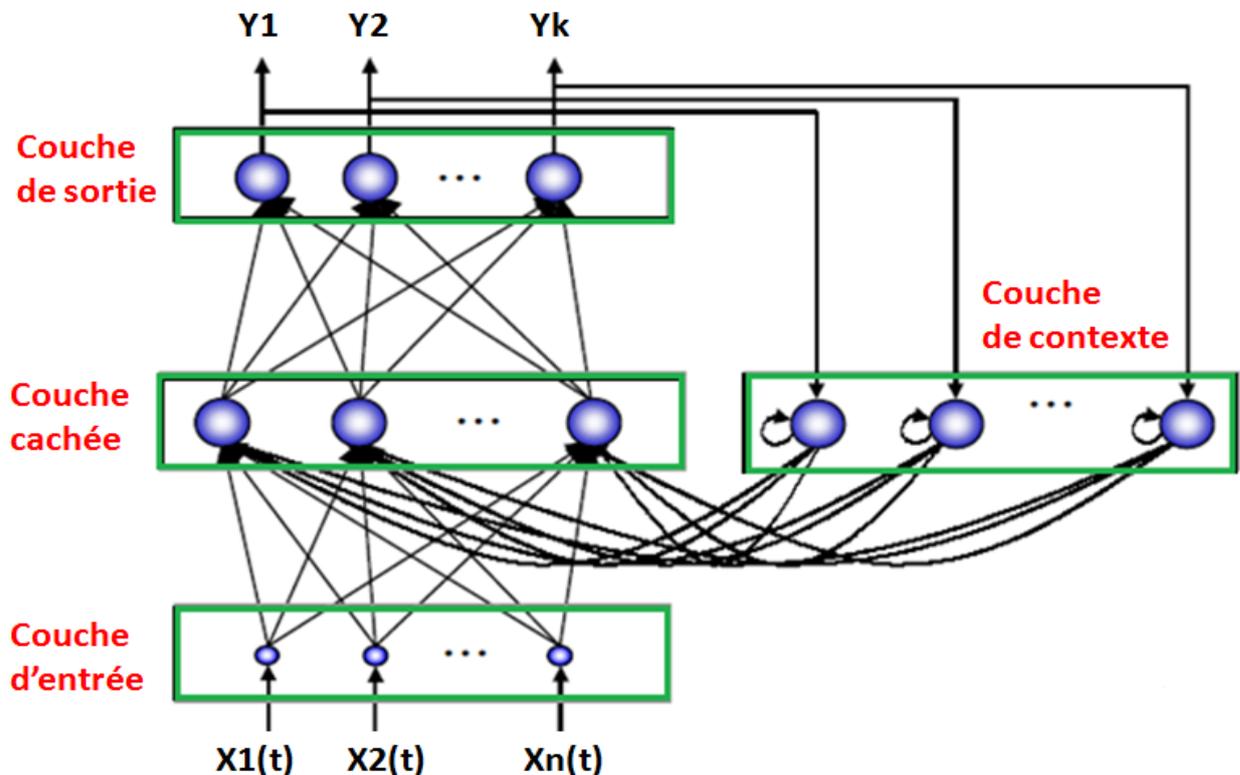


Fig. 2.6 Architecture de Jordan.

Elman (Elman, 1990) s'est inspiré en grande partie du réseau de Jordan pour proposer son architecture (Figure 2.7). Cette fois-ci, ce sont les unités de la couche cachée qui sont dupliquées dans la couche contexte avec un poids unitaire.

L'apprentissage s'effectue par l'algorithme de rétropropagation et ne concerne que les poids de propagation avant.

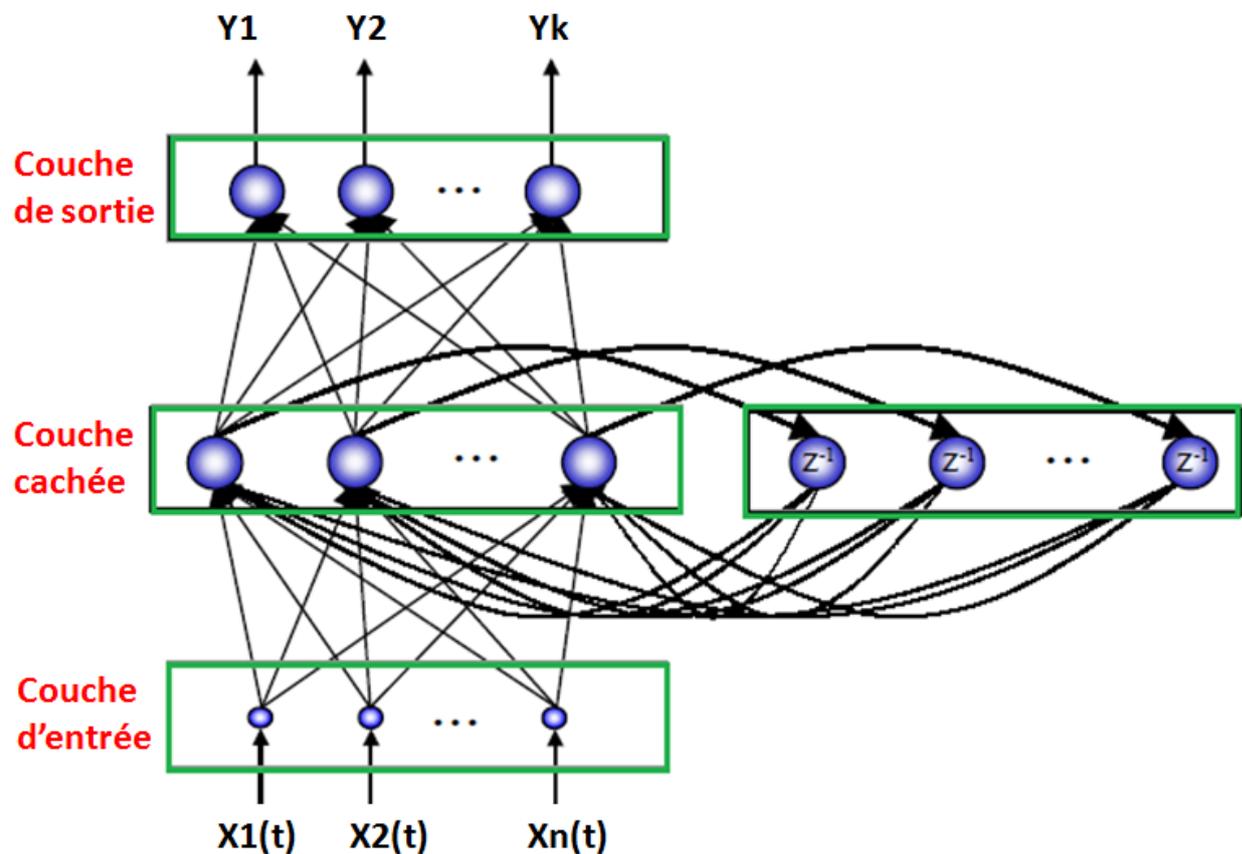


Fig. 2.7 Architecture de Elman.

Plusieurs variantes de l'architecture de Elman étaient proposées dans la littérature. On peut citer les travaux de Kalinli et al. (Kalinli et Sagirolu, 2006), Stagge et al. (Stagge et Sendhoff, 1997), Gao et al. (Gao et al., 1996) et de Kwok et al. (Kwok et al., 1994).

Applications en traitement d'images animées

Tai et al. (Tai et Huang, 2009) proposent un système de reconnaissances des expressions de visages dans des séquences d'images par classification en utilisant les réseaux récurrents d'Elman. La performance des résultats obtenus est évaluée par la base de séquences de Cohn-Kanade avec un taux de reconnaissance de 92.80%.

Sato et al. (Sato et al., 2003) utilisent un réseau récurrent partiellement connecté pour la détection des véhicules à partir des images issue du rétroviseur d'une voiture. Les résultats trouvés sont satisfaisantes.

Tan et al. (Tan et De Silva, 2003) proposent l'utilisation du réseau de neurones d'Elman pour reconnaître 10 activités humaines à partir d'une base de 200 séquences d'activités dans un environnement de bureau.

Psarrou et al. (Psarrou et al., 1995) abordent le problème de la prédiction de trajectoire en vision industrielle. Les auteurs utilisent un contexte dynamique pour construire les représentations apprises par le réseau de Elman afin d'explorer les caractéristiques de différentes représentations en entrée. La stabilité du réseau et la généralisation à partir de la base d'apprentissage sur des trajectoires complexes 2D sont testés. Une application est mise en œuvre pour la reconnaissance des trajectoires faciales.

c. Réseau de neurones récurrents bidirectionnel

Principe du fonctionnement

Les réseaux de neurones récurrents standards traitent les entrées dans l'ordre temporel, leurs résultats ont tendance à s'appuyer principalement sur le contexte et ils sont connus pour avoir des difficultés à apprendre des dépendances de temps entre le passé et le futur. Une solution à cette situation est fournie par les réseaux de neurones récurrents bidirectionnels (BRNN) (Schuster et Paliwal, 1997).

L'idée de base de réseaux de neurones récurrents bidirectionnels est de présenter chaque contexte d'entrée en avant et en arrière à plusieurs réseaux récurrents qui se suivent. Cela fournit un contexte complet, symétrique avec un passé et un futur pour chaque vecteur d'entrée (Figure 2.8).

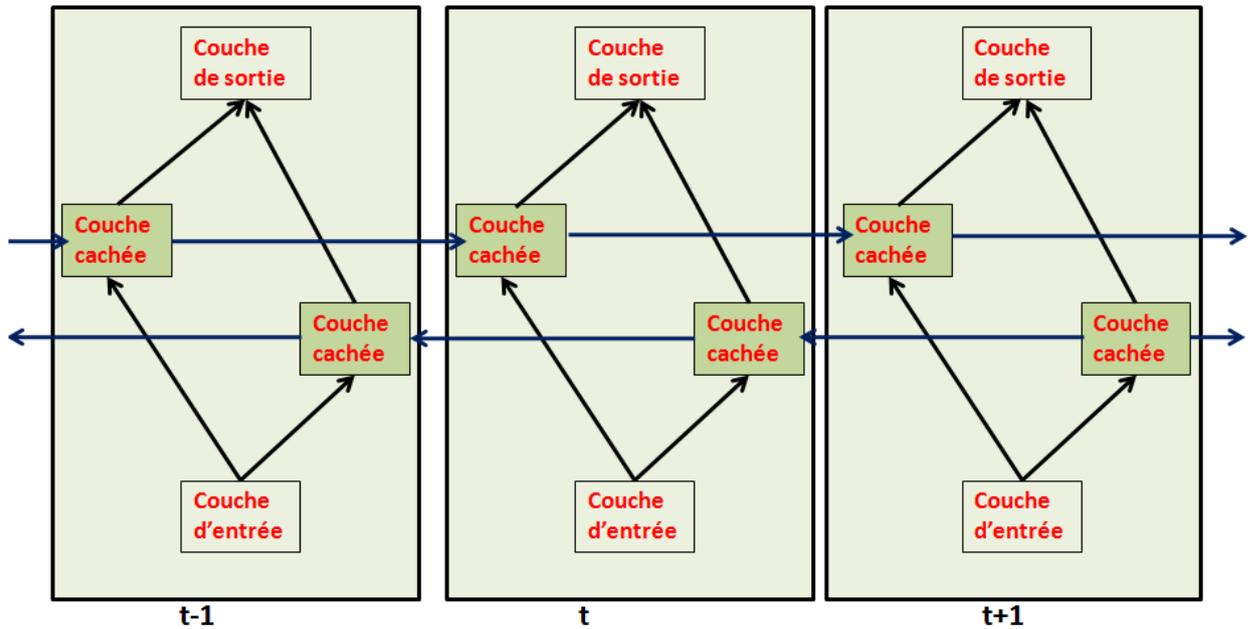


Fig. 2.8 Architecture d'un réseau de neurones récurrents bidirectionnel.

Application en traitement d'images

Eyben et al. (Eyben et al., 2010) ont appliqué les réseaux de neurones récurrents bidirectionnels en reconnaissance des émotions. Les résultats trouvés confirment que les réseaux bidirectionnels donnent de meilleurs résultats que l'architecture unidirectionnelle.

d. Réseau de neurones récurrents multidimensionnel

Principe du fonctionnement

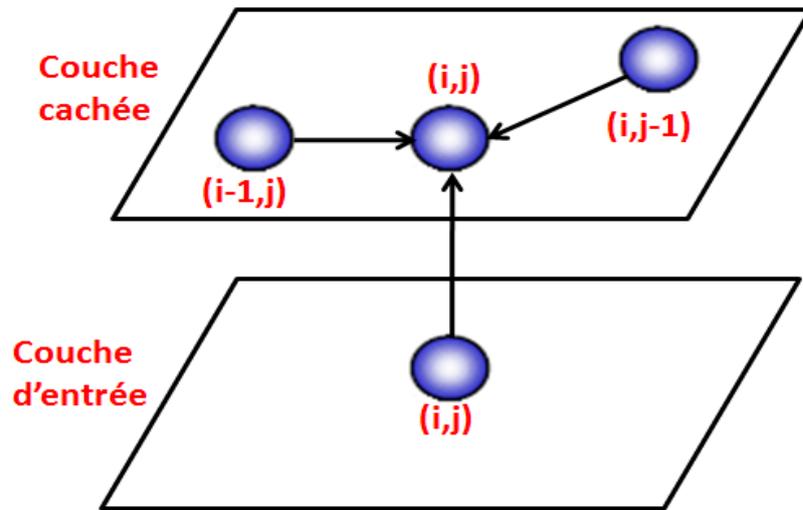
Les architectures récurrentes sont explicitement unidimensionnelles, ce qui signifie que pour les utiliser pour des tâches multidimensionnelles, les données doivent être prétraitées à une seule dimension, par exemple en présentant les données d'une image ligne par ligne successivement à un réseau.

Les réseaux de neurones récurrents multidimensionnels (MDRNN) (Graves et al., 2007) apportent les avantages de réseaux de neurones récurrents aux données multidimensionnelles, sans souffrir des problèmes de mise à l'échelle.

L'idée de base des réseaux MDRNN est de remplacer la connexion unique récurrente dans RNR standard avec autant de connexions récurrentes qu'il y a de dimensions dans les données.

Au cours de la propagation en avant, à chaque point de la séquence de données, la couche cachée du réseau reçoit à la fois une entrée externe et ses

propres activations d'un pas en arrière le long de toutes les dimensions. La [figure 2.9](#) illustre le cas à deux dimensions.



[Fig. 2.9](#) Propagation en avant de 2D RNN.

De toute évidence, les données doivent être traitées de telle sorte que lorsque le réseau atteint un point dans une séquence à n -dimensions, il a déjà passé par tous les points à partir de laquelle il recevra ses activations précédentes.

Chaque point dans la séquence d'entrée sera en général un vecteur à valeurs multiples. Par exemple, dans une image couleur en deux dimensions, les entrées peuvent être représentées par des pixels individuels RGB ou des blocs de pixels.

Le gradient d'erreur d'un MDRNN peut être calculé avec une extension à n -dimensions de l'algorithme de rétropropagation à travers le temps (BPTT). La [Figure 2.10](#) illustre la rétropropagation à deux dimensions.

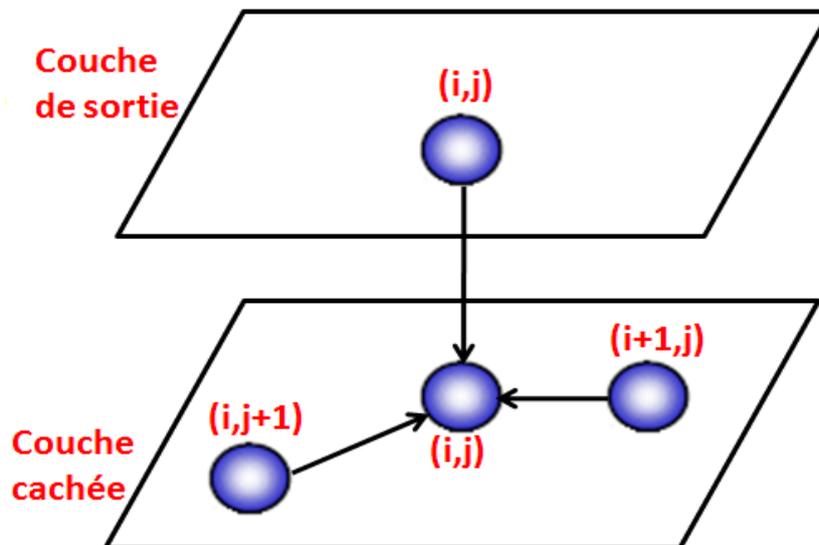


Fig. 2.10 Propagation en arrière de 2D RNN.

Application en traitement d'images animées

Graves et al (Graves et al., 2007) ont utilisé un MDRNN composé de 4 couches cachées et chaque couche composée de 25 blocs de mémoire pour une segmentation 3D de la base Air Freight. L'objectif était de reconnaître les différentes classes que contient cette base. L'apprentissage est effectué online par l'algorithme de descente de gradient. L'erreur de classification était de 7.1% sur la base de test.

e. Réseaux de longue mémoire à court terme (LSTM)

Principe du fonctionnement

Les réseaux récurrents standards peuvent difficilement apprendre les dépendances à long terme entre les données d'entrée. En effet, ils ne sont capables de mémoriser que le passé dit «proche», et commencent à «oublier» au bout d'une cinquantaine d'itérations environ.

Ce phénomène appelé dans la littérature « Vanishing Gradient Problem » et a été mis en évidence par Hochreiter dans (Hochreiter, 1998). L'auteur a étudié plusieurs algorithmes d'apprentissage (BPTT, RTRL...) pour les RNN et montre que l'erreur rétro-propagée liée à une entrée du réseau à l'instant t décroît de manière exponentielle après un certain nombre d'itérations.

Le diagramme dans la Figure 2.11 représente le déroulement d'un réseau de neurones récurrents dans le temps. Les unités sont ombrées en fonction de la façon dont ils sont sensibles à l'entrée en fonction de temps.

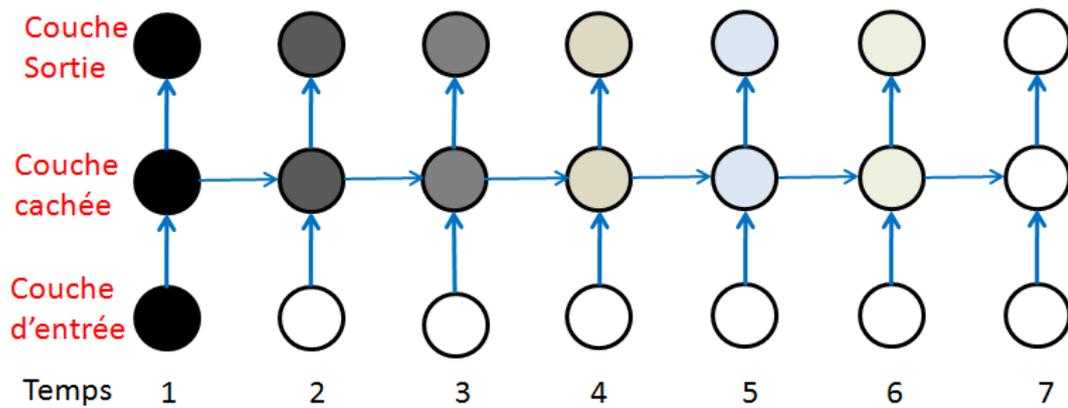


Fig. 2.11 Illustration de vanishing gradient problem.

Pour remédier le problème de vanishing, Hochreiter et Schmidhuber (Hochreiter et Schmidhuber, 1997) ont mis au point des neurones particuliers, à longue mémoire court-terme : les LSTM (Long Short-Term Memory).

Les réseaux de longue mémoire à court terme ont une architecture beaucoup plus complexe. Tout d'abord, plutôt que d'avoir plusieurs couches, ces réseaux ont une seule couche cachée formée de plusieurs blocs mémoires parallèles (Figure 2.12).

Les signaux d'entrées sont distribués vers chaque bloc mémoire en deux types d'entrées différentes : l'entrée principale, et l'entrée des portes. De la même façon, les sorties des blocs mémoires sont distribuées vers la couche de sortie, mais aussi par boucle rétroactive vers les entrées des blocs mémoires (entrées principales, et entrées des portes).

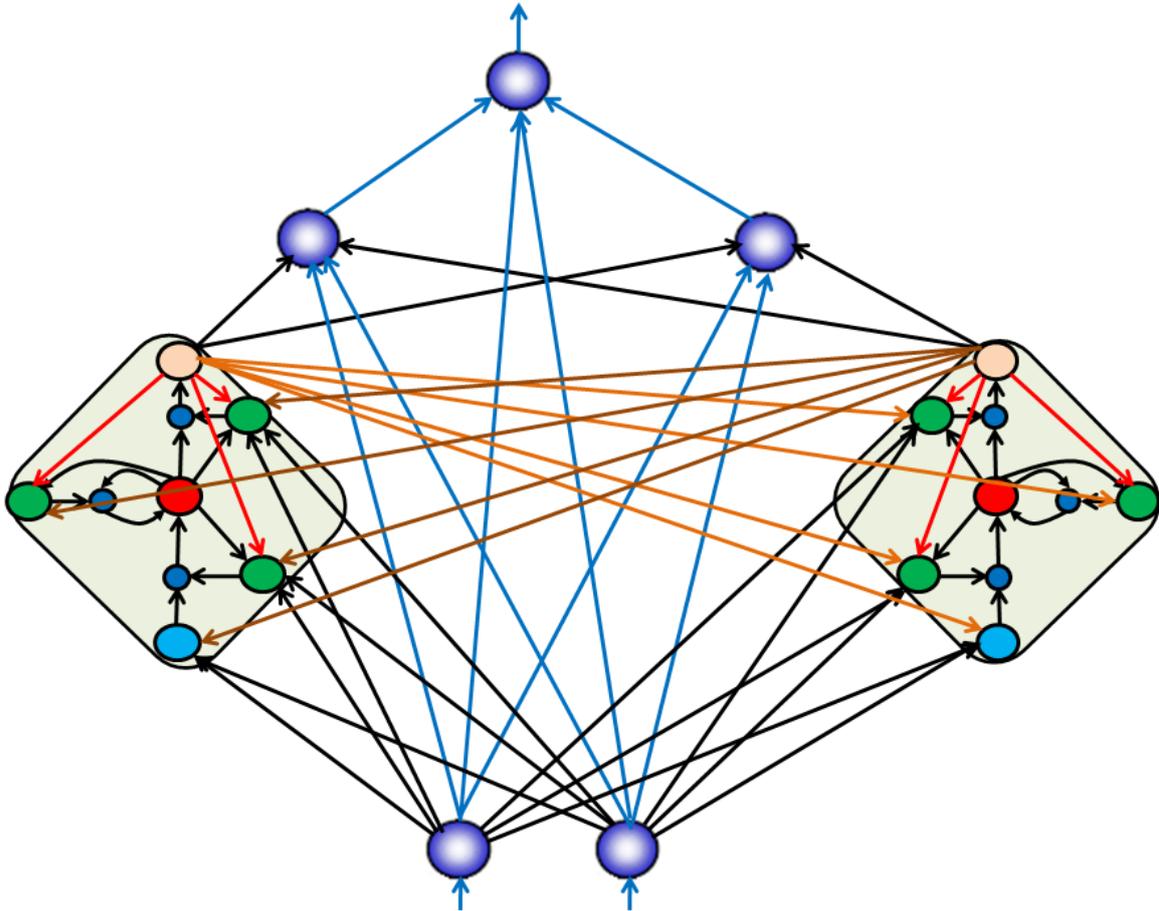


Fig. 2.12 Réseau de longue mémoire à court terme. Ce réseau a deux blocs mémoires d'une cellule mémoires chacun.

Le cœur du réseau réside dans l'architecture des blocs mémoires (Figure 2.13). Les portes calculent chacune une sommation pondérée des entrées qui est ensuite passée par une fonction d'activation (dont la sortie est dans $[0,1]$) comme un neurone artificiel ordinaire.

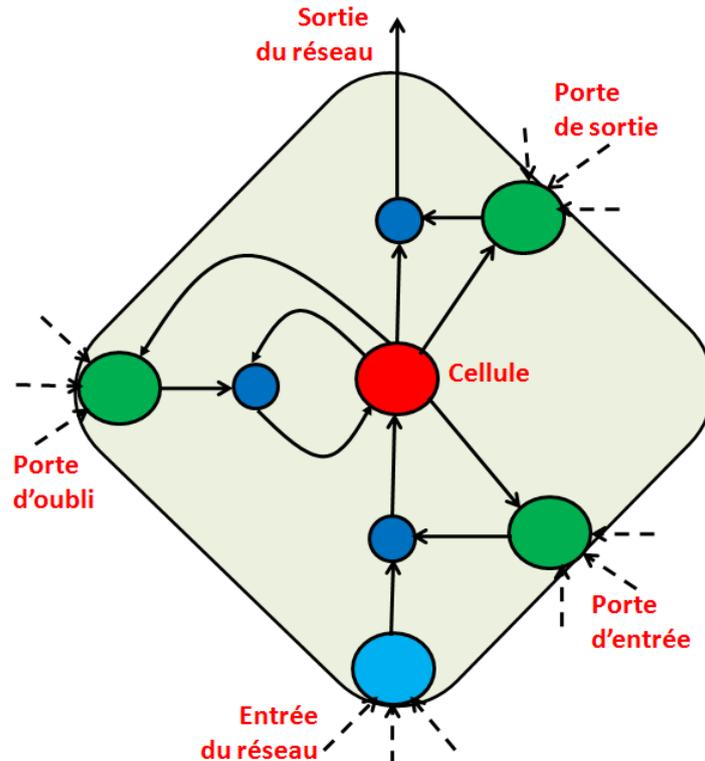


Fig. 2.13 Exemple d'un bloc de mémoire. Les cellules bleues représentent les unités multiplicatives.

Cependant, les sorties de ces neurones servent de gains à différents endroits dans le bloc agissant un peu comme des interrupteurs. Pour chaque cellule mémoire du bloc, une sommation pondérée des entrées est aussi calculée et passée par une fonction d'activation. Ce signal est ensuite multiplié par la porte d'entrée. Si la porte d'entrée est ouverte, alors le signal entre dans la cellule mémoire. La valeur de la cellule mémoire dépend de deux termes : du signal d'entrée multiplié par la porte d'entrée et de la valeur de la cellule mémoire au temps précédent multiplié par la porte d'oubli.

Finalement, le signal de la cellule mémoire peut encore une fois être passé par une fonction d'activation (optionnelle) avant d'être multiplié par la porte de sortie. Si cette dernière est fermée, le bloc produit une sortie de zéro, si elle est ouverte, alors la valeur des cellules mémoires peut-être lu à la sortie du bloc (soit par les autres blocs au temps suivant, soit directement par la couche de sortie. Chacune des sommations pondérées du réseau est faite de poids différents qui sont appris par rétropropagation.

Les concepteurs de cette architecture (Hochreiter et Schmidhuber, 1997) ont cependant proposé ultérieurement une méthode d'apprentissage qu'ils considèrent très performante, conseillant de n'employer le modèle LSTM que lorsque le nombre de poids est grand ou lorsque les poids nécessitent d'être

déterminés avec une grande précision. Cette méthode, qu'ils ont baptisé *guessing* (littéralement "divination") (Hochreiter et Schmidhuber, 1996) et qu'ils limitent à des tâches d'apprentissage de séquences, consiste à effectuer successivement plusieurs apprentissages avec un réseau dont les poids sont à chaque fois initialisés au hasard de manière différente et à conserver l'ensemble des poids qui aura eu les meilleurs résultats à la fin du processus d'apprentissage.

Applications en traitement d'images animées

Baccouche et al. (Baccouche et al., 2010) proposent une approche de classification automatique de séquences vidéo d'actions de sport. Pour cela, ils extraient de chaque action des caractéristiques du contenu visuel, en utilisant deux approches, l'une par sac de mots, et l'autre par le mouvement dominant de la scène à chaque instant. La classification de l'évolution temporelle de ces caractéristiques extraites est gérée dynamiquement par un modèle neuronal, basé sur les réseaux de neurones récurrents à longue mémoire court-terme (LSTM). Les expérimentations faites sur la base « MICCSoccer-Actions-4 » montrent que l'approche neuronale de classification permet d'obtenir des résultats supérieurs à l'état de l'art (76% de bonne classification), et que la combinaison des caractéristiques (information visuelle et mouvement dominant) permet un taux de bonne classification de 92%.

Delakis a utilisé avec succès dans des applications de structuration de vidéos de tennis en utilisant les réseaux de neurones récurrents à longue mémoire court-terme pour modéliser l'évolution temporelle des transitions entre les plans de la vidéo sans en analyser le contenu (Delakis, 2006).

f. Recursive SOM

Principe du fonctionnement

Ce modèle proposé par Thomas Voegtlin (Voegtlin, 2002) représente une application de l'algorithme non supervisé de Kohonen sur un simple réseau récurrent de Elman.

Le SOM récursif est donc un SOM avec une auto-référence, ce qui entraîne l'ajout de connexions récurrentes à l'architecture originale du SOM d'une manière qu'il soit compatible avec le principe d'auto-organisation (Figure 2.14).

Les connexions associées avec le contexte et la couche d'entrée sont homogènes définies par la même équation de l'activité neuronale. La nouvelle entrée et la

copie des dernières activités de chaque neurone de la carte sont considérées comme un seul vecteur d'entrée pour l'algorithme SOM.

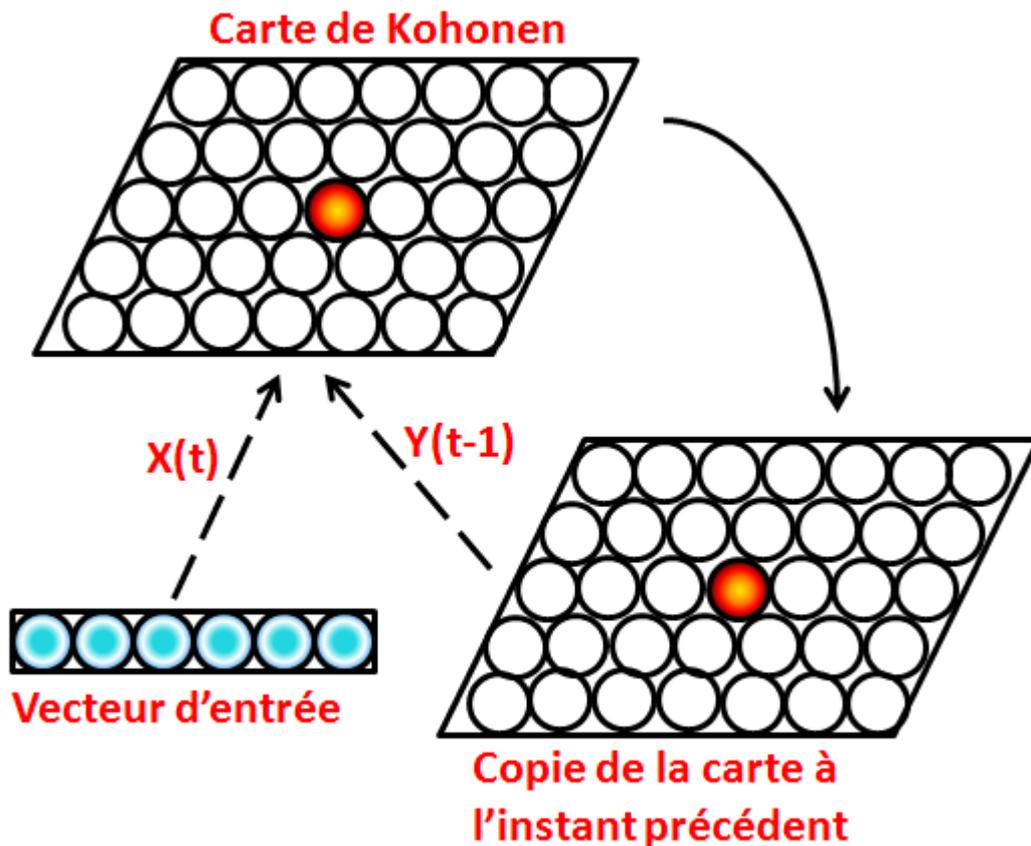


Fig. 2.14 Schéma illustratif du fonctionnement du réseau SOM récursif.

Application en traitement d'images animées

Huang et Wu présentent une approche pour reconnaître les trajectoires du mouvement humain en utilisant SOM récursif. Les informations contextuelles issues de chaque pose d'un mouvement humain de la séquence sont extraites pour être injecté au réseau pour apprentissage puis pour reconnaissance. L'approche est testée sur plusieurs bases de séquences, le résultat était impressionnant (Huang et Wu, 2009).

g. Réservoir de neurones

Jaeger et Maass ont proposé, indépendamment, un réservoir de neurones. Ce type de réseaux a peu à peu émergé dans la littérature sous différentes formes telles que les « ESN : Echo State Networks » (Jaeger, 2002) ou bien, « LSM : Liquid State Machine » (Maass et al., 2002) et l'on parle maintenant de « Reservoir Computing » pour désigner ces approches. Ces réseaux, et plus généralement les

réseaux de type réservoir sont des réseaux récurrents, dont les topologies sont basées sur des probabilités de connexions entre les neurones du réseau de manière aléatoire et clairsemée (ex. seulement 10% des connexions possibles ont des poids non nuls).

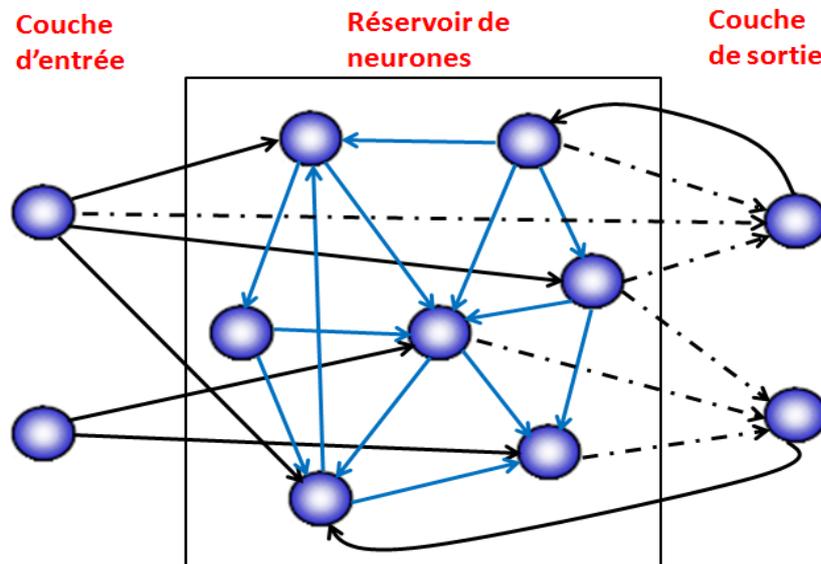


Fig. 2.15 Schéma d'un réservoir de neurones. Les lignes en gras indiquent des connexions aléatoires fixées et les lignes pointillées indiquent les connexions qui vont être modifiées par apprentissage.

Pour une revue détaillée, nous referons le lecteur aux travaux de Gibbons (Gibbons, 2010), Lukosevicius et Jaeger (Lukosevicius et Jaeger, 2009) et Schrauwen et al. (Schrauwen et al., 2007).

Il existe de nombreux domaines où l'application de cette thématique est limitée ou inexistante. On peut citer néanmoins le domaine du traitement d'images animées relatif aux mouvements et la reconnaissance d'objets (Gibbons, 2010).

2.2.2.2 Représentation explicite du temps dans les réseaux de neurones

Représentation explicite du temps au niveau des connexions

a. Réseaux de neurones à délais temporels adaptatifs (ATNN)

Principe du fonctionnement

Le choix manuel qui dirige l'extraction des caractéristiques présentées à l'entrée du réseau TDNN suscite des difficultés d'initialisation du paramètre délai propre aux connexions de ce réseau de type convolutif. Ces délais, qui représentent les décalages temporels responsables de l'extraction du signal sur un champ de vision limité appelé fenêtre de spécialisation, sont initialement fixés sur les mêmes positionnements pendant et après l'apprentissage.

Les ATNNs proposés par (Lin, 1994) apparaissent comme une solution via cette initialisation hasardeuse des délais, qui toute au long de l'apprentissage, admettent des ajustements par le biais d'un algorithme proche de celui de la rétropropagation du gradient. Ainsi, nous obtenons une nouvelle combinaison de poids et de délais optimaux pour la tâche requise du réseau.

L'architecture d'un ATNN ressemble de très près à l'architecture d'un TDNN où entre deux unités de deux couches successives s'engendre un bloc de retards (Figure 2.16). Chaque bloc entre une unité i et unité j , appartenant respectivement au couche L_h et L_{h-1} , contient un ensemble de liaison superposées $K_{ji,h-1}$ ayant deux propriétés : les poids synaptiques $W_{ji,h-1}$ et les retards $\tau_{jik,h-1}$ ou appelés aussi des décalages temporels. La Figure 2.17 montre un exemple d'un réseau de neurones avec des délais internes au niveau des connexions.

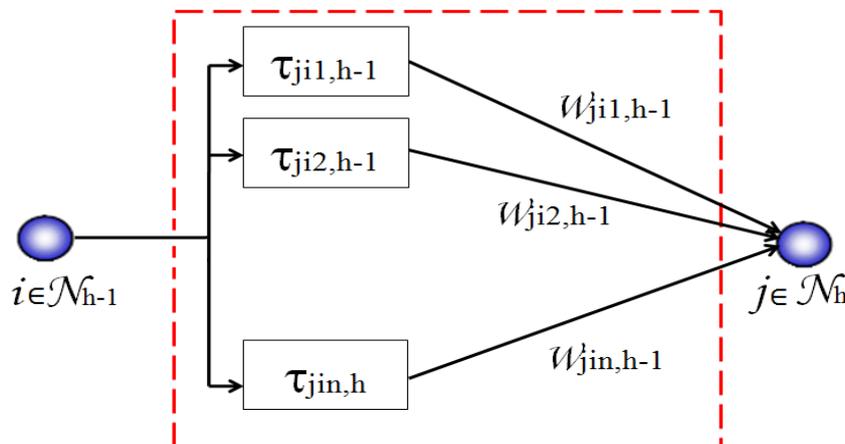


Fig. 2.16 Représentation d'un bloc de retards entre deux neurones appartenant à deux couches successives avec n est le nombre de délais.

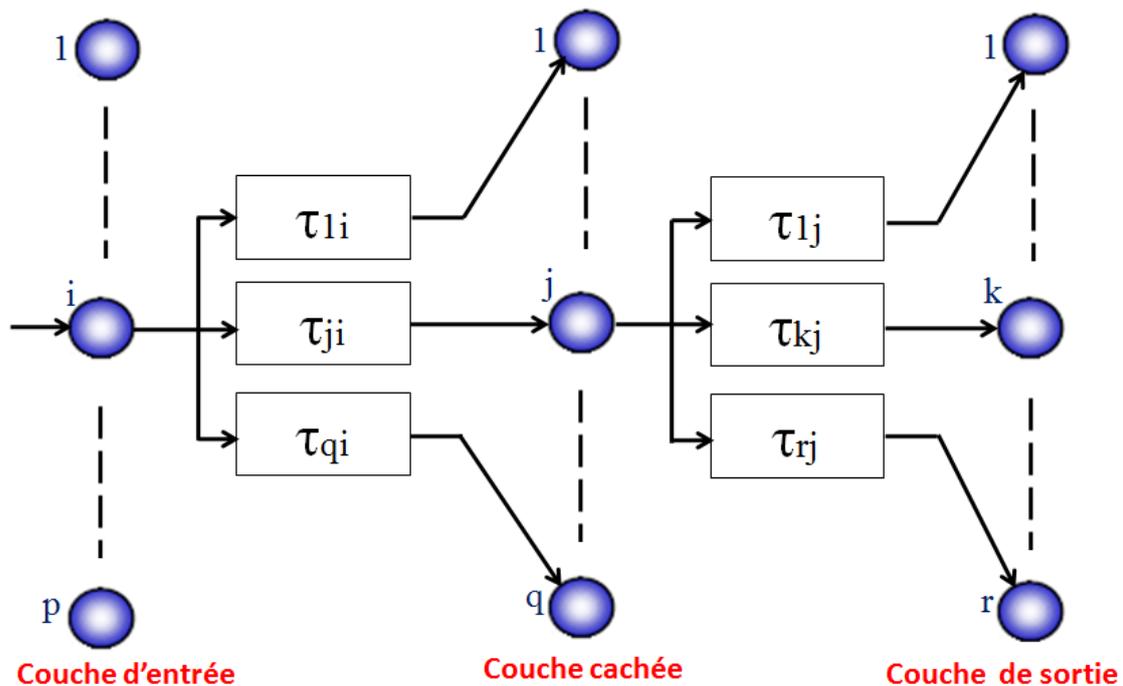


Fig.2.17 Exemple d'un ATNN à trois couches.

L'ATNN est théoriquement fondé sur une fonction d'approximation universelle et un algorithme d'apprentissage construit autour du principe de la descente du gradient, tout en préservant l'aspect temporel externe que renferme ce type de réseaux dans sa conception de base.

Cet algorithme d'apprentissage effectue un ajustement parallèle des poids synaptique et des retards des connexions. A chaque adaptation, le réseau arrive à une nouvelle combinaison de poids et de délais jugée optimale selon le critère de la minimisation d'une fonction d'erreur.

Applications en traitement d'images animées

Wohler et Anlauf (Wohler et Anlauf, 1999b) présentent un algorithme basé sur les ATNN pour l'analyse des séquences d'images. L'adaptation manuelle fastidieuse de la taille temporelle des champs récepteurs peut être évitée en utilisant une méthode pour adapter le retard correspondant et les paramètres du réseau pendant la phase d'apprentissage.

Dans leur article, Wohler et al. (Wohler et al., 2000), abordent le problème de l'analyse de séquences d'images par classification, qui est de plus en plus important dans le contexte de systèmes de vision en temps réel pour la reconnaissance d'objets et l'analyse du mouvement. Ils utilisent les ATNN pour la reconnaissance des piétons dans des scènes de circulation complexes.

Représentation explicite du temps au niveau des neurones

Le temps peut également être traité localement au niveau du neurone, ce qui permet d'avoir une certaine robustesse temporelle (Chappelier et al., 2001). Ce modèle de neurone temporel peut être réalisé de deux manières différentes : soit en gardant les propriétés biologiques d'un neurone soit en construisant un modèle pour résoudre un problème «d'ingénierie» sans autre type de préoccupations (ignorant complètement l'aspect biologique). Le premier cas conduit à des modèles plus ou moins compliqués. Il existe de nombreux modèles d'inspiration biologique. Ces modèles possèdent ainsi des propriétés temporelles inspirées des neurones biologiques tout en permettant un calcul rapide. Le second point de vue consiste à temporaliser les modèles existants performants du point de vue statique. On peut citer par exemple l'approche de Vaucher (Vaucher, 1993) qui a temporalisé un modèle statique d'une façon purement algébrique. Le corps de représentation d'un MLP n'est plus dans le domaine des réels (\mathcal{R}) mais dans celui des complexes (\mathbb{C}).

Modèles neuronales bio-inspirés

Bien que les réseaux de neurones temporels présentés précédemment utilisant le taux de décharge des neurones comme élément de calcul se sont révélés être très puissants pour certains types de problèmes, certaines propriétés des réseaux de neurones biologiques sont ignorées.

Grâce à des expériences biologiques, il est devenu évident que certains effets et calculs du cerveau humain ne peuvent pas être effectués en utilisant simplement le taux de décharge des neurones respectifs en tant que transmetteur d'informations. En outre, le moment exact de décharge d'impulsion doit achever au moins une partie des informations transmises.

Plusieurs modèles impulsifs bio-inspirés sont développés dans ce contexte de décharge impulsif. Les modèles de Hodgkin & Huxley (HH), intégrateur à fuite (IF) et le modèle à réponse impulsif (SRM) feront l'objet d'une étude détaillée au chapitre suivant. Le réseau de neurones à Impulsion-couplés (PCNN) est présenté dans le paragraphe qui suit.

a. Réseaux de neurones à Impulsion-couplés (PCNN)

Inspiré d'un modèle de neurone biologique (Eckhorn et al., 1990), Johnson a proposé l'architecture du réseau de neurones à Impulsion-couplés (PCNN) (Johnson et al., 1999).

Un PCNN est un réseau bidimensionnel dont la structure est montrée sur la **Figure 2.18**. Le réseau est constitué d'une couche d'entrée, d'une couche de liaison et d'un générateur d'impulsions. Ces modèles ne possèdent pas d'algorithme d'apprentissage adapté. Ainsi, les poids sont souvent fixés a priori par l'utilisateur pour chaque tâche particulière.

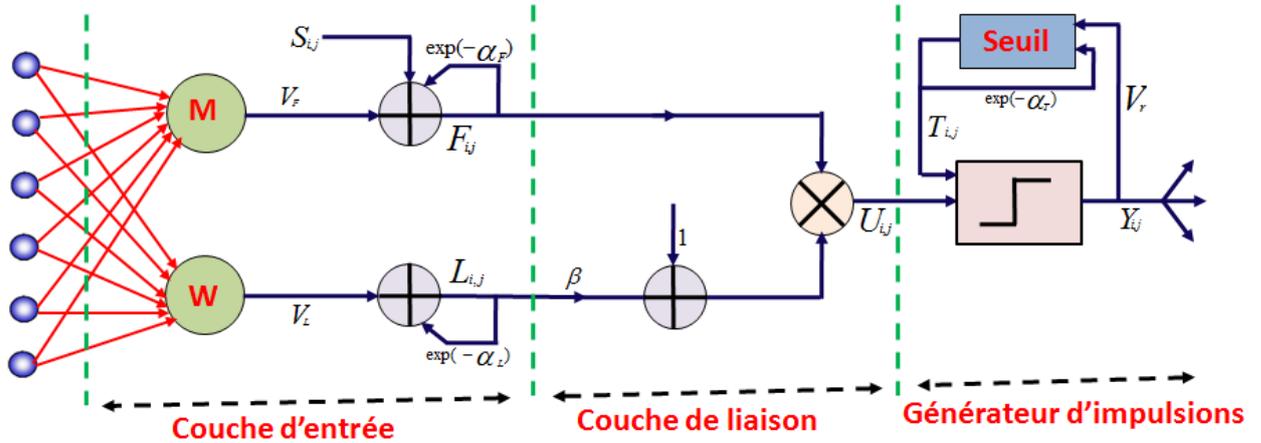


Fig.2.18 Architecture du réseau PCNN.

Le modèle PCNN est régi par une suite d'équations :

$$F_{i,j}[n] = e^{\alpha_F} F_{i,j}[n-1] + V_F \sum_{k,l} W_{i,j,k} Y_{i,j}[n-1] + S_{i,j} \quad (2.1)$$

$$L_{i,j}[n] = e^{\alpha_L} L_{i,j}[n-1] + V_L \sum_{k,l} W_{i,j,k} Y_{i,j}[n-1] \quad (2.2)$$

$$U_{i,j}[n] = F_{i,j}[n](1 + \beta L_{i,j}[n]) \quad (2.3)$$

$$Y_{i,j}[n] = \begin{cases} 1 & U_{i,j}[n] > T_{i,j}[n] \\ 0 & \text{otherwise} \end{cases} \quad (2.4)$$

$$T_{i,j}[n] = e^{\alpha_T} T_{i,j}[n-1] + V_T Y_{i,j}[n] \quad (2.5)$$

Où $S_{i,j}$ représente le stimuli d'entrée, tel le niveau de gris d'un pixel (i, j) dans une image,

n est le numéro d'itération,

$F_{i,j}[n]$ est la rétroaction du neurone dans la position (i, j) ,

$L_{i,j}[n]$ est le point de liaison,

$U_{i,j}[n]$ est l'activité interne du neurone,

$T_{i,j}[n]$ le seuil dynamique,

$Y_{i,j}[n]$ est l'impulsion neuronale générée en sortie dont la valeur est 0 ou 1,

M et W représentent matrices synaptiques constantes de poids. Généralement elles correspondent à une distribution gaussienne locale,

β est le coefficient de liaison,

α_F , α_L et α_T représentent des constantes d'atténuation temporelles de $F_{i,j}[n]$, $L_{i,j}[n]$ et $T_{i,j}[n]$ respectivement,

V_F , V_L et V_T dénotent le potentiel inhérent de tension de $F_{i,j}[n]$, $L_{i,j}[n]$ et $T_{i,j}[n]$ respectivement.

Le PCNN a été utilisé pour une variété d'applications de traitement d'images, incluant : segmentation d'images, extraction de visage, détection de mouvement, croissance de région, réduction de bruit, etc. (Lindblad et Kinser, 2005).

Une revue complète sur les PCNN et ces applications depuis 1999 à 2009 est exposée dans l'article de Wang et al. (Wang et al., 2010).

b. Applications des modèles neuronales bio-inspirés en traitement d'images animées

Dans leur article, Bandala et al. (Bandala et al., 2010) proposent une architecture bio-inspirés pour détecter, décrire et distinguer les objets en mouvement. En utilisant des mécanismes neuronaux et physiologiques dans le cortex visuel primaire (V1), l'aire temporale médiane (MT) et l'aire inféro-temporel (IT). Ils commencent à isoler les objets de leur environnement, puis suivre, étiqueter et distinguer les humains à partir des images non-humains en mouvement. Enfin, ils représentent la silhouette de la personne pour obtenir une meilleure compréhension de la structure du corps.

Escobar et al. (Escobar et al., 2009) proposent un modèle impulsif générique V1-MT (aires cérébrales) qui peut être utilisé pour des tâches de reconnaissance de mouvements humains. Le modèle prend en entrée une séquence d'images.

L'activité des neurones dans la couche V1 sont transformées en trains d'impulsions en utilisant un modèle LIF qui seront injectés par la suite à la couche suivante du réseau impulsif (couche MT).

A partir des activités des neurones dans la couche MT, les auteurs ont défini deux types de cartes de mouvement : carte moyenne des mouvements et la carte de synchronisation des mouvements. Enfin, une étape de classification supervisée sur la base de ces cartes est exécutée.

Inspiré par le comportement du système visuel humain, un réseau de neurones impulsifs est proposé pour détecter les objets en mouvement dans

une séquence d'images visuelles (Xiang Wu et al., 2008). Les résultats de simulation montrent que le réseau est capable d'effectuer la détection de mouvement à partir de séquences d'images visuelles dynamiques

Le système prend en entrée une image de la séquence d'images à chaque instant de temps qui sera transférée en niveaux de gris. Les contours des objets en mouvement sont extraits d'un groupe de neurones actifs. À partir de ces contours, un filtre est généré pour extraire le déplacement des objets de l'arrière-plan.

Risinger et Kaikhah (Risinger et Kaikhah, 2008) proposent un système inspiré biologiquement capable de détecter et suivre les mouvements en utilisant un modèle de neurones impulsionsnels intègre et tire à fuite (cf. Section 3.4.2). Les résultats expérimentaux confirment leur approche.

SpikeNet Technology est une spin-off ³ qui travaille en reconnaissance et suivi d'objets en temps réel. Le SpikeNet a permis de montrer qu'une architecture simple de traitement ascendant de l'information visuelle est capable d'exécuter des tâches difficiles telles que la détection de visages dans des images naturelles (VanRullen et al., 1998) ou l'identification de visages de manière indépendante au point de vue (Delorme et Thorpe, 2001). SpikeNet utilise un modèle neuronal de type intègre et tire (cf. Section 3.4.2) dont les règles hebbiennes de l'apprentissage synaptiques sont complétées par l'intégration d'une dynamique synaptique au niveau du neurone.

2.3. Conclusion et Analyse des Architectures Neuronales Temporelles

Nous avons donné dans ce chapitre un état de l'art aussi large que possible des différentes façons de prendre en compte l'aspect temporel par les réseaux de neurones, les différentes architectures de réseaux de neurones temporels. Cette étude nous a permis de conclure que les façons d'aborder le temps par les réseaux de neurones sont nombreuses de même que les travaux et publications concernant l'application de ces réseaux de neurones temporels en traitement d'images animées.

Néanmoins, à travers la liste de références non exhaustives mais représentatives que nous avons consulté, nous remarquons deux représentations temporelles des réseaux de neurones : une représentation spatiale ou externe et

³ Filiale dédiée à une activité nouvelle pour permettre son décollage plus rapide, hors des contraintes d'organisation de la maison mère ; démarche fréquemment adoptée par les entreprises qui souhaitent muter vers la Net économie, et notamment l'e-business.

une représentation dynamique ou interne. Les topologies des réseaux de neurones pour chaque représentation temporelle sont complètement différentes et chacune possède ses avantages et ses inconvénients.

2.3.1. Limites de la représentation spatiale du temps

Cette représentation présente plusieurs désavantages. Tout d'abord elle suppose l'existence d'une interface avec le monde extérieur dont le rôle est de retarder ou de retenir les données jusqu'au moment de leur utilisation dans le réseau: comment connaître l'instant où les données doivent être traitées c'est-à-dire comment définir la taille de la fenêtre ?

Ensuite, le fait d'utiliser une fenêtre temporelle (ou des retards) de longueur finie et déterminée a priori soit par la plus longue information à traiter, soit en supposant la même longueur pour toutes les données, présente un second désavantage.

Enfin, c'est dans la nature même de la représentation spatiale que se pose la difficulté de différencier une position temporelle relative d'une position temporelle absolue. On peut dire qu'il n'y a aucune exploitation du caractère temporel dans le fonctionnement interne des réseaux.

Le choix manuel qui dirige l'extraction des caractéristiques présentées à l'entrée du réseau TDNN suscite des difficultés d'initialisation du paramètre délai propre aux connexions de ce réseau de type convolutif. Ces délais, qui représentent les décalages temporels responsables de l'extraction du signal sur un champ de vision limité appelé fenêtre de spécialisation, sont initialement fixés sur les mêmes positionnements pendant et après l'apprentissage.

Des modifications sont portées sur l'architecture TDNN tels : Adaptive time delay radial basis function, Adaptive time-delay neural networks, Memory neural networks, Multi state TDNN et Cascaded TDNN.

2.3.2. Limites de la représentation dynamique du temps

Les réseaux récurrents peuvent présenter deux types de comportements : le premier est que ces réseaux sont capables de se stabiliser dans un certain nombre de points de leur espace d'état, appelés points d'équilibre. Ces points constituent la réponse du réseau en présence d'une donnée et l'apprentissage consiste à affecter aux poids des connexions les valeurs permettant cette relaxation vers un point d'équilibre. L'architecture la plus connue et la plus ancienne ayant ce type de comportement est le modèle de Hopfield ([Hopfield, 1982](#)).

L'inconvénient des réseaux récurrents provient du fait qu'ils favorisent la corrélation d'événements en provenance du passé récent par rapport au passé plus lointain (seul l'expérimentation permet d'estimer le temps de prise en compte).

En pratique les réseaux récurrents demeurent peu répandus. En effet, ils sont le plus souvent associés à des algorithmes d'apprentissage basés sur le calcul du gradient, comme la rétropropagation à travers le temps (Back Propagation Through Time (BPTT)), plus consommateurs de temps de calcul que leurs équivalents pour les réseaux à propagation en avant, à nombre de paramètres égal. Par ailleurs, ces algorithmes éprouvent des difficultés dans la prise en compte des dépendances à long terme.

Une revue critique sur les réseaux de neurones récurrents est présentée dans le rapport de Szilas (Szilas, 1997).

2.3.2.1. Modèle de Hopfield

Bien que ce réseau soit généralement utilisé pour apprendre un certain nombre d'associations statiques pour des problèmes d'optimisation multicritères, le réseau de Hopfield possède une certaine dynamique de convergence, c'est-à-dire que le réseau passe par un certain nombre de points de son espace d'états avant d'atteindre un point d'équilibre. Ce type de réseau peut donc générer une séquence finie à partir d'un point donné de son espace d'état.

En tant que modèle du fonctionnement biologique, le modèle de Hopfield a l'avantage de mettre en lumière, après d'autres modèles plus anciens mais moins connus, le rôle de la dynamique dans les fonctions cognitives des réseaux de neurones et le lien établi par la règle de Hebb entre apprentissage et corrélation. En tant que prototype de mémoires associatives, et malgré le développement, dans les années 1980, de nouvelles variantes (réseaux de Hopfield de champ moyen à fonctions d'activation continues, réseaux de Hopfield stochastiques et machines de Boltzmann), et la publication d'une littérature considérable, les performances faibles des réseaux de Hopfield ont entraîné, à juste titre, l'abandon des recherches à leur sujet, notamment celles qui concernent leur applications potentielles en reconnaissance des formes (Dreyfus et al., 2010).

2.3.2.2. Réseaux à longues mémoires court-terme (LSTM)

Bien que cette architecture semble extrêmement artificielle, elle reste plus plausible biologiquement que plusieurs des autres réseaux récurrents présentés dans ce chapitre. Tout d'abord, le gradient de chaque poids ne requiert que des signaux disponibles au temps présent ou précédent (t ou $t-1$), aucun gradient ni aucune activation de neurones à un temps $t-k$, $k>1$ n'est requis. Ensuite, l'ajustement des poids à l'intérieur d'un bloc ne requiert aucun autre signal en

provenance des autres blocs que celui qui y arrive déjà (pas de gradient, pas d'activité de neurones locaux à d'autres blocs). Seuls quelques signaux non locaux de la couche de sortie sont nécessaires à la mise à jour des poids.

En contre partie, le réseau à longues mémoires court-terme se limite à sa capacité de mémoire en nombre de blocs et qui reste un problème qui peut être surmontée en augmentant la taille du réseau. En matière de modularité, comment les modules peuvent être apprises ou exécutés de manière efficace et comment ils devraient être interconnectés n'est généralement pas clair.

2.3.2.3. Réseau de neurones récurrents bidirectionnel

Un inconvénient majeur de cette architecture est que la séquence d'entrée dans son ensemble doit être disponible à l'avance, ce qui rend cette architecture limitée en traitement temps réel (Graves et al., 2008).

2.3.2.4. Réseaux de neurones à délais temporels adaptatifs (ATNN)

L'algorithme d'apprentissage des ATNN effectue un ajustement parallèle des poids synaptique et des retards des connexions. À chaque adaptation, le réseau arrive à une nouvelle combinaison de poids et de délais jugée optimale selon le critère de la minimisation d'une fonction d'erreur. Cependant cet algorithme, comme toute méthode d'optimisation déterministe, est plus limité dans les problèmes de grande taille, où la probabilité de converger vers un minimum globale, en un temps raisonnable dépend essentiellement de la bonne connaissance du problème. Dans ces cas, en manque de bonnes connaissances, il est conseillé de se tourner vers des méthodes d'optimisations stochastiques (Ourdighi et Benyettou, 2010).

2.3.3. Représentation du temps à base de SOMs

Plusieurs méthodes ont été proposées pour représenter le temps avec les SOMs. Ceci inclut des représentations explicites (Wiemer, 2003 ; Vesanto, 1997 ; Chappelier et Grumbach, 1996b), des méthodes basées sur des connexions latérales (Kopecz, 1995 ; Siroh, 1997), sur des connexions récurrentes (McQueen et al., 2003 ; Voegtlin, 2002 ; Varsta et al., 1997), sur des intégrateurs à fuite (Privitera et Morasso, 1994 ; Chappell et Taylor, 1993) ou sur différentes combinaisons de ces principes (Strickert et al., 2005 ; Zehraoui et Bennani, 2004 ; Koskela et al., 1998 ; James et Miikkulainen, 1995).

Des revues sur la représentation du temps à base de SOM sont présentés par Salhi ([Salhi et al., 2009](#)), Hammer ([Hammer et al., 2004](#)), Guimarães ([Guimarães et al., 2003](#)) et Barreto ([Barreto et al., 2003](#) ; [Barreto et Araújo, 2001](#)).

2.3.4. Liste d'applications

Nous avons récapitulé sur le tableau ci-dessous une liste non exhaustive d'application des réseaux de neurones temporels en matière de traitement d'images animées, déjà présentés dans ce chapitre. Cette liste est ordonnée par année de manière descendante selon les références bibliographiques.

RN Année	TDNN	TDRBF	Hopfield	Jordan & Elman	RNR Multidim..	RNR bidirection .	LSTM	ATNN	Bio-Inspiré
2010						(Eyben et al., 2010)	(Baccouche et al., 2010)		(Bandala et al., 2010)
2009				(Tai et Huang, 2009)					(Escobar et al., 2009)
2008			(Sulehria et Zhang, 2008 ; Pervej et al., 2008)						(Xiang Wu et al., 2008)
2007			(Sulehria et al., 2007)		(Graves et al., 2007)				(Risinger et Kaikhah, 2008)
2006	(Le Callet et al., 2006)		(Pajares, 2006)				(Delakis, 2006)		
2005	(Zou et Bhanu, 2005)								
2003				(Sato et al., 2003 ; Tan et De Silva, 2003)					
2002	(Ming et al., 2002)	(Howell et Buxton, 2002)							

RN Année	TDNN	TDRBF	Hopfield	Jordan & Elman	RNR Multidim.	RNR bidirection.	LSTM	ATNN	Bio-Inspiré
2001		(Howell et Buxton, 2002)							(Delorme et Thorpe, 2001)
2000	(Cutler et Davis, 2000)		(Sun, 2000)					(Wohler et al., 2000)	
1999	(Wohler et Anlauf, 1999a)	(Howell et Buxton, 1999)						(Wohler et Anlauf, 1999b)	
1998		(Howell et Buxton, 1998)							
1997			(Young et al., 1997)						
1995	(Ambellouis et Cabestaing, 1995)			(Psarrou et al., 1995)					

Réseaux de neurones impulsionnels

Sommaire

3.1 Introduction	70
3.2 Les bases biologiques	71
3.2.1 Anatomie d'un neurone biologique	71
3.2.2 Physiologie d'un neurone biologique	75
3.2.2.1 Le potentiel d'action	75
3.2.2.2 Le potentiel de repos	77
3.2.2.3 Potentiel postsynaptique	77
3.3 Codage neuronal de l'information	78
3.3.1 Coder l'information avec un seul neurone	79
3.3.1.1 Qu'est-ce qu'un code neuronal ?	79
3.3.1.2 Le codage par taux de décharge	80
3.3.1.3 Le codage temporel	81
3.3.1.4 Synthèse	82
3.3.2 Coder l'information avec une population de neurones	83
3.3.2.1. Le taux de décharge dans une population	83
3.3.2.2. Codage temporel dans les populations	84
3.3.2.3. Alors, quel codage ?	86
3.4 Modélisation des neurones impulsionnels	86
3.4.1 Modèle de Hodgkin et Huxley	87
3.4.2 Modèle Intègre et Tire	91
3.4.3 Modèle d'Izhikevich	95
3.4.4 Modèle à réponse impulsionnelle	95
3.5 Architecture des réseaux de neurones impulsionnels	99

3.5.1 Réseaux feed-forward	100
3.5.2 Réseaux récurrents.....	100
3.5.3 Réseaux hybrides	101
3.5.3.1 Chaîne de décharge synchrone (Synfire chain)	101
3.5.3.2 Machine à états liquides	102
3.6 Quelques propriétés des réseaux de neurones impulsionnels	103
3.6.1 Réponses synaptiques	103
3.6.2 Délais de transmission dans les réseaux de neurones impulsionnels.....	106
3.7 Apprentissage des réseaux de neurones impulsionnels	107
3.7.1 Apprentissage supervisé de réseaux de neurones impulsionnels basé sur le gradient.....	108
3.7.2 Apprentissage non supervisé de réseaux de neurones impulsionnels basé sur la STDP.....	110
3.7.2.1 Plasticité synaptique:base cellulaire de la mémoire et l'apprentissage ..	110
3.7.2.2 Modèles formels d'implémentations de la STDP	112
3.7.2.3 Mise en œuvre de la STDP	113
3.8 Conclusion	114

3.1 Introduction

L'évolution des modèles connexionnistes est bien capturée dans la catégorisation proposée par Maass (Maass, 1997), qui classe les modèles neuromimétiques en trois générations selon les types de neurones qu'ils utilisent :

- les réseaux basés sur des neurones de McCulloch&Pitts à sortie binaire sont dit de première génération. Ils peuvent calculer n'importe quelle fonction booléenne avec un seul niveau d'unités cachées. Les réseaux de première génération incluent aussi les réseaux de Hopfield et la Machine de Boltzmann qui sont de type récurrent.
- les réseaux basés sur des neurones à sortie analogique. Introduits par le groupe de Rumelhart et al. (Rumelhart et al., 1986b).
- les réseaux utilisant des neurones qui pulsent (réseaux de neurones impulsionnels) et qui peuvent modéliser des phénomènes spatio-temporels. Ces réseaux sont neurobiologiquement les plus réalistes (VanRullen et al., 2005 ; Gerstner et Kistler, 2002a ; Maass, 1997).

Les réseaux de neurones impulsionnels peuvent transmettre et recevoir des quantités importantes d'informations grâce à la synchronisation relative de quelques impulsions. Il a été démontré que les réseaux de neurones impulsionnels peuvent être appliqués à tous les problèmes solvables par les réseaux de neurones classiques. D'autre part, il a été prouvé qu'ils sont plus puissants que les perceptrons (Maass, 1997). En raison de ces propriétés intéressantes, ils font l'objet de plusieurs études (Ghosh-Dastidar et Adeli, 2009 ; Paugam-Moisy et Bohte, 2009 ; Stratton et Wiles, 2007 ; Belatreche et al., 2007). Bohte et Kok présentent une liste d'applications des réseaux de neurones impulsionnels (Bohte et Kok, 2005).

Ce chapitre dresse un panorama sur les réseaux de neurones impulsionnels toute en introduisant les sources biologiques et computationnelles en Section 3.2, le codage neuronal en Section 3.3. Les définitions et fonctionnements des modèles les plus utilisés seront détaillés en Section 3.4 ainsi que les architectures en Section 3.5. Les modèles de réponses synaptiques les plus courants et la prise en compte des délais de transmission des potentiels d'action entre les neurones en Section 3.6, les méthodes d'apprentissage dans les réseaux de neurones impulsionnels sont étudiés en Section 3.7 et enfin, nous terminons par une conclusion.

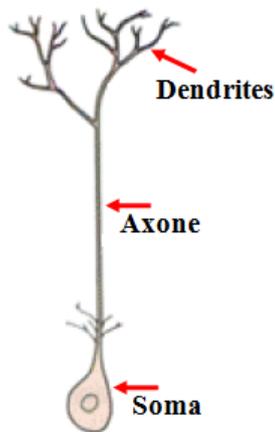
3.2 Les bases biologiques

La présentation des notions biologiques dans cette section est non exhaustive et grandement simplifiée. Pour plus de détail voir le livre de Kandel ([Kandel et al., 2000](#)) et celui d'Arbid ([Arbid, 2003](#)).

3.2.1 Anatomie d'un neurone biologique

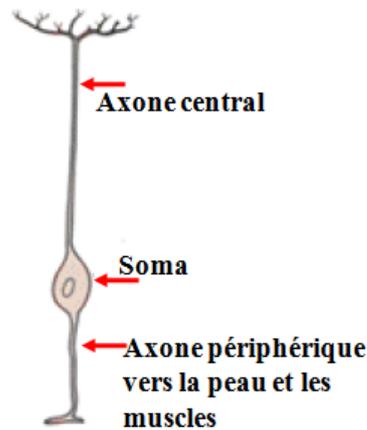
Il existe une grande diversité de neurones ([Figure 3.1](#)). Ils ont été classés selon leur morphologie qui reflète à la fois l'espèce animale, leur localisation dans le système nerveux et leur fonction ([Kandel, 2000](#)). Il a cependant été dégagé une architecture globale ([Figure 3.2](#)), commune à tous les types de neurones chez lesquels on retrouve toujours dendrite(s), axone, synapses et corps cellulaire, appelé également soma.

A) Cellule unipolaire



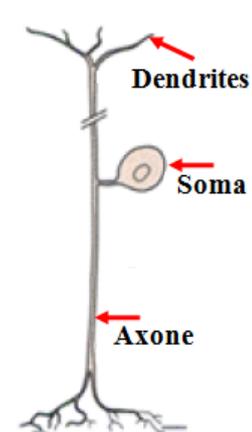
Neurone d'invertébré

B) Cellule pseudo unipolaire



Cellule ganglionnaire de l'épine dorsale

C) Cellule bipolaire



Cellule bipolaire de la rétine

D) Trois types de cellules multipolaires

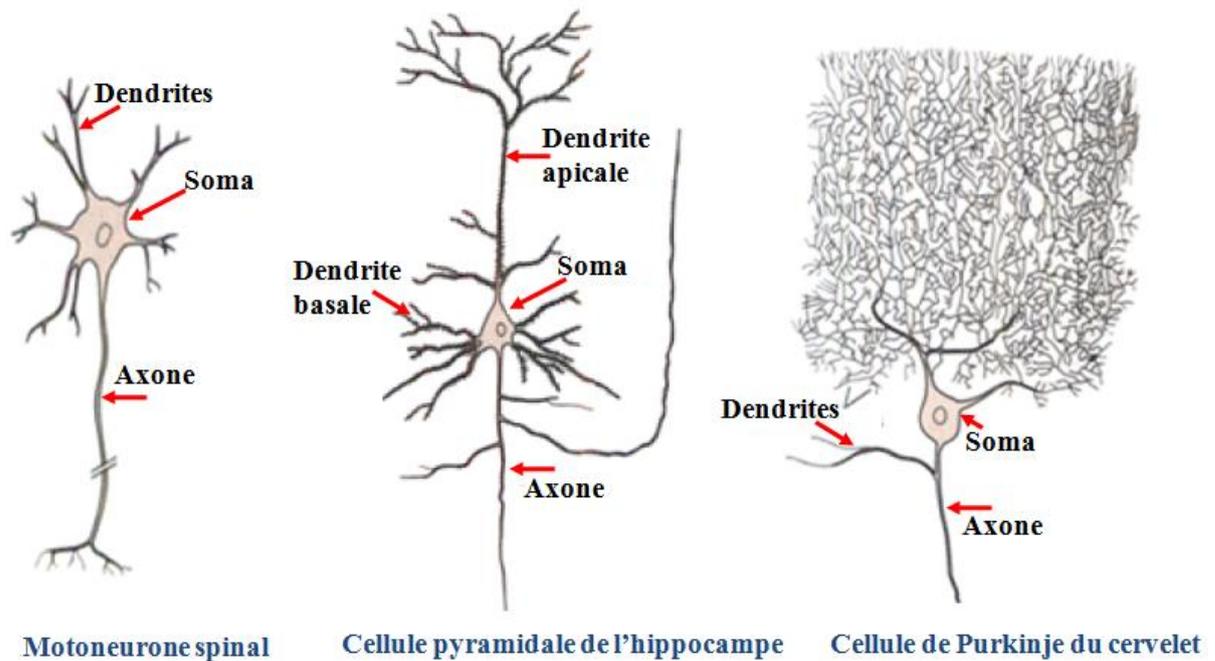


Fig. 3.1 Morphologie d'arbres dendritiques (Kandel, 2000).

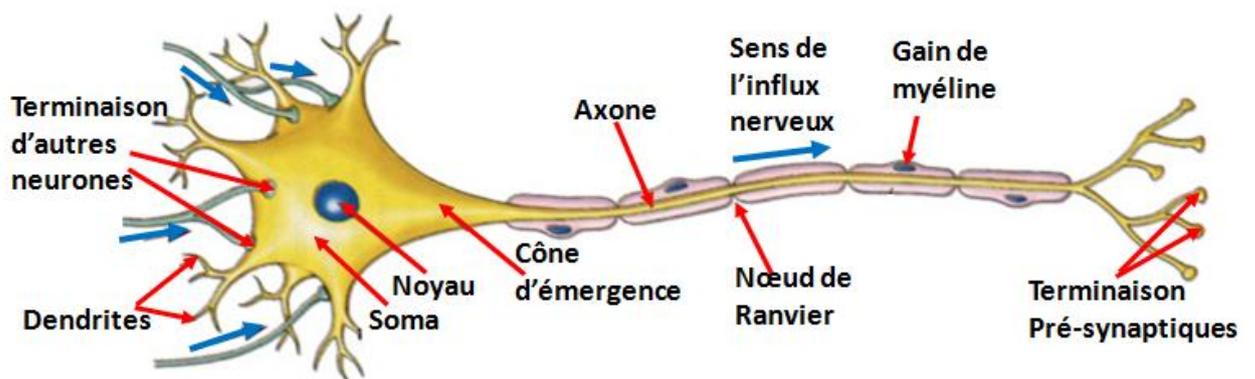


Fig. 3.2 Schéma d'un neurone biologique.

Les différentes classes de neurones se distinguent essentiellement par la forme et la taille du soma et le nombre, la longueur et la répartition des dendrites. La description ci-après détaille le rôle de chacun de ces éléments.

- **Les dendrites**

Les dendrites sont des ramifications issues du corps cellulaire et qui permettent les connexions synaptiques afférentes au neurone. Elles assurent ainsi la propagation de l'information entre les différents neurones

impliqués dans la communication en transmettant les informations générées au niveau synaptique préférentiellement vers le soma. Morphologiquement, elles se différencient de l'axone par des contours irréguliers (présence de ribosomes, petites protubérances synthétisant des protéines) et un diamètre diminuant au fur et à mesure qu'elles s'éloignent du soma. L'essentiel de l'information transmise d'un neurone à l'autre est reçu par les dendrites.

- **L'axone**

L'axone, généralement unique, est un prolongement long et cylindrique du soma. Son diamètre est inférieur à celui d'une dendrite et, chez l'homme, sa longueur peut être de l'ordre du mètre. Il conduit vers d'autres neurones un signal, le potentiel d'action, qui naît dans son segment initial. L'axone peut être entouré d'une gaine de myéline (substance essentiellement constituée de lipides, isolant et protégeant l'axone) qui accélère la propagation du signal. On trouve, le long de l'axone, des points où la gaine de myéline s'interrompt ; ces points sont appelés nœuds de Ranvier. C'est au niveau de ces nœuds que le signal électrique se régénère.

- **Les synapses**

Les synapses sont les points de contact entre un neurone et sa cible, un autre neurone, une fibre musculaire, etc. On les retrouve aussi bien sur le soma que sur l'arbre dendritique du neurone post-synaptique. En revanche, concernant le neurone pré-synaptique, elles se situent uniquement au niveau de ses terminaisons axonales.

On distingue dans le vivant deux catégories de synapses : les synapses dites « chimiques » et les synapses « électriques ». Les premières (Figure 3.3), rencontrées le plus fréquemment, nécessitent un intermédiaire pour transmettre l'information : le neurotransmetteur. En effet, l'activité électrique arrivant dans la terminaison axonale pré-synaptique stimule des vésicules synaptiques. Ces vésicules libèrent alors des neurotransmetteurs dans la fente synaptique et des récepteurs commandent l'ouverture de canaux ioniques particuliers induisant un courant post-synaptique dans les dendrites. Ces dernières transmettent ensuite à leur tour l'information au soma.

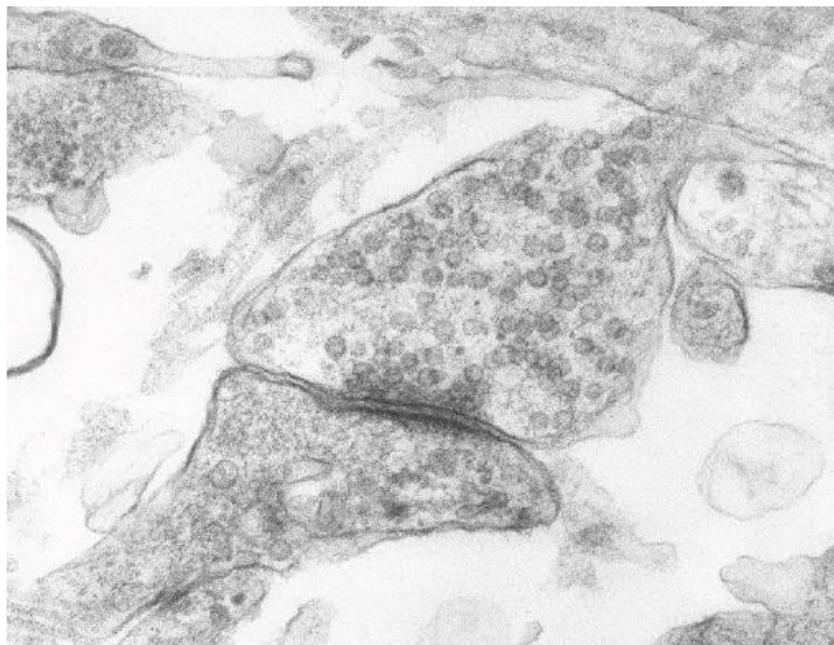


Fig. 3.3 Microphotographie électronique d'une synapse issue d'une culture de neurones du rat (**Laboratoire de neurobiologie d'Orléans**).

Dans le cas d'une synapse électrique, le signal est transmis directement au neurone post-synaptique sans délai de transmission. On estime qu'environ 40% de la surface membranaire de certains types de neurones (comme les cellules de Purkinje) sont couverts de synapses.

- **Le soma et la membrane plasmique**

Le soma d'un neurone du cortex humain a un diamètre de l'ordre de 20 μ m pour une épaisseur de membrane d'environ 5nm. Les signaux arrivant sur les dendrites sont intégrés au niveau du soma qui génère à son tour une réponse se propageant dans l'axone.

La membrane plasmique d'une cellule en délimite le contour. Elle est la frontière entre les milieux intra- et extracellulaire et se compose d'une bicouche lipidique traversée par plusieurs types de protéines. Les milieux intracellulaire et extracellulaire possèdent les mêmes espèces ioniques en des concentrations différentes de part et d'autre de la membrane. Le milieu extracellulaire est plus riche en ions sodium Na⁺ mais plus pauvre en ions potassium K⁺ que le milieu intracellulaire. Il existe d'autres types d'ions, notamment calciques Ca²⁺. Les différences de concentration calcique sont souvent proportionnelles aux variations de potentiel du neurone et peuvent être ainsi utilisées comme marqueur de l'activité électrique neuronale.

3.2.2 Physiologie d'un neurone biologique

Explicitons maintenant l'aspect fonctionnel des éléments décrits précédemment.

3.2.2.1 Le potentiel d'action

Le neurone, comme toute cellule, est doté d'une membrane qui isole les composants internes de l'extérieur. Il contient, notamment, un ensemble d'ions chargés (pour l'essentiel sodium, potassium, calcium et chlorure). Les concentrations de ces ions peuvent varier et surtout être différentes d'un côté et de l'autre de la membrane. Cette différence de charge crée une différence de potentiel au voisinage membranaire.

La membrane contient des canaux (portes) qui laissent passer les ions. D'une part, ces portes sont ions-spécifiques : elles ne laissent passer qu'un certain type d'ions. D'autre part elles sont voltage-dépendantes puisque l'ouverture de la porte est une fonction de la différence de potentiel membranaire. Afin de maintenir à long terme la concentration des différents ions de part et d'autre de la membrane, des pompes transportent les ions à l'encontre de leur gradient de concentration.

Du fait de l'action des ces pompes, au repos, la concentration des ions K^+ n'est pas à l'équilibre, ils sont plus fortement concentrés à l'intérieur de la membrane. Il y a donc une force nette de sortie. La situation est exactement inverse pour les ions Na^+ . Ces différentes concentrations ioniques créent, au repos, un potentiel membranaire non nul (environ $-70mV$).

Lors d'une augmentation de la polarité (dépolariation), du fait de la voltage-dépendance, les canaux Na^+ (respectivement K^+) voient leur efficacité augmenter (respectivement diminuer). La cinétique de ces deux variations est cependant très différente. En effet les canaux Na^+ sont composés de deux portes (les portes m et les portes h). Les portes m s'ouvrent lors d'une dépolariation. Lorsque le potentiel dépasse un seuil (celui où les courants induits le sodium se révèlent plus forts que ceux du potassium), un rétrocontrôle positif se met en place car l'afflux de sodium dans le neurone augmente cette dépolariation. Celle-ci s'accélère jusqu'au potentiel de repos du sodium (qui est positif). La polarité du neurone s'inverse alors (Figure 3.4).

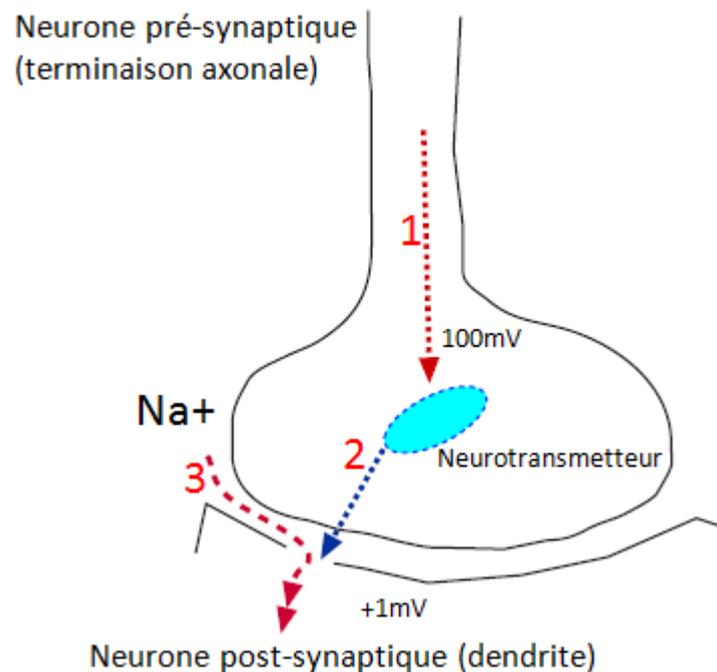


Fig. 3.4 Propagation du potentiel d'action dans la synapse : suite à l'excitation reçue par l'axone du neurone pré-synaptique (1), la bulle de neurotransmetteurs libère sa substance qui traverse la membrane du neurone émetteur et se fixe sur les chémorécepteurs sur la membrane du neurone hôte (2). Des canaux à ions s'ouvrent dans la membrane du neurone post-synaptique et les ions de l'espace intercellulaire affluent dans le neurone post-synaptique en modifiant sa polarité(3).

Cependant, cette augmentation du potentiel n'est que transitoire. En effet, deux phénomènes viennent concurrencer cette dépolarisation : premièrement, les portes h du sodium activées en sens inverse mais plus lentement que les portes m, on dit alors que le canal sodium est inactivé. Parallèlement, la porte h du canal potassium s'active, laissant sortir les ions K^+ . Cela entraîne une repolarisation puis une hyperpolarisation, lorsque le potentiel du neurone est ramené en dessous de son potentiel de repos par l'excès de K^+ sortis. Comme les portes h du sodium sont activées, les ions Na^+ ne peuvent plus entrer. Le neurone ne peut plus exercer sa boucle positive pendant un certain temps. Cette période s'appelle la période réfractaire.

Cette phase de dépolarisation/repolarisation/hyperpolarisation s'appelle un potentiel d'action (PA). On dit que le neurone, à ce moment là, décharge. Cette phase est très courte (de l'ordre de la milliseconde) et très stéréotypée. Une version schématisée est présentée sur la [Figure 3.5](#).

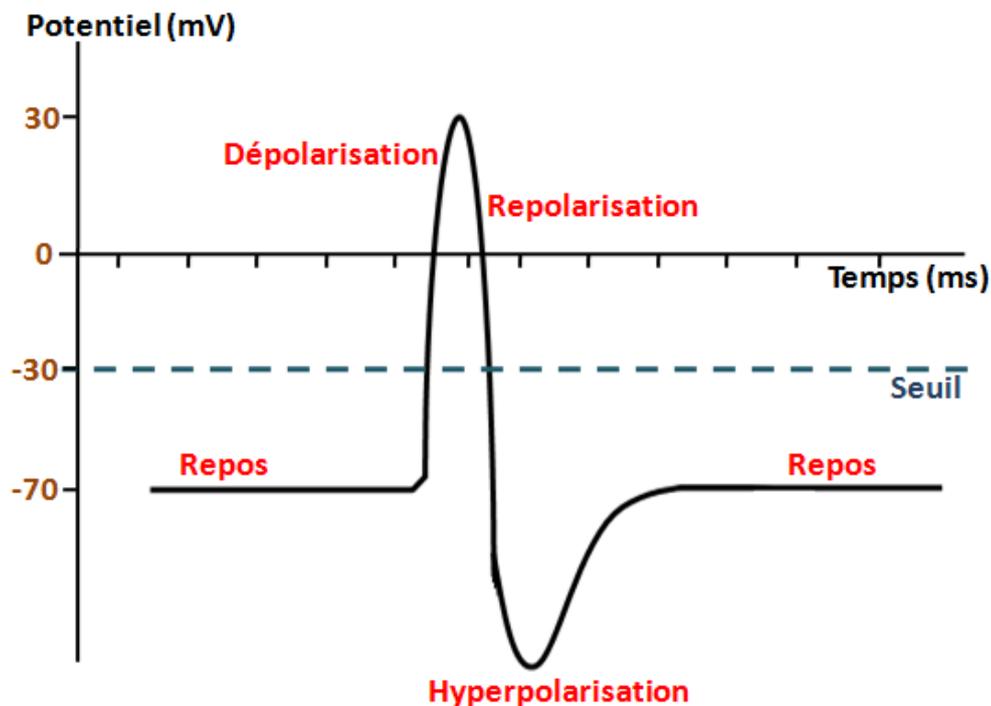


Fig. 3.5 Description d'un potentiel d'action.

Le potentiel d'action s'initie en principe dans une zone membranaire, puis se propage tout le long de la membrane tandis que la période réfractaire empêche toute rétropropagation. Ainsi, l'émission d'un potentiel d'action se fait dans un seul sens. Le potentiel d'action constitue un signal nerveux qui se propage du corps cellulaire vers l'axone qu'il suit jusqu'aux synapses, transmettant ainsi l'information vers les autres neurones.

3.2.2.2 Le potentiel de repos

Lorsque celle-ci n'est soumise à aucune excitation, le système constitué de la membrane et des milieux extracellulaire et intracellulaire est à l'équilibre, mais il existe malgré tout une différence de potentiel due à l'inégalité de concentration des ions de part et d'autre de la membrane. Cette différence de potentiel prend toujours pour référence le milieu extracellulaire et porte le nom de potentiel de repos. Ce dernier est généralement de l'ordre de -70mV .

3.2.2.3 Potentiel post-synaptique

Un potentiel post-synaptique (PPS) est le signal unitaire produit en aval d'une synapse. Il s'agit d'un changement transitoire de la différence de potentiel

électrochimique établie de part et d'autre de la membrane. D'un point de vue informationnel, il y a deux types de PPS (Figure 3.6) :

- les potentiels post-synaptiques excitateurs (PPSE) correspondent à une dépolarisation de la membrane, ce qui augmente la probabilité que le neurone décharge un potentiel d'action. Ils correspondent à une réduction temporaire du potentiel de membrane post-synaptique provoqué par un flux d'ions dont le potentiel de repos est positif (sodium et calcium principalement) entrant dans la cellule post-synaptique.
- les potentiels post-synaptiques inhibiteurs (PPSI) correspondent à une hyperpolarisation de la membrane, ce qui diminue la probabilité de la décharge. Ils correspondent à une augmentation temporaire du potentiel de membrane post-synaptique provoqué par un flux d'ions chargés négativement entrant dans la cellule post-synaptique.

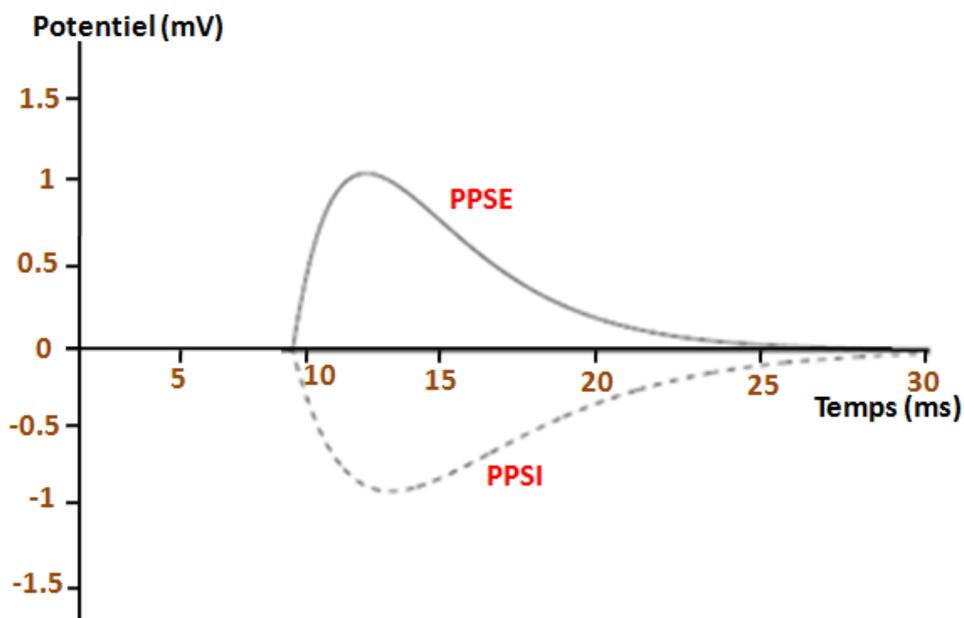


Fig. 3.6 Exemple de potentiels post-synaptiques.

3.3 Codage neuronal de l'information

L'observation de l'activité électrique d'un neurone sensoriel permettait à Edgar Douglas Adrian (Adrain, 1926) d'établir que les neurones sensoriels émettent des impulsions à la présentation d'un stimulus adéquat. Il est depuis cette époque largement admis que les neurones communiquent entre eux, sur des

distances allant de quelques centaines de microns à quelques centimètres, à l'aide de ces impulsions. Dès lors, expérimentalistes et théoriciens ont pu se poser la question de savoir comment lire les trains d'impulsions pour comprendre la communication neuronale, et jusqu'à aujourd'hui, plusieurs propositions existent (Averbeck et Lee, 2004 ; Cariani, 2004 ; Pouget et al., 2003 ; Gerstner et Kistler, 2002a ; Maass et Bishop, 2001 ; Diesmann, 1999 ; Borst et Theunissen, 1999 ; Abbott et Sejnowski, 1998 ; Gautrais et Thorpe, 1998 ; Victor et Purpura, 1996 ; Rieke et al., 1996 ; Shadlen et Newsome, 1994).

Dans cette section nous exposerons différentes propositions de codages neuronaux, leurs principes et propriétés. Nous les aborderons sous un angle unitaire, celui du neurone seul. Nous distinguerons les codes basés sur un comptage d'impulsions, et ceux basés sur le temps de décharge. Sur la base de ces codes possibles pour un neurone, nous considérerons ensuite leurs transpositions au niveau de population de neurones.

3.3.1 Coder l'information avec un seul neurone

Les impulsions sont des signaux électriques uniformes, d'une durée de 1 à 2 millisecondes, aussi appelés « potentiels d'action » ou « spikes », générés quand le potentiel de membrane du neurone atteint une valeur seuil (Figure 3.5). Ils se propagent ensuite le long de l'axone vers les synapses d'autres neurones, où leurs arrivées initient des courants post-synaptiques qui vont à leur tour déclencher, ou inhiber, de nouveaux potentiels d'action. L'impulsion, brève et indivisible, forme à ce titre l'élément de base de la communication neuronale à distance. Mais comment l'information est-elle représentée par ces impulsions ? Quels sont les aspects des trains de décharge, quelles sont les combinaisons d'impulsions, qui vont permettre à l'observateur de prédire la présence de tel stimulus à partir de la réponse du neurone ? Comme le suggère la formulation de cette interrogation, nous répondrons à la question de la représentation neuronale de l'information du point de vue de l'observateur.

3.3.1.1 Qu'est-ce qu'un code neuronal ?

L'information va être représentée par ce qui se répète dans l'émission des impulsions par le neurone, à la présentation d'un même stimulus pour un neurone sensoriel par exemple, ou dans l'exécution d'un même geste pour un neurone moteur. Autrement dit, un code neuronal peut se définir ainsi : une dimension de

l'activité neuronale dont les changements vont refléter une différence dans l'information transmise au reste du système (Perkell et Bullock, 1968).

Nous pouvons identifier plusieurs types de codage, selon les régularités constatées dans les réponses d'un neurone seul. Ainsi, le nombre d'impulsions est une variable très fiable, et simple, de codage de l'information.

3.3.1.2 Le codage par taux de décharge

Un stimulus est présenté à un neurone. En réaction, celui-ci émet des impulsions. En règle générale, plus le stimulus est intense, plus le neurone décharge. Le premier code neuronal est simple, il est basé sur le nombre de potentiels d'action émis. Et puisque ce nombre est mesuré sur une fenêtre de temps finie, nous parlerons d'un taux de décharge. Par code en taux de décharge, nous entendrons donc tout code basé sur un comptage des impulsions émises par un neurone durant une période donnée (Figure 3.7). C'est une moyenne temporelle, un taux v exprimé en Hertz calculé d'après le nombre n d'impulsions observées sur une période T exprimée en secondes, typiquement 0.1 ou 0.5s:

$$v = n/T \quad (3.1)$$

Le taux de décharge est une mesure fiable du codage neuronal de l'information, qui permet de s'affranchir du caractère temporel irrégulier de certaines réponses corticales (Shadlen et Newsome, 1994). Son grand avantage réside donc dans l'affranchissement de la dimension temporelle à travers une moyenne, particulièrement pratique pour rendre compte d'un phénomène complexe, se manifestant à des échelles spatiale et temporelle minuscules dans un environnement qui plus est bruité.

C'est en même temps un inconvénient : si l'expérimentateur bénéficie d'autant de temps qu'il désire pour observer l'activité neuronale, il n'en va pas toujours de même pour le neurone, ou pour l'organisme. La mouche est par exemple capable de s'adapter à l'apparition d'un nouveau stimulus en changeant sa trajectoire en 30 à 40 millisecondes (Rieke et al., 1996). Cela ne laisse pas assez de temps pour compter les impulsions d'un neurone et calculer une moyenne sur une longue période. De fait, la mouche doit répondre alors qu'un neurone post-synaptique n'aura reçu qu'un ou deux potentiels d'action. Ceci limite sérieusement la capacité du taux de décharge à porter l'information, à moins de rendre compte de ce taux de manière instantanée.



Fig. 3.7 Codage par taux de décharge. Le nombre d'impulsions émis par le neurone est mesuré sur une période donnée.

3.3.1.3 Le codage temporel

Nous qualifierons de temporel les codes où les variations des dates précises de décharge accompagnent une variation de l'information contenue dans le stimulus. Dit autrement, les codes temporels font l'hypothèse d'une reproductibilité des temps de décharges : un neurone utilise un code temporel quand un même stimulus entraînera la production d'une ou de plusieurs impulsions à des dates précises (de l'ordre de la milliseconde).

Pour un même neurone, l'information pourra alors être portée par le temps de décharge lui-même (Bialek et al., 1991), mais aussi par le délai, conséquemment constant, entre deux ou plusieurs décharges (Figure 3.8.A) (Berry et al., 1997) ou bien encore la latence (Figure 3.8.B) (Petersen et al., 2001).

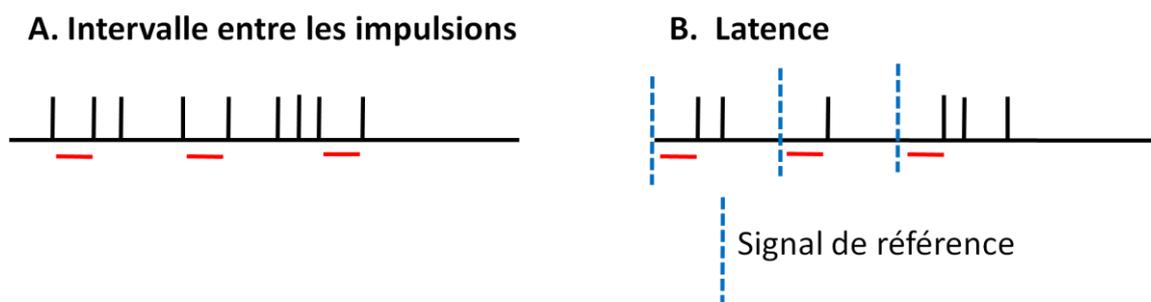


Fig. 3.8 Codage temporel. (A) Codage par délai entre deux impulsions. (B) Codage par latence. Le délai d'apparition de la première impulsion est mesuré par rapport à un signal de référence.

Le codage par latence

Le codage par latence est une forme de codage temporel qui utilise le temps mis par un neurone pour émettre son premier potentiel d'action (Figure 3.7.B). La latence (aussi appelée lag) désigne le délai entre le moment où une information est envoyée et celui où elle est reçue. De façon plus générale, la latence peut aussi désigner l'intervalle entre la fin d'un événement et le début de la réaction à celui-

ci. Ainsi, on y considérera que plus l'intensité du stimulus est grande, plus la latence sera courte. La mesure des latences de décharge présente cependant un problème, inhérent à la prise en compte de la dimension temporelle : celui du signal de référence.

3.3.1.4 Synthèse

Le codage par taux de décharge, dans sa version classique, permet une évaluation simple, pratique et efficace de la quantité d'information transmise par un neurone au reste du système : plus le neurone décharge, plus la probabilité de présence de son stimulus préféré dans son champ récepteur est grande (Barlow, 1972). Le codage par taux de décharge s'affranchit donc de la dimension temporelle par une moyenne sur une période donnée. En réduisant la résolution temporelle de comptage des potentiels d'action, et en répétant les observations, le taux de décharge peut varier dans le temps, il devient instantané. Il s'assimile alors à une probabilité instantanée de décharge au cours du temps : la probabilité d'observer l'émission d'une impulsion à un pas de temps donné sera d'autant plus grande que le neurone aura déchargé souvent à ce pas de temps dans les observations précédentes.

Un codage temporel s'appuie quant à lui sur les dates précises des impulsions, puisque certaines des réponses évoquées dans différentes zones du cerveau présentent une précision de l'ordre de la milliseconde. Quand le temps de décharge est mesuré à partir d'un signal de référence, telle que la présentation du stimulus par exemple, le codage temporel devient alors codage par latence : plus tôt le neurone décharge, plus la probabilité de présence de son stimulus préféré dans son champ récepteur est grande (Thorpe, 1990).

En fin de compte, le codage temporel met l'accent sur l'importance de chaque impulsion dans la représentation de l'information, plutôt que dans leur nombre. Pourvu que les réponses soient alors reproductibles, c'est-à-dire précises de l'ordre de la milliseconde, la quantification de l'information véhiculée par chacun de ces vecteurs montrent un très net avantage en faveur de l'approche temporelle. De plus, sa capacité à suivre fidèlement le déroulement temporel des stimulations exogènes lui donne un avantage écologique incontestable. Mais est-il pour autant meilleur que le codage par taux de décharge?

3.3.2 Coder l'information avec une population de neurones

L'information émise par un neurone, sous la forme d'impulsions, peut être appréhendée selon deux angles différents, temporel ou atemporel, séparés par le rapport au temps qu'installe l'observateur dans sa mesure, pour expliquer tel ou tel phénomène. Le neurone lui-même est susceptible de lire différents aspects des entrées synaptiques. Mais ces entrées sont elles-mêmes induites par les décharges émises vers le neurone concerné. Cela veut dire qu'un neurone reçoit de l'activité d'un grand nombre d'autres neurones, que par simplification nous considérerons organisés en population de neurones. En effet, le cerveau est constitué en partie d'un très grand nombre de neurones. Et tous ne vont pas réagir à la présentation d'un stimulus, mais certains d'entre eux peuvent décharger suite à l'événement, ensemble, en assemblée de cellules (Hebb, 1949).

3.3.2.1. Le taux de décharge dans une population

La définition d'un taux de décharge dans une population est très proche de celle d'un taux de décharge instantané. Le taux de population mesure les variations de décharge dans la population afférente au cours du temps (Figure 3.9).

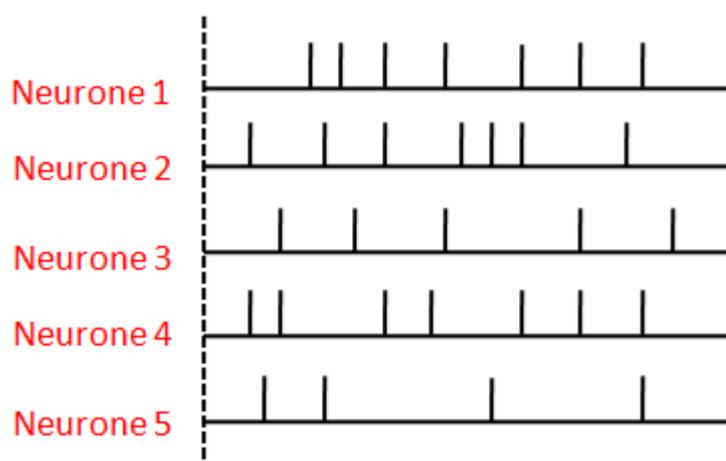


Fig. 3.9 Codage par modulation de l'activité de population. Le nombre d'impulsions émis par chaque neurone de la population est mesuré sur une période donnée. Les variations du taux des décharges émises par la population représentent l'information reçue.

Le temps t est mesuré dès le début de la stimulation et le nombre d'impulsions observées rapporté pour chaque pas de temps Δt , typiquement entre 1 et quelques millisecondes, et chaque neurone. Le nombre d'impulsions à chaque

pas de temps, $n_{act}(t; t + \Delta t)$, est ensuite sommé sur l'ensemble des neurones enregistrés, pour être ensuite divisé par le nombre N de neurones. On obtient ainsi une mesure de l'activité moyenne d'un neurone de la population entre t et $t + \Delta t$. En divisant par la longueur de l'intervalle Δt , on obtient finalement le taux de décharge instantanée de la population, exprimé en Hertz :

$$A(t) = \frac{(n_{act}(t; t + \Delta t)/N)}{\Delta t} \quad (3.2)$$

3.3.2.2. Codage temporel dans les populations

Le codage temporel de l'information au niveau d'un neurone se transpose aisément au niveau d'une population : des neurones, considérés indépendants par simplicité, déchargent à des dates plus ou moins précises. Comment l'information à propos du stimulus peut-elle y être représentée ? Nous allons ici répondre à cette question en considérant deux cas : celui où la place précise des décharges les unes par rapport aux autres n'a pas d'importance (synchronie) et celui où la structure spatio-temporelle précise est importante pour le codage (asynchronie).

a. Codage synchrone

Nous parlerons de code synchrone quand plusieurs neurones vont émettre des impulsions à peu près en même temps, de l'ordre de la milliseconde (Figure 3.10). Le caractère synchrone des décharges peut alors représenter un aspect important de la stimulation. Les neurones sélectifs aux différents sous-éléments constitutifs d'un objet déchargeraient ensemble pour représenter le stimulus en entier.

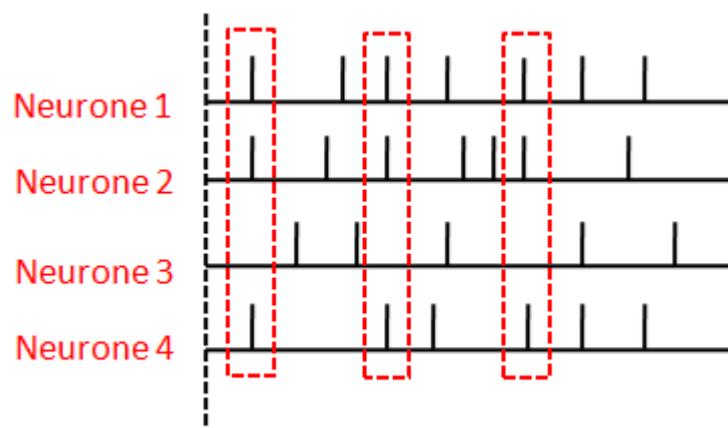


Fig. 3.10 Codage temporel synchrone. L'information est représentée par la synchronisation des réponses de certaines cellules.

b. Codage asynchrone

Le codage temporel asynchrone met l'accent sur la structuration temporelle des impulsions à travers une population (Figure 3.11). Pour cette forme de code, les neurones ont tendance à décharger en même temps, l'affiliant de ce point de vue au codage temporel. Mais ici, contrairement au codage synchrone, les temps relatifs des impulsions émises à travers la population sont considérés comme porteurs d'information.

Evidemment, ce type de code exige une structuration temporelle vraiment précise des impulsions pour en tirer bénéfice : à quelques millisecondes près d'imprécision, l'asynchronie peut être perturbée à tel point que le message reçu différera qualitativement de celui envoyé. Si ce type de codage paraît très fragile, et de fait reste encore très théorique, des évidences commencent à percer en sa faveur. Dans le cas de la localisation de sources auditives, un décalage de 0,5 ms seulement entre les arrivées, sur l'oreille droite et l'oreille gauche, d'un même stimulus sonore permet de déterminer le côté de la stimulation.

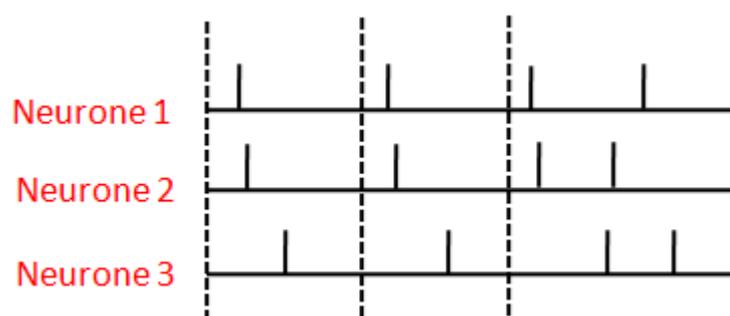


Fig. 3.11 Codage temporel asynchrone. Les premières impulsions de chaque neurone sont émises à peu près au même moment d'une présentation d'un même stimulus à l'autre.

c. Synthèse

Codage synchrone et asynchrone diffèrent selon que la structure temporelle des impulsions les unes par rapport aux autres porte de l'information ou non. Dans le premier cas, peu importe qu'une cellule décharge avant ou après une autre, l'information est plutôt portée par quelles cellules déchargent à peu près ensemble, en réaction à une stimulation. Le second cas, celui de l'asynchronie, insiste sur les temps précis de chacune des impulsions pour encoder l'information et ; potentiellement ; l'ordre avec lequel les neurones déchargent. Ces deux codes sont-ils vraiment différents ? Pas nécessairement. Après tout, le codage

asynchrone répond aux critères de la synchronie, pourvu que les délais entre décharges ne soient pas trop importants.

Symétriquement, les évidences expérimentales en faveur du codage synchrone ne désavouent pas plus l'asynchronie ; la plupart rapportent l'existence d'une activité synchrone dans différents systèmes sans aller spécifiquement jusqu'à interroger la question de l'asynchronie dans les chaînes observées. De ce point de vue-là, nous pouvons même concevoir l'asynchronie comme une sophistication de la synchronie, un moyen supplémentaire donné aux décharges synchrones de convoier de l'information.

Quoiqu'il en soit, le codage temporel par population, basé sur des impulsions isolées corrélées à travers une assemblée de neurones, est particulièrement adapté à un traitement rapide de l'information (Hopfield, 1995).

3.3.2.3. Alors, quel codage ?

Le codage par taux de décharge dans une population s'apparente directement au codage unitaire par modulation de décharge : les fluctuations d'activité constatées au niveau de la population déterminent l'information propagée d'une population à une autre.

Les codes temporels tirent parti de la reproductibilité des réponses, pour en dériver deux modes d'exploitations : un premier, dit synchrone, où différents neurones s'organisent en assemblée en déchargeant à peu près au même moment; un second, dit asynchrone, où la reproductibilité temporelle des réponses unitaires se traduit au niveau des populations en délai quasi constant entre différentes cellules, pour représenter un même stimulus.

La question du codage neuronal apparaît comme celle de l'échelle temporelle à laquelle varient les taux de décharge instantanée. À une résolution temporelle suffisamment fine, on capturera dans les modulations d'activité de population les corrélations correspondant à des décharges hautement synchrones. L'asynchronie augmenterait quant à elle la puissance computationnelle d'un codage purement synchrone.

3.4 Modélisation des neurones impulsionnels

La modélisation consiste à identifier et mettre en équation les différents mécanismes physiques et chimiques qui sont en cause dans la génération du potentiel d'action. La variable importante ici est le potentiel de membrane V_m . Elle

est régie par, d'une part, les propriétés électriques passives de la membrane, et, d'autre part, les courants ioniques. Le but d'un modèle biophysique est de comprendre les mécanismes qui constituent le comportement neuronal : génération du potentiel d'action, computation dendritique, propagation dans l'axone, etc.

3.4.1 Modèle de Hodgkin et Huxley

Bien avant l'arrivée des modèles impulsionnels, Hodgkin & Huxley (Hodgkin et Huxley, 1952) avaient proposé un modèle de neurone, issu d'études sur les mécanismes responsables de la génération d'un potentiel d'action (PA) au sein d'un neurone. Ce modèle (HH) est le modèle de neurone le plus complexe mais également le plus précis existant à ce jour. Il est composé de quatre équations différentielles couplées exprimant la dynamique du potentiel de membrane V_m du neurone. Celui-ci correspond à la différence de potentiel entre l'intérieur du neurone et le milieu extérieur.

V_m est fonction du courant d'entrée I appliqué au neurone lorsque celui-ci est stimulé, des courants I_k et I_{Na} générés par la circulation d'ions K^+ et Na^+ au travers de la membrane et d'un courant de fuite I_l représentant notamment les mouvements d'ions Cl^- . Le schéma électrique correspondant est présenté sur la [Figure 3.12](#).

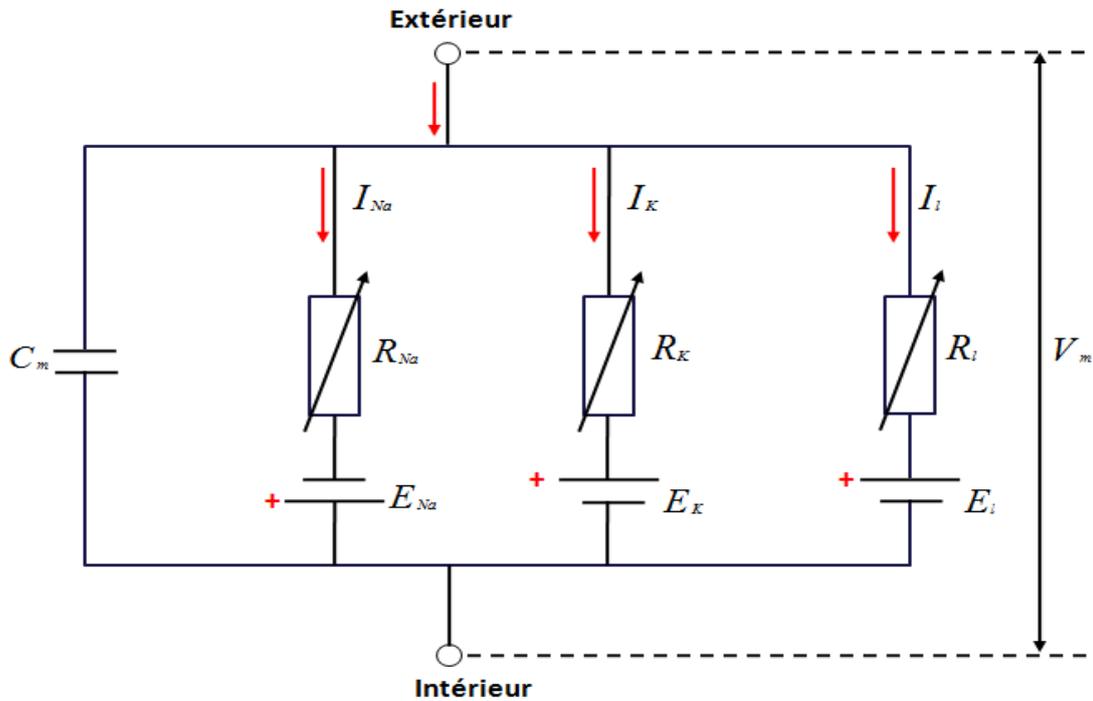


Fig. 3.12 Schéma électrique représentant les flux ioniques au travers de la membrane d'un neurone selon le modèle de Hodgkin et Huxley. C est la capacité membranaire, R la résistance, V le potentiel de membrane, E le potentiel d'inversion et I le courant injecté.

Chacun de ces courants est fonction de l'écart entre le potentiel de membrane V_m et les potentiels d'inversion E_{Na} , E_K et E_l . Le potentiel de membrane du neurone est régi par l'équation :

$$C_M \frac{dV_M}{dt} = I - g_{Na} m^3 h (V_M - E_{Na}) - g_K n^4 (V_M - E_K) - g_l (V_M - E_l) \quad (3.3)$$

où les $g_i, i \in \{K, Na, l\}$ sont des constantes et les paramètres h , m et n décrivent les probabilités d'ouverture/fermeture des canaux ioniques : sodium pour h et m , potassium pour n . Pour chacun de ces paramètres on a :

$$\frac{dx}{dt} = \alpha_x (1 - x) - \beta_x x \quad (3.4)$$

où $x \in \{h, m, n\}$ et où α_x et β_x correspondent respectivement aux probabilités d'entrée de particules à l'intérieur de la membrane et de sortie de particules vers le milieu extérieur. α_x et β_x ne dépendent pas directement du temps, mais dépendent du potentiel de membrane selon les équations :

$$\alpha_n = 0.01(V_M + 10)/(e^{(\frac{V_M+10}{10})}) - 1 \quad (3.5)$$

$$\beta_n = 0.125e^{\frac{V_M}{80}} \quad (3.6)$$

$$\alpha_m = 0.01(V_M + 25)/(e^{(\frac{V_M+25}{10})}) - 1 \quad (3.7)$$

$$\beta_m = 4e^{\frac{V_M}{18}} \quad (3.8)$$

$$\alpha_h = 0.07e^{(\frac{V_M}{20})} \quad (3.9)$$

$$\beta_{hH} = 1/(e^{(\frac{V_M+30}{10})} + 1) \quad (3.10)$$

Ces équations décrivent le comportement d'un élément de surface du neurone. Dans les travaux de Hodgkin & Huxley, la description ne s'arrête pas au comportement d'un élément de surface. Ils tiennent aussi compte de la propagation des courants ioniques et de la température afin de décrire la transmission des PA d'un élément de surface au suivant. Mais la description du comportement d'un élément de surface est utilisée comme approximation dans la plupart des simulations, principalement pour des raisons de puissance de calcul nécessaire (par exemple à la simulation de la propagation d'un PA le long d'un axone).

A partir de ce modèle, de nombreux comportements neuronaux peuvent être reproduits. Le gros inconvénient de ce modèle est sa complexité. En effet, le couplage des équations différentielles (Eq. 3.3) et (Eq.3.4) avec $x \in \{h, m, n\}$ rend extrêmement lourde la construction de réseaux à partir de tels modèles de neurones. En conséquence les simulations informatiques utilisant des modèles de type HH sont, la plupart du temps, des simulations unitaires ou de couples de neurones (Wang et Buzsaki, 1996). Il est nécessaire de simplifier ce modèle pour être en mesure de simuler des réseaux d'au moins quelques centaines de neurones et pouvoir ainsi étudier leur dynamique ou la plausibilité d'interactions présumées entre différentes populations.

Réductions du modèle d'Hodgkin-Huxley

Le modèle HH connaît des simplifications diminuant de trois à deux le nombre de paramètres représentant les dynamiques ioniques conservant un seul

paramètre en plus du potentiel de membrane qu'ils proposent d'assimiler à l'activation des canaux K^+ (paramètre n dans l'équation (Eq. 3.3)). Ceci peut se justifier par le fait que les dynamiques des trois paramètres du modèle HH ont des échelles de temps différentes : on peut se permettre de ne conserver que celui des trois qui a le plus d'influence pour approximer les équations (Eq. 3.3) et (Eq. 3.4). On peut citer les variantes de Fitz Hugh (FitzHugh, 1961), Nagumo (Nagumo et al., 1962), Morris et Lecar (Morris et Lecar, 1981), Hindmarsh et Rose (Hindmarsh et Rose, 1984) ou Doya (Doya et al., 1993). Nous présentons brièvement les équations de quelques unes d'entre elles :

1) Equations du modèle de FitzHugh-Nagumo :

$$\frac{dV_M(t)}{dt} = V_M(t) - V_M^3(t) - w(t) - I(t) \quad (3.11)$$

$$\frac{dw(t)}{dt} = V_M(t) - a - bw(t) \quad (3.12)$$

avec w , variable homogène à une tension, I un courant de stimulation extérieur, a et b des constantes à déterminer.

2) Le modèle de Morris et Lecar est un modèle à conductances qui combine le formalisme original d'Hodgkin et Huxley et le modèle de FitzHugh-Nagumo; il a plus de paramètres à ajuster que FitzHugh-Nagumo, mais moins qu'Hodgkin-Huxley. Il est préféré pour la modélisation d'une cellule ou de réseaux de taille moyenne avec des comportements dynamiques variés. Il se présente sous la forme suivante:

$$C_M \frac{dV_M}{dt} = \sum_{ion} I_{ion}(V_M) \quad (3.13)$$

$$\frac{dw}{dt} = \phi \frac{w_\infty - w}{\tau_w(V_M)} \quad (3.14)$$

$$\sum_{ion} I_{ion} = g_{c_m} \cdot m_\infty (V_M - V_{Ca}) + g_k \cdot w (V_M - E_k) + g_{fuite} (V_M - E_{fuite}) \quad (3.15)$$

- 3) Enfin, le modèle Hindmarsh-Rose, quant à lui, est construit à partir du modèle de FitzHugh-Nagumo et combine trois équations différentielles du premier ordre:

$$\frac{dx}{dt} = y + 3x^2 - x^3 - z + I \quad (3.16)$$

$$\frac{dy}{dt} = 1 - 5x^2 - y \quad (3.17)$$

$$\frac{dz}{dt} = r \left[4 \left(x + \frac{8}{5} \right) - z \right] \quad (3.18)$$

Le paramètre à estimer est ici la constante r . Ce modèle simple est capable de reproduire des comportements dynamiques variés sans nécessiter beaucoup de réglages.

3.4.2 Modèle Intègre et Tire

Lapicque ([Lapicque, 1907](#)), très en avance sur les connaissances du fonctionnement neuronal, avait proposé un modèle ne tenant compte que du potentiel de membrane V_M pour représenter l'activité du neurone. Ces travaux furent repris ([Abbott, 1999](#)) pour lui donner le nom de neurone IF « Intègre et Tire » en anglais: Integrate and Fire.

Par rapport au modèle HH décrivant précisément le décours temporel du PA, le neurone IF propose une représentation du PA par une impulsion instantanée. Dans le modèle de Lapicque, la dynamique sous le seuil est donnée par l'équation différentielle linéaire du premier ordre :

$$C_M \frac{dV_M}{dt} = -g_M(V_M - V_r) + I \quad (3.19)$$

Le potentiel de membrane V_M du neurone est constant et égal à son potentiel de repos V_r lorsqu'il n'est pas stimulé. Les PPS (potentiel post-synaptique) reçus par le neurone (et/ou le courant appliqué au neurone) sont intégrés dans le courant I . Ils génèrent une variation de V_M qui revient ensuite à son potentiel de repos selon le coefficient de conductance g_M ([Figure 3.13](#)). C_M est la capacité (de charge) du neurone. Une non-linéarité est introduite dans la

dynamique illustrée sur la [Figure 3.14](#). Lorsqu'à $t = t_0$, $V_M(t_0)$ franchit une valeur seuil ϑ , un PA est alors émis à t_0 et le potentiel de membrane V_M est réinitialisé à un potentiel d'hyperpolarisation V_h qui est parfois le même que le potentiel de repos V_r .

Ce modèle est plus communément nommé Intègre et Tire à Fuite, (LIF pour Leaky Integrate and Fire) du fait que le potentiel de membrane est ramené progressivement à son potentiel de repos comme une fuite le ferait (pour vider un récipient par exemple). Un neurone IF au sens strict est alors un LIF sans fuite, c'est-à-dire qu'il intègre les entrées et émet un PA en dépassant le seuil. Mais il conservera un potentiel de membrane constant, quel que soit le temps de repos du neurone et son état d'excitation, jusqu'à la prochaine stimulation. Le neurone ainsi défini correspond alors au schéma électrique de la [Figure 3.15](#).

Après émission d'un potentiel d'action (PA), le neurone sera maintenu à la valeur de son potentiel V_h pendant une période dite « réfractaire » τ_r pendant laquelle les stimulations n'auront pas d'impact sur le neurone. Il est également possible de modéliser cette période réfractaire par une augmentation arbitraire de la valeur seuil ϑ (ou par une fonction seuil appropriée). Cette période réfractaire peut se diviser en deux phases ([Figure 3.14](#)) :

1. Une période réfractaire absolue qui maintient V_M à son potentiel d'hyperpolarisation V_h .
2. Une période réfractaire relative, durant laquelle le potentiel de membrane rejoint son potentiel de repos V_r selon l'équation ([Eq. 3.19](#)).

Le modèle LIF réduit de 4 à 1 le nombre d'équations différentielles à résoudre ; il ne modélise plus les variations du potentiel de membrane d'un neurone à un instant donné mais réduit les émissions de PA à une impulsion instantanée et décrit modélise l'évolution du potentiel de membrane en l'absence de PA. Malgré cette simplification, la puissance computationnelle ([Maass, 1997](#)) et les capacités de stockage ([Izhikevich, 2006](#)) d'un réseau utilisant de tels modèles de neurones sont bien supérieures à celles des modèles de neurone à seuil ou sigmoïdaux.

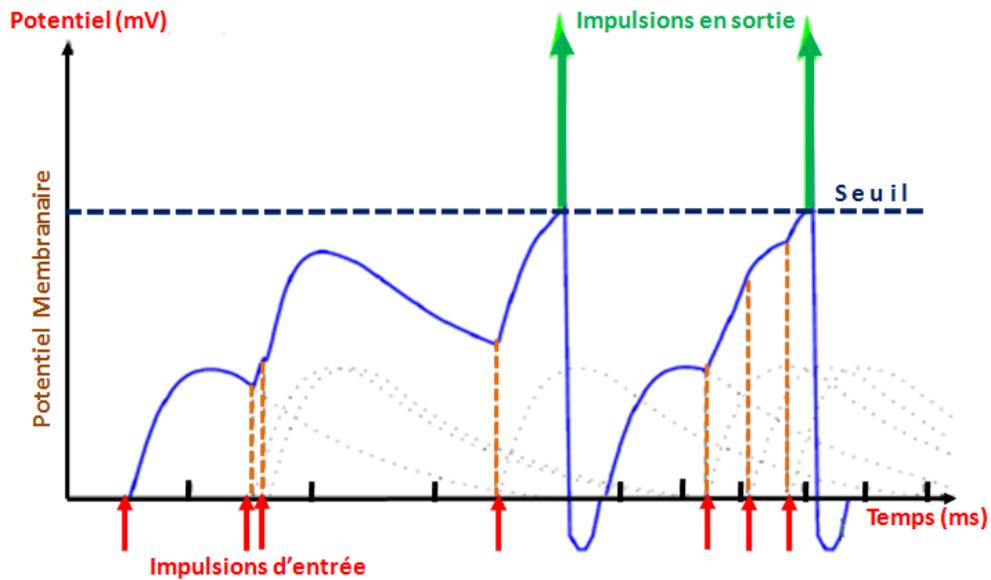


Fig. 3.13 Schéma d'intégration de PPS par un neurone de type LIF. Les flèches rouges indiquent la réception d'un PPS. Les flèches vertes représentent les PA émis par le neurone. Lorsque l'arrivée de PPS augmente le potentiel de membrane jusqu'à dépasser le seuil, alors un PA est émis. (Paugam-Moisy, 2006).

La simplification existant entre les modèles HH et LIF permet d'accéder à des informations sur les comportements neuronaux à des échelles allant du neurone unitaire au réseau de très grande taille. Les temps de simulation et les consommations mémoire sont significativement réduits par l'utilisation de modèles plus simples que le HH. On trouve entre ces deux modèles différents niveaux de complexité. Chaque simplification entraîne de nouvelles approximations par rapport au comportement de neurones biologiques.

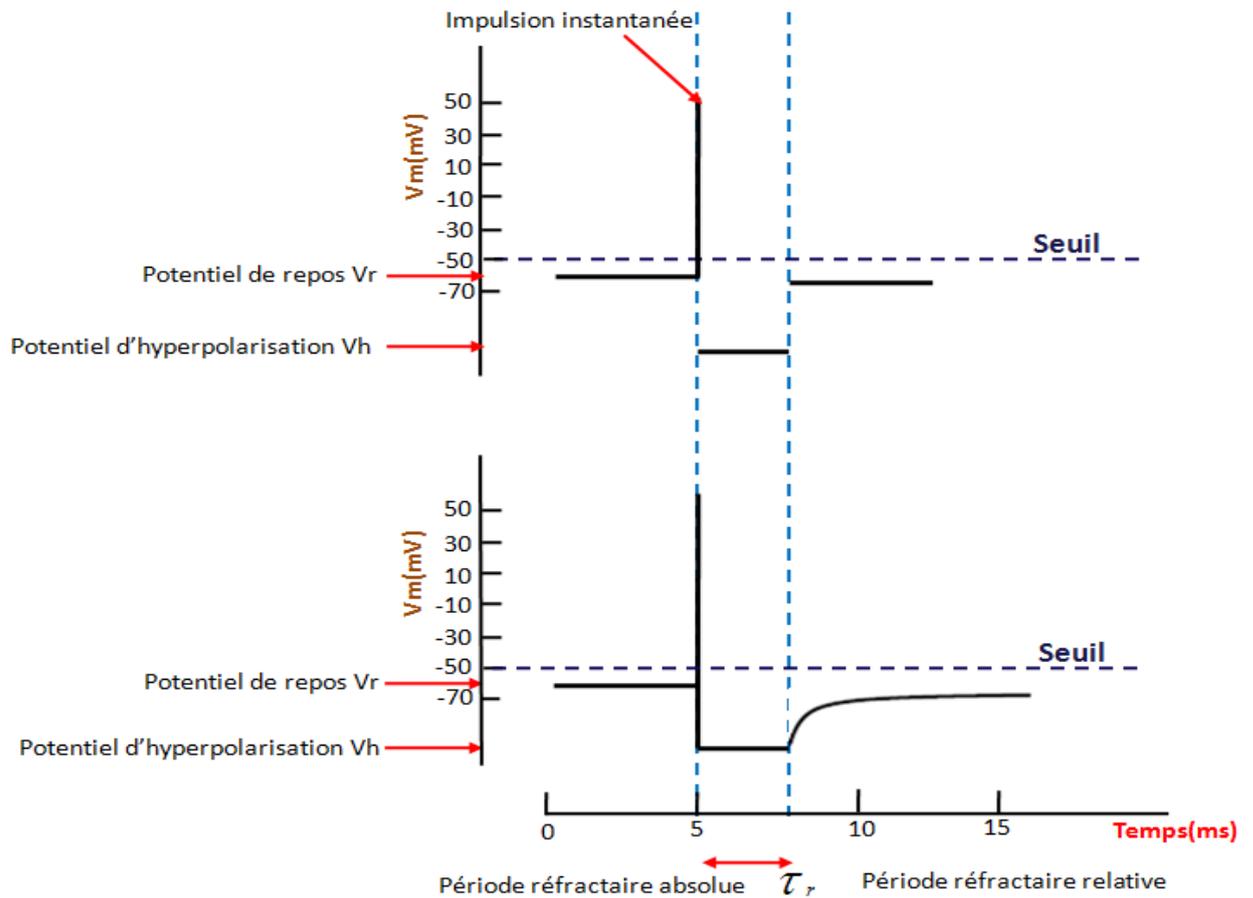


Fig. 3.14 Courbe du haut : forme du PA pour un neurone IF avec période réfractaire absolue. Courbe du bas : forme du PA pour un neurone IF avec période réfractaire absolue et relative.

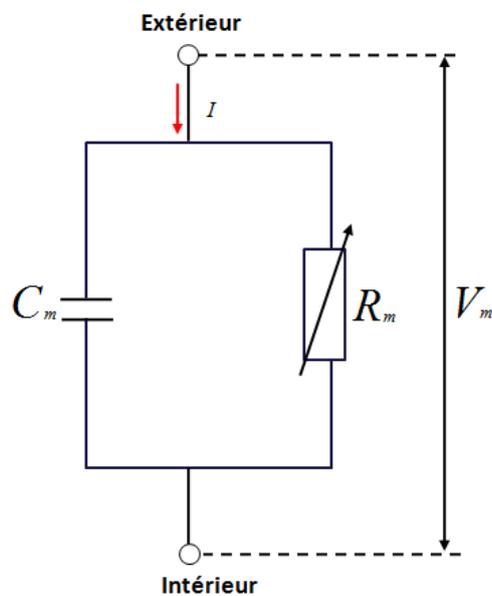


Fig. 3.15 Schéma électrique du neurone Intègre et Tire à fuite. Un condensateur de capacité C_M est branché en parallèle avec une résistance $R_M = 1/g_M$.

3.4.3 Modèle d'Izhikevich

Un autre type de modèle (Izhikevich et al., 2003a), basé sur deux équations différentielles couplées et inspiré du modèle de Fitzhugh-Nagumo (FitzHugh, 1961 ; Nagumo et al., 1962), a ouvert l'accès à la reproduction de nombreux comportements observés sur les neurones biologiques. Ce modèle est décrit par les équations :

$$v' = 0.004v + 5v + 140 - u + I \quad (3.20)$$

$$u' = a(bv - u) \quad (3.21)$$

$$\text{Si } v \geq 30\text{mV}, \text{ alors } \left\{ \begin{array}{l} v \leftarrow c \\ u \leftarrow u + d \end{array} \right\} \quad (3.22)$$

où v et u sont des variables sans dimension et qui correspondent au potentiel de membrane et à la récupération de la membrane selon (Izhikevich, 2003a).

En faisant varier les quatre paramètres a , b , c et d on peut obtenir la plupart des réponses neuronales connues, ce qui rend ce modèle particulièrement intéressant. De plus, malgré le couplage de ses deux équations différentielles, ce modèle reste relativement simple en comparaison des modèles de type HH et il est particulièrement adapté aux simulations de réseaux de neurones impulsionnels (Izhikevich, 2004a).

3.4.4 Modèle à réponse impulsionnelle

Deux différences majeures distinguent les modèles précédents du modèle SRM (pour Spike Response Model) décrit par Gerstner & Kistler (Gerstner et Kistler, 2002a). Tout d'abord, les paramètres influençant le potentiel de membrane ne sont plus des fonctions de ce potentiel, ils dépendent de la date \hat{t}_i du dernier PA émis par ce neurone. De plus, le potentiel de membrane n'est plus obtenu à l'aide d'équations différentielles ; il s'agit désormais d'une intégrale tenant compte du passé du neurone pour décrire l'évolution de son potentiel de membrane.

Nous avons besoin de définir quelques notations et symboles qui seront utilisés. Un neurone pré-synaptique est désigné par i et un neurone post-synaptique par j . L'ensemble des neurones pré-synaptiques des neurones j est alors définie comme $\Gamma_j = \{i | i \text{ presynaptic à } j\}$. De même, l'ensemble des neurones post-synaptiques de neurone i est défini comme $\Gamma_i = \{j | j \text{ postsynaptic à } i\}$. Le moment où un neurone donné j émet n potentiel d'action sera dénommé le temps

d'impulsion de ce neurone et sera désigné par $t_j^{(f)}$, avec f varie de 1 à n . Le train d'impulsions du neurone j est alors caractérisé par l'ensemble des temps de tirs indiqués par l'équation (Eq. 3.23).

$$F_j = \{t_j^{(1)}, t_j^{(2)}, \dots, t_j^{(n)}\} \quad (3.23)$$

avec $t_j^{(n)}$ est l'impulsion la plus récente du neurone j , également désigné par \hat{t} . Le potentiel post-synaptique (PPS) est la réponse générée par une impulsion du neurone pré-synaptique i vers les neurones post-synaptiques j . Il est important de noter que, en raison des distances inter-neurones et les temps de transmission axonale, il peut y avoir un délai entre le début d'émission d'une impulsion et son PPS correspondant. L'équation (Eq. 3.24) donne une formulation mathématique d'un PPS (Gerstner et Kistler, 2002a), ou dans une forme plus simplifiée l'équation (Eq. 3.25) (Bohte et al., 2002a).

$$\varepsilon_{ij}(t) = \left[\exp\left(-\frac{t - t_i^{(f)} - \Delta_{ij}^{ax}}{\tau_m}\right) - \exp\left(-\frac{t - t_i^{(f)} - \Delta_{ij}^{ax}}{\tau_s}\right) \right] H\left(t - t_i^{(f)} - \Delta_{ij}^{ax}\right) \quad (3.24)$$

$$\varepsilon_{ij}(t) = \frac{t - t_i^{(f)} - \Delta_{ij}^{ax}}{\tau} \exp\left(1 - \frac{t - t_i^{(f)} - \Delta_{ij}^{ax}}{\tau}\right) H\left(t - t_i^{(f)} - \Delta_{ij}^{ax}\right) \quad (3.25)$$

avec τ_m , τ_s et τ sont des constantes de temps, Δ_{ij}^{ax} est le délai de transmission axonale, et $H(\cdot)$ est la fonction de Heaviside. La Figure 3.16 est un exemple de l'équation (Eq. 3.25). Malgré les deux équations ont été largement utilisés pour simuler le PPS, dans une approche plus formelle et biologiques, le noyau ε_{ij} devrait être également une fonction du dernier tir du neurone j . Le PPS dépend de l'état du potentiel de membrane, qui, en son tour, dépend de \hat{t}_j . Le noyau $\varepsilon_{ij}(t - \hat{t}_j, t - t_i^{(f)} - \Delta_{ij}^{ax})$ peut être interprétée comme l'évolution dans le temps d'un potentiel post-synaptique évoquée par le tir d'un neurone pré-synaptique à $t_i^{(f)}$.

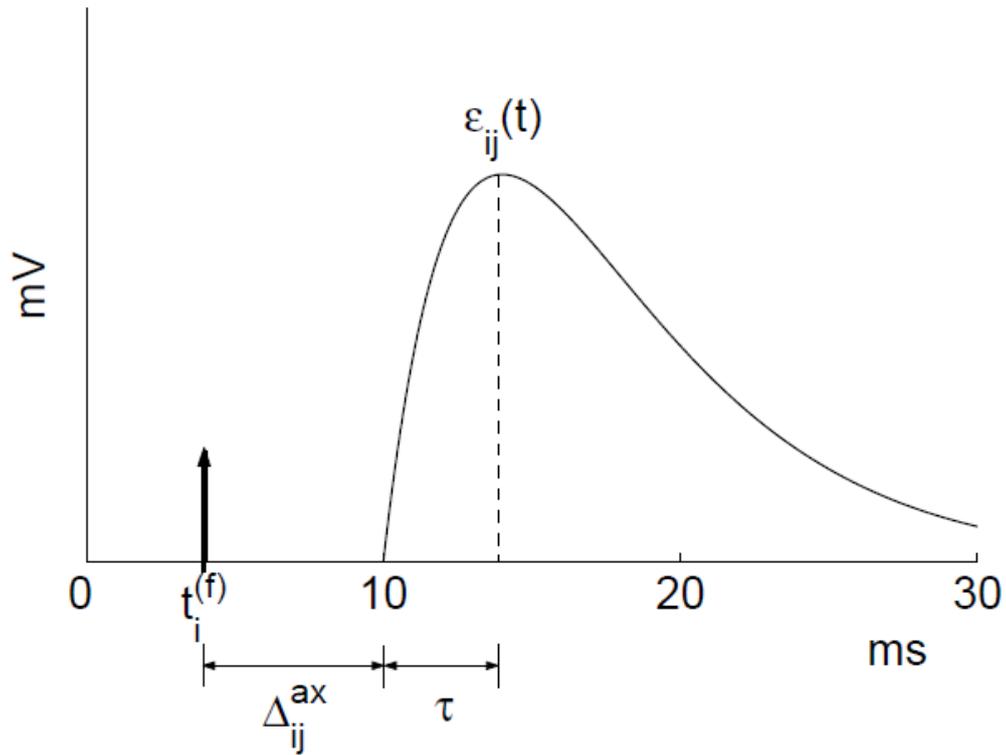


Fig. 3.16 Exemple d'un PPSE utilisant l'équation 3.25, avec $t_i^{(f)}=4\text{ms}$, $\Delta_{ij}^{ax}=6\text{ms}$ et $\tau=4\text{ms}$.

Le potentiel de membrane u_j d'un neurone i est alors donné par l'équation :

$$u_j = \eta(t - \hat{t}_j) + \sum_i w_{ij} \sum_{t_i^{(f)}} \varepsilon_{ij}(t - \hat{t}_j, t - t_i^{(f)}) + \int_0^\infty k(t - \hat{t}_j, s) I^{ext}(t - s) ds \quad (3.26)$$

Un PA est émis par le neurone j à la date t lorsque :

$$u_j(t) = \vartheta(t - \hat{t}_j) \text{ et } \frac{du_j(t)}{dt} > 0 \quad (3.27)$$

Le terme $\eta(t - \hat{t}_j)$ décrit le comportement du neurone lors de l'émission d'un PA et durant la période réfractaire qui suit. \hat{t}_j est la date du dernier PA émis par le neurone j . w_{ij} est le poids de la synapse entre un neurone pré-synaptique i et le neurone post-synaptique j . L'impact de l'arrivée d'un PA d'un neurone i sur le neurone j est décrit par le terme $\varepsilon(t - \hat{t}_j, t - t_i^{(f)})$ et dépend de la date du dernier PA post-synaptique \hat{t}_j et de celle d'un ou plusieurs PA pré-synaptiques $t_i^{(f)}$. Dans l'équation (Eq. 3.26) la double somme est effectuée sur tous

les neurones pré-synaptiques et sur tous leurs PA passés. Le dernier terme correspond à l'impact d'un éventuel courant extérieur appliqué au neurone. k décrit la forme d'une impulsion appliquée au neurone. Le seuil ϑ au-dessus duquel un PA est émis peut également être variable et dépendre de \hat{t}_i ; il s'exprime alors selon le terme $\vartheta(t - \hat{t}_i)$. La dépendance temporelle du seuil peut être évitée en fixant un seuil constant et en effectuant un changement de variable dans le terme η .

La [Figure 3.17](#) fournit un exemple d'évolution du potentiel de membrane d'un neurone n_j que l'on peut obtenir grâce au modèle SRM.

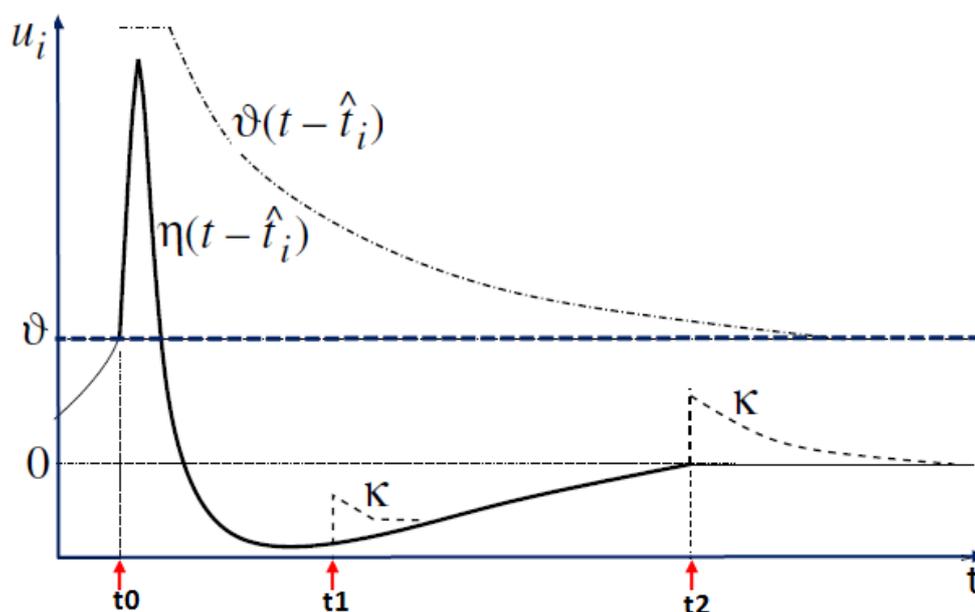


Fig. 3.17 Comportement du modèle SRM. Le neurone j émet un PA à $t = t_0 = \hat{t}_j$. $\eta(t - \hat{t}_j)$ décrit l'évolution du potentiel de membrane après un PA. k décrit l'impact d'un PPS sur le potentiel de membrane. Cet impact dépend du temps écoulé depuis le dernier PA ($t - \hat{t}_j$). ϑ décrit l'évolution du seuil après l'émission d'un PA. (Gerstner et Kistler, 2002a).

Le modèle ainsi défini permet d'obtenir de nombreux comportements neuronaux. Le modèle LIF est alors un cas particulier du modèle SRM. Le modèle de Hodgkin et Huxley peut être reproduit avec une précision remarquable (Kistler et al., 1997). Cependant, le modèle SRM reste complexe dans sa forme complète du

fait qu'il est nécessaire de prendre en compte tous les impulsions précédemment reçus par le neurone (Eq. 3.26). Les auteurs proposent alors une variante beaucoup plus légère : le SRM₀. Dans ce modèle, la forme des impacts ne dépend plus du neurone pré-synaptique impliqué ($\varepsilon_{ij} \rightarrow \varepsilon$). Ce modèle ne tient plus compte de la date du dernier PA post-synaptique dans la forme des PPS (terme $\varepsilon(t - \hat{t}_j, t - t_i^{(f)}) \rightarrow \varepsilon(t - t_i^{(f)})$) et dans la forme des impulsions de courant extérieur (terme $k(t - \hat{t}_j, t - t_i^{(f)}) \rightarrow k(t - t_i^{(f)})$). Le terme η représentant la forme du PA peut être exprimé comme un Dirac suivi d'une réinitialisation à un potentiel inférieur au potentiel de repos puis d'un retour vers le potentiel de repos avec une constante de temps τ_r .

$$\eta(t - \hat{t}_i) = \delta(t - \hat{t}_i) - \eta_0 e^{\left(-\frac{t - \hat{t}_i}{\tau_r}\right)} \text{ avec } \eta_0 > 0 \quad (3.28)$$

Ceci exprime l'hyperpolarisation post potentiel d'action. L'équation (Eq. 3.26) devient alors :

$$u_j = \eta(t - \hat{t}_j) + \sum_i w_{ij} \sum_{t_i^{(f)}} \varepsilon_{ij}(t - t_i^{(f)}) + \int_0^\infty k(s) I^{ext}(t - s) ds \quad (3.29)$$

Du fait de sa simplicité pour le calcul du potentiel de membrane (u_i), le modèle SRM₀ est un très bon candidat pour les simulations de grands réseaux. Il permet, comme le modèle SRM, d'approcher le modèle LIF (Gerstner et al., 1996b). Le modèle SRM définit en fait le cadre d'un modèle de neurone. Ce sont surtout les fonctions choisies pour la forme du PA, du courant d'entrée et des réponses synaptiques qui détermineront la complexité réelle du modèle.

3.5 Architecture des réseaux de neurones impulsionnels

L'étude du fonctionnement des neurones biologiques permet de définir des modèles de neurones. Mais lorsque les neurones se trouvent interconnectés en réseaux, des phénomènes (non définis explicitement par les modèles de neurones) émergent tels que des synchronisations transitoires, des oscillations qui sont autant d'interactions entre populations. Depuis le début des années 90, de nombreux modèles de réseaux, basés sur des modèles de neurones non-impulsionnels, ont été développés et étudiés. Ces modèles de réseaux peuvent s'appliquer aux neurones impulsionnels. D'une manière générale on distingue les

réseaux «feed-forward» des réseaux récurrents et la plupart des modélisations d'inspiration biologique sont basées sur des réseaux récurrents.

3.5.1 Réseaux feed-forward

Cette topologie est introduite par Gerstner (Gerstner et al., 1996a) et utilisé dans (Natschläger et Ruf, 1998 ; Bohte et al., 2002a, b ; Adibi et al., 2005 ; Meftah et al. 2008 ; Meftah et al. 2010a, b). C'est une architecture de réseau feed-forward classique, dont chaque connexion se compose de plusieurs liaisons synaptiques. Chacune d'elles avec des poids différents, des retards et des constantes de temps synaptique. Un neurone post-synaptique est connecté à des neurones pré-synaptiques par plusieurs liaisons et des paramètres différents (Figure 3.18).

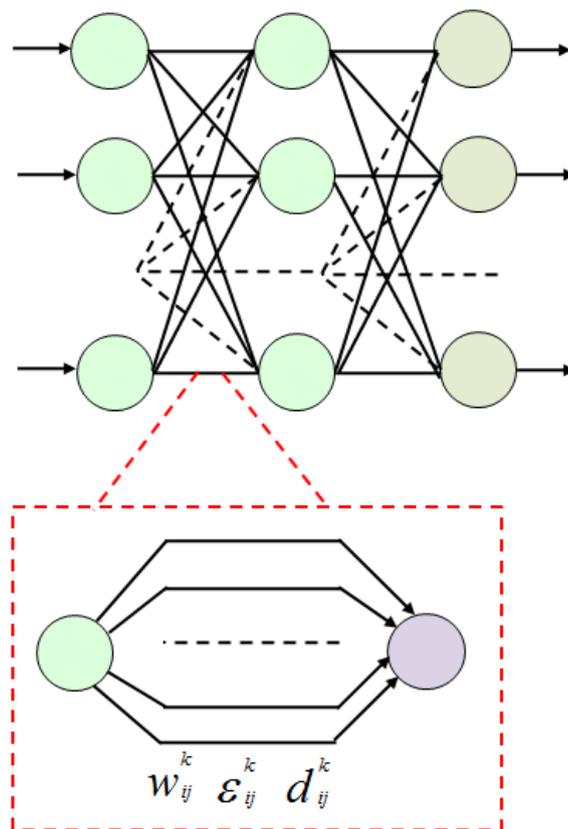


Fig. 3.18 Architecture d'un réseau feed-forward avec multiples synapses.

3.5.2 Réseaux récurrents

Les réseaux récurrents sont bien adaptés à des tâches de modélisation, d'analyse et de contrôle d'objets dynamiques. Les réseaux de neurones

impulsionnels récurrents sont souvent considérés comme des modèles pour étudier la théorie de l'activité neuronale (Gerstner et Kistler, 2002a).

Dans de nombreuses régions du cerveau, les neurones sont structurés en zones de réseaux récurrents peu connectés, dites microcircuits neuronaux (Markram, 2002). Les modèles impulsionnels des microcircuits neuronaux, leurs propriétés et leurs capacités de calcul ont été étudiés dans (Legenstein et al., 2003 ; Maass et al., 2004).

3.5.3 Réseaux hybrides

3.5.3.1 Chaîne de décharge synchrone (Synfire chain)

L'apprentissage humain consiste souvent à associer deux événements, ou en pointant un signal et une action ultérieure dans une relation de cause à effet. Les événements sont souvent séparés dans le temps, mais, néanmoins, les humains peuvent les lier, ce qui leur permet de prédire avec précision le bon moment pour une action particulière. La chaîne de décharge synchrone (Synfire chain) (Abeles, 1982) est considérée comme un mécanisme possible pour représenter les relations entre ces événements retardés.

La chaîne de décharge synchrone est une architecture multi-couches, sous forme d'une chaîne, dans lequel l'activité impulsionnelle peut se propager comme une onde synchrone des décharges neuronales (un paquet d'impulsions) d'une couche (sous-population) aux autres couches qui se suivent (Figure 3.19). Cette définition suggère l'architecture feed-forward. Toutefois, les sous-populations peuvent contenir des connexions récurrentes.

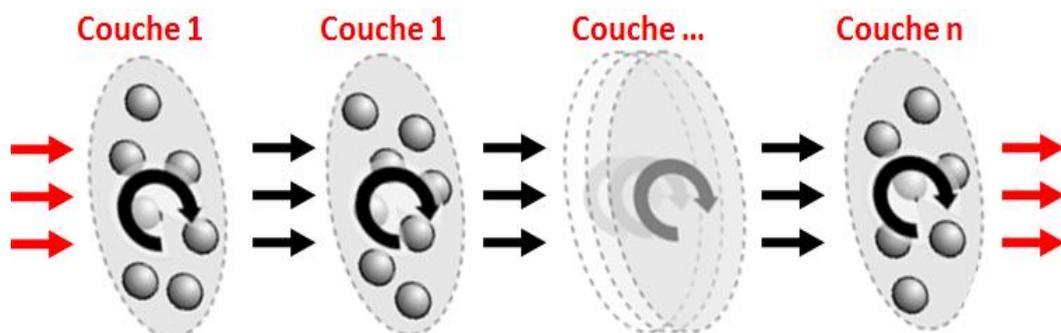


Fig. 3.19 Architecture multi-couches de la chaîne de décharge synchrone.

3.5.3.2 Machine à états liquides

Maass et ses collègues (Maass et al., 2004 ; Maass et al., 2002 ; Natschlaeger et al., 2002) ont mis en place un nouveau cadre conceptuel, connu sous le nom machine à états liquides (Liquid State Machine (LSM)), qui tire parti des propriétés des réseaux récurrents, tout en évitant les difficultés liées à leur apprentissage.

Les machines à états liquides sont des systèmes dynamiques astables dont on peut exploiter la puissance computationnelle en entraînant un perceptron ou un neurone en sortie; sous certaines conditions sur les connexions synaptiques dynamiques. Un réseau neuronal récurrent à pulses de grande taille en est une bonne approximation.

Pour calculer n'importe quelle fonction-objectif sur des entrées, une LSM tire parti d'un ensemble de nombreuses fonctions sur ces entrées générées par un filtre adéquat, et une fonction de sortie combine ensuite « simplement » ces fonctions, sans devoir connaître le codage de l'information au niveau du filtre (Figure 3.20). Ainsi une LSM peut être vue comme un « calculateur universel » en temps réel. Le terme « états liquides » réfère au fait que le système ne fonctionne pas avec des états stables mais de manière dynamique.

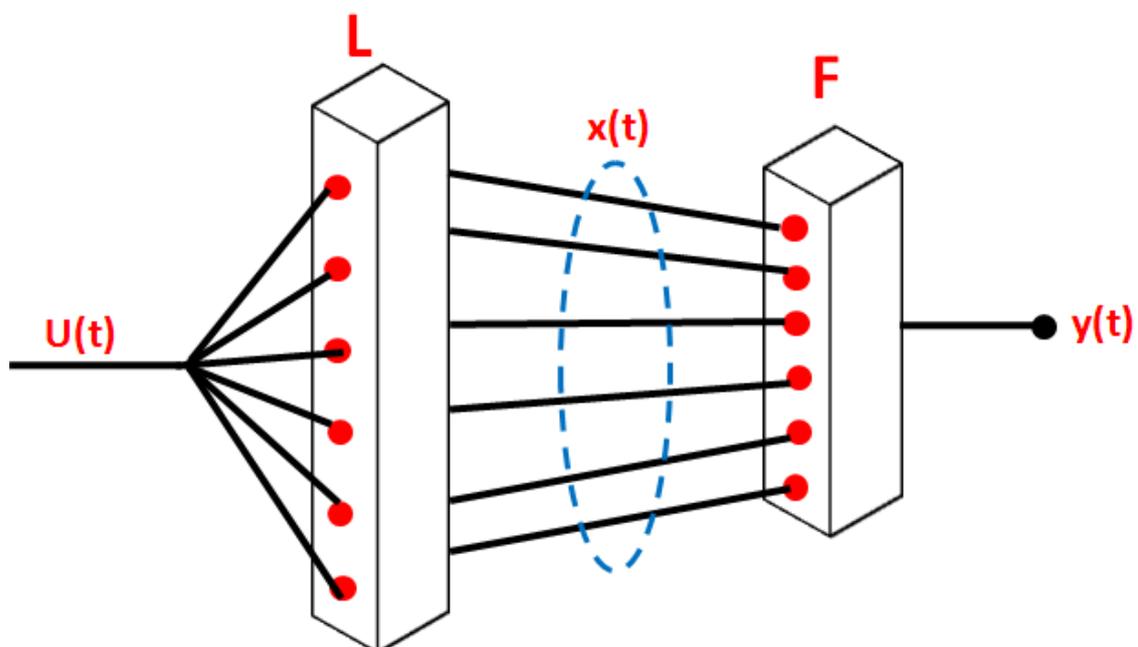


Fig. 3.20 Schéma général d'une LSM à une entrée et une sortie (Maass et al., 2002).

3.6 Quelques propriétés des réseaux de neurones impulsionnels

3.6.1 Réponses synaptiques

Dans de nombreuses simulations, la forme du PA est réduite à une impulsion et les courants d'entrée appliqués ont des formes simples. Pour tous les modèles de neurones cités, lorsque ceux-ci sont placés dans un réseau, les impacts des PA émis sur les membranes post-synaptiques jouent un rôle important. Enrichir les modèles de réponses synaptiques est un autre bon moyen de se rapprocher du fonctionnement biologique. De plus, la mise en place, à une échelle locale, de règles de fonctionnement, même simples, permet l'émergence de comportements complexes à l'échelle d'un réseau.

D'un point de vue computationnel, dans un réseau de neurones en fonctionnement normal, les fonctions modélisant les réponses synaptiques (ε) sont plus utilisées que les équations modélisant la forme des PA et de la période réfractaire (η) (puisque chaque PA génère un grand nombre de PPS) ce qui implique que la complexité des fonctions mises en place pour modéliser les réponses synaptiques a plus d'impact sur la charge computationnelle requise pour la simulation.

La forme de réponse la plus simple ([Figure 3.21](#)) est représentée par un créneau constitué d'une impulsion à la date d'impact t_0 du PPS, d'amplitude donnée g et maintenue pendant une durée τ . Ce modèle a permis à Maass de déterminer la puissance computationnelle d'un neurone impulsionnel simple ([Maass et al., 2004](#)). On peut utiliser une impulsion instantanée à t_0 , suivie d'une descente linéaire ou exponentielle. Des réponses synaptiques, plus proches des réponses biologiques, basées sur une ou plusieurs exponentielles sont également couramment utilisées. À $t_0 = 0$, l'impact génère une montée exponentielle, d'amplitude g suivie d'une descente exponentielle avec une constante de temps τ . Si les constantes de temps de la montée et de la descente de la réponse sont les mêmes ([Figure 3.22, courbe du haut](#)), alors une fonction α peut être utilisée :

$$\alpha(t) = \frac{t}{\tau} e^{(1-\frac{t}{\tau})} \quad (3.30)$$

Dans le cas où les constantes de temps sont différentes, on utilise fréquemment une forme en double exponentielle selon :

$$\phi(t) = \frac{\exp\left(-\frac{t}{\tau_1}\right) - \exp\left(-\frac{t}{\tau_2}\right)}{\tau_1 - \tau_2} \quad (3.31)$$

Dans les modèles de neurones présentés, les réponses synaptiques peuvent être introduites via une conductance synaptique ou un courant synaptique exprimées grâce aux équations (Eq. 3.30) ou (Eq. 3.31). Il est également possible de faire varier les constantes de temps (τ) en fonction du temps ou du potentiel de membrane, de même que les amplitudes : variation de conductance (g) ou de courant (I) dans les réponses afin de modéliser les phénomènes d'adaptation observés dans les synapses biologiques (Brette, 2007).

Les conductances synaptiques unitaires peuvent être approchées à l'aide d'une seule conductance, de deux conductances distinctes représentant les influences excitatrices et inhibitrices avec des dynamiques différentes, ou encore à l'aide d'une conductance par synapse ce qui alourdit considérablement le coût computationnel et l'espace mémoire nécessaire. En contrepartie, utiliser une conductance spécifique de chaque synapse donne la possibilité d'observer des dynamiques complexes au niveau de synapses individuelles et augmente ainsi la puissance computationnelle de tels réseaux.

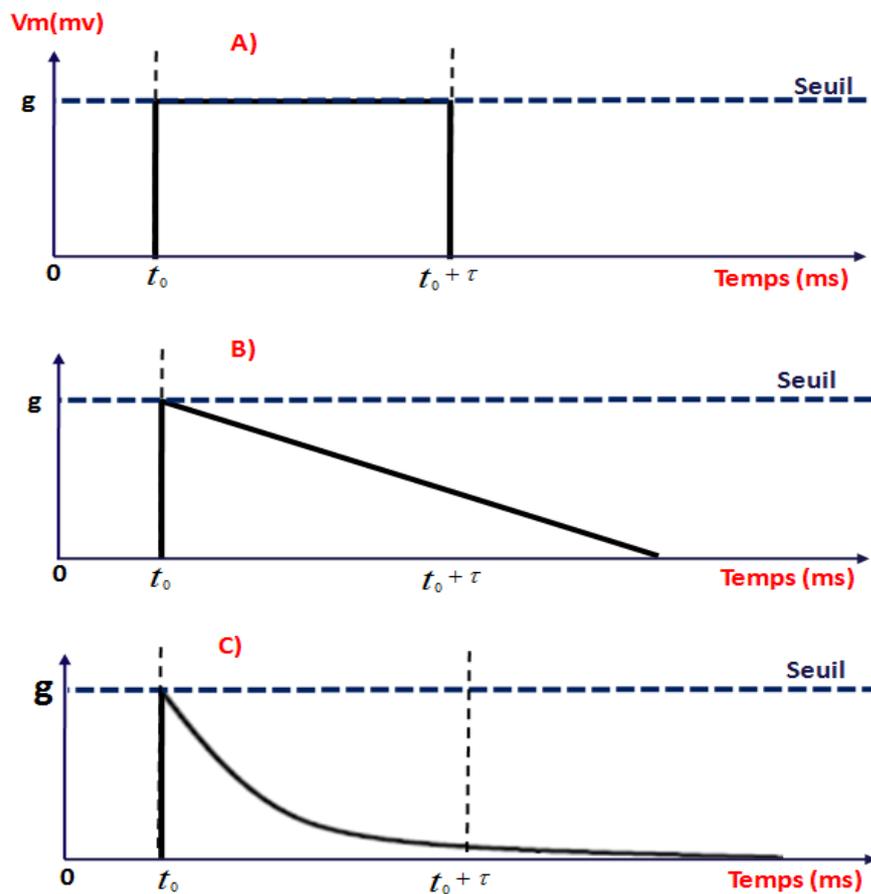
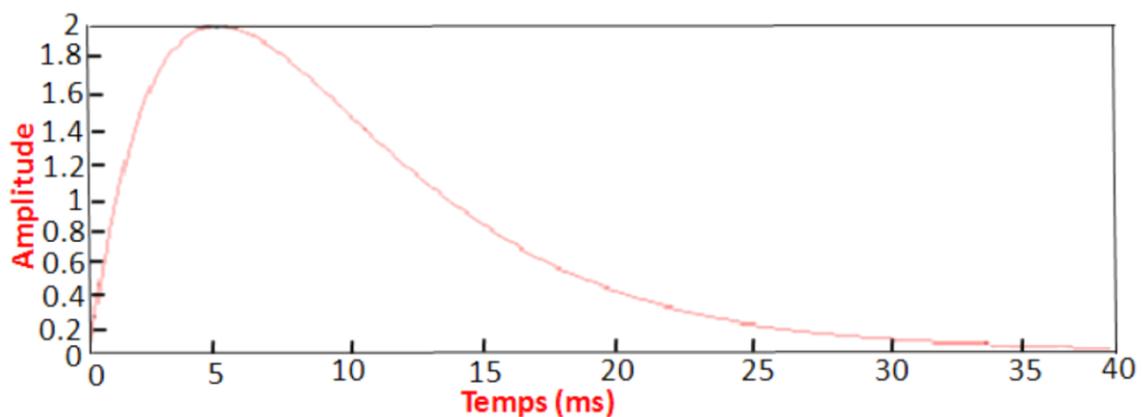


Fig. 3.21 Trois formes de réponse synaptique les plus simples. Les courbes correspondent à la réponse synaptique à l'arrivée d'un PPS, où g représente l'amplitude de la dépolarisation et τ la constante de temps. **A)** Impulsion en créneau. **B)** Impulsion instantanée suivie d'une fuite linéaire. **C)** Impulsion instantanée suivie d'une fuite exponentielle.



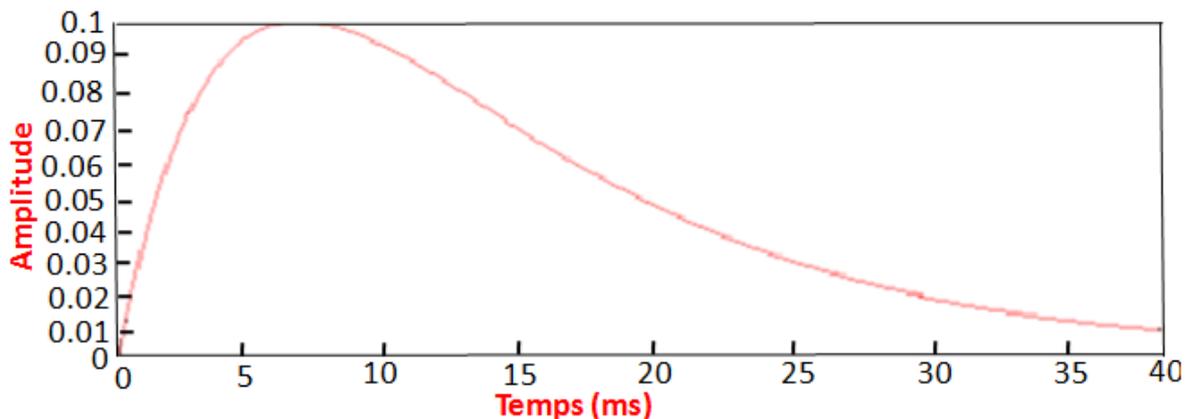


Fig. 3.22 Formes de réponses synaptiques en « fonction α » avec une constante de temps $\tau_1 = 5$ (courbe du haut) ou double exponentielle avec les constantes de temps $\tau_1 = 5$ et $\tau_2 = 10$ (courbe du bas).

3.6.2 Délais de transmission dans les réseaux de neurones impulsionnels

La modulation des réponses synaptiques est un des outils des modélisateurs qui étudient les réseaux de neurones impulsionnels. Source d'une grande richesse de comportements, elle implique cependant un coût computationnel important, c'est pourquoi un nombre minimum de paramètres est souvent recherché afin d'exploiter la puissance de calcul pour agrandir la taille des réseaux étudiés plus que pour affiner le comportement des neurones.

La richesse des comportements neuronaux est également due à la temporalité des émissions de PA au sein des réseaux de neurones. La temporalité des émissions de PA dépend étroitement des délais de transmission des PA entre les neurones. Dans le cerveau, la transmission de l'influx nerveux d'un neurone vers un autre n'est pas instantanée. Le délai temporel entre l'émission d'un PA par un neurone et la réception de celui-ci par les différents neurones cibles est variable. Ce délai inclut la propagation du signal dans l'axone, le franchissement de la synapse et la propagation dans la dendrite jusqu'au soma du neurone post-synaptique. En moyenne, le délai de transmission d'un PA entre deux neurones est de l'ordre de la milliseconde même s'il peut atteindre plus de 40 millisecondes (Swadlow, 1988). L'intégration temporelle réalisée par les neurones biologiques, ainsi que par les modèles de neurones impulsionnels, implique que l'ordre chronologique des impacts des PPS sur la membrane post-synaptique influence l'instant d'émission d'un PA.

On modélise les délais de transmission à l'aide d'un vecteur contenant un délai par connexion. L'existence et les bases biologiques des modifications des délais en fonction des entrées ne sont pas encore bien claires. Cependant des résultats récents (Bakkum et al., 2008) montrent que des modifications s'opèrent sur les délais axonaux en fonction du pattern de stimulation. Certains travaux ont également montré qu'il est possible de faire varier les délais en s'appuyant sur la variation des poids synaptiques (Senn et al., 2002).

3.7 Apprentissage des réseaux de neurones impulsionnels

Deux grands types d'apprentissage sont à distinguer : apprentissage supervisé et apprentissage non supervisé. Dans le premier cas on dispose de la bonne réponse et on l'utilise pour mesurer le taux d'erreur dans la réponse du réseau, puis on modifie les connexions en fonction de cette erreur. Dans le deuxième cas le réseau explore l'espace des réponses à l'aide d'un algorithme stochastique et on cherche à renforcer les patrons d'activité similaires pour former des catégories de réponses pour les différents stimuli. Dans les deux cas, le seul moyen dont on dispose pour améliorer les réponses du système est le réglage du poids des connexions.

Plusieurs algorithmes d'apprentissage pour les réseaux de neurones impulsionnels ont été proposés. Hopfield (Hopfield, 1995) a adapté la méthode d'apprentissage des RBF pour les réseaux de neurones impulsionnels en fonction des instants de tirs. Ruf et Schmitt (Ruf et Schmitt, 1997) ont proposé une méthode basée sur le modèle d'apprentissage de Hebb qui modifie les poids de connexion en fonction de la différence de temps entre les impulsions pré-et post-synaptiques. Plus tard, Ruf et Schmitt (Ruf et Schmitt, 1998) ont proposé un schéma d'auto-organisation dans un réseau de neurones impulsionnels basé sur les événements de tirs impulsionnels. L'auto-organisation dans les réseaux de neurones impulsionnels a également été élaborée par Natschläger et Ruf (Natschläger et Ruf, 1998) et Bohte et al. (Bohte et al., 2002b). Des méthodes d'apprentissages non supervisées tels que LTP (Long Term Depression), LTD (Long Term Potentiation) et STDP (Spike Timing Dependent Plasticity) ont déjà été largement étudiées et décrites dans la littérature (Kepecs et al., 2002 ; Kistler, 2002). Belatreche et al. ont étudié la possibilité des stratégies évolutives pour l'apprentissage supervisé de réseaux de neurones impulsionnels (Belatreche et al., 2003). Ponulak a présenté une méthode d'apprentissage supervisée baptisée Resume (Ponulak et kasiński, 2010).

Nous présentons deux méthodes d'apprentissage supervisé et non supervisé de réseaux de neurones impulsionnels.

3.7.1 Apprentissage supervisé de réseaux de neurones impulsionnels basé sur le gradient

L'apprentissage classique supervisé de réseaux de neurones artificiels est généralement effectuée par la technique de descente du gradient. Cependant, la mise en œuvre de cette technique est impossible à cause de la nature discontinue des signaux. Des simplifications spécifiques pour adapter la descente du gradient aux réseaux de neurones impulsionnels doivent être mises en place.

Bohte et ses collègues (Bohte et al., 2000 ; Bohte et al., 2002b) ont présenté une de ces approches. Leur méthode, appelée SpikeProp, est analogue à l'algorithme de rétropropagation du gradient classique (Rumelhart et al., 1986a).

L'objectif du SpikeProp est d'apprendre un ensemble de tirs impulsionnels, dénoté t_j^d , d'un neurone j à partir d'un ensemble de patterns en entrée $S^{in}(t)$. Toutefois, il est supposé que chaque neurone dans un réseau simulé est autorisé à tirer une seule fois au cours d'un cycle de simulation unique.

L'algorithme SpikeProp a été établi à partir du modèle à réponse impulsionnelle (SRM) (cf. Section 3.4.4). Le potentiel de membrane du neurone j peut être défini comme:

$$u_{mj}(t) = \sum_{i \in N_j^{pre}} \sum_k w_{ij}^k \varepsilon(t - t_i^{out} - d_{ij}^k) \quad (3.32)$$

où N_j^{pre} représente l'ensemble de tous les neurones pré-synaptiques du neurone j . Le terme w_{ij}^k est le poids du terminal synaptique k de la liaison entre le neurone i et j . Le paramètre t_i^{out} est l'instant de tir du neurone i et d_{ij}^k est le délai synaptique et $\varepsilon(t) = \frac{t}{\tau} (1 - t/\tau)$. L'équation (Eq. 3.32) est valable pour une région où les changements de potentiel membranaire sont essentiellement déterminés par les potentiels post-synaptiques $\varepsilon(t)$.

La méthode d'apprentissage est basée explicitement sur l'évaluation du gradient de l'erreur $E = \frac{1}{2} \sum_j (t_j^{out} - t_j^d)^2$.

$$\frac{dE}{dw_{ij}^k} = \frac{dE}{dt_j} (t_j^{out}) \frac{dt_j}{dw_{ij}^k} (t_j^{out}) = \frac{dE}{dt_j} (t_j^{out}) \frac{dt_j}{du_{mj}} (t_j^{out}) \frac{du_{mj}}{dw_{ij}^k} (t_j^{out}) \quad (3.33)$$

Où t_j est exprimée en fonction de la membrane potentiel $u_{mj}(t)$ autour de $t = t_j^{out}$. Pour simplification de l'évaluation du gradient, il est assumé pour une petite région autour de t_j^{out} la fonction $u_{mj}(t)$ est traités come une fonction temporelle linéaire. Par conséquent, la dérivation locale de t_j est constante.

L'équation de la rétropropagation d'erreur dérivée pour un réseau entièrement connecté feedforward est :

$$\frac{dE}{dw_{ij}^k} = y_{ij}^k(t) \delta_j \quad (3.34)$$

Où y_{ij}^k est le signal propagé du neurone i à j à travers la connexion k .

Pour les neurones de la couche de sortie δ_j est :

$$\delta_j = \frac{-(t_j^{out} - t_j^d)}{\sum_{i \in N_j^{pre}} \sum_k w_{ij}^k \frac{dy_{ij}^k(t)}{dt}} \quad (3.35)$$

Et pour les neurones de la couche cachée δ_j est :

$$\delta_j = \frac{\sum_{i \in N_j^{post}} \delta_i \sum_k w_{ij}^k \frac{dy_{ij}^k(t)}{dt}}{\sum_{i \in N_j^{pre}} \sum_k w_{ij}^k \frac{dy_{ij}^k(t)}{dt}} \quad (3.36)$$

Finalement les poids sont modifiés par :

$$\Delta w_{ij}^k = -\eta \frac{dE}{dw_{ij}^k} = -\eta y_{ij}^k(t) \delta_j \quad (3.37)$$

Où η est le taux d'apprentissage. L'erreur est minimisée par changement du poids due à la valeur négative du gradient.

SpikeProp souffre de plusieurs faiblesses. Tout d'abord, la performance de l'algorithme d'apprentissage est très sensible à l'initialisation des paramètres. Si un neurone ne tire pas après l'initialisation, aucun apprentissage n'est possible pour le poids de ses connexions entrantes. En conséquence, le neurone ne sera jamais en

mesure de produire des impulsions. Deuxièmement, SpikeProp est applicable uniquement au codage par latence. Troisièmement, SpikeProp ne fonctionne que pour réseau impulsif où les neurones tirent qu'une seule fois dans le temps de simulation. Quatrièmement, SpikeProp a été développé pour l'apprentissage poids seulement. Pour remédier à ces faiblesses, quelques extensions de l'algorithme SpikeProp ont été proposées (Booij et Nguyen, 2005 ; Tiño et Mills, 2005 ; Schrauwen et Campenhout, 2004 ; Moore, 2002 ; Xin et Embrechts, 2001).

3.7.2 Apprentissage non supervisé de réseaux de neurones impulsif basé sur la STDP (Spike Timing Dependent Plasticity)

3.7.2.1 Plasticité synaptique : base cellulaire de la mémoire et l'apprentissage

La plasticité synaptique correspond aux modifications morphologiques, chimiques et fonctionnelles qui interviennent au cours du temps au niveau de la synapse. Les synapses sont des zones spécialisées dans la transmission nerveuse. Elles évoluent avec le temps, certaines disparaissent, d'autres se créent mais toutes se modifient et tendent, soit à renforcer, soit à affaiblir, la communication entre deux neurones.

La plasticité synaptique serait ainsi à la base des processus d'apprentissage et de mémorisation. L'acquisition de réflexes conditionnés responsables d'associations stimulus-réponse correspondrait à une diminution ou à une augmentation de certaines connexions synaptiques.

Mais comment se forment (et se déforment) les connexions entre neurones ? Selon Hebb (Hebb, 1949), si deux neurones reliés par une synapse sont à un instant donné simultanément actifs, cette synapse voit son efficacité de transmission accrue. La loi de Hebb constitue depuis un modèle élémentaire de mémoire invoqué pour rendre compte de la manière dont un réseau de neurones peut garder la trace de ses expériences passées.

Expérimentalement, des changements d'efficacité synaptique furent enregistrés dans différentes structures centrales aboutissant à une potentialisation (LTP) ou une dépression (LTD) à long-terme.

- À l'échelle de la seconde, on parle alors de «Short Term Potentiation STP» ou de «Short Term Depression STD» selon qu'il s'agit d'une augmentation ou une diminution de l'impact synaptique. À cette échelle, il s'agit

principalement de prendre en compte l'adaptation des synapses aux activités pré et post-synaptiques.

- À l'échelle de plusieurs heures, on parle alors de «Long Term Potentiation LTP» ou «Long Term Depression LTD». Dans ce cas les modifications induites dans les synapses subsistent plus longtemps et on considère alors qu'une forme de mémorisation est effectuée.

Alors que la plupart des études sur la plasticité à long terme (LTP et LTD) s'intéressaient avant tout à l'évolution de l'efficacité d'une synapse en fonction de fréquences d'excitation, l'idée de la STDP est que l'évolution de l'efficacité synaptique dépend des dates précises de décharge entre le neurone pré-synaptique et le neurone post-synaptique.

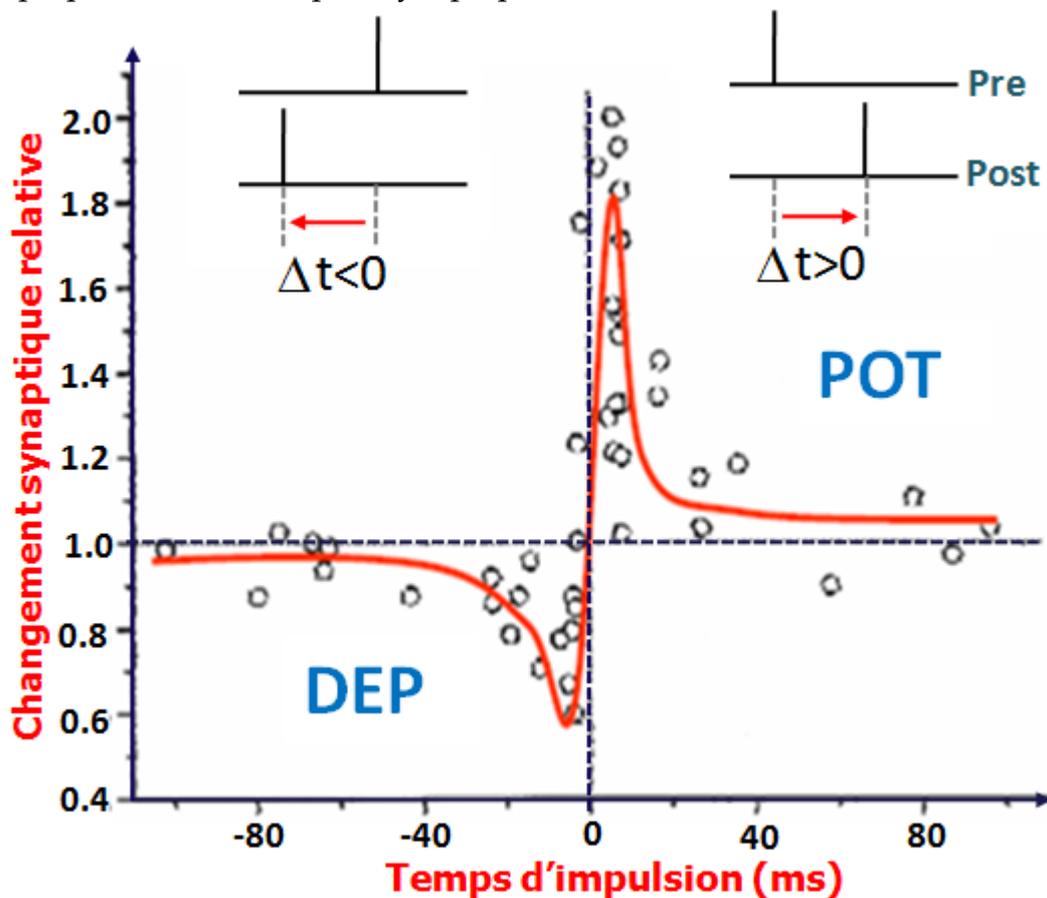


Fig. 3.23 Variation d'amplitude du courant post-synaptique (potentialisation et dépression) en fonction des temps d'arrivée des PPS par rapport au PA d'un neurone post-synaptique (Bi, 2002).

Des études (Song et Abbott, 2001 ; Bi et Poo, 1998 ; Markram et al., 1997) ont permis de prouver l'existence de la règle STDP, avec une potentialisation (POT) lorsque le neurone pré-synaptique décharge légèrement avant le neurone post-

synaptique et une dépression (DEP) dans le cas opposé (Figure 3.23). Cette forme correspond à une règle causale qui est cohérente avec le postulat initial de Hebb. Pour une revue récente sur le principe de la STDP voir l'article de Morrison (Morrison et al., 2007).

3.7.2.2 Modèles formels d'implémentations de la STDP

Il existe plusieurs variantes d'implémentation de la loi STDP, qui peuvent amener à des changements au niveau de la répartition des poids.

La première solution (Song et al., 2000) consiste à prendre un taux de modification constant quel que soit le poids de la synapse et égal à :

$$F(\Delta t) = \begin{cases} A_+ \exp\left(\frac{\Delta t}{\tau_+}\right) \text{ si } \Delta t < 0 \\ A_- \exp\left(\frac{\Delta t}{\tau_-}\right) \text{ si } \Delta t \geq 0 \end{cases} \quad (3.38)$$

Où Δt correspond à la différence de temps entre la décharge du neurone pré-synaptique et du neurone post-synaptique et où A_+ et A_- correspondent à la valeur maximale de modification, respectivement pour la LTP et la LTD.

D'après Song (Song et al., 2000), il est nécessaire, pour la stabilité, d'avoir $A_- \tau_- > A_+ \tau_+$ afin que des impulsions non corrélés entre deux neurones aient pour effet un affaiblissement des synapses entre ces deux neurones.

Dans cette implémentation de la STDP, il y a compétition entre les synapses car les entrées sont en « concurrence » pour l'impulsion post-synaptique. Lorsqu'une entrée commence à « diriger » le neurone post-synaptique, son poids augmente et elle le dirige encore plus. Les autres entrées sont donc moins corrélées, entraînant une réduction de leur poids.

Une autre variante (van Rossum et al., 2000) consiste, au lieu d'utiliser une modification constante (en pourcentage) du poids de la synapse, à diminuer la force relative de la LTP pour les synapses fortes, alors que celle de la LTD va rester constante. Une force relative de LTP inversement proportionnelle au poids est utilisée, ce qui revient à décrire l'évolution des synapses comme $\Delta w = w_p$ pour la LTP et $\Delta w = w_d$ pour la LTD avec

$$w_p = c_p \exp\left(\frac{-\Delta t}{\tau}\right) \quad (3.39)$$

$$w_d = -c_d w \exp\left(\frac{-\Delta t}{\tau}\right) \quad (3.40)$$

Où c_p et c_d sont les constantes de potentiation et de dépression. On voit donc que lors d'une dépression, la diminution du poids de la synapse est relative, alors que l'augmentation du poids de la synapse lors d'une potentialisation est absolue.

Dans ce modèle, il y a très peu de compétition entre les synapses : le fait que la potentialisation et la dépression des synapses soient relativement équilibrées limite la concurrence entre les entrées pour contrôler la décharge du neurone post-synaptique.

3.7.2.3 Mise en œuvre de la STDP

Parallèlement à la fonction modélisant la STDP proprement dite, se pose la question des couples d'impulsions auxquels l'appliquer. En l'occurrence, il est possible d'appliquer la STDP pour tous les l'impulsions émises par les neurones pré-synaptiques et post-synaptiques (Song et al., 2000 ; kempster et al., 1999) ou au contraire de ne considérer que les « plus proches voisins » (modèles dits « nearest-neighbour » (Izhikevich et al., 2004b ; Bi, 2002 ; van Rossum et al., 2000)), c'est-à-dire que seul le première impulsion pré-synaptique après la décharge d'un neurone post-synaptique entraînera une dépression, et que seul le première impulsion post-synaptique après la décharge d'un neurone pré-synaptique entraînera une potentialisation.

Si le modèle « all-to-all » peut sembler plus logique, l'avantage de ne considérer que les impulsions les plus proches est, d'un point de vue computationnel, qu'il n'est pas nécessaire de garder en mémoire tous les événements.

Cependant une différence d'implémentation va entraîner là encore des différences dans le comportement des réseaux, même si, à cause de la décroissance exponentielle, c'est l'impulsion la plus proche qui va entraîner l'essentiel de la modification synaptique.

Ainsi il a été montré (Izhikevich et al., 2003b) que des schémas de couplage différents entraînent une différence dans la dynamique des poids en fonction du taux de décharge post-synaptique. Morrison montre également que la règle « all-to-all » va réduire les efficacités des synapses menant aux neurones post-synaptiques dont on augmente le taux de décharge, et inversement pour les neurones dont on le réduit (Morrison et al., 2007). À l'inverse, la règle « nearest-

neighbour » va faire l'inverse. Ainsi le schéma « all-to-all » semble avoir un certain effet régulateur, contrairement au « nearest-neighbour ».

Kistler (Kistler et al., 2000) propose par ailleurs un modèle « all-to-all » d'application de la STDP qui, en plus de prendre en compte l'influence des paires d'impulsions, ajoute une légère contribution non-Hebbienne : l'efficacité synaptique est aussi modifiée lorsque seul le neurone pré-synaptique ou post-synaptique décharge (de manière potentiatrice lorsque c'est le neurone pré-synaptique et dépréciatrice lorsque c'est le neurone post-synaptique). Cette contribution qui correspond à des observations biologiques (Nelson et al., 1993 ; Alonso et al., 1990) permet de faire en sorte qu'un neurone post-synaptique qui ne décharge jamais voit ses connexions augmenter pour qu'il finisse par décharger, ou au contraire de réguler l'activité post-synaptique d'un neurone qui déchargerait « trop ».

3.8 Conclusion

Ce chapitre a 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. On se rapportera également à Rochel (Rochel, 2004) pour une liste des simulateurs disponibles pour implémenter de tels réseaux.

Les deux prochains chapitres mettront en relation l'aspect théorique des réseaux de neurones impulsionnels avec leur implémentation en traitement d'images fixes et animées, en abordant la segmentation, la détection de contour et le suivi d'objets.

CHAPITRE 4

Segmentation et détection de contours par réseaux de neurones impulsionnels

Sommaire

4.1 Introduction	117
4.2. Segmentation des images couleurs	117
4.2.1 Définition d'un classifieur	118
4.2.2 Classification à partir d'histogrammes	119
4.2.3 Classification par clustering	120
4.2.3.1 Classification de pixels non supervisée	120
4.2.3.2 Classification de pixels supervisée	120
4.2.4 Précautions liées à l'utilisation de la classification pixellaire	120
4.2.5 Mesure de distorsion	121
4.3 Bases d'images utilisées	122
4.4 Architecture, codage et apprentissage du réseau de neurones impulsionnels	124
4.4.1 Architecture du réseau des neurones impulsionnels	124
4.4.2 Codage des entrées	126
4.4.3 Apprentissage du réseau des neurones impulsionnels	127
4.5 Influence des différents paramètres du réseau de neurones impulsionnels sur la segmentation	129
4.5.1 Influence du nombre de classes sur le résultat de la segmentation	130
4.5.2 Influence du nombre de sous-synapses sur le résultat de la segmentation	132
4.5.3 Influence du nombre de champs récepteurs sur le résultat de la segmentation	134
4.5.4 Influence de la taille du corpus sur le résultat de la segmentation	136

4.5.5 Influence du seuil de déclenchement téta sur le résultat de la segmentation	137
4.5.6 Influence du paramètre Tau sur le résultat de la segmentation	139
4.5.7 Influence des paramètres d'apprentissage b et c sur le résultat de la segmentation.....	141
4.5.8 Meilleurs paramètres du réseau de neurones impulsionnels.....	144
4.6 Segmentation des cellules tumorales par réseaux de neurones impulsionnels	145
4.6.2 Résultats et évaluation de la segmentation cellulaire par les réseaux de neurones impulsionnels	147
4.7 Comparaison entre la segmentation par K-means, carte de Kohonen et les réseaux de neurones impulsionnels.....	149
4.8 Segmentation des images en contours	151
4.7 Conclusion	155

4.1 Introduction

Nous présentons dans ce chapitre la méthode de segmentation par classification pixellaires et la détection de contours. Nous présenterons également les précautions à prendre dans le choix de certains paramètres de la classification pixellaire ainsi que dans la construction des différentes bases d'images. Les paramètres du réseau de neurones impulsionnels pour la segmentation et la détection du contour sont ainsi exposés dans ce chapitre. Enfin nous illustrerons certains résultats de segmentation et de détection de contours sur des images issues de la base Berkeley et sur des images microscopiques de cellules de tumeurs bronchiques.

4.2. Segmentation par classification des images couleurs

L'emploi de la couleur en segmentation d'images est devenu très populaire depuis l'avènement du multimédia grand public. La couleur étant une information que l'on attache directement à une région, de nombreux auteurs ont pensé qu'une segmentation couleur rendrait ces régions plus significatives et apporterait des améliorations en terme de segmentation d'objets.

La segmentation consiste à partitionner une image en régions homogènes c'est-à-dire en un ensemble connexe de points de l'image ayant des propriétés communes. Cette étape primordiale dans le processus de traitement et d'analyse d'images couleur fait référence à l'extraction de primitives suivant deux approches que l'on nomme usuellement : approche région et approche contour. Nous rappelons brièvement ci-dessous les différences entre ces deux notions :

1. La notion de « contour » est associée à la recherche de discontinuités locales, de transitions entre différentes zones où certains paramètres comme la couleur, l'intensité, la texture ne sont pas identiques. Mais cette notion de contour ne permet pas d'obtenir directement une segmentation de l'image. En effet, il est assez rare d'obtenir directement après traitement, des contours entièrement connectés et il est alors indispensable de fermer ces contours pour pouvoir ensuite assimiler les régions aux zones à l'intérieur d'une ligne fermée;
2. La notion de « région » est associée à des regroupements de pixels connexes ayant des attributs communs (couleur, texture, etc). Ces méthodes aboutissent directement à une partition de l'image où chaque pixel est affecté à une région unique. Selon la définition de Zucker ([Zucker, 1976](#)), la segmentation d'une image A définie sur un ensemble de sites S est une partition de S en sous-ensembles disjoints non vides $\{R_i\}$ pour $i = \{1, 2, \dots, M\}$ appelés régions contenant des sites connexes au sens d'un prédicat et tels que:

$$S = \bigcup_{i=1}^M R_i \quad (4.1)$$

Nous présentons ci-dessous différentes techniques de classification pixellaire utilisées dans un cadre de segmentation d'images couleur. De nombreux travaux sur la classification de pixels qui consiste à affecter à chaque pixel de l'image une classe qui définit les régions à extraire de l'image, ont vu le jour (Cinque et al. 2004; Madhubanti et Amitava, 2008 ; Wang et Sun, 2010). Nous présentons dans cette section trois techniques de classification pixellaire permettant la construction de classes de pixels. Les deux premières techniques correspondent à une classification par histogrammes (1D et 3D) et la troisième correspond à une classification par clustering.

Une étude bibliographique détaillée sur les méthodes de segmentation est exposé dans les travaux de Cheng (Cheng et al., 2001) et de Freixenet (Freixenet et al., 2002).

4.2.1 Définition d'un classifieur

Un classifieur désigne tout outil de reconnaissance qui pour un vecteur reçu en entrée, donne des informations sur sa classe d'appartenance. Cet outil peut s'écrire sous la forme d'une fonction e , qui à l'aide de descripteurs d'un vecteur x à reconnaître, attribue à x la classe C_i parmi k classes possibles ($i \in [1, k]$).

Nous pouvons alors définir un classifieur par la relation suivante où l'ensemble $k = \{C_1, \dots, C_k\}$:

$$e: x \in \mathbb{R}^n \rightarrow K \quad (4.2)$$

Les réponses fournies par un classifieur e qu'il soit non-supervisé ou supervisé peuvent être classées en trois catégories :

1. Type classe : $e(x) = C_i$, indique que la classifieur a attribué la classe C_i à x ;
2. Type rang : $e(x) = [r_1^j, \dots, r_k^j]$ où r_i^j est le rang attribué à la classe i par le classifieur j ;
3. Type mesure : $e(x) = [m_1^j, \dots, m_k^j]$ où m_i^j est la mesure attribuée à la classe i par le classifieur j .

Dans tout ce chapitre, la réponse fournie directement en sortie de classification pixellaire sera considérée de type classe. Nous obtenons alors via une classification

de pixels, pour chaque pixel de l'image, sa classe d'appartenance parmi les classes disponibles.

4.2.2 Classification à partir d'histogrammes

La classification par histogrammes 1D est une technique couramment utilisée dans la segmentation d'images couleur car elle présente l'avantage de ne pas nécessiter de connaissance à priori sur l'image. Les méthodes d'analyse d'histogrammes se différencient par l'espace couleur choisi ou la composante couleur la plus significative, la détection des modes principaux et le critère d'arrêt de l'analyse récursive d'histogramme (Lambert et Macaire, 2000). Les deux approches utilisées pour cette analyse sont l'analyse récursive des histogrammes 1D des composantes couleurs et l'analyse hiérarchique des histogrammes 1D.

L'analyse récursive des histogrammes met en évidence au fil des itérations, les modes correspondant aux classes de pixels. La plupart des méthodes que l'on peut trouver dans la littérature supposent que la distribution des couleurs de chaque région est de type gaussienne. Ces méthodes se différencient par le choix des composantes couleurs utilisées, la détermination de la composante la plus représentative, les critères d'arrêt, ainsi que la méthode et les critères d'extraction des modes. Nous pouvons citer les travaux relatifs à ce type d'analyse de Celenk (Celenk, 1990).

La segmentation d'image couleur peut également se faire par analyse hiérarchique des histogrammes 1D (Arifin et Asano, 2006). A ce titre, nous présentons brièvement la méthode de Cheng (Cheng, 2000). L'auteur définit une mesure d'homogénéité prenant en compte l'information colorimétrique locale et globale et effectuent ensuite une analyse spatiale de l'image segmentée. Une division de chaque région en sous-régions de teinte uniforme est effectuée dans le cas où le nombre de modes de l'histogramme est supérieur à un. Une étape de fusion de sous-régions adjacentes possédant des attributs proches est ensuite effectuée de sorte à éviter toute sur-segmentation.

L'utilisation de classification par histogrammes 3D semble plus appropriée à la segmentation d'images couleur que celle par histogrammes 1D (Petrou et al., 1998). En effet, l'information couleur étant tridimensionnelle, une classe de pixels séparable dans un espace couleur à trois dimensions peut ne pas l'être par la projection selon une composante. Mais à contrario, la couleur d'un pixel étant généralement codée sur 2563 valeurs, l'histogramme 3D d'une image couleur requiert alors une place mémoire importante. Ceci explique que peu d'auteurs se soient intéressés à une analyse sur ce type de classification.

4.2.3 Classification par clustering

Les méthodes de classification par clustering présentent l'espace couleur en sous espaces homogènes selon un critère de ressemblance des couleurs de pixels. Ces méthodes analysent, dans l'espace couleur considéré, les nuages formés par les points associés aux pixels et tentent de retrouver k nuages correspondant aux k classes de pixels de l'image. Différents algorithmes de classification de pixels existent dans la littérature.

4.2.3.1 Classification de pixels non supervisée

La classification de pixels non supervisée appelée aussi classification de pixels sans apprentissage consiste à découper l'espace de représentation en zones homogènes selon un critère de vraisemblance entre les individus. Cette approche est utilisée pour effectuer une classification de pixels en aveugle c'est-à-dire sans connaissance a priori sur l'image et ne nécessite donc pas de phase d'apprentissage.

Nous pouvons donc citer des algorithmes de classification de pixels non-supervisés comme les k -moyennes proposée par Mac Queen ([Macqueen, 1967](#)), C -moyennes floues ([Lim et Lee, 1990](#)) ou encore le Mean shift ([Cheng, 1995](#)).

4.2.3.2 Classification de pixels supervisée

La classification de pixels supervisée appelée aussi classification de pixels avec apprentissage consiste à définir une fonction de discrimination effectuant un découpage de l'espace de représentation à partir d'une connaissance a priori de l'image. Ce type de classification nécessite la création d'une base d'apprentissage faisant intervenir une segmentation de référence. La taille de cette base d'apprentissage doit être choisie avec précaution.

Nous pouvons donc citer des algorithmes de classification de pixels supervisés comme celui de Bayes ([Cocquerez et Philipp, 1995](#)), les réseaux de neurones ([Dong et Xie, 2005](#) ; [Meftah et al., 2008](#)), les machines à supports de vecteurs (SVM) ([Wang et al., 2011](#)) ou encore les k plus proches voisins (k -PPV).

4.2.4 Précautions liées à l'utilisation de la classification pixellaire

La difficulté de la classification pixellaire réside en partie dans le choix du nombre de classes pour extraire de façon précise les différents objets jugés importants d'une image.

Une connaissance à priori de l'image ou de l'application est donc un facteur intéressant pour la détermination de ce paramètre. Il est toujours possible de choisir

un nombre de classes plus important et de fusionner ensuite certaines classes afin de revenir au nombre d'objets extraits désiré par une classification ascendante hiérarchique.

Une attention particulière devra également être apportée à la création de la base d'apprentissage principalement rencontrée dans le cadre d'une classification de pixels supervisée. En effet, ce type de classification nécessite deux types de bases : une base d'apprentissage et une base de test.

1. La base d'apprentissage est généralement construite à partir de plusieurs images contenant des objets d'une grande variabilité et segmentés manuellement par un ou plusieurs experts. Cette expertise permet d'attribuer une classe à chaque pixel de l'image de sorte que le système puisse effectuer un apprentissage de l'image soit en mesure d'associer de façon plus sûre une classe à un pixel non appris ;
2. La base de tests est généralement construite à partir de plusieurs images contenant des objets d'une grande variabilité. Cette base de tests est utilisée dans le but de tester les algorithmes de classification de pixels. Bien qu'elle soit utilisée dans le cadre de classification pixellaire non-supervisée, cette base est particulièrement adaptée à la classification de pixels supervisée puisqu'elle permet de tester les algorithmes sur des images qui n'ont pas été apprises par le système.

Il est alors conseillé pour cette démarche d'utiliser des images représentatives mais variées de sorte à ce que le système puisse apprendre tous les types d'objets sans en omettre et permettre ainsi d'obtenir de bons taux de reconnaissance sur des images inconnues.

4.2.5 Mesures de distorsion

La distorsion est l'erreur introduite par l'opération de clustering, due au fait qu'éventuellement l'image reconstruite n'est pas exactement identique à l'image originale. Les mesures de distorsion utilisées généralement sont :

- Erreur quadratique moyenne MSE (Mean Square Error)

$$MSE = \frac{1}{M \cdot N} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} [f(i, j) - f'(i, j)]^2 \quad (4.3)$$

- Erreur absolue moyenne MEA (Mean Absolute Error)

$$MEA = \frac{1}{M \cdot N} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} |f(i,j) - f'(i,j)| \quad (4.4)$$

- Erreur quadratique moyenne normalisée NMSE (Normalized Mean Square Error)

$$NMSE = \frac{\sum_{i=0}^{N-1} \sum_{j=0}^{M-1} [f(i,j) - f'(i,j)]^2}{\sum_{i=0}^{N-1} \sum_{j=0}^{M-1} [f(i,j)]^2} \quad (4.5)$$

- Rapport signal sur bruit crête PSNR (Peak Signal to Noise Ratio)

$$PSNR = 20 \log_2 \left(\frac{255}{\sqrt{MSE}} \right) \quad (4.6)$$

Où les images ont $N \times M$ pixels, $f(i,j)$ représente l'image originale et $f'(i,j)$ représente l'image reconstruite.

4.3 Bases d'images utilisées

Nous avons utilisé deux bases d'images pour les opérations de segmentation et de détection de contours. La première est la base d'images de Berkeley ([Martin et al., 2001](#)). Elle regroupe un grand nombre d'images et dispose de segmentation de référence produites par plusieurs experts. Cette base devient de plus en plus utilisée dans le domaine du traitement d'images (compression, segmentation, indexation) car elle peut permettre non seulement de tester des algorithmes par rapport à une segmentation de référence, mais aussi de pouvoir comparer les performances de méthodes de segmentation proposées dans la littérature sur des images identiques. La figure suivante présente quelques images de cette base.

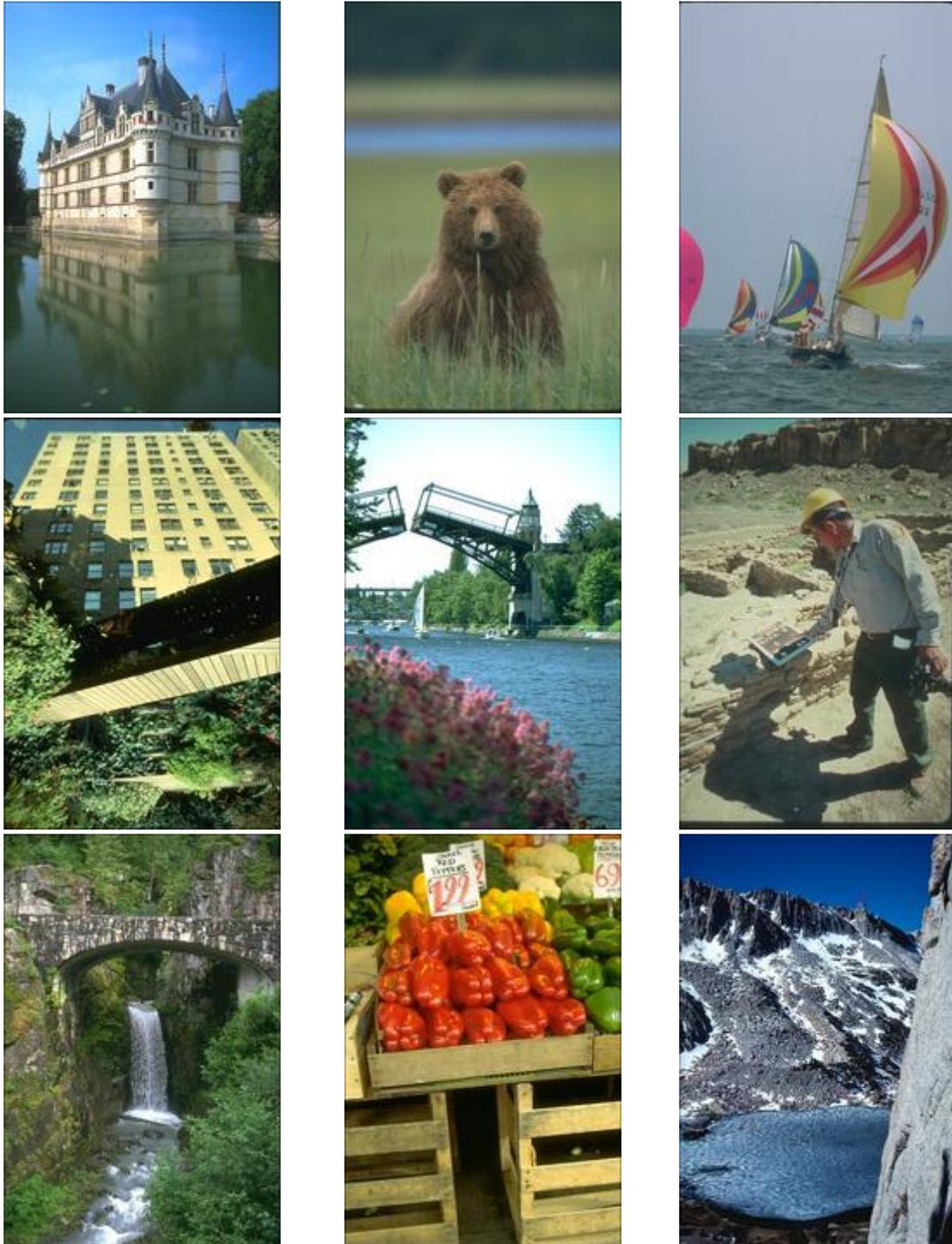


Fig. 4.1 Quelques images de la base de Berkeley.

La deuxième base se réfère à la base d'images microscopiques biomédicales¹ (Meurie et al., 2005). Les pixels de ces images doivent être classés dans l'une des trois catégories suivantes : fond, cytoplasme de la cellule et le noyau cellulaire. La figure suivante présente quelques images de cette base.

¹ <http://users.info.unicaen.fr/~lezoray/database.php>

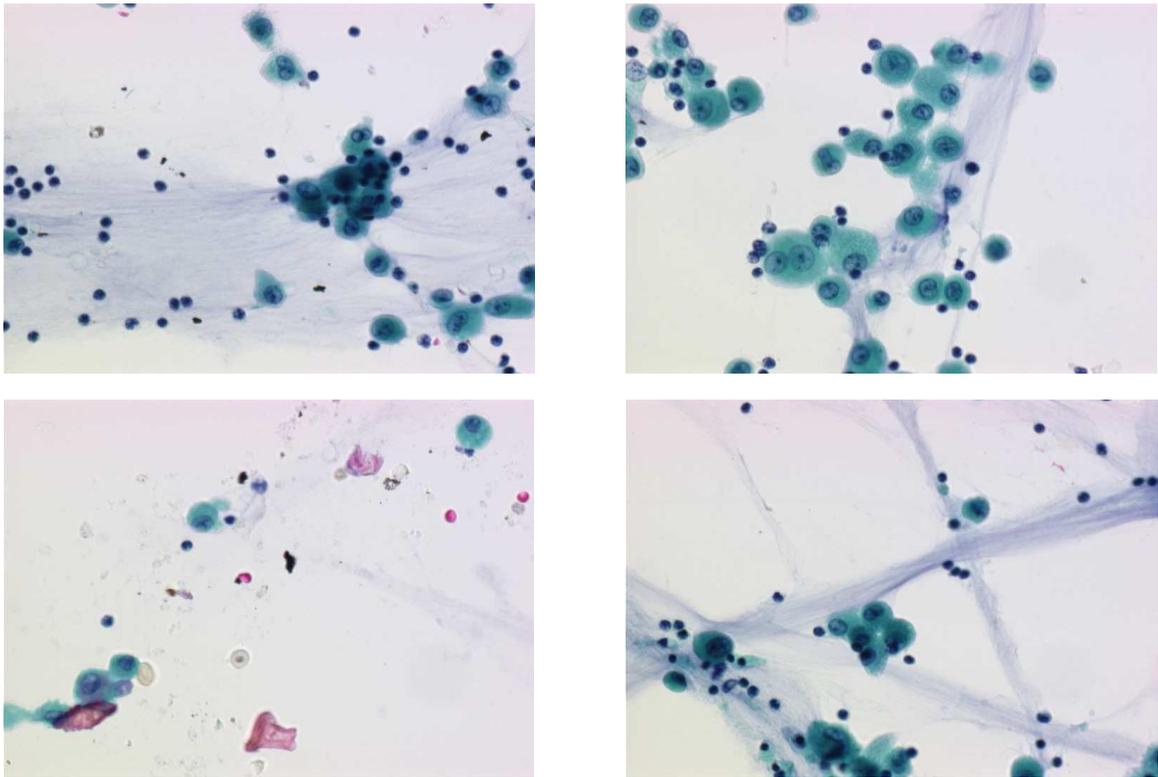


Fig. 4.2 Quelques images microscopiques biomédicales.

4.4 Architecture, codage et apprentissage du réseau de neurones impulsionnels

Avant de procéder à la mise en œuvre des réseaux de neurones impulsionnels en segmentation et en détection de contours, nous devons explorer trois questions importantes concernant le paramétrage de ce dernier. A savoir définir l'architecture du réseau, le codage de l'information en entrée et la méthode d'apprentissage.

4.4.1 Architecture du réseau des neurones impulsionnels

Le réseau impulsionnel est constitué de trois couches (Figure 4.3).

La première couche est constituée de trois neurones représentant les composantes colorimétriques RVB d'un pixel dans un espace à trois dimensions. La totalité des pixels de l'image ainsi représentées dans l'espace RVB constituent l'échantillon pour l'apprentissage du réseau. Lors de la phase d'apprentissage, les observations sont présentées séquentiellement une par une à l'entrée du réseau de manière aléatoire et sans remise.

La deuxième couche est constituée d'un ensemble de champs récepteurs effectuant l'opération de codage (cf. Section 4.4.2), transformant chaque entrée en un

ensemble d'impulsions. Ces impulsions sont injectées aux neurones de la couche de sortie. La liaison du neurone pré-synaptique de la deuxième couche au neurone post-synaptique de la couche de sortie est effectuée à travers un ensemble de sous synapses possédant chacun son délais de transmission d^k (cf. [Section 3.6.2](#)), variable dans un intervalle de 1ms, et son poids w_{ij}^k ([Figure 4.3.b](#)).

Chaque neurone de la couche de sortie présente une classe. Il est donc caractérisé par sa position et par son vecteur poids dans l'espace de représentation des observations. L'apprentissage du réseau consiste, à chaque présentation à l'itération t d'une observation à l'entrée du réseau, à sélectionner le neurone gagnant à l'aide d'une mesure de similarité et de mettre à jour les poids des neurones concernés, autrement dit celui dont le vecteur poids est le plus proche de cette observation. A la fin de cette étape, chaque neurone devient sensible à une zone de l'espace de représentation des observations et son vecteur poids converge vers le barycentre des observations présentes dans cette zone. Il existe des connexions latérales inhibitrices entre les neurones de la couche de sortie empêchant l'activation de deux neurones ou plus au même temps pour induire une compétition du type « le vainqueur prend tous » (cf. [Section 4.4.3](#)).

Le modèle mathématique des neurones qu'on a utilisé pour modéliser la réponse impulsionnelle est le SRM (Spike Response Model) (cf. [Section 3.4.4](#)).

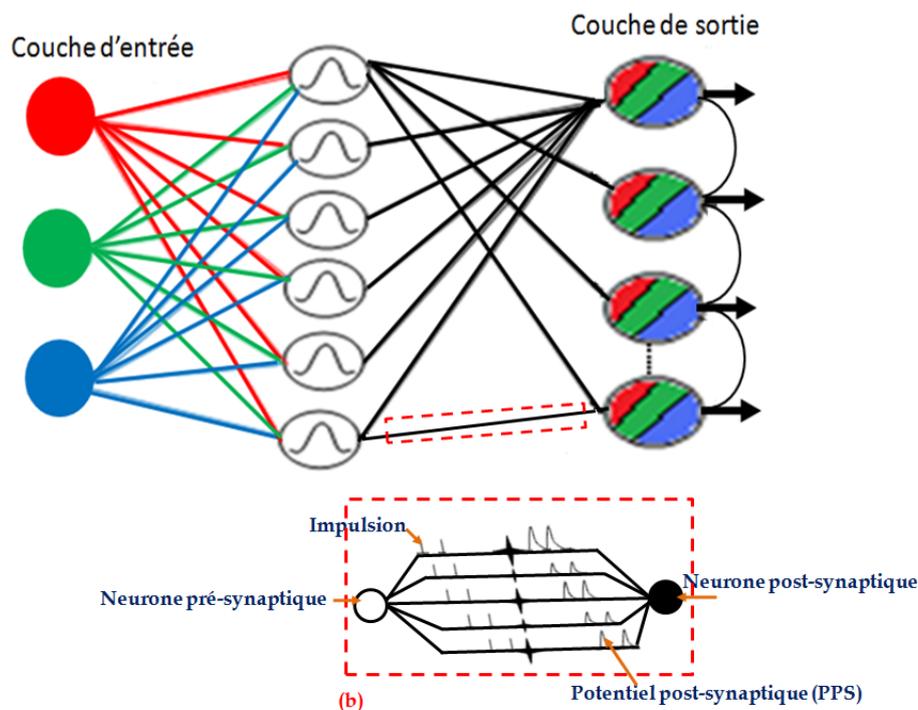


Fig. 4.3 Architecture du réseau de neurones impulsionnels.

4.4.2 Codage des entrées

Le codage impulsif est le processus de transformation des valeurs sensorielles en une représentation temporelle sous forme de train d'impulsions. Bohte et al. (Bohte et al., 2002a) ont présenté une méthode de codage par population de neurones. Chaque neurone d'entrée est modélisé par un champ récepteur local (RF). Pour une variable limitée dans un intervalle de [Max.. Min] un ensemble de m neurones récepteurs gaussiens sont utilisés. Le centre C_i et la largeur σ_i de chaque neurone i sont déterminés par les équations suivantes:

$$C_i = I_{min} + \left(\frac{2i - 3}{2}\right) \left(\frac{I_{max} - I_{min}}{m - 2}\right) \quad (4.7)$$

$$\sigma_i = \frac{1}{\gamma} \left(\frac{I_{max} - I_{min}}{m - 2}\right) \quad (4.8)$$

Où m est le nombre de champs récepteurs dans chaque population et une valeur de 1,5 est utilisé pour la variable γ . Pour chaque modèle d'entrée de dimension n , les résultats de codage sont dans une matrice $n \times m$ de valeurs comprises entre 0 et 1. Ces valeurs sont ensuite converties en retard temporels.

Pour la conversion des valeurs en entrée en temps d'impulsions, un seuil a été imposé sur la valeur d'activation. Un champ récepteur qui donne une valeur d'activation inférieure à ce seuil sera marqué comme non-mise à feu et le neurone d'entrée correspondant ne contribuera pas au potentiel post-synaptique.

Une illustration de ce schéma de codage est indiquée sur la [Figure 4.4](#), qui montre les temps d'impulsion résultant de l'encodage de la valeur réelle "0.3" en utilisant six champs récepteurs locaux. Dans cet exemple, en supposant que l'unité de temps est la milliseconde, la valeur 0.3 a été codée avec six neurones en retardant l'activation des neurones. Le neurone 1 donne 5.564 ms, le neurone 2 donne 1.287 ms, le neurone 3 donne 0.250 ms, le neurone 4 donne 3.783 ms et le neurone 5 donne 7.741 ms. Le neurone 6 ne se déclenche pas du tout, puisque le retard est supérieur au seuil 9 ms et reste dans la zone de non tir.

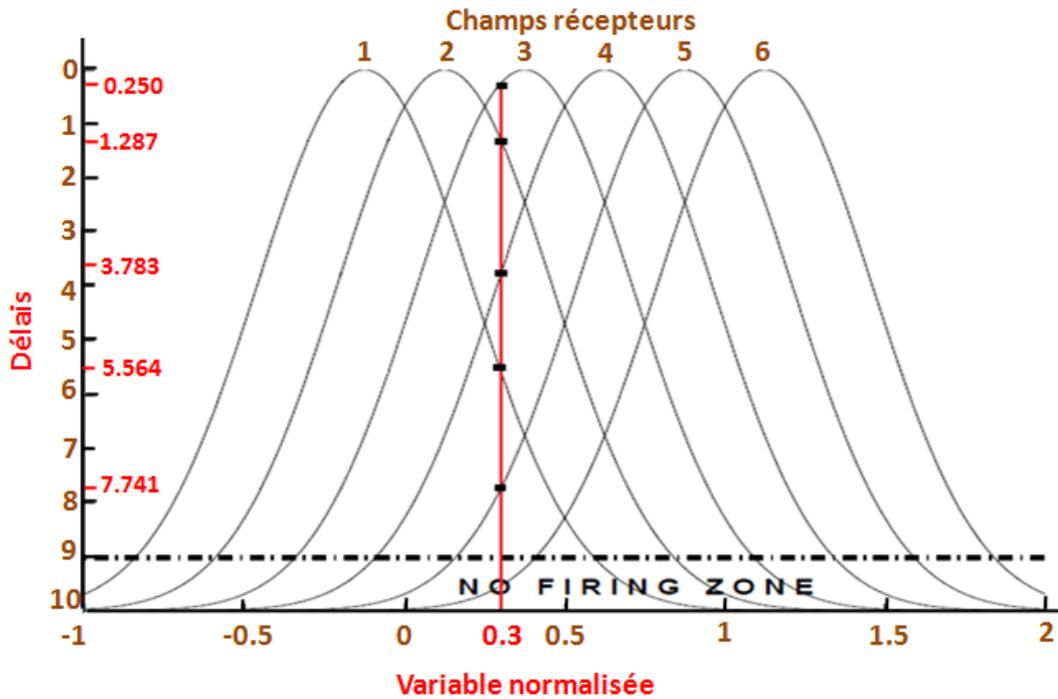


Fig. 4.4 Codage d'une valeur réelle.

4.4.3 Apprentissage du réseau des neurones impulsionnels

La capacité d'apprentissage d'un réseau de neurones impulsionnels réside dans l'ajustement des poids existants entre les différents neurones. L'approche implémentée est l'apprentissage Hebbien à travers l'algorithme du vainqueur prend tous (Winner Takes All) a été proposée pour la première fois par Rosenblatt (Rosenblatt, 1958).

Les réseaux WTA (Winner Takes All) consistent en un groupement de neurones impulsionnels qui entrent en concurrence mutuelle comme réponse à un stimulus. Les neurones possédant la plus forte réponse éliminent les autres et gagnent la compétition. La compétition est réalisée à travers des vecteurs de connectivité récursifs impliquant des connexions synaptiques inhibitrices et excitatrices. La compétition et la coopération entre les neurones font que l'état de chaque neurone dépend des activités des autres et non plus de sa propre stimulation.

Dans une couche de neurones régie par un mécanisme d'activation compétitive, seul un neurone unique doit être actif à tout moment donné en utilisant une règle d'apprentissage Hebbien $L(\Delta t)$. Pour cela, le réseau comporte, en plus des connexions entrantes qui aboutissent à toutes les unités du réseau, des connexions entre toutes les unités d'une même couche (Figure 4.3). Ces dernières connexions, dites « latérales », sont de nature inhibitrice et matérialisent la concurrence entre les neurones d'une même couche : un seul neurone est actif et inhibe tous les autres. Il a

été montré que, sous la condition que les poids de liens soient normalisés, c'est-à-dire que la somme des poids d'un neurone reste constante pendant la phase d'apprentissage et que tous les neurones ont un poids global équivalent, le neurone présentant la plus grande valeur d'activation sera le « vainqueur » (Kohonen, 1984). Une classe supplémentaire est créée dans le cas où plusieurs neurones en sortie possèdent la grande valeur d'activation, prenant comme valeurs d'intensité la couleur noire.

La fonction d'apprentissage (Figure 4.5) est de forme gaussienne et définie par l'équation (Eq. 4.8). Elle renforce la connexion entre le neurone i et j si $\Delta t_{ij} < \nu$ et affaiblit la connexion si $\Delta t_{ij} > \nu$ (Leibold et Hemmen, 2001 ; Gerstner et Kistler, 2002b).

$$\Delta w_{ij}^k = \eta L(\Delta t_{ij}) \quad (4.9)$$

$$L(\Delta t) = (1 + b) e^{\frac{(\Delta t - c)^2}{2(k-1)} - b} \quad (4.10)$$

Avec

$$k = 1 - \frac{\nu^2}{2 \ln \frac{b}{1+b}}$$

Où $L(\cdot)$ est la fonction d'apprentissage; η est le taux d'apprentissage; ν détermine la largeur de la fenêtre d'apprentissage; Δt est la différence entre l'arrivée de l'impulsion et l'activation du neurone j ; b détermine la mise à jour négative donnée à un neurone; c fixe le pic de la fonction d'apprentissage; w_{ij}^k est l'augmentation de la $k^{\text{ième}}$ connexion entre les neurones i et j . Les poids sont limités dans l'intervalle 0 à w_{max} (la valeur maximale que peut prendre un poids).

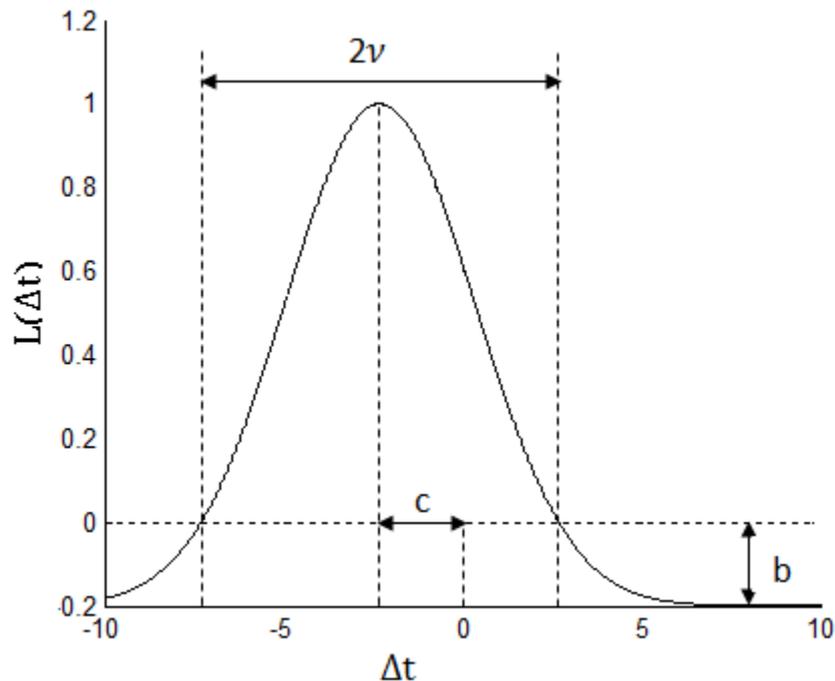


Fig. 4.5 Fonction d'apprentissage gaussienne avec $b = -0.2$, $c = -2.3$ et $\nu = 5$.
L'algorithme d'apprentissage se présente comme suit :

1. Initialisation : choix aléatoire des vecteurs poids.
2. Codage : codage de toutes les entrées en valeurs impulsives.
3. Présentation : tirage aléatoire d'une entrée x dans la base d'apprentissage afin de le présenter au réseau.
4. Réponse impulsive : calcul de la réponse impulsive de chaque entrée suivant l'équation (Eq. 3.25)
5. Recherche du neurone gagnant dont la réponse impulsive est maximale.
6. Modification des poids à ajuster: ajuster les poids du neurone gagnant suivant l'équation (Eq. 4.9).
7. Continuation : retour à l'étape 3 si le changement des poids demeure non négligeable.

4.5 Influence des différents paramètres du réseau de neurones impulsives sur la segmentation

Plusieurs expériences sont faites sur la base de Berkeley (Martin et al., 2001) pour montrer l'influence des différents paramètres du réseau de neurones impulsives (le nombre de classes, le nombre des champs récepteurs, le nombre de sous-synapses, la taille du corpus d'apprentissage, les paramètres d'apprentissage b et c , seuil de déclenchement τ et θ) sur le résultat de la segmentation. L'objectif est de trouver les meilleurs paramètres du réseau impulsif.

4.5.1 Influence du nombre de classes sur le résultat de la segmentation

Pour montrer l'influence du nombre de neurones de la couche de sortie (le nombre de classes) sur le résultat de la segmentation, nous avons fixé le nombre de sous-synapses de chaque connexion entre deux neurones à 8, le nombre de champs récepteurs à 12 pour chaque composante de couleur, le pas d'apprentissage à 0.35, la taille du corpus d'apprentissage est de 20% pris aléatoirement de l'image originale. Pour 3 itérations nous avons varié le nombre des neurones de la couche de sortie. Les résultats de segmentation obtenus sont montrés par la [Figure 4.6](#). La fidélité de l'image segmentée augmente en correspondance avec le nombre de classes en sortie [Figure 4.7](#).



Image originale



Segmentation avec 3 classes



Segmentation avec 6 classes



Segmentation avec 8 classes



Segmentation avec 10 classes



Segmentation avec 12 classes

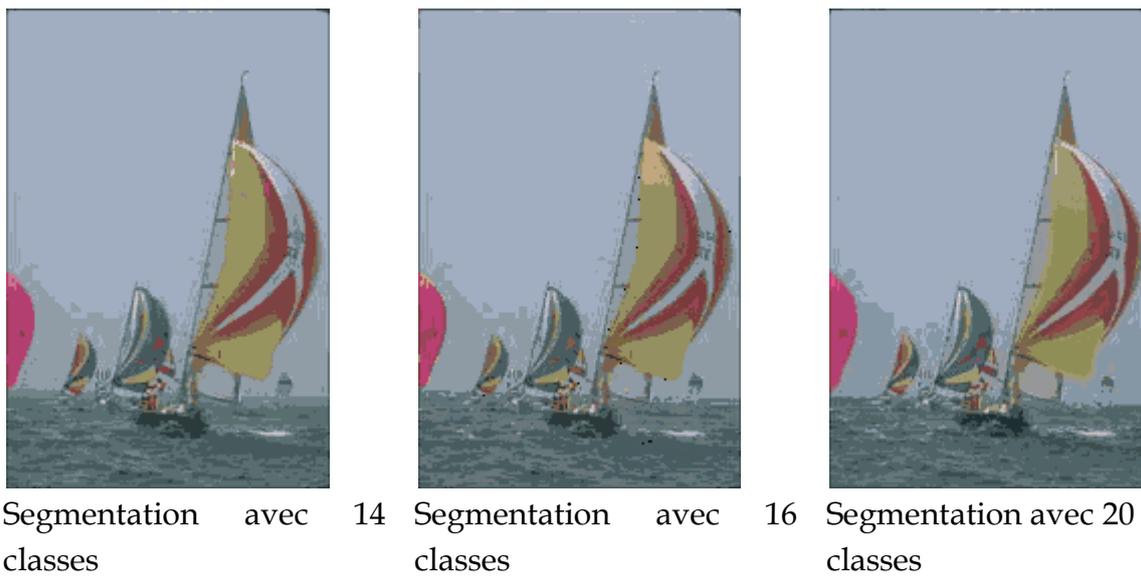


Fig. 4.6 Résultat de segmentation par variation du nombre de classes en sortie.

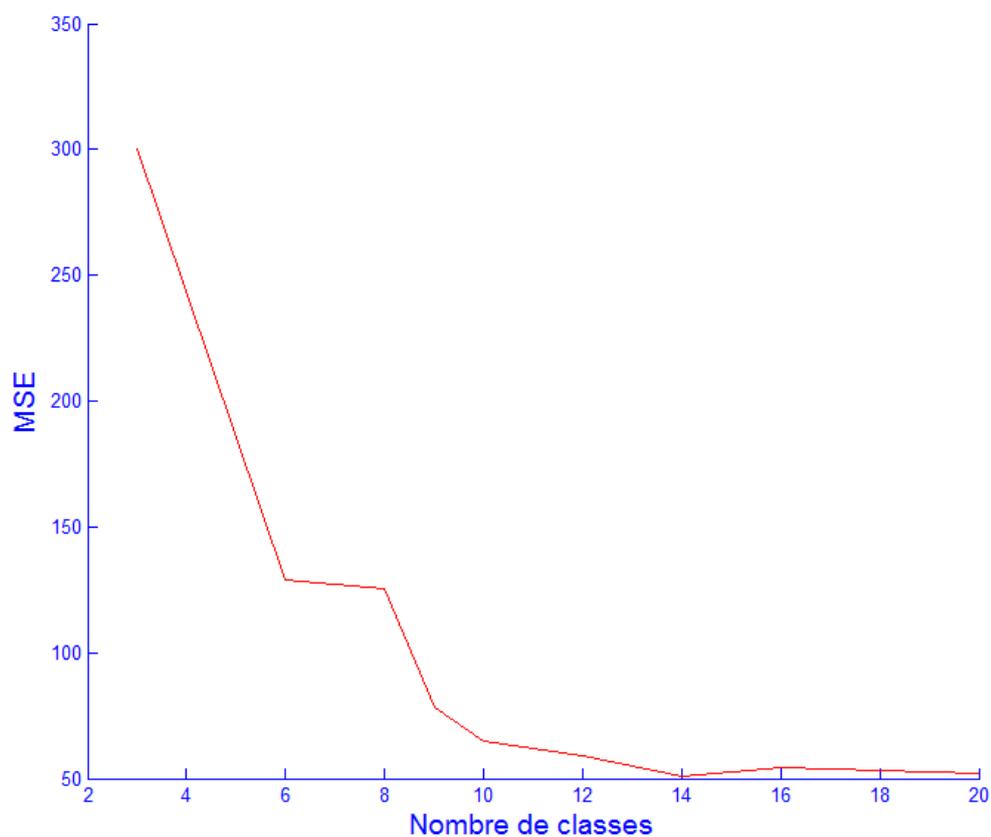


Fig. 4.7 Evaluation de la segmentation en variant le nombre de classes par MSE.

4.5.2 Influence du nombre de sous-synapses sur le résultat de la segmentation

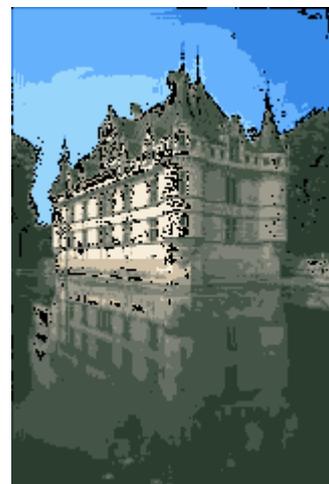
Pour montrer l'influence du nombre de sous-synapses sur le résultat de la segmentation, nous avons fixé le nombre de classes à 10, le nombre de champs récepteurs à 12 pour chaque composante de couleur, le pas d'apprentissage à 0.35, la taille du corpus d'apprentissage est de 20% pris aléatoirement de l'image originale. Pour 3 itérations nous avons varié le nombre des sous-synapses. Les résultats de segmentation obtenus sont montrés par la [Figure 4.8](#). L'évaluation par MSE est montrée par la [Figure 4.9](#). Malgré le nombre important de sous synapses allant à 50, la variation de l'erreur quadratique reste inchangeable. Pour 4 sous synapses une classe supplémentaire est ajoutée vu que plusieurs neurones en sortie possèdent la plus grande valeur d'activation. Un nombre de 8 à 10 sous synapses donne une bonne segmentation.



Image originale



Segmentation avec 4 sous synapses



Segmentation avec 6 sous synapses



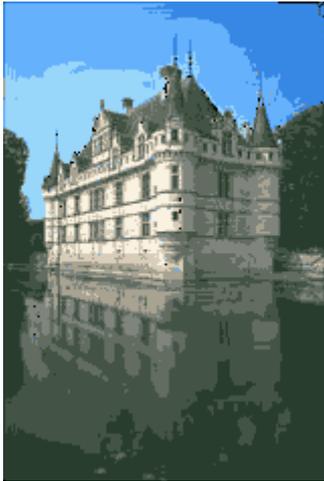
Segmentation avec 8 sous synapses



Segmentation avec 10 sous synapses



Segmentation avec 12 sous synapses



Segmentation avec 14 sous synapses



Segmentation avec 16 sous synapses



Segmentation avec 20 sous synapses

Fig. 4.8 Résultat de segmentation par variation du nombre de sous synapses.

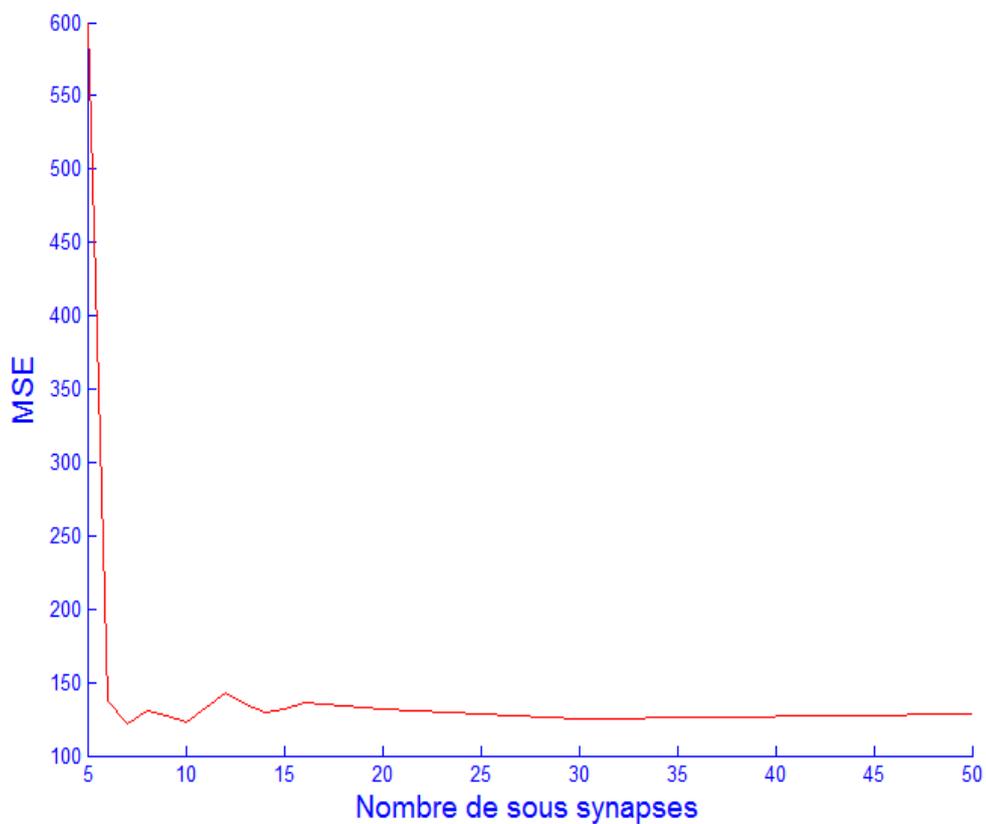


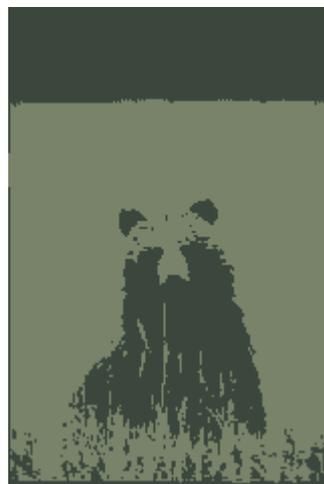
Fig. 4.9 Evaluation de la segmentation en variant le nombre de sous synapses par MSE.

4.5.3 Influence du nombre de champs récepteurs sur le résultat de la segmentation

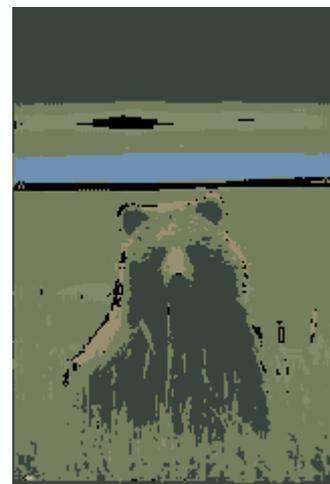
Pour montrer l'influence du nombre des champs récepteurs sur le résultat de la segmentation, nous avons fixé le nombre de classes à 10, le nombre de sous-synapses de chaque connexion entre deux neurones à 8, le pas d'apprentissage à 0.35, la taille du corpus d'apprentissage est de 20% pris aléatoirement de l'image originale. Pour 3 itérations nous avons varié le nombre des champs récepteurs. Les résultats de la segmentation obtenus sont montrés par la [Figure 4.10](#). L'évaluation par MSE est montrée par la [Figure 4.11](#). Au-delà de 12 champs récepteurs plusieurs neurones en sortie possèdent la plus grande valeur d'activation. Une classe supplémentaire est ajoutée prenant comme valeurs d'intensité la couleur noire. Un nombre de 10 à 12 champs récepteurs donne une bonne segmentation.



Image originale



Segmentation avec 4 champs récepteurs



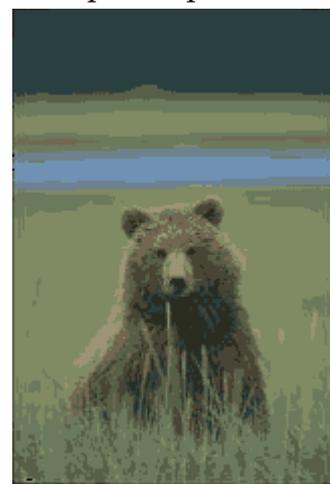
Segmentation avec 6 champs récepteurs



Segmentation avec 8 champs récepteurs



Segmentation avec 10 champs récepteurs



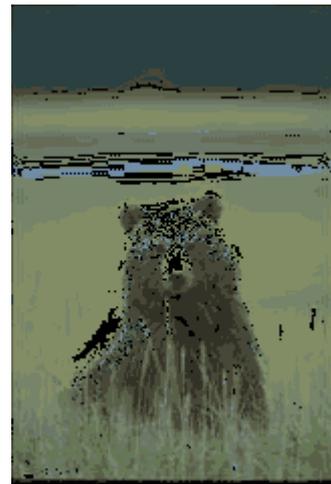
Segmentation avec 12 champs récepteurs



Segmentation avec 14 champs récepteurs



Segmentation avec 16 champs récepteurs



Segmentation avec 20 champs récepteurs

Fig. 4.10 Résultat de segmentation par variation du nombre de champs récepteurs.

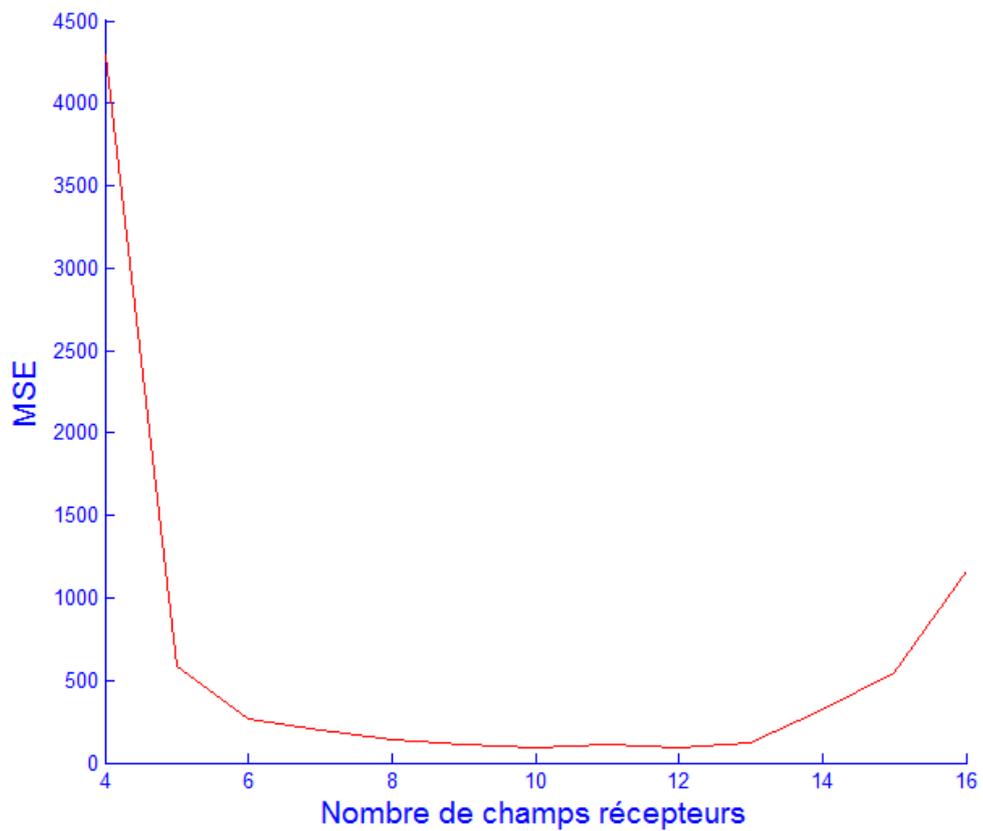


Fig. 4.11 Evaluation de la segmentation en variant le nombre de champs récepteurs par MSE.

4.5.4 Influence de la taille du corpus sur le résultat de la segmentation

Pour montrer l'influence de la taille du corpus d'apprentissage sur le résultat de la segmentation, nous avons fixé le nombre de classes à 10, le nombre de sous-synapses de chaque connexion entre deux neurones à 8, le nombre de champs récepteurs à 12 pour chaque composante de couleur, le pas d'apprentissage à 0.35. Pour 3 itérations nous avons varié la taille du corpus d'apprentissage. Les résultats de segmentation obtenus sont montrés par la [Figure 4.12](#). On peut remarquer qu'un petit corpus d'apprentissage (5%) donne une segmentation représentative.



Image originale



Segmentation avec 5 %



Segmentation avec 10 %



Segmentation avec 15 %



Segmentation avec 20 %



Segmentation avec 30 %



Segmentation avec 40%



Segmentation avec 50 %



Segmentation avec 80 %

Fig. 4.12 Résultat de segmentation par variation du corpus d'apprentissage.

4.5.5 Influence du seuil de déclenchement téta sur le résultat de la segmentation

Pour montrer l'influence du seuil téta le résultat de la segmentation, nous avons fixé le nombre de classes à 10, le nombre de sous-synapses de chaque connexion entre deux neurones à 8, le nombre de champs récepteurs à 12 pour chaque composante de couleur, le pas d'apprentissage à 0.35, la taille du corpus d'apprentissage est de 20% pris aléatoirement de l'image originale. Nous avons varié la valeur du seuil téta. Les résultats de la segmentation obtenus sont montrés par la [Figure 4.13](#). L'évaluation par MSE est montrée par la [Figure 4.14](#). Pour un seuil de 1 et au-delà de 20 plusieurs neurones en sortie possèdent la plus grande valeur d'activation. Une classe supplémentaire est ajoutée prenant comme valeurs d'intensité la couleur noire. Un seuil de déclenchement de 9 donne une bonne segmentation.



Image originale

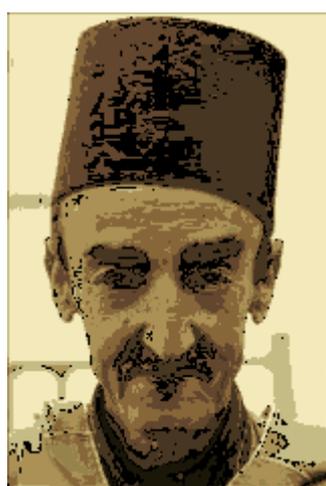
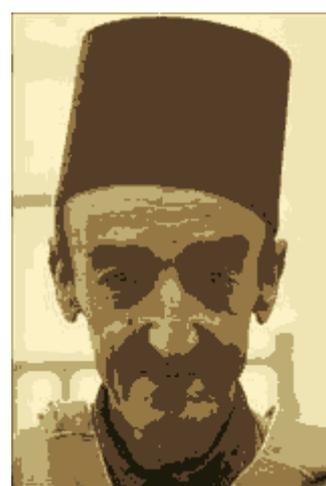
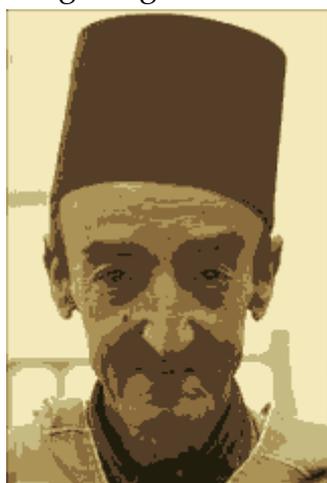
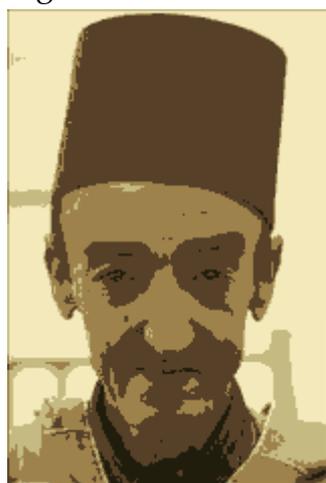
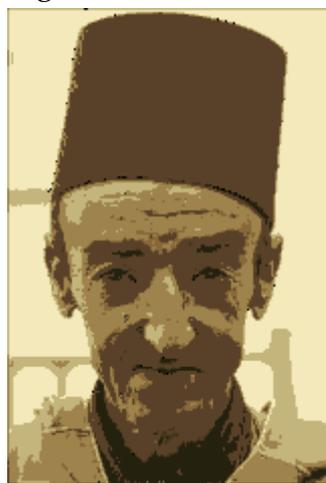
Segmentation avec $\theta=1$ Segmentation avec $\theta=2$ Segmentation avec $\theta=5$ Segmentation avec $\theta=7$ Segmentation avec $\theta=9$ Segmentation avec $\theta=12$ Segmentation avec $\theta=16$ Segmentation avec $\theta=20$

Fig. 4.13 Résultat de segmentation par variation du seuil θ .

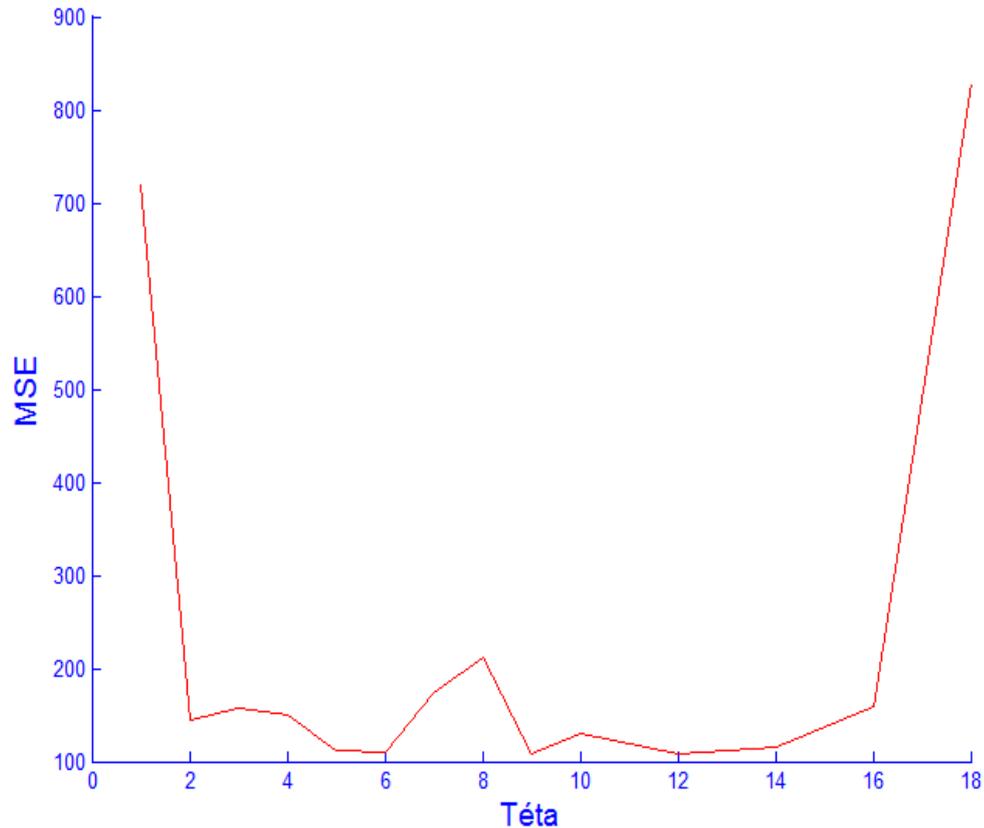


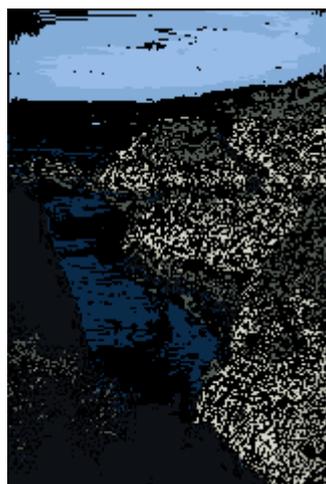
Fig. 4.14 Evaluation de la segmentation en variant téta par MSE.

4.5.6 Influence du paramètre Tau sur le résultat de la segmentation

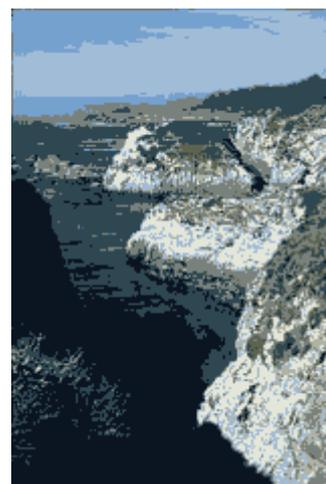
Pour montrer l'influence du paramètre Tau sur le résultat de la segmentation (Eq. 3.25), nous avons fixé le nombre de classes à 10, le nombre de sous-synapses de chaque connexion entre deux neurones à 8, le nombre de champs récepteurs à 12 pour chaque composante de couleur, le pas d'apprentissage à 0.35, la taille du corpus d'apprentissage est de 20% pris aléatoirement de l'image originale. Pour 3 itérations nous avons varié la valeur de Tau. Les résultats de la segmentation obtenus sont montrés par la Figure 4.15. L'évaluation par MSE est montrée par la Figure 4.16. Tau n'apporte pas une grande influence sur le processus de segmentation au-delà de 3. Une bonne segmentation est obtenue par une valeur de 3.



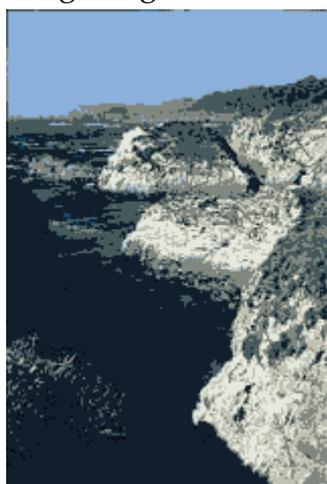
Image originale



Segmentation avec tau=1



Segmentation avec tau=2



Segmentation avec tau=3



Segmentation avec tau=5



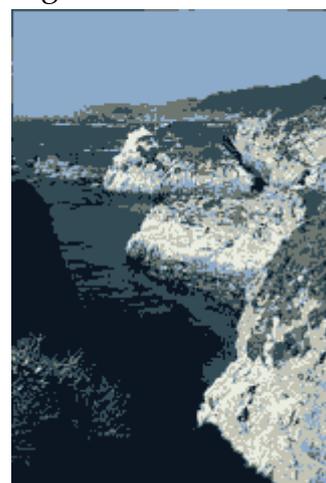
Segmentation avec tau=9



Segmentation avec tau=12



Segmentation avec tau=16



Segmentation avec tau=20

Fig. 4.15 Résultat de segmentation par variation du paramètre tau.

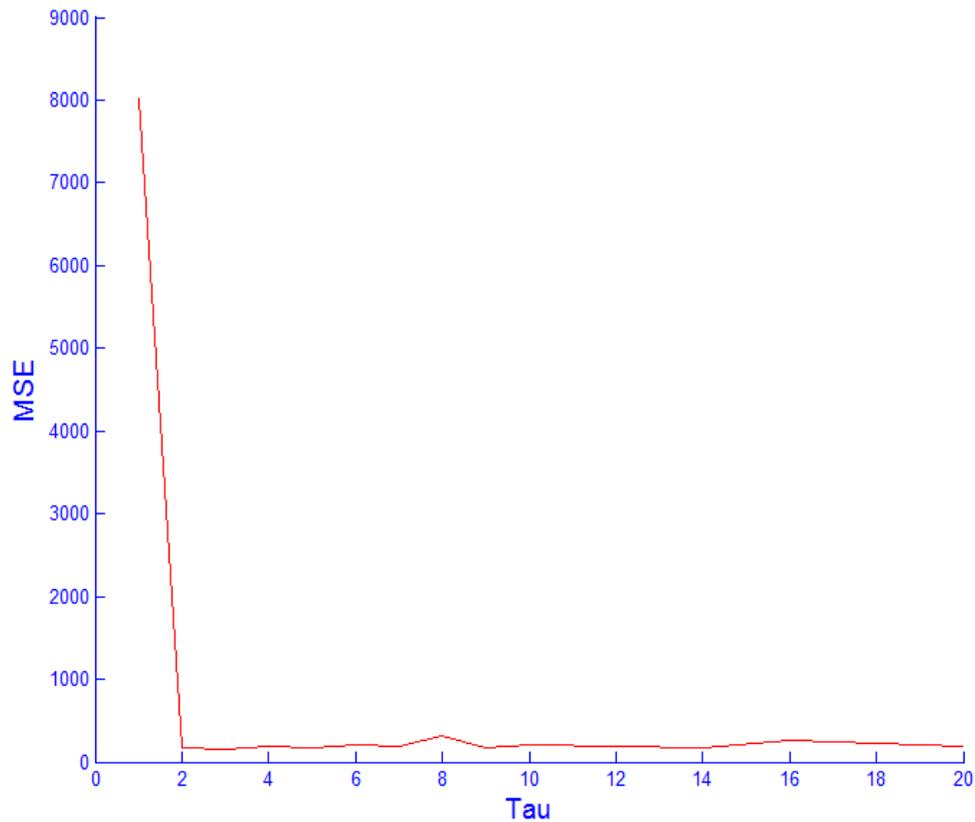


Fig. 4.16 Evaluation de la segmentation en variant tau par MSE.

4.5.7 Influence des paramètres d'apprentissage b et c sur le résultat de la segmentation

Pour montrer l'influence des paramètres d'apprentissage b et c sur le résultat de la segmentation, nous avons fixé le nombre de classes à 10, le nombre de sous-synapses de chaque connexion entre deux neurones à 8, le nombre de champs récepteurs à 12 pour chaque composante de couleur, le pas d'apprentissage à 0.35, la taille du corpus d'apprentissage est de 20% pris aléatoirement de l'image originale. Pour 3 itérations nous avons varié le paramètre b et c. Les résultats de segmentation obtenus sont montrés par les Figures 4.17 et 4.19. L'évaluation par MSE est montrée par les Figure 4.18 et Figure 4.20. Le paramètre b doit être choisi dans un intervalle de -0.01 à 0.1 et c est choisi entre [-3,0] et [2,3]. Quand c varie de 0 à 2 plusieurs neurones en sortie possèdent la plus grande valeur d'activation. Une classe supplémentaire est ajoutée prenant comme valeurs d'intensité la couleur noire.



Image originale

Segmentation avec $b=0.005$ Segmentation avec $b=0.05$ Segmentation avec $b = 0.1$ Segmentation avec $b = 0.5$ Segmentation avec $b = 1$

Fig. 4.17 Résultat de segmentation par variation du paramètre d'apprentissage b .

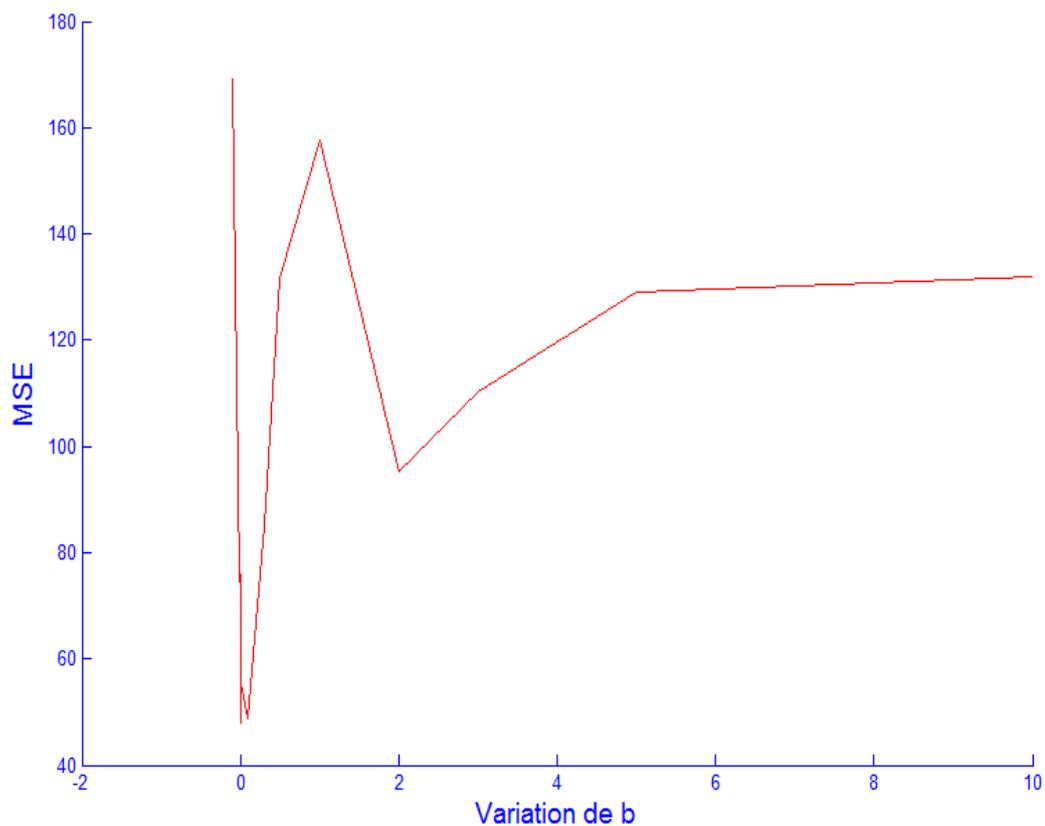
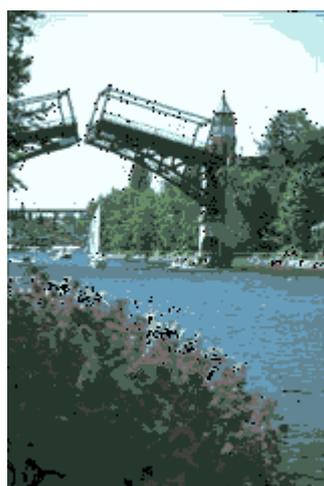


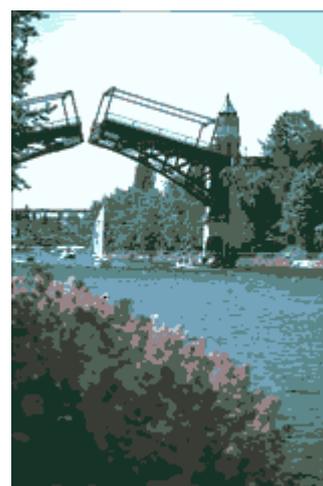
Fig. 4.18 Evaluation de la segmentation en variant le paramètre b par MSE.



Image originale



Segmentation avec $c = -5$



Segmentation avec $c = -2.3$

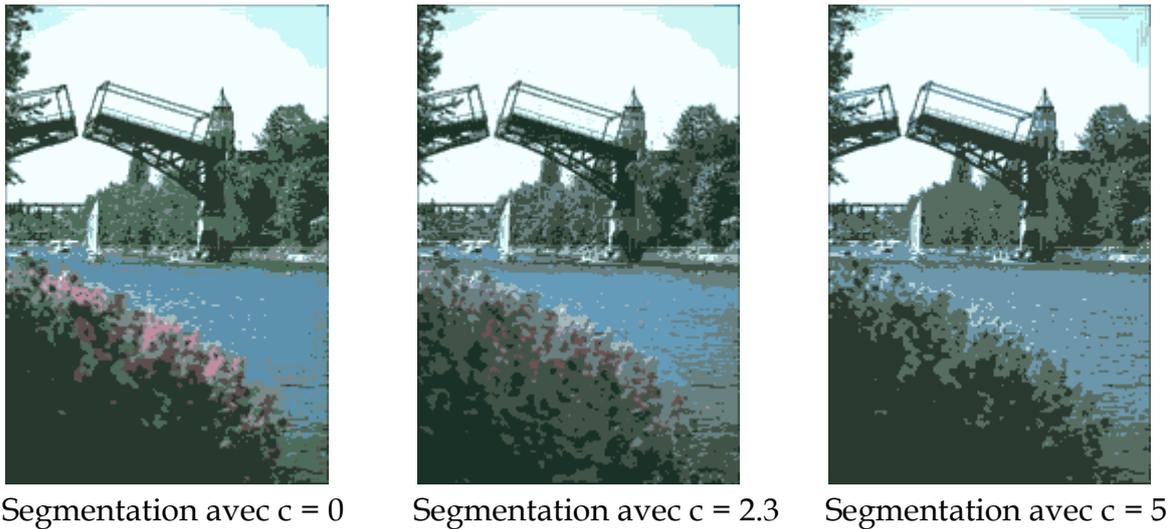


Fig. 4.19 Résultat de segmentation par variation du paramètre d'apprentissage c .

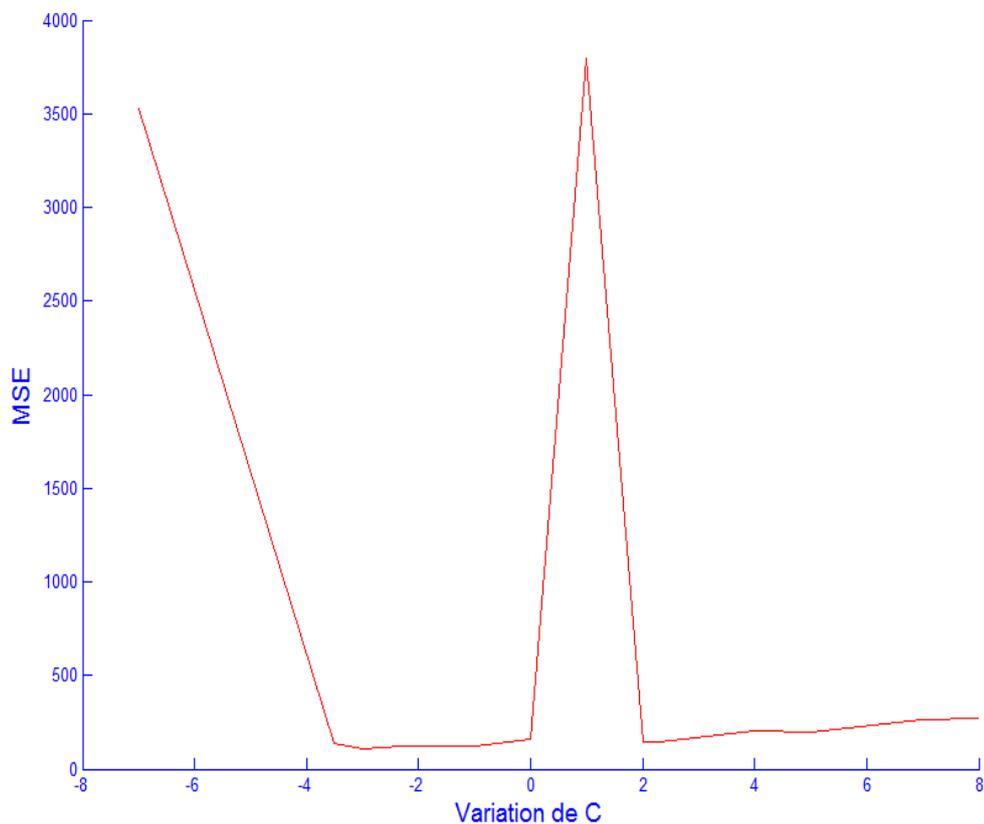


Fig. 4.20 Evaluation de la segmentation en variant le paramètre c par MSE.

4.5.8 Meilleurs paramètres du réseau de neurones impulsionnels

Plusieurs expériences sont effectuées sur 50 images de la base de Berkeley, tout en variant le nombre de sous synapses, le nombre de champs récepteurs, la taille du corpus d'apprentissage et les paramètres b , c et τ (Mefteh et al., 2010a). Les

meilleurs paramètres du réseau de neurones impulsionnels pour une erreur quadratique de $87.352 \pm [28.747, 39.319]$ sont montrés sur le tableau suivant.

Champs récepteurs	Sous synapses	Seuil	Corpus d'apprentissage	b	c	η	τ	Nb. itération
8	12	9	20%	0.007	-2.3	0.0025	3	3

Tableau 4.1 Meilleurs paramètres du réseau de neurones impulsionnels.

4.6 Segmentation des cellules tumorales par réseaux de neurones impulsionnels

L'analyse d'images dans le domaine du dépistage du cancer est un outil important pour la cytopathologie (Knesek, 1999). Deux raisons principales peuvent être soulignées. Premièrement, l'analyse quantitative de la forme et la structure des noyaux provenant des images microscopiques couleur apporte de l'information précieuse pour le pathologiste et aide au diagnostic. Deuxièmement, la quantité d'informations que le pathologiste doit traiter est importante, en particulier lorsque le nombre de dépistage du cancer augmente.

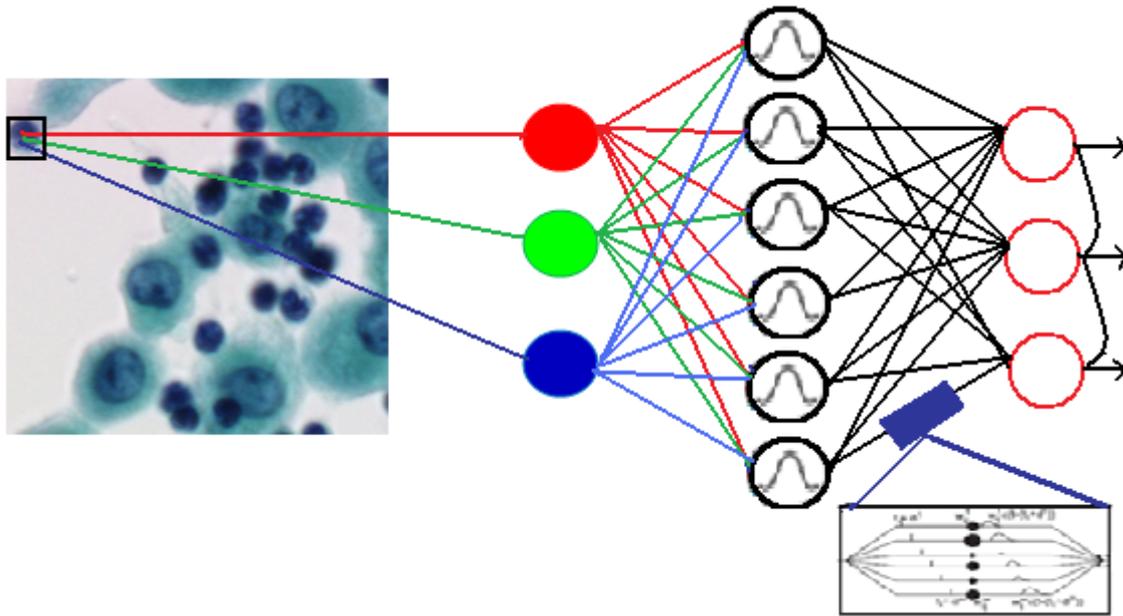
De nombreuses méthodes de segmentation cellulaire ont été présentées à ce jour (Di Rubeto et al., 2000). Ils comprennent la ligne de partage des eaux (Lezoray et Cardot, 2002), segmentation basés région (Mouroutis et al. 1998) et les méthodes à seuil (Wu et al., 2000). L'application des contours actifs a été largement étudiée pour la segmentation des cellules (Murashov, 2004).

Les images de cytologie bronchique sur lesquelles nous appliquons la segmentation par réseaux de neurones impulsionnels sont des images cellulaires microscopiques, constituées de cytoplasmes, noyaux, et parfois d'objets indésirables comme le mucus. La difficulté supplémentaire sur certaines de ces images est d'extraire ce mucus qui ne nous apporte aucune information afin de l'inclure dans le fond.

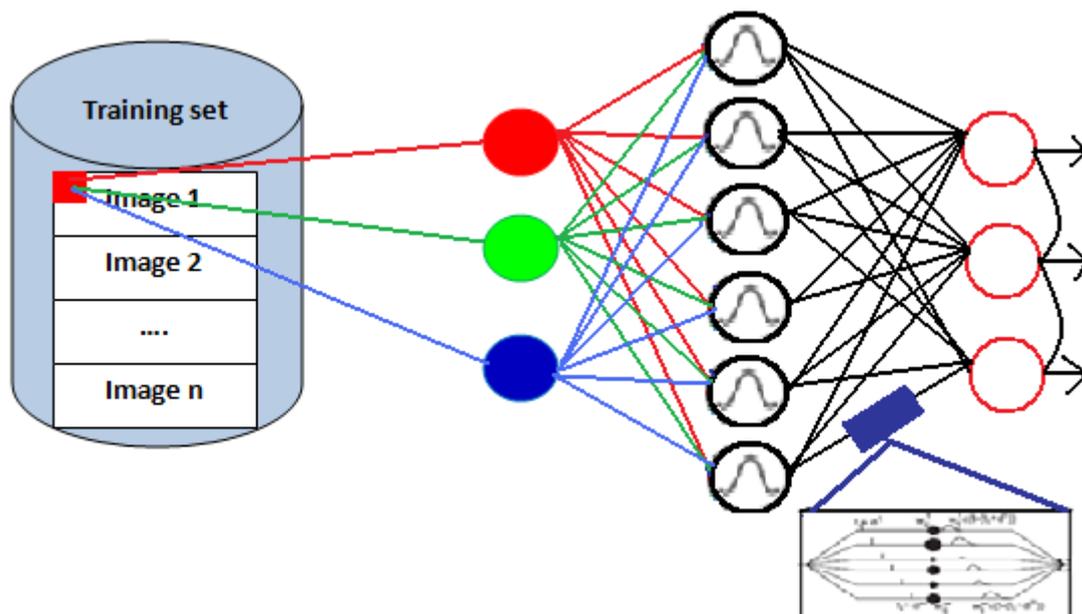
Dans cette partie, nous présentons les résultats que nous avons obtenus par réseaux de neurones impulsionnels sur une base de 12 images couleur de cytologie de tumeurs bronchiques présentant chacune des dizaines à des centaines de cellules (Meftah et al., 2010b). Les images testées sont des images couleur 24 bits de 574x752 pixels acquises par une plateforme normalisée.

Le modèle utilisé est représenté par les Figure 4.21 et 4.22 Il s'agit d'un réseau à couches, dont l'apprentissage est régi par la règle du vainqueur prend tous. La sélection de la base d'apprentissage est obtenue par deux méthodes. Soit par

sélection aléatoire [Figure 4.21](#) ou par sélection supervisée en correspondance avec la vérité terrain [Figure 4.22](#).



[Fig. 4.21](#) Topologie du réseau de neurones impulsionnels pour une sélection aléatoire.



[Fig. 4.22](#) Topologie du réseau de neurones impulsionnels pour une sélection supervisée.

4.6.2 Résultats et évaluation de la segmentation cellulaire par les réseaux de neurones impulsionnels

La **Figure 4.23** montre le résultat obtenu par réseaux de neurones impulsionnels dont la sélection est aléatoire ou supervisée. Le nombre de neurone en sortie est fixée à 3 (fond, cytoplasme et noyau), le pas d'apprentissage à 0.35, la taille du corpus d'apprentissage à 20%, le nombre des champs récepteurs à 8 et le nombre de sous-synapses à 12.

Dans le cadre de l'étude de cellules anormales, dont le but est d'assurer une sécurisation du diagnostic, l'étape d'évaluation est très importante. La qualité de la segmentation cellulaire doit être évaluée automatiquement. Pour cela, nous utilisons plusieurs taux de classification. Ces taux sont exprimés comme suit:

$$R_0 = \frac{\text{Nombre de pixels bien classés}}{\text{Nombre total de pixels de l'image}} \quad (4.11)$$

$$R_1 = \frac{\text{Nombre de pixels du noyau bien classés}}{\text{Nombre total de pixels du noyau de l'image}} \quad (4.12)$$

$$R_2 = \frac{\text{Nombre de pixels du fond bien classés}}{\text{Nombre total de pixels du fond de l'image}} \quad (4.13)$$

$$R_3 = \frac{R_1 + R_2}{2} \quad (4.14)$$

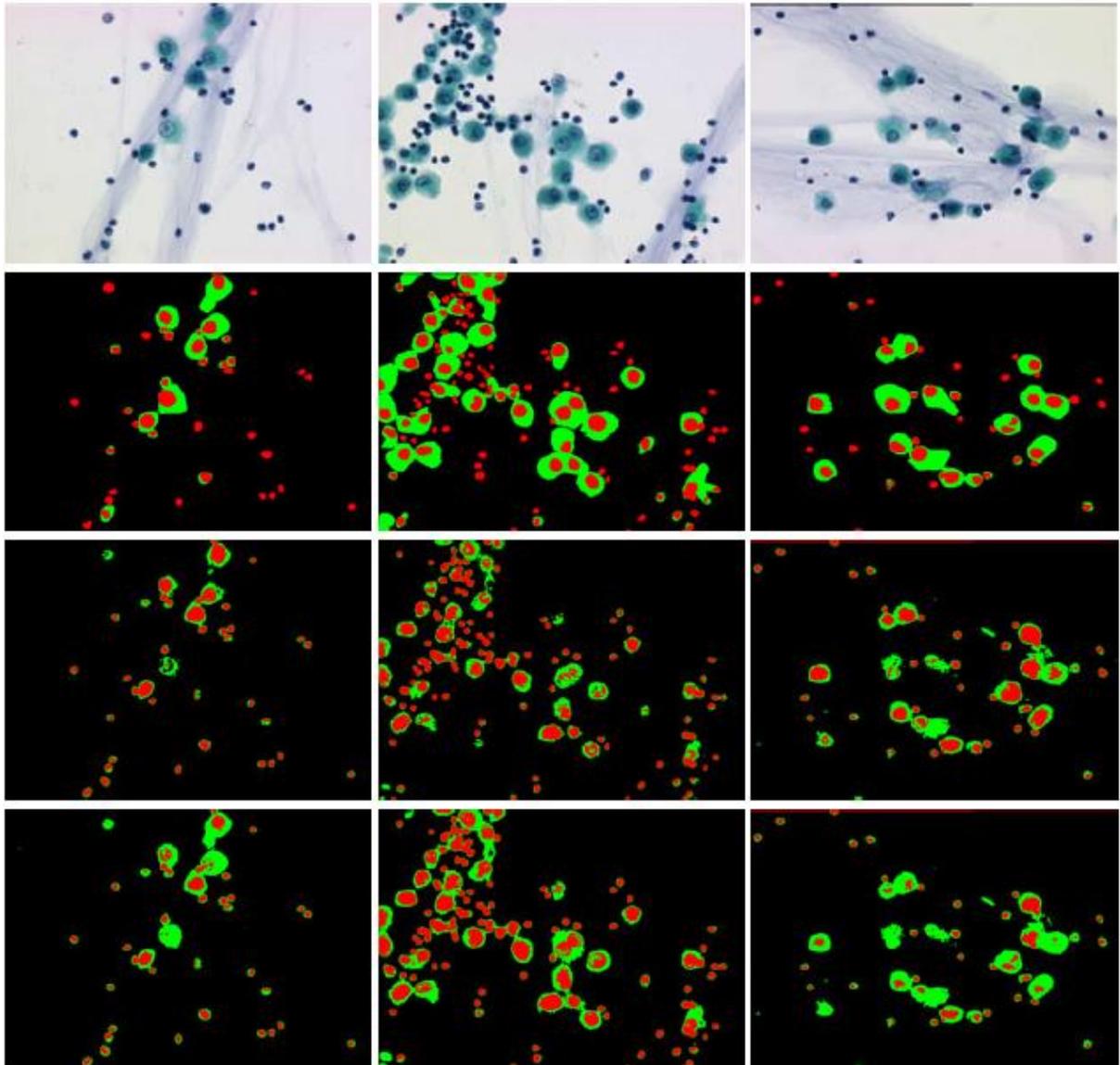


Fig. 4.23 Images microscopiques cellulaires (première ligne), segmentation par expert (second ligne), segmentation par sélection aléatoire (troisième ligne) et segmentation supervisée (dernière ligne).

Les résultats d'évaluation sur le **Tableau 4.2** montrent que l'approche par sélection supervisée est meilleure par rapport à l'approche par sélection aléatoire.

	Classification avec sélection aléatoire	Classification avec sélection supervisée
R ₀	89.07%	94.27%
R ₁	69.57%	80.37%
R ₂	64.55%	99.06%
R ₃	82.03%	89.71%

Tableau 4.2 Taux de classification (meilleurs taux en gras).

Le **Tableau 4.3** présente une comparaison des taux de segmentation sur la base des travaux de Meurie (Meurie et al., 2005) avec l'approche impulsionnel supervisée.

Classifieur	R ₁
SVM	74.2%
Bayes	74.6%
K-Means	74.4%
MLP	73%
Fisher 1	72.6%
K-ppv	70%
SNN supervisé	80.37%

Tableau 4.3 Taux de segmentation en comparaison avec les travaux de Meurie (Meurie et al., 2005).

4.7 Comparaison entre la segmentation par K-means, carte de Kohonen et les réseaux de neurones impulsionnels

K-means est l'algorithme de clustering qui partitionne les données d'une image en K classes. C'est un algorithme itératif qui minimise la somme des distances entre chaque objet et le centroïde de son cluster. La position initiale des centroïdes conditionne le résultat final, de sorte que les centroïdes doivent être initialement placés le plus loin possible les uns des autres de façon à optimiser l'algorithme. K-means change les objets de cluster jusqu'à ce que la somme ne puisse plus diminuer. Le résultat est un ensemble de clusters compacts et clairement séparés, sous réserve qu'on ait choisi la bonne valeur K du nombre de classes.

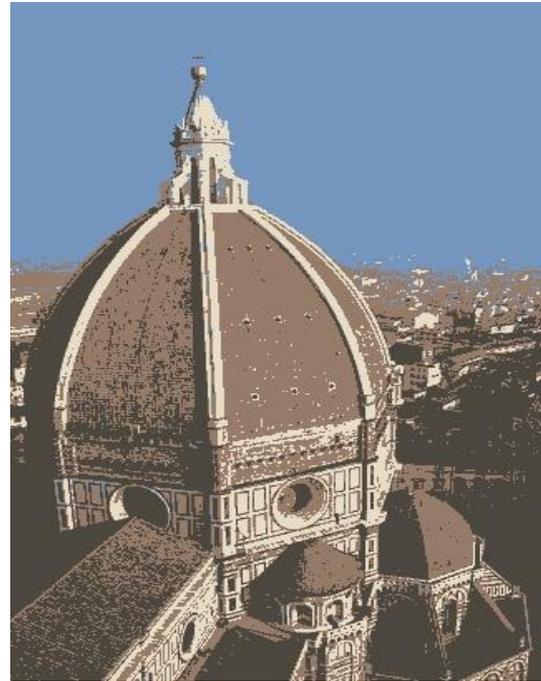
Les cartes de Kohonen (Self-Organizing Maps ou SOM) dans la littérature (Kohonen, 1982) peuvent être définies comme un algorithme de classification non supervisé issu du domaine des réseaux de neurones artificiels. L'algorithme des "cartes auto-organisées", ou "cartes topologiques" a été introduit par Kohonen au début des années 80 pour modéliser un phénomène, couramment observé dans le cerveau : la formation de « cartes ». L'algorithme est une variante de l'algorithme des K-means qui, lors d'une itération, modifie non seulement un centre sélectionné comme étant le plus proche d'une donnée, mais aussi les centres voisins pour un graphe de voisinage fixé. Le graphe implique des interactions latérales entre centres qualifiés de voisins.

Pour illustrer les performances des réseaux de neurones impulsionnels par rapport à d'autres méthodes de segmentation pixellaire (K-means et SOM). Nous avons appliquée la segmentation à une image réelle de la base de Berkeley de taille 321x481 codée sur 24 bits. Le résultat de segmentation par K-means, SOM et les réseaux de neurones impulsionnels est montré dans la **Figure 4.24**.

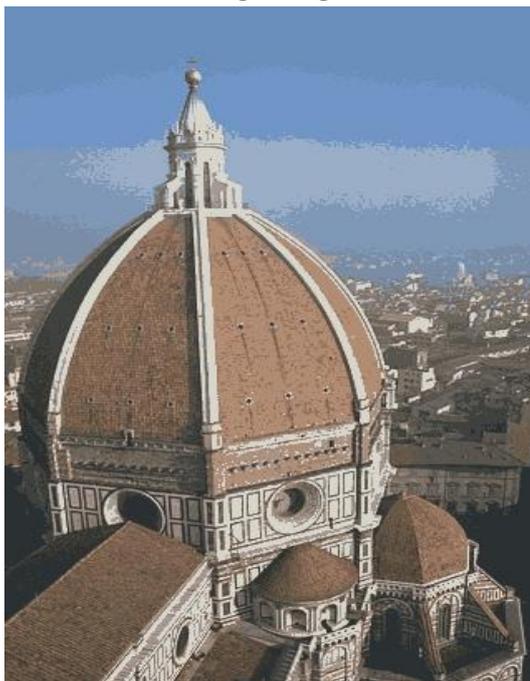
Pour la segmentation par réseaux de neurones impulsionnels, nous avons fixé le nombre de neurone de sortie à 10, le pas d'apprentissage à 0.35, la taille du corpus d'apprentissage à 80%, le nombre des champs récepteurs à 10 et le nombre de sous-synapses à 10.



(a) Image originale



(b) Segmentation par K-means (K=10)



(c) Segmentation par SOM (10 classes)



(d) Segmentation par réseaux de neurones impulsionnels (10 classes)

Fig. 4.24 Résultat de segmentation par différentes approches pixellaires.

Le **Tableau 4.4** présente les taux de segmentation des différentes approches pixellaires. La segmentation par réseaux de neurones impulsionnels donne des meilleurs résultats.

	K-means	SOM	Réseaux de neurones impulsionnels
MSE	385.375	124.596	92.4165
PSNR	51.2831	62.5745	65.5622
MEA	16.6438	7.96044	7.66665

Tableau 4.4 Taux de segmentation par différentes approches pixellaires.

4.8 Segmentation des images en contours

Le modèle utilisé est représenté par la **Figure 4.25**. Il s'agit d'un réseau à couches impulsionnels, dont l'apprentissage est régi par la règle du vainqueur prend tout. La première étape consiste à segmenter l'image en entrée. L'activité neuronale de chaque classe en sortie est extraite on enregistrant l'activité de chaque neurone de sortie qui donne pour chaque pixel une sortie binaire : 1 si le neurone est actif ou 0 si le neurone est inactif. Le résultat binaire des matrices d'activation des neurones de sortie peuvent être représentés par des images binaires contenant les contours détectés par ces neurones.

Une phase de fusion est réalisée par superposition d'images segmentée en contours de toutes les classes en sortie.

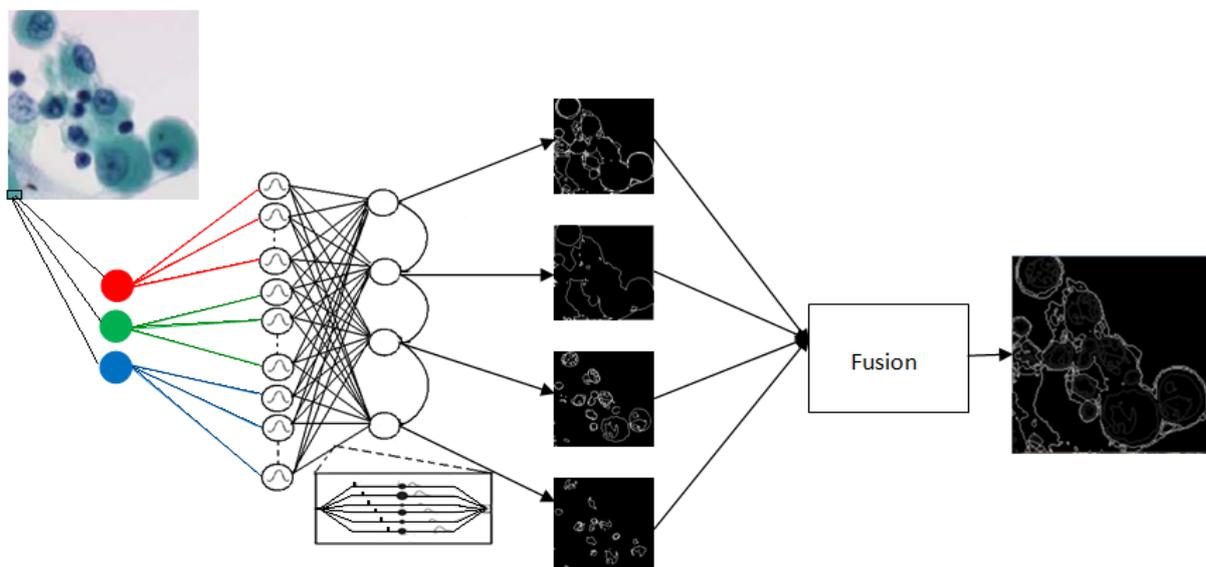


Fig. 4.25 Architecture du réseau de neurones impulsionnels pour la détection de contours.

La [Figure 4.26](#) montre le résultat obtenu par réseaux de neurones impulsionnels, dont le nombre de neurone en sortie est fixée à 8, le pas d'apprentissage à 0.35, la taille du corpus d'apprentissage à 20%, le nombre des champs récepteurs à 8 et le nombre de sous-synapses à 12. On remarque bien que les neurones 4 et 7 donnent des contours de l'avion.

Une comparaison entre la segmentation en contour par réseaux de neurones impulsionnels et d'autres méthodes de détection de contours, tels que Prewitt, morphologie mathématique et Canny, est montrée sur la [Figure 4.27](#).



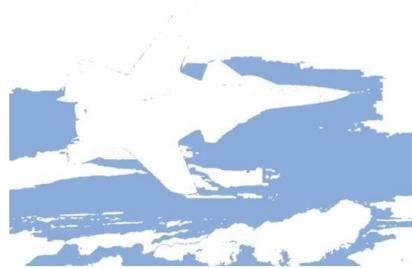
Image originale



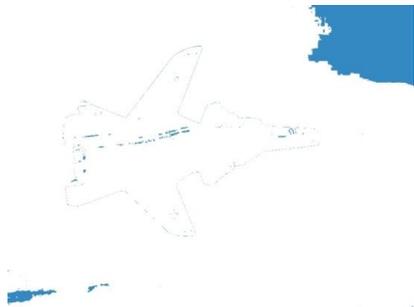
Image segmenté



Zone d'activation du neurone 1



Zone d'activation du neurone 2



Zone d'activation du neurone 3



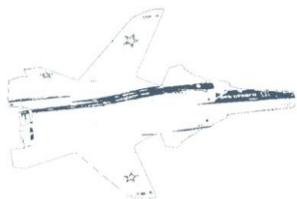
Zone d'activation du neurone 4



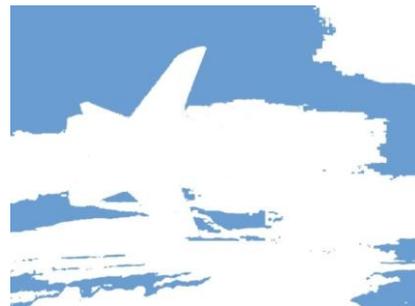
Zone d'activation du neurone 5



Zone d'activation du neurone 6



Zone d'activation du neurone 7



Zone d'activation du neurone 8

Fig. 4.26 les zones d'activation des neurones de la couche de sortie.

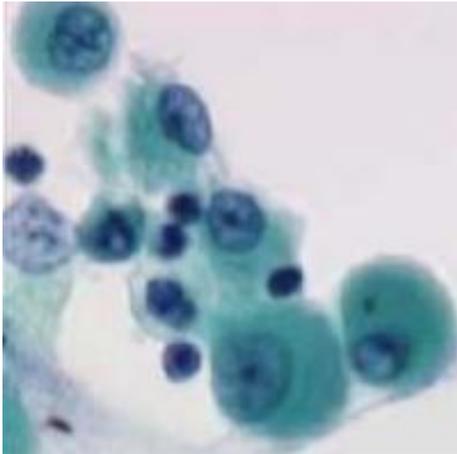


Image originale



Détection de contours par Prewitt



Détection de contours par morphologie



Détection de contours par Canny

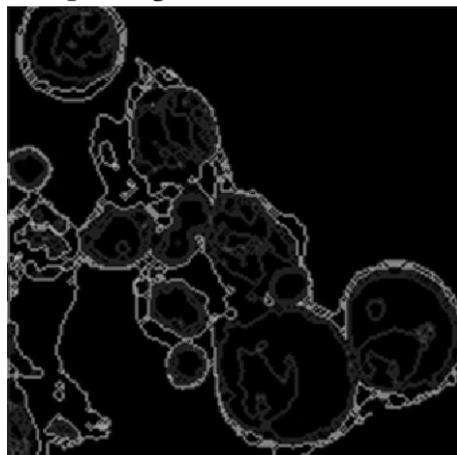


Image segmentée en contours par réseaux de neurones impulsionnels

Fig. 4.27 Détection de contours par différentes approches.

4.7 Conclusion

Dans ce chapitre on a montré comment un réseau de neurones impulsionnels peut segmenter des images visuelles en régions et en contours tout en détectant des détails qui sont difficiles à détecter par d'autres méthodes classiques. Les réseaux de neurones impulsionnels nécessitent une petite base de données d'apprentissage par rapport à d'autres méthodes neuronales pour donner de meilleurs résultats de segmentation.

Citons qu'aucune méthode de segmentation n'est valable pour tout type d'image et que la connaissance à priori de l'image et du but recherché sont deux critères prépondérants dans le choix de la méthode à utiliser (présence ou non de texture ou de bruit, choix de l'espace couleur et de la métrique).

La difficulté de la classification pixellaire par clustering réside en partie dans le choix du nombre de classes pour extraire de façon précise les différents objets jugés importants d'une image. Une connaissance à priori de l'image ou de l'application est donc un facteur intéressant pour la détermination de ce paramètre.

CHAPITRE 5

Suivi d'objets basé sur les réseaux de neurones impulsionnels

Sommaire

5.1 Introduction	157
5.2 Modèle d'un réseau de neurones impulsionnels pour la détection de contours ..	158
5.2.1 Architecture du réseau de neurones impulsionnels	158
5.2.2 Description Fonctionnelle du réseau	159
5.2.3 Modèle formel du réseau de neurones impulsionnels	160
5.2.4 Résultat de détection de contours.....	163
5.3 Extraction d'un modèle initial pour le suivi	164
5.4 Correspondance par la distance de Hausdorff	166
5.4.1 Distance de Hausdorff.....	166
5.4.1.1 Propriétés générales de la distance de Hausdorff	166
5.4.1.2 Distance de Hausdorff totale	166
5.4.1.3 Distance Hausdorff partielle.....	167
5.4.2 Mise en correspondance de blocs	167
5.5 Résultats de suivi sur une séquence d'images	168
5.5.1 Séquences d'images utilisées.....	168
5.5.2 Résultats de suivi	169
5.6 Conclusion	172

5.1 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. 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, la vidéosurveillance, la compression vidéo, l'imagerie médicale, 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.

Le problème de suivi d'objets dans une séquence d'images peut être divisé en deux sous problèmes, l'extraction de l'objet cible et son suivi le long de la séquence. En général ces deux tâches ne sont pas simples en raison de présence de milieux encombrés et 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, le flot optique, les contours, etc. (Yilmaz et al, 2006).

Notre travail dans ce chapitre consiste à suivre un objet dans une séquence vidéo (Yedjour et al., 2011). Le processus du suivi est montré dans la Figure 5.1. Une fois la séquence présentée au réseau impulsional (Wu et al., 2007), des cartes de fréquences de décharges reflétant les contours des images en entrées sont obtenues. 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 avec celle du système visuel humain (Thorpe et Fabre-Thorpe, 2002). Chaque carte de contours est ensuite segmentée séparément afin d'extraire les objets en mouvement. Une fois la segmentation faite, nous obtenons des cartes binaires de mouvement sous forme de régions connexes ou blocs, chacune correspondant à un mouvement particulier et donc à un objet. L'étape de suivi d'un objet mobile se fait quand à elle par une approche basée sur l'emploi de la distance de Hausdorff pour comparer et déterminer la nouvelle position de l'objet localisé. L'approche de comparaison d'images utilisant la distance de Hausdorff permet de donner une mesure définie entre deux ensembles de points. 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, de mise en échelle et de translation entre deux images en appliquant le processus en réduisant au minimum la distance de Hausdorff sur les deux ensembles de points (Teng, 2002).

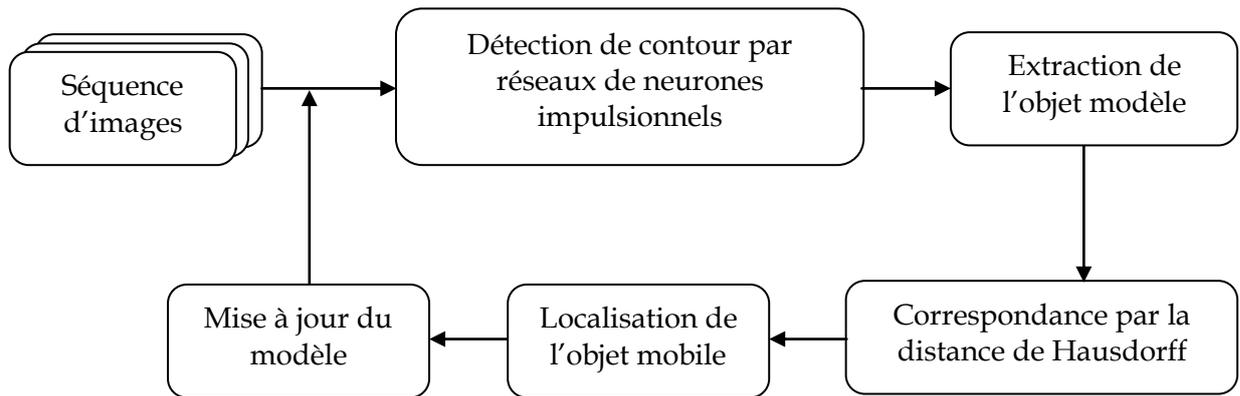


Fig. 5.1 Processus de suivi.

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

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 (Kandel et al. 2000) 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.

5.2.1 Architecture du réseau de neurones impulsionnels

Le réseau de neurones impulsionnels sur lequel nous avons travaillé se présente comme le montre la Figure 5.2. Il possède 3 couches : une couche d'entrée, une couche intermédiaire et une couche de sortie (Xiang Wu et al. 2007).

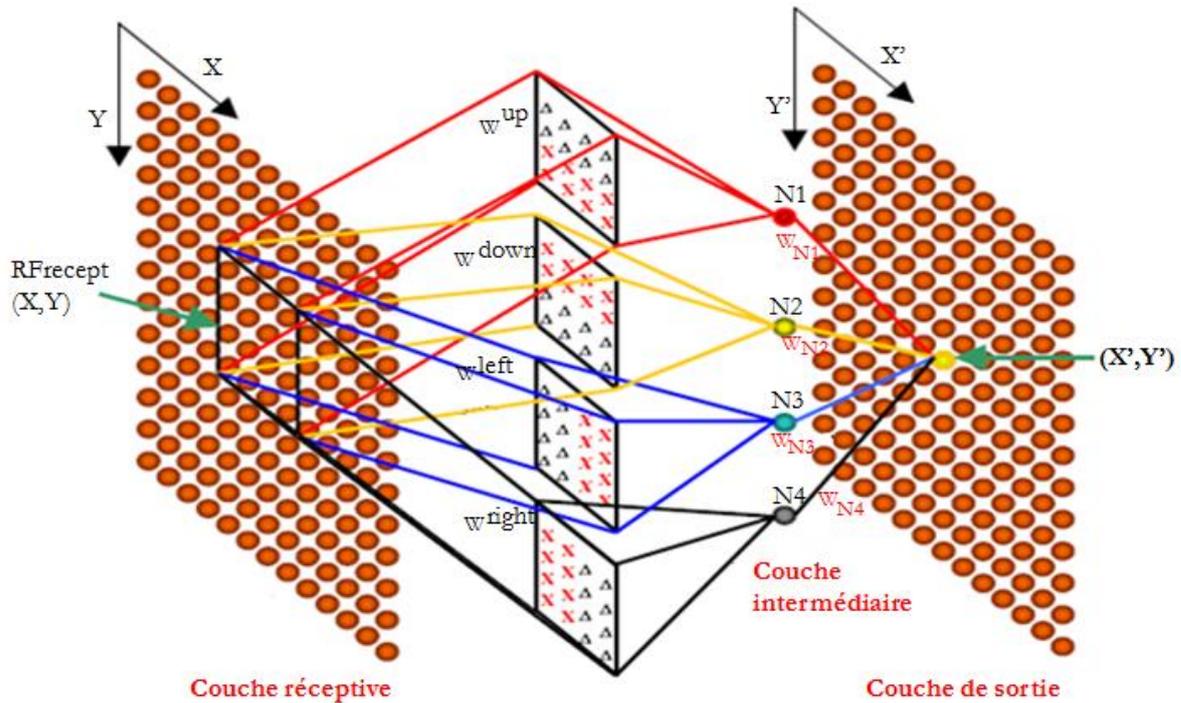


Fig. 5.2 Modèle de réseau de neurones impulsionnels pour la détection de contours.

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.

5.2.2 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 $N4$. 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 en haut, en bas, à gauche et à droite 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. Par exemple dans la Figure 5.2, le neurone $N1$ est relié au champ récepteur $RFrecept$ de la couche d'entrée à travers une matrice w^{up} de distribution des poids synaptiques et répond à un contour haut dans ce champ.

Si une image uniforme dans RFrecept donne une sortie uniforme, les sorties par les synapses w^{up} 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 RFrecept apparaît sur les récepteurs de moitié inférieure avec un signal fort et les récepteurs de moitié supérieure avec un signal très faible, alors un signal fort renforcera (en raison des synapses excitatrices) le neurone $N1$, mais un 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 w^{up} 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 w^{down} peut mieux répondre à un contour bas dans le champ récepteur, le neurone $N3$ avec la distribution de poids synaptiques w^{left} peut mieux répondre à un contour gauche, et le neurone $N4$ avec la distribution de poids synaptiques w^{right} 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 RFrecept.

5.2.3 Modèle formel du réseau de neurones impulsionnels

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 (Koch, 1999 ; Xiang Wu et al., 2006). La conductance basée sur le modèle intègre-et-tire est appliquée au modèle du réseau présenté ci-dessus.

On pose que $G_{x,y}$ représente le niveau de gris du point $(x,y) \in \text{RFrept}$, $q_{x,y}^{ex}$ 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 champ récepteur peut transformer une valeur en niveau de gris à une conductance maximale par les expressions suivantes :

$$q_{x,y}^{ex} = \alpha G_{x,y} ; q_{x,y}^{ih} = \beta G_{x,y} \quad (5.1)$$

Où α et β sont des constantes. Selon le modèle à base de conductance intègre et tire (Xiang Wu et al., 2005 ; Xiang Wu et al., 2006), le neurone $N1$ est régi par les équations suivantes :

$$\frac{g_{x,y}^{ex}(t)}{dt} = -\frac{1}{\tau_{ex}} g_{x,y}^{ex}(t) + \alpha G_{x,y} \quad (5.2)$$

$$\frac{g_{x,y}^{ih}(t)}{dt} = -\frac{1}{\tau_{ih}} g_{x,y}^{ih}(t) + \beta G_{x,y} \quad (5.3)$$

$$c_m \frac{dv_{N1}(t)}{dt} = g_l(E_l - v_{N1}(t)) + \sum_{(x,y) \in \text{ERF}_{\text{rept}}} \frac{w_{x,y}^{\text{up-ex}} g_{x,y}^{ex}(t)}{A_{\text{ex}}} (E_{\text{ex}} - v_{N1}(t)) \\ + \sum_{(x,y) \in \text{ERF}_{\text{rept}}} \frac{w_{x,y}^{\text{up-ih}} g_{x,y}^{ih}(t)}{A_{\text{ih}}} (E_{\text{ih}} - v_{N1}(t)) \quad (5.4)$$

$g_{x,y}^{ex}(t)$ 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,

$v_{N1}(t)$ 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,

$w_{x,y}^{\text{up-ex}}$ représente le poids des synapses excitatrices et $w_{x,y}^{\text{up-ih}}$ 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 biologique des champs récepteurs (Kandel et al., 2000), les valeurs de $w_{x,y}^{\text{up-ex}}$ et de $w_{x,y}^{\text{up-ih}}$ sont données par les expressions suivantes :

$$w_{x,y}^{up_ex} = \begin{cases} 0 & \text{si } (y - y_c) \leq 0 \\ w_{emax} e^{-\frac{(x-x_c)^2}{\delta_x^2} - \frac{(y-y_c)^2}{\delta_y^2}} & \text{si } (y - y_c) > 0 \end{cases} \quad (5.5)$$

$$w_{x,y}^{up_ih} = \begin{cases} 0 & \text{if } (y - y_c) > 0 \\ w_{imax} e^{-\frac{(x-x_c)^2}{\delta_x^2} - \frac{(y-y_c)^2}{\delta_y^2}} & \text{if } (y - y_c) \leq 0 \end{cases} \quad (5.6)$$

Où $(x - x_c)$ est le centre du champ récepteur $RF_{recept}(x, y) \in RF_{rept}$, 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 v_{th} , le neurone génère une impulsion, et 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 $N1$.

$$S_{N1}(t) = \begin{cases} 1 & \text{si le neurone } N1 \text{ décharge à l'instant } t \\ 0 & \text{si le neurone } N1 \text{ ne décharge pas à l'instant } t \end{cases} \quad (5.7)$$

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 $N_{x',y'}$ dans la couche de sortie est régi par les équations suivantes :

$$\begin{aligned} \frac{g_{x',y'}^{ex}(t)}{dt} = & -\frac{1}{\tau_{ex}} g_{x',y'}^{ex}(t) + (w_{N1}S_{N1}(t) + w_{N2}S_{N2}(t) \\ & + w_{N3}S_{N3}(t) + w_{N4}S_{N4}(t)). \end{aligned} \quad (5.8)$$

$$c_m \frac{dv_{x',y'}(t)}{dt} = g_l (E_l - v_{x',y'}(t)) + \frac{g_{x',y'}^{ex}(t)}{A_{ex}} (E_{ex} - v_{x',y'}(t)) \quad (5.9)$$

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

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

$$r_{x',y'} = \frac{1}{T} \sum_t^{t+T} S_{x',y'}(t) \quad (5.10)$$

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

5.2.4 Résultat de détection de contours

Le modèle du réseau utilise un ensemble de paramètres : $v_{th} = -60$ mv. $v_{res} = -70$ mv. $E_{ex} = 0$ mv. $E_{ih} = -75$ mv. $E_l = -70$ mv. $g_l = 1.0$ μ s/mm². $c_m = 10$ nF/mm². $\tau_{ex} = 4$ ms. $\tau_{ih} = 10$ ms. $\tau_{ref} = 6$ ms. $A_{ih} = 0.028953$ mm². $A_{ex} = 0.014103$ mm². Ces paramètres sont conformes à ceux des neurones biologiques (Thorpe et Fabre-Thorpe, 2002).

w_{emax} et w_{imax} représentent les poids synaptiques. w_{emax} est initialisé à 0.7093 pour les synapses excitatrices, et w_{imax} à 0.3455 pour les synapses inhibitrices.

Les valeurs des niveaux de gris de l'image sont normalisées dans un intervalle de 0 et 1. Par conséquent α et β sont initialisés à 1/valeur_max_image. La taille de RFrept peut être prise dans le rang 2×2 à 6×6.

Les paramètres δ_x et δ_y peuvent être appliqués pour contrôler la sensibilité des contours. Des tests pour différentes valeurs de δ_x , δ_y ainsi que la taille de RFrecept ont été faites. Les résultats montrent que plus les valeurs de δ_x , δ_y ainsi que la taille de RFrecept augmentent, moins est la sensibilité au bruit pour la détection. D'autre part l'augmentation des valeurs de δ_x , δ_y ainsi que la taille de RFrecept, rendent les contours plus flous. Il y a une différence dans le choix de ces valeurs.

Pour $\delta_x = 6$, $\delta_y = 2$, et la taille de RFrept est de 5×5, la carte de la fréquence de décharge dans la couche de sortie est obtenue, en reflétant les contours de l'image présentée en entrée (Figure 5.3.d).

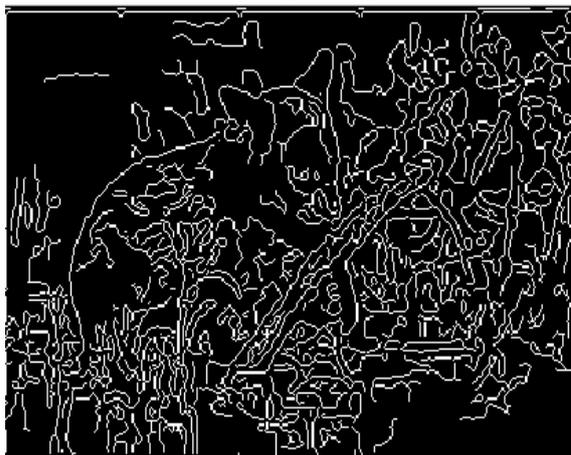
La Figure 5.3 montre différentes approches de détection de contours. La méthode de détection par réseaux de neurones impulsionnels donne des contours fins et plus de détails par rapport aux autres méthodes. C'est pour cela que l'approche impulsionnelle est utilisée dans la suite du chapitre.



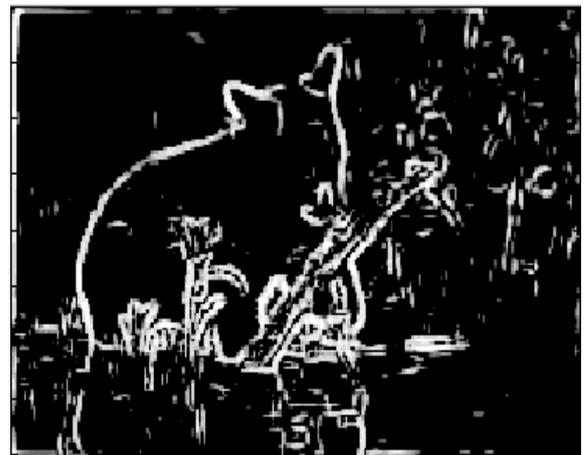
a. Image originale



b. Détection de contours par Prewitt



c. Détection de contours par Canny



d. Carte de fréquence de décharge de la couche de sortie

Fig. 5.3 Différentes approches de détection de contours.

5.3 Extraction de l'objet modèle pour le suivi

Le suivi d'objet nécessite une connaissance à priori de la forme et l'apparence de l'objet d'intérêt, pour cela nous avons besoin d'un modèle de l'objet. Plusieurs approches existent pour l'extraction d'un modèle initial pour le suivi. Kim et Hwang (Kim et Hwang, 2002) présentent une méthode pour la détection des objets en mouvements dans une séquence vidéo. Dans leurs travaux, ils emploient la différence de cartes de contours entre deux images consécutives pour l'extraction des objets en mouvement. Ce modèle s'appelle la segmentation d'objet mobile à partir d'une séquence, il est ensuite poursuivi dans les images suivantes. Pour simplifier et contraindre le problème d'obtenir un modèle initial, certaines conditions ont été posées, qui sont les suivantes : le fond est stationnaire, il y a un seul objet mobile et qu'il n'y a aucune occlusion.

Le but principal de l'extraction d'objet est de délimiter la région de l'objet d'intérêt dans une image donnée. L'objet cible est extrait en utilisant une différence

seuillée entre deux cartes de contours obtenues par réseaux de neurones impulsionnels, la première représentant l'image de fond et la deuxième une image de la séquence telle que montrée dans la [Figure 5.4](#) pour différentes valeurs de seuil.

Dans cette approche, le modèle d'objet, qui est habituellement sous forme de carte de contours détectée par le réseau de neurones impulsionnels, est mis à jour continuellement le long de la séquence après que l'objet soit détecté. Cette mise à jour est conseillée afin de ne considérer que les petits mouvements d'un objet non rigide entre deux images consécutives. Le principe de cette technique est qu'elle permet un suivi d'objet qui maintient la correspondance temporelle des objets à travers la séquence vidéo.

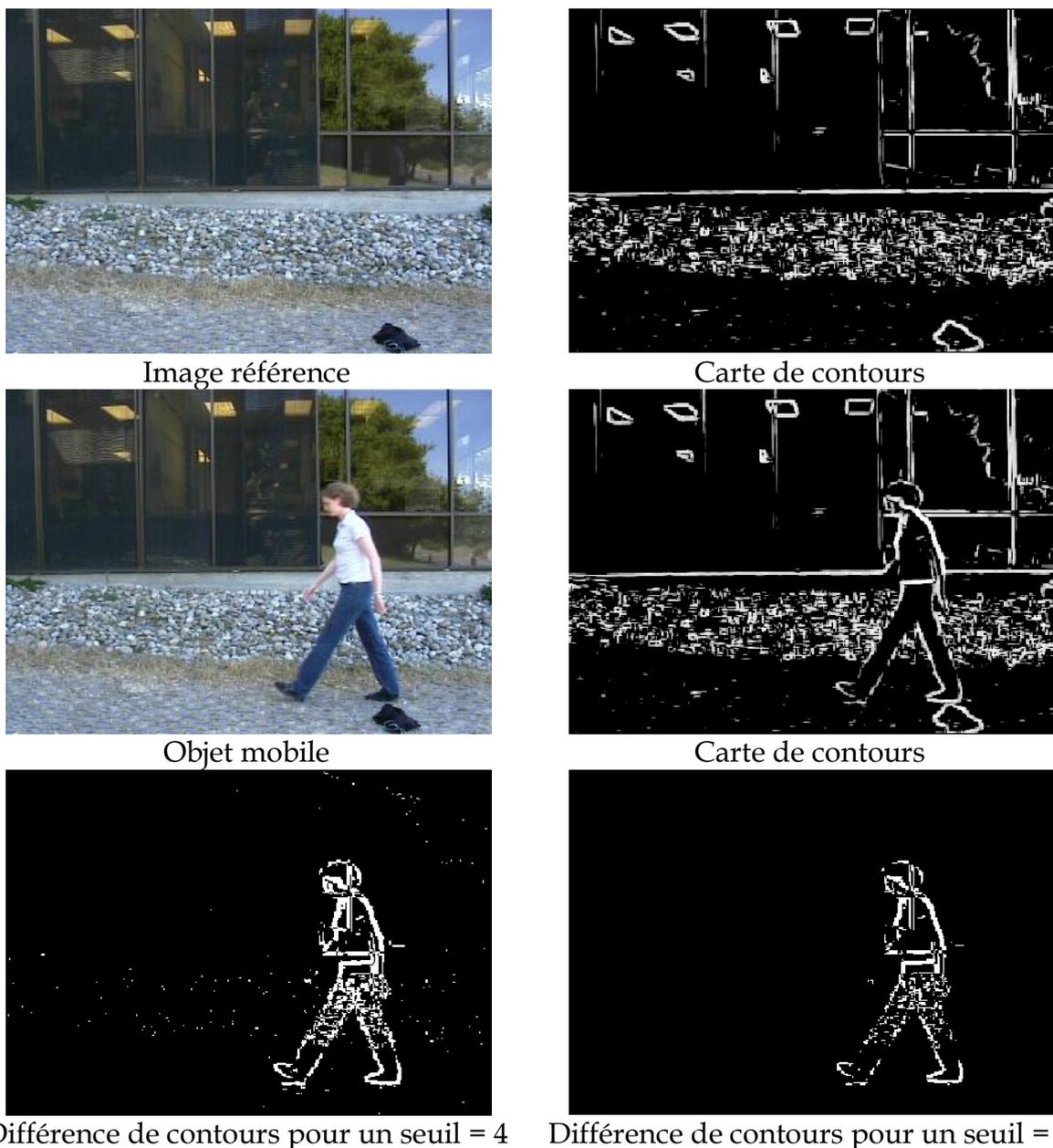


Fig. 5.4 Soustraction de fond en utilisant les cartes de contours.

5.4 Correspondance par la distance de Hausdorff

5.4.1 Distance de Hausdorff

La distance de 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 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 de 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.

5.4.1.1 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 distance de Hausdorff provient de la propriété de la translation. Cela implique que pour de petites translations, la distance de Hausdorff sera petite, ce qui correspond à notre attente pour une mesure de dissimilarité.

5.4.1.2 Distance de Hausdorff totale

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

$$H(A, B) = \max (h(X_i, R_i), h(R_i, X_i)) \quad (5.11)$$

Où

$$h(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\| \quad (5.12)$$

$\|\cdot\|$ 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 de 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 de Hausdorff inverse $h(B, A)$.

Les distances de Hausdorff directe et inverse ne sont pas forcément égales. Pour créer une mesure de distance symétrique, on définit la distance de 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 .

5.4.1.3 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 (Eq. 5.12) 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 $k^{\text{ième}}$ plutôt que de prendre la distance maximale. Ainsi la distance Hausdorff généralisée ou partielle est définie par:

$$h_k(A, B) = k_{a \in A}^{\text{th}} \min_{b \in B} \|a - b\| \quad (5.13)$$

Où k^{th} dénote la $k^{\text{ième}}$ 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 impulsif.

5.4.2 Mise en correspondance de blocs

La caractéristique région se caractérise par l'extraction dans l'image courante des régions dénommées « blocs », un ensemble de pixels connexes, ces régions homogènes seront considérées comme des éléments de base pour le suivi au cours de la séquence.

La mise en correspondance de blocs est réalisée par la distance de Hausdorff partielle entre deux ensembles de points, si cette distance est inférieure à un certain

seuil θ alors nous dirons que la cible existe. La comparaison de régions est réalisée en utilisant une représentation basée sur les contours. Deux régions sont deux ensembles de points de contours extraits à partir du même objet dans deux images de la séquence. 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 la trame courante. La recherche est réalisée en calculant la similitude entre le modèle cible tirée de la trame courante avec le modèle référence tirée de la trame précédente.

Les objets extraits sous forme de « blocs » à l'instant $t-1$ sont mis en correspondance avec ceux détectés à l'instant t . Tout d'abord, nous avons pris la plus grande région au sens mouvement, i.e. 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 d'extraction de régions cibles I_k , il faut trouver la région $R_i \in I_k$ possédant la plus grande ressemblance avec une région $X_i \in I_{ref}$ pour assurer un meilleur appariement. Une fois trouvée, la région R_i devient alors la référence (I_{ref}). 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.

Le modèle d'objet, qui est habituellement sous forme de carte de contours détecté par le réseau de neurones impulsionnels, est mis à jour continuellement dans chaque image de la séquence après que l'objet soit localisé, i.e. 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.

5.5 Résultats sur une séquence d'images

5.5.1 Séquences d'images utilisées

Afin d'étudier les différentes manières dont on peut se représenter une séquence vidéo, considérons l'exemple d'une scène comportant une personne traversant le champ de vision de la caméra de la droite vers la gauche. La séquence vidéo correspondant à cette scène serait constituée des images prises à des instants successifs telles que représentées sur la [Figure 5.5](#). Une séquence avec occultations est montrée sur la [Figure 5.6](#).

Dans toutes les méthodes de détection de mouvement par modélisation de l'arrière-plan de la scène, les données vidéo sont traitées comme une telle succession d'images bidimensionnelles que l'on confronte à un modèle de l'arrière-plan afin de décider en tout point si celui-ci dénote ou non un mouvement apparent. Ce formalisme quasi commun s'explique par plusieurs facteurs.

- La manière dont les données vidéo sont acquises incite naturellement à considérer les différentes images comme des entités indépendantes. En effet, une caméra vidéo statique fonctionne comme un appareil photographique prenant des vues d'une même scène à intervalles réguliers (25 fois par seconde), et les données obtenues sont transmises et stockées sous la forme d'une succession d'images numériques.
- Une coupe dans le plan (x, y) est une image statique qui est facile à se représenter, tandis qu'un échantillon de valeurs prises le long de l'axe du temps, ou une région du plan (x, t) par exemple, sont des entités plus abstraites auxquelles il est difficile d'attribuer du sens.
- Les deux dimensions spatiales d'une séquence vidéo sont connues et finies, tandis que la dimension temporelle est a priori infinie.



Fig. 5.5 Exemple de séquence sans occultations vue comme une succession d'images 2D.

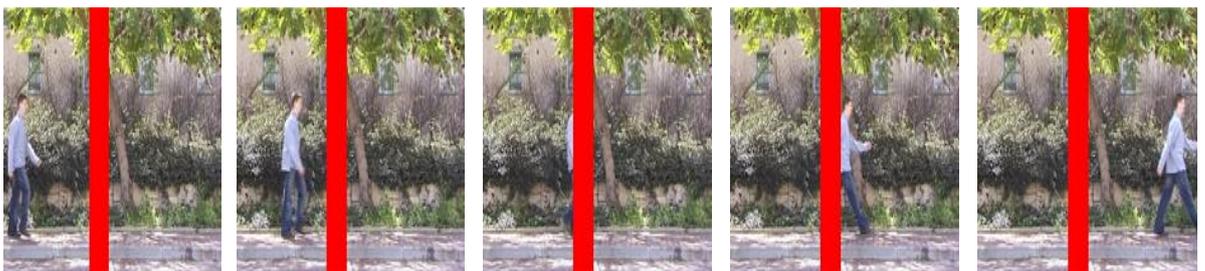


Fig. 5.6 Exemple de séquence avec occultations.

5.5.2 Résultats de suivi

Le suivi d'un objet mobile est évalué sur un PC Pentium 4, avec une vitesse CPU de 3 Ghz et une RAM de 2 Go, la durée d'exécution est d'environ de 100ms pour chaque trame de la séquence de 150 images.

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 trames 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. L'application fonctionne très bien aussi longtemps que l'objet se déplace le long de la séquence. Le résultat d'exécution de quelques trames de la

séquence vidéo est montré dans la [Figure 5.7](#). Un autre exemple de suivi dans une séquence avec occultations est montré sur la [Figure 5.8](#).



Fig. 5.7 Résultat de suivi d'une personne sans occultations.



Fig. 5.8 Résultat de suivi d'une personne avec occultations.

Le taux de suivi est défini par la formule (Eq 5.14). Le tableau 5.1 montre les taux de suivi pour des séquences avec et sans occultations.

$$\text{Tr} = \frac{\text{Nombre de trames bien suivi}}{\text{Nombre total des trames}} \times 100 \quad (5.14)$$

	Nombre de trames	Nombre de trames bien suivi	Tr
Séquence sans occultations	120	119	99.16%
Séquence avec occultations	61	44	72.13%

Tableau 5.1 Taux de suivi pour différentes séquences.

5.6 Conclusion

Dans ce chapitre, nous avons présenté une application de suivi d'objet en se basant sur les contours détectés avec un réseau de neurones impulsionnels.

Le processus de détection d'objet dans une scène est plus difficile pour un objet mobile 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.

Bien que le fonctionnement du neurone dans le cerveau pour la détection de contours ne soit pas encore clair, le modèle du réseau utilisé dans ce chapitre 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.

L'application de suivi avec mise en correspondance utilisant comme mesure de similarité la distance Hausdorff. Cette application peut être implémentée sur d'autres machines plus sophistiquées pour réduire le temps d'exécution. Dans l'avenir nous pouvons 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. Nous pouvons aussi étendre l'idée de ce travail pour utiliser des applications en temps réel pour le suivi d'objets.

Conclusion générale

Conclusion Générale

Le succès des réseaux de neurones s'exprime principalement dans les domaines de traitement d'informations statiques, atemporels, tels que la reconnaissance des formes, la classification ou la quantification vectorielle. Nous avons donné dans cette thèse un état de l'art aussi large que possible des différentes façons de prendre en compte l'aspect temporel par les réseaux de neurones, les différentes architectures de réseaux de neurones temporels. Cette étude nous a permis de conclure que les façons d'aborder le temps par les réseaux de neurones sont nombreuses de même que les travaux et publications concernant l'application de ces réseaux de neurones temporels en traitement d'images animées.

Nous remarquons deux représentations temporelles des réseaux de neurones : une représentation spatiale ou externe et une représentation dynamique ou interne. Les topologies des réseaux de neurones pour chaque représentation temporelle sont complètement différentes et chacune possède ses avantages et ses inconvénients. Parmi les réseaux de neurones temporels nous distinguons les réseaux de neurones impulsionnels.

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 (Abott et Regehr, 2004). 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.

Les réseaux de neurones impulsionnels ont montré ses capacités de résoudre des problèmes déjà résolus par les réseaux de neurones classiques. En plus ils ont l'avantage d'être plus proche des neurones biologiques et de leur capacité de traiter des informations temporelles, ce qui donne une bonne chance à eux pour être un domaine de recherche très riche avec la tendance de créer des machines vraiment intelligentes.

Quoi que des questions soient exposées est ouvertes tels que : Quel est le code utilisé par les neurones pour transmettre une information? Comment modéliser l'apprentissage humain ? Quel type d'apprentissage est adéquat pour les réseaux de neurones impulsionnels ? Il existe une importante littérature scientifique dans le domaine de l'apprentissage pour des réseaux de neurones « classiques ». Le théorème d'approximation universelle (Cybenko, 1989) permet au moins

théoriquement de considéré qu'on « peut tout faire ». Cependant, les résultats dans les réseaux de neurones impulsionnels restent limités.

L'application des réseaux de neurones impulsionnels, dans notre travail, s'articule sur la segmentation d'images et la détection de contours utilisant le modèle à réponse impulsionnelle et sur le suivi d'objets utilisant le modèle intègre et tire.

De ce travail de thèse ont émergé de nombreuses pistes de recherche qui feront l'objet de futurs travaux. Les perspectives concernent plus particulièrement le suivi robuste de plusieurs cibles puisque la mesure Hausdorff a prouvée sa robustesse face au bruit ainsi qu'à l'occlusion. Nous pouvons aussi étendre l'idée de ce travail pour segmenter des vidéos en temps réel.

Publication et citations de l'auteur en relation avec la thématique

Publication et citations de l'auteur

1. Revues internationales

- a) H. Yedjour, **B. Meftah**, D. Yedjour et A. Benyettou. Combining Spiking Neural Network with Hausdorff Distance Matching for Object Tracking. *Asian Journal of Applied Sciences*, 4(1): 63-71, 2011.
- b) **B. Meftah**, O. Lezoray et A. Benyettou. Segmentation and Edge Detection Based on Spiking Neuron Networks. *Neural Processing Letters*, 32(2): 131-146, 2010.

2. Communications internationales avec actes et comité de lecture

- a) **B. Meftah**, O. Lezoray, M. Lecluse et A. Benyettou. Cell Microscopic Segmentation with Spiking. *20th International Conference on Artificial Neural Networks (ICANN 2010)*, Thessaloniki, Greece, LNCS 6352: 117-126, September 15-18, 2010.
- b) **B. Meftah**, A. Benyettou, O. Lezoray and M. Debakla, Image Segmentation with Spiking Neuron Network, *1st Mediterranean Conference on Intelligent Systems and Automation (CISA'08)*, AIP Conf. Proc. Annaba, Algérie, 1019: 15-19, June 30-July 2, 2008, ISBN 978-0-7354-0540-0.
- c) **B. Meftah**, A. Benyettou, O. Lezoray et W. Qing Xiang. Image Clustering with Spiking Neuron Network, *International Joint Conference on Neural Networks (IJCNN 2008)*, part of the IEEE World Congress on Computational Intelligence (WCCI 2008), 681-685, Hong Kong, ISBN 978-1-4244-1821-3, June 2008.

3. Communications nationales avec actes et comité de lecture

- a) H. Yedjour, A. Souahlia, **B. Meftah** et A. Benyettou. Approche de détection de contour basée sur les réseaux de neurones impulsionsnels, *1st International Conference on Image and Signal Processing and their Applications (ISPA'09)*, Mostaganem, Algérie, 19-21 Octobre 2009,.
- b) A. Souahlia, H. Yedjour, **B. Meftah** et A. Benyettou. Approche Hybride de Segmentation des Images Couleurs en Régions et en Contours par l'Utilisation des Réseaux de Neurones Impulsionsnels, *1st International Conference on Image and Signal Processing and their Applications (ISPA'09)*, Mostaganem, Algérie, 19-21 Octobre 2009.

4. Exposés nationaux avec actes et sans comité de lecture

- a) **B.Meftah**, Le Temps dans les Réaux de Neurones, *1re journée d'étude sur les grilles de calcul et intelligence artificielle*, Mascara, Algérie, 11-12 Juin 2006.

5. Citations d'articles

- **Image Segmentation with Spiking Neuron Network** est cité par :

[[Aysha, et al., 2010](#)] V. Aysha, K. Balakrishnan et S.B Sundar Parallel Genetic Algorithm for Document Image Compression Optimization. In *proceedings of International Conference on Electronics and Information Engineering (ICEIE'2010)*, Kyoto, Japan, August, 1-3, 2010.

- **Image clustering with spiking neuron network** est cité par :

[[Nuño-Maganda et Torres-Huitzil, 2010](#)] M. Nuño-Maganda et C. Torres-Huitzil. A temporal coding hardware implementation for spiking neural networks. *ACM SIGARCH Computer Architecture News*, 38(4), September 2010

[[Dolotov et Bodyanskiy, 2009](#)] A. Dolotov et Y. Bodyanskiy. Analog-Digital Self-Learning Fuzzy Spiking Neural Network in Image Processing Problems. In *Image processing, Chapter 20: 357-380*. Yung-Sheng Chen. InTech. ISBN 978-953-307-026-1. 2009.

- **Combining Spiking Neural Network with Hausdorff Distance Matching for Object Tracking** est cité par :

[[Farhan, et al, 2011](#)] H. A. Farhan, H. H. Owaied et S. I. Al-Ghazi. Finding Shortest Path for Developed Cognitive Map Using Medial Axis. *World of Computer Science and Information Technology Journal (WCSIT)*, 1(2):17-25, 2011.

Bibliographie

Bibliographie

- [[Abbott et Sejnowski, 1998](#)] L. F. Abbott et T. J. Sejnowski. *Neural coding and distributed representations*. MIT Press, Cambridge, MA, 1998.
- [[Abbott, 1999](#)] L. F. Abbott. Lamicque's introduction of the integrate-and-fire neuron model, *Brain Research Bulletin*, 50(5-6): 303-304, 1999.
- [[Abeles, 1982](#)] M. Abeles. *Local cortical circuits: an electrophysiological study*. Springer, Berlin, 1982.
- [[Adibi et al., 2005](#)] P. Adibi, M. Meybodi et R. Safabakhsh. Unsupervised learning of synaptic delays based on learning automata in an RBF-like network of spiking neurons for data clustering. *Neurocomputing*, 64:335-357, 2005.
- [[Adrain, 1926](#)] E. Adrian. The impulses produced by sensory nerve endings. *Journal of Physiology*, 1926.
- [[Ahissar et Arieli 2001](#)] E. Ahissar et A. Arieli, "Figuring space by time", *Neuron*, Volume 32, pages 185-201, 2001.
- [[Alonso et al., 1990](#)] A. Alonso, M. de Curtisand et R. Linás. Postsynaptic hebbian and non-Hebbian long-term potentiation of synaptic efficacy in the entorhinal cortex slices and in the isolated adult guinea pig brain. In *Proceedings of the National Academy of Sciences, USA*, 87:9280-9284, 1990.
- [[Ambellouis et Cabestaing, 1996](#)] S. Ambellouis et F. Cabestaing. Motion analysis with a time delayed neural network. In *Proceedings of the IEEE Symposium on Robotics and Cybernetics (CESA'1996)*, Lille, France, 328-332, July 9-12, 1996.
- [[Arbid, 2003](#)] M. A. Arbid. *Handbook of brain theory and neural networks*. MIT press, second edition, 2003.
- [[Arifin et Asano, 2006](#)] A. Z. Arifin et A. Asano. Image segmentation by histogram thresholding using hierarchical cluster analysis. *Pattern Recognition Letters*, 27 : 1515-1521, 2006.
- [[Averbeck et Lee, 2004](#)] B. B. Averbeck et D. Lee. Coding and transmission of information by neural ensembles. *Trends Neuroscience*, 27:225-230, 2004.
- [[Baccouche et al., 2010](#)] M. Baccouche, F. Mamalet, C. Wolf, C. Garcia, et A. Baskurt. Action classification in soccer videos with long short-term memory recurrent neural networks. In *Proceedings of the 20th International Conference on Artificial Neural Networks (ICANN'2010)*, Thessaloniki, Greece, Part II, LNCS 6353: 154-159, September 15-18, 2010.
- [[Bakkum et al., 2008](#)] D. J. Bakkum, Z. C. Chao et S. M. Potter. Long-term activity dependent plasticity of action potential propagation delay and amplitude in cortical networks. *PLoS ONE*, 3(5), 2008.

- [Bandala et al., 2010] D. A. Bandala, P. L. Orellana et C. C. Sánchez. Bio-inspired Architecture for Human Detection. In *Proceedings of the 20th International Conference on Artificial Neural Networks (ICANN'2010)*, Thessaloniki, Greece, Part I, LNCS 6352: 188–192, September 15-18, 2010.
- [Barlow, 1972] H. B. Barlow. Single units and sensation: a neuron doctrine for perceptual psychology?. *Perception*, 1:371-94, 1972.
- [Barreto et Araújo, 2001] G. A. Barreto et A. F. R. Araújo. Time in self-organizing maps: An overview of models. *International Journal of Computer Research*, 10(2):139-179, 2001.
- [Barreto et al., 2003] G A. Barreto, A. F. R. Araújo et S. C. Kremer. A taxonomy of spatiotemporal connectionist networks revisited: The unsupervised case. *Neural Computation*, 15(6):1255-1320, 2003.
- [Belatreche et al., 2003] A. Belatreche, L. P. Maguire, M. McGinnity et Q. Xiang Wu. A method for supervised training of spiking neural networks. In *Proceedings of the IEEE Cybernetics Intelligence - Challenges and Advances (CICA'2003)*, Reading, UK: 39–44, September 17, 2003.
- [Belatreche et al., 2007] A. Belatreche, L. P. Maguire et M. McGinnity. Advances in design and application of spiking neural networks. *Soft Computing*, 11: 239–248, 2007.
- [Berry et al., 1997] M. J. Berry, D. K. Warlan et M. Meister. The structure and precision of retinal spike trains. In *Proceedings of the National Academy of Sciences of the USA*, 94(10):5411–5416, May 13, 1997.
- [Berthold, 1994a] M. R. Berthold. A time delay radial basis function network for phoneme recognition. In *Proceedings of the IEEE World Congress on Computational Intelligence (WCCI'94)*, Orlando, Florida, USA : 4470 - 4472, June 27-July 2, 1994.
- [Berthold, 1994b] M. R. Berthold. The TDRBF : A shift invariant radial basis function network. In *Proceedings of the Fourth Irish Neural Networks Conference (INNC'1994)*, Dublin, Irland: 7-12, September 11-12, 1994.
- [Bi et Poo, 1998] G.-Q. Bi et M.-M. Poo. Synaptic modifications in cultured hippocampal neurons: Dependence on spike timing, synaptic strength, and postsynaptic cell type. *The Journal of Neuroscience*, 18(24):10464-10472, 1998.
- [Bi, 2002] G.-Q. Bi. Spatiotemporal specificity of synaptic plasticity: cellular rules and mechanisms. *Biological Cybernetics*, 87:319–332, 2002.
- [Bialek et al., 1991] W. Bialek, F. Rieke, de R. R. Ruyter van Steveninck et D. Warland. Reading a neural code. *Science*, 252:1854-1857,1991.
- [Bogdanov, 2009] I. Bogdanov et G. Bogdanov. *Au commencement du temps*. Flamarion, Paris, 2009.

- [Bohte et al., 2000] S. Bohte, J. Kok et H. La Poutre. Spike-prop: error-backpropagation in multi-layer networks of spiking neurons. In *Proceedings of the European Symposium on Artificial Neural Networks (ESANN'2000)*, M. Verleysen, editor, Bruges, Belgium: 419–425, April 26–28, 2000.
- [Bohte et al., 2002a] S.M. Bohte, H. La Poutre et J. N. Kok. Unsupervised clustering with spiking neurons by sparse temporal coding and multilayer RBF networks. *IEEE Transactions on Neural Networks*, 13(2):426–435, 2002.
- [Bohte et al., 2002b] S. Bohte, J. Kok et H. L. Poutre. Error-backpropagation in temporally encoded networks of spiking neurons. *Neurocomputing*, 48:17–37, 2002.
- [Bohte et Kok, 2005] S. Bohte et J. Kok. Applications of Spiking Neural Networks. *Information Processing Letters - Special issue on applications of spiking neural networks*, 95(6):519–520, 2005.
- [Booij et Nguyen, 2005] O. Booij et H. Nguyen. A gradient decent rule for spiking neurons emitting multiple spikes. *Information Processing Letters*, 95(6):552–558, 2005.
- [Borst et Theunissen, 1999] A. Borst et F. E. Theunissen. Information theory and neural coding. *Natural Neuroscience*, 11:947–957, 1999.
- [Brennan, 2008] J. H. Brennan, *Voyage à travers le temps, un guide pour les débutants*. traduction M. H. Therrien, Bibliothèque national de canada, 2008.
- [Brette, 2007] R. Brette. Exact simulation of integrate-and-fire models with exponential currents. *Neural Computation*, 19(10):2604–2609, 2007.
- [Cariani, 2004] P. A. Cariani. Temporal codes and computations for sensory representation and scene analysis. *IEEE Transactions on Neural Networks*, 5:1100–1111, 2004.
- [Celenk, 1990] M. Celenk. A color clustering technique for image segmentation. *Computer vision Graphics and Image Processing*, 52:145–170, 1990.
- [Cinque et al. 2004] L. Cinque, G. Foresti et L. Lombardi. A clustering fuzzy approach for image segmentation. *Pattern Recognition*, 37:1797–1807, 2004.
- [Chappell et Taylor, 1993] G. J. Chappell et J. G. Taylor. The temporal Kohonen map. *Neural Networks*, 6(3):441–445, 1993.
- [Chappelier, 1996a] J. Chappelier. *RST : une architecture connexionniste pour la prise en compte de relations spatiales et temporelles*. Thèse de doctorat, Ecole nationale Supérieure des Télécommunications de Paris, Janvier 1996.
- [Chappelier et Grumbach, 1996b] J. Chappelier et A. Grumbach. A Kohonen Map for Temporal Sequences. In *Proceedings of the Conference Neural Networks and their Applications*, Marseille, France: 104–110, 1996.

- [Chappelier et al., 2001] J. Chappelier, M. Gori et A. Grumbach. Time in connectionist models. In *Sequence Learning: Paradigms, Algorithms and Applications*. Chapter 5. R. Sun and L. Giles editors, Lecture Notes in Artificial Intelligence 1828: 105-134, Springer, 2001.
- [Chardin, 2007] G. Chardin. *Qu'est-ce que la flèche du temps?*. Le Pommier (Eds), 2007.
- [Cheng, 1995] Y. Cheng. Mean Shift, Mode Seeking, and Clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(8): 790-799, 1995.
- [Cheng, 2000] H. Cheng. A hierarchical approach to color image segmentation using homogeneity. *IEEE Transactions on Image Processing*. 9(12):2071-2082, 2000.
- [Cheng et al., 2001] H.D. Cheng, X.H. Jiang, Y. Sun et J. Wang. Color image segmentation: advances and prospects. *Pattern Recognition*, 34:2259-2281, 2001.
- [Chung, 1998] A. Chung Tsoi. Recurrent neural network architectures: An overview. *Adaptive Processing of Sequences and Data Structures*, LNCS 1387:1-26, 1998.
- [Cocquerez et Philipp, 1995] J. Cocquerez et S. Philipp. Analyse d'images : filtrage et segmentation. Paris, Masson, 1995.
- [Culer et Davis, 2000] R. Cutler et L. Davis. Look who's talking: Speaker detection using video and audio correlation. In *Proceedings of the IEEE International Conference on Multimedia and Expo. (ICME'2000)*, New York, USA: 1589-1592, July 30-August 2, 2000.
- [Cybenko, 1989] G. Cybenko. Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems*, 2(4):303-314, 1989.
- [Delakis, 2006] E. Delakis. *Structuration multimodale des vidéos de tennis en utilisant des modèles segmentaux*. Thèse de doctorat, Université de Rennes 1, France, 2006.
- [Delorme et Thorpe, 2001] A. Delorme et S. J. Thorpe. Face identification using one spike per neuron: resistance to image degradations. *Neural Networks*, 14 : 795-803, 2001.
- [Diesmann et al., 1999] M. Diesmann, M. O. Gewaltig et A. Aertsen. Stable propagation of synchronous spiking in cortical neural networks. *Nature*, 6761: 529-533, 1999.
- [Di Rubeto et al., 2000] C. Di Rubeto, A. Dempster, S. Khan et B. Jarra. Segmentation of blood image using morphological operators. In *Proceeding of the 15th Int. Conference on Pattern Recognition*, Barcelona, Spain, 3: 397-400, September 3-8, 2000.
- [Dong et Xie, 2005] G. Dong et M. Xie. Color Clustering and Learning for Image Segmentation Based on Neural Networks. *IEEE Transactions on Neural Networks*, 16(4): 925-936, 2005.

- [Doya et al., 1993] K. Doya, A. I. Selverston et P. F. Rowat. A Hodgkin-Huxley type neuron model that learns slow non-spike oscillations. In *Proceedings of NIPS*: 566-573, 1993.
- [Dreyfus et al., 2005] G. Dreyfus, M. Samuelides, J.M. Martinez, M.B. Gordon, F. Badran, S. Thiria et L. Héroult. *Neural networks, methodology and applications*. Springer-Verlag Berlin Heidelberg, 2005.
- [Dreyfus, et al. 2010] G. Dreyfus, M. Samuelides, J.M. Martinez, M.B. Gordon, F. Badran et S. Thiria. *Apprentissage statistique, réseaux de neurones, cartes topologiques, machine à vecteurs support*. Eyrolles, 2010.
- [Eckhorn et al., 1990] R. Eckhorn, H.J. Reitboeck, M. Arndt et et P. Dicke. Feature linking via synchronization among distributed assemblies: Simulations of results from cat visual cortex. *Neural Computation*, 2(3):293-307, 1990.
- [Elman, 1990] J. Elman. Finding structure in time. *Cognitive Science*, 14:179-211, 1990.
- [Escobar et al., 2009] M. Escobar, G. S. Masson, T. Vieville et P.Kornprobst. Action Recognition Using a Bio-Inspired Feedforward Spiking Network. *International Journal of Computer Vision*, 82: 284-301, 2009.
- [Eyben et al., 2010] F. Eyben, M. Wöllmer, A. Graves, B. Schuller, E. Douglas Cowie et R. Roddy Cowie. On-line emotion recognition in a 3-D activation-valence-time continuum using acoustic and linguistic cues. *Journal on Multimodal User Interfaces*, 3:7-19, 2010.
- [FitzHugh, 1961] R. FitzHugh. Impulses and physiological states in theoretical models of nerve membrane. *Biophysical Journal*, 1:445-466, 1961.
- [Freixenet et al., 2002] J. Freixenet, X. Muñoz, D. Raba, J. Martí et X. Cufí. Yet Another Survey on Image Segmentation: Region and Boundary Information Integration. In *Proceeding of ECCV 2002*. A. Heyden et al. (Eds.), LNCS 2352: 408-422, 2002.
- [Gao et al., 1996] X. Z. Gao, X. M. Gao et S. J. Ovaska. A modified Elman neural network model with application to dynamical systems identification. In *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, 2: 1376-1381, October 14-17, 1996.
- [Gautrais et Thorpe, 1998] J. Gautrais et S. Thorpe. Rate coding versus temporal order coding: a theoretical approach. *Biosystems*, 48:57-65, 1998.
- [Gers et al., 2002] F. A. Gers, N. N. Schraudolph et J. Schmidhuber. Learning precise timing with LSTM recurrent networks. *Journal of Machine Learning Research*, 3:115-143, 2002.

- [Gerstner et al., 1996a] W. Gerstner, R. Kempter, J. L. Van Hemmen et H. Wagner. A neuronal learning rule for sub-millisecond temporal coding. *Nature*, 384:76-78, 1996.
- [Gerstner et al., 1996b] W. Gerstner, J. L. Van Hemmen et J. D. Cowan. What matters in neuronal locking. *Neural Computation*, 8: 1653-1676, 1996.
- [Gerstner et Kistler, 2002a] W. Gerstner et W. Kistler. *Spiking neuron models. single neurons, populations, plasticity*. Cambridge University Press, Cambridge, 2002.
- [Gerstner et Kistler, 2002b] W. Gerstner et W. Kistler. Mathematical formulations of Hebbian learning. *Biol Cybernet*, 87: 404-415, 2002.
- [Ghosh-Dastidar et Adeli, 2009] S. Ghosh-Dastidar et H. Adeli. Third generation neural networks: spiking neural networks. In *Advances in Computational Intelligence*, 61:167-178, 2009.
- [Gibbons, 2010] T. E. Gibbons. Reservoir computing: a rich area for undergraduate research. *Midwest Instruction and Computing Symposium*, University of Wisconsin-eau Claire, April 16-17, 2010.
- [Graves et al., 2007] A. Graves, S. Fernandez et J. Schmidhuber. Multi-dimensional recurrent neural networks. In *Proceedings of the International Conference on Artificial Neural Networks (ICANN'2007)*, Porto, Portugal, Part I, LNCS 4668:549-558, September 9-13, 2007.
- [Graves, 2008] A. Graves. *Supervised sequence labelling with recurrent neural networks*. Thèse de doctorat, Université Technique de Munich, 2008.
- [Guimarães et al., 2003] G. Guimarães, V. Lobo et F. Moura-Pires. A Taxonomy of self-organizing maps for temporal sequence processing. *Intelligent Data Analysis*, 7(4):269-290, 2003.
- [Guyon et al., 1991] I. Guyon, Y. Le cun, J. Denker et W. Hubbard. Design of a neural network character recogniser for a touch terminal. *Pattern Recognition*, 24(2):105-119, 1991.
- [Hammer et al., 2004] B. Hammer, A. Micheli, M. Strickert et A. Sperduti. A general framework for unsupervised processing of structured data. *Neurocomputing*, 57:3-35, 2004.
- [Hawking, 2009] S. Hawking. *Une belle histoire du temps*. Traduction B. Commengé, Flammarion (Eds), 2009.
- [Haykin, 1999] S. Haykin. *Neural networks: a comprehensive foundation*. second edition, Prentice-Hall, 1999.

- [Hebb, 1949] D. O. Hebb. *The organisation of behavior: a neuropsychological theory*. Wiley (Eds), NY, 1949.
- [Hindmarsh et Rose, 1984] J. L. Hindmarsh et R. M. Rose. A model of neuronal bursting using three coupled first order differential equations. In *Proceedings of the Royal Society*, London, 221:87–102, 1984.
- [Hochreiter et Schmidhuber, 1996] S. Hochreiter et J. Schmidhuber. *Bridging long time lags by weight guessing and Long Short-Term Memory*. In F. L. Silva, J. C. Principe, L. B. Almeida, Eds., *Frontiers in Artificial Intelligence and Applications*, 37:65-72, IOS Press, Amsterdam, Netherlands, 1996.
- [Hochreiter et Schmidhuber, 1997] S. Hochreiter et J. Schmidhuber. Long short-term memory. *Neural Computation*, 9(8):1735-1780, 1997.
- [Hochreiter, 1998] S. Hochreiter. The vanishing gradient problem during learning recurrent neural nets and problem solutions. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 6(2):1-10, 1998.
- [Hodgkin et Huxley, 1952] A. L. Hodgkin et A. F. Huxley. Currents carried by sodium and potassium ions through the membrane of the giant axon of *Loligo*. *Journal of Physiology*: 449-472, 1952.
- [Hopfield, 1982] J. Hopfield. Neural networks and physical Systems with emergent collective computational abilities. *National Academy of Science*, 79:2554–2558, 1982.
- [Hopfield, 1984] J. Hopfield. Neurons with graded response have collective computational properties like those of two-states neurons. In *Proceedings of the National Academy of Sciences*, 81:3088–3092, 1984.
- [Hopfield, 1995] J. Hopfield. Pattern recognition computation using action potential timing for stimulus representation. *Nature*, 376:33-6, 1995.
- [Howell et Buxton, 1998] A. J. Howell et H. Buxton. Learning identity with radial basis function networks. *Neurocomputing*, 20:15-34, 1998.
- [Howell et Buxton, 1999] A. J. Howell et H. Buxton. Gesture recognition for visually mediated interaction. In *Proceedings of the International Gesture Workshop (GW'1999)*, Yvette, France, LNAI 1739:141-151, March 17-19, 1999.
- [Howell et Buxton, 2001] A. J. Howell et H. Buxton. Visually mediated interaction using learnt gestures and camera control. In *Proceedings of the International Gesture Workshop (GW'2001)*, London, UK, LNCS 2298:272-284, April 18-20, 2001.
- [Howell et Buxton, 2002] A. J. Howell et H. Buxton. Active vision techniques for visually mediated interaction. *Image and Vision Computing*, 20:861-871, 2002.
- [Huang et Wu, 2009] W. Huang et J. Wu. Human action recognition using Recursive Self Organizing map and longest common subsequence matching. In *Proceedings of*

- the Workshop on Applications of Computer Vision (WACV'2009)*, Snowbird, USA: 1-6, December 7-8, 2009.
- [Izhikevich, 2003a] E. M. Izhikevich. Simple model of spiking neurons. *IEEE Transactions on Neural Networks*, 14(6):1569–1572, 2003.
- [Izhikevich et al., 2003b] E. M. Izhikevich, N. S. Desai, E. C. Walcott et F. C. Hoppensteadt. Bursts as a unit of neural information: selective communication via resonance. *Trends in Neuroscience*, 26:161-167, 2003.
- [Izhikevich, 2004a] E. M. Izhikevich. Which model to use for cortical spiking neurons?. *IEEE Transactions on Neural Networks*, 15(5): 1063–1070, 2004.
- [Izhikevich et al., 2004b] E. M. Izhikevich, J. A. Gally et G. M. Edelman. Spike-timing dynamics of neuronal groups. *Cerebral Cortex*, 14:933-944, 2004.
- [Izhikevich, 2006] E. M. Izhikevich. Polychronisation : Computation with spikes. *Neural Computation*, 18:245–282, 2006.
- [Jaeger, 2002] H. Jaeger. *Tutorial on training recurrent neural networks, covering BPPT, RTRL, EKF and the echo state network approach*. GMD Report 159, German National Research Center for Information Technology, 2002.
- [James et Miikkulainen, 1995] D. L. James et R. Miikkulainen. Sardnet: A self-organising feature map for sequences. In D.S. Touretzky G. Tesauro and T.K. Leen, editors, *Advances in Neural Processing Systems*, 7:577–584, Cambridge, MA: MIT Press, 1995.
- [Johnson et al., 1999] J. L. Johnson, M. L. Padgett et O. Omidvar. Overview of pulse coupled neural networks (PCNN). *IEEE Transactions on Neural Networks*, 10(3):461-463, 1999.
- [Jordan, 1986] M. Jordan. Attractor dynamics and parallelism in a connectionist sequential machine. In *Proceedings of the 8th annual conference on Cognitive Science Society*: 531–546, 1986.
- [Kalinli et Sagioglu, 2006] A. Kalinli et S. Sagioglu. Elman network with embedded memory for system identification. *Journal of Information Science and Engineering*, 22: 1555-1568, 2006.
- [Kandel et al., 2000] E. R. Kandel, J.H. Schwartz et T.M. Jessell. *Principles of neural science*. 4th edition McGraw-Hill, New York, 2000.
- [kempster et al., 1999] R. Kempster, W. Gerstner et J. van Hemmem. Spike-based compared to rate-based hebbian learning. In *Proceedings of the Conference on Advances in neural information processing systems II*, 1999.
- [Kepecs et al., 2002] A. Kepecs, M. Van Rossum, S. Song et J. Tegner. Spike-timing-dependent plasticity: common themes and divergent vistas. *Biological Cybernetics*, 87:446–458, 2002.

- [Kim et Hwang, 2002] C. Kim et J-N Hwang. Fast and Automatic Video Object Segmentation and Tracking for Content-Based Applications. *IEEE Transactions on Circuits and Systems for Video Technology*, 12(2):122-129, 2002.
- [Kistler et al., 1997] W. M. Kistler, W. Gerstner et J. L. Van Hemmen. Reduction of the Hodgkin-Huxley equations to a single-variable threshold model. *Neural Computation*, 9(5):1015-1045, 1997.
- [Kistler et van Hemmen, 2000] W. M. Kistler et J. van Hemmen. Modeling synaptic plasticity in conjunction with the timing of pre-and postsynaptic action potentials. *Neural Computation*, 12:385-405, 2000.
- [Kistler, 2002] W. Kistler. Spike-timing dependent synaptic plasticity: a phenomenological framework. *Biological Cybernetics*, 87:416-427, 2002.
- [Knesek, 1999] E.A. Knesek. Roche image analysis system. *Acta Cytologica*, 40(1):60-66, 1996.
- [Kohonen, 1982] T. Kohonen. Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43(1): 59-69, 1982
- [Kopecz, 1995] K. Kopecz. Unsupervised learning of sequences on maps with lateral connectivity. In *Proceedings of the International Conference on Artificial Neural Networks (ICANN'1995)*, Nanterre, France, Fogelman-Soulié and P. Gallinari, ed., 2:431-436, 1995.
- [Koskela et al., 1998] T. Koskela, J. Heikkonen et K. Kaski. Time series prediction using recurrent SOM with local linear models. *International Journal of Knowledge-Based Intelligent Engineering Systems*, 2(1):60-68, 1998.
- [Kwok et al., 1994] D. P. Kwok, P. Wang et K. Zhou. Process identification using a modified Elman neural net. In *Proceedings of the International Symposium on Speech, Image Processing and Neural Networks (ISSIPNN'1994)*: 499-502, 1994.
- [Lambert et Macaire, 2000] P. Lambert et L. Macaire. Filtering and segmentation: the specificity of color images. *Color in Graphics and Image Processing*, 1: 57-71, 2000.
- [Lapicque, 1907] L. Lapicque. Recherches quantitatives sur l'excitation électrique des nerfs traitée comme une polarization. *J. Physiol. Pathol. Gen.* 9:620-635, 1907.
- [Le Callet et al., 2006] P. Le Callet, C. Viard-Gaudin et D. Barba, A convolutional neural network approach for video quality assessment, *IEEE Transaction on Neural Networks*, 17(5):1316-1327, 2006.
- [Legenstein et al., 2003] R. A. Legenstein, H. Markram et W. Maass. Input prediction and autonomous movement analysis in recurrent circuits of spiking neurons. *Reviews in the Neurosciences*, 14(1-2):5-19, 2003.
- [Leibold et Hemmen, 2001] C. Leibold et J.L. Hemmen. Temporal receptive fields, spikes, and Hebbian delay selection. *Neural Networks*, 14(6-7):805-813, 2001.

- [Lezoray et Cardot, 2002] O. Lezoray et H. Cardot. Cooperation of pixel classification schemes and color watershed: a study for microscopical images. *IEEE Transactions on Images Processing*, 11(7): 738–789, 2002.
- [Lim et Lee, 1990] Y. Lim et S.U. Lee. On the color image segmentation algorithm based on the thresholding and the fuzzy c-means techniques. *Pattern Recognition*, 23(9): 935–952, 1990.
- [Lin, 1994] D. T. Lin. *The adaptable time delay neural network characterisation and application to pattern recognition, prediction and signal processing*. PhD thesis, Institute for Systems Research (ISR), University of Maryland, USA, 1994.
- [Lin et al., 1995] D. L. Lin, J. E. Dayhoff et P. A. Ligomenides. Trajectory production with the adaptive time-delay neural network. *Neural Networks*, 8(3):447–461, 1995.
- [Lindblad et Kinser, 2005] T. Lindblad et J. M. Kinser. *Image Processing Using Pulse-Coupled Neural Networks*. 2nd ed., Springer, 2005.
- [Little, 1974] W. A. Little. The existence of persistent states in the brain. *Mathematical Biosciences*, 19: 101–120, 1974.
- [Lukosevicius et Jaeger, 2009] M. Lukosevicius et H. Jaeger. Reservoir computing approaches to recurrent neural network training. *Computer Science Review*, 3(3):127-149, 2009.
- [Maass, 1997] W. Maass. Network of spiking neurons: the third generation of neural network models. *Neural Networks*, 10(9):1659–1671, 1997.
- [Maass et Bishop, 2001] W. Maass et C. M. Bishop. *Pulsed neural networks*. MIT Press, Cambridge, MA, 2001.
- [Maass et al., 2002] W. Maass, T. Natschläger et H. Markram. Real-time computing without stable states: a new framework for neural computation based on perturbations. *Neural Computation*, 14:2531-2560, 2002.
- [Maass et al., 2004] W. Maass, T. Natschlaeger et H. Markram. *Computational models for generic cortical microcircuits*. In J. Feng, editor, *Computational Neuroscience: A Comprehensive Approach*, Chapman and Hall/CRC, Boca Raton, 2004.
- [Macqueen, 1967] J. Macqueen. Some methods for classification and analysis of multivariate observations. In *Berkeley Symposium on Mathematical Statistics and Probability*, 1:281–297, 1967.
- [Madhubanti et Amitava, 2008] M. Madhubanti et C. Amitava. A hybrid cooperative-comprehensive learning based PSO algorithm for image segmentation using multilevel thresholding. *Expert Systems with Applications*, 34(2): 1341-1350, 2008.

- [Markram et al., 1997] H. Markram, J. Lübke, M. Frotscher et B. Sakmann. Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science*, 275(5297): 213–215, 1997.
- [Markram, 2002] H. Markram. Structural and functional principles of neocortical microcircuits. *Special issue on Foundations of Information Processing of TELEMATIK*, 8(1), 2002.
- [Martin et al., 2001] D. Martin, C. Fowlkes, D. Tal et L. Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *Proceedings of the 8th International Conference On Computer Vision (ICCV'2001)*, Vancouver, British Columbia, Canada. 2:416–423, 7–14 July 2001, 2001.
- [McQueen et al., 2003] T. A. McQueen, A. A. Hopgood, J. A. Tepper et T. J. Allen. A recurrent self-organizing map for temporal sequence processing. In *Proceedings of the 4th International Conference on Recent Advances in Soft Computing (RASC2002)*, Nottingham, United Kingdom: 43-49, December 12-13, 2002.
- [Meftah et al., 2010a] B. Meftah, O. Lezoray et A. Benyettou. Segmentation and edge detection based on spiking neuron networks. *Neural Processing Letters*, 32(2):131-146, 2010.
- [Meftah et al., 2010b] B. Meftah, O. Lezoray, M. Lecluse, et A. Benyettou. Cell microscopic segmentation with spiking neuron networks. In *Proceedings of the 20th International Conference on Artificial Neural Networks (ICANN'2010)*, Thessaloniki, Greece, K. Diamantaras, W. Duch, L.S. Iliadis (Eds.), Part I, LNCS 6352:117–126, September 15-18, 2010.
- [Meftah et al., 2008] B. Meftah, A. Benyettou, O. Lezoray et W. Qing Xiang. Image clustering with spiking neuron network. In *Proceedings of the International Joint Conference on Neural Networks (IJCNN'2008), part of the IEEE World Congress on Computational Intelligence (WCCI'2008)*, Hong Kong: 681-685, IEEE Press, Piscataway, June, 1-6, 2008.
- [Meurie et al., 2005] C. Meurie, O. Lezoray, C. Carrier et A. Elmoataz. Combination of Multiple Pixel Classifiers for Microscopic Image Segmentation. *International Journal of Robotic and Automation, Special Issue on Color Image and Analysis for Machine Vision*, 20(2):63–69, 2005.
- [Ming et al., 2002] Y. Ming-Hsuan, N. Ahuja et M. Tabb. Extraction of 2D motion trajectories and its application to hand gesture recognition. *IEEE Transactions on Pattern analysis and Machine Intelligence*, 24(8):1061-1074, 2002.
- [Moore, 2002] S. C. Moore. *Back-propagation in spiking neural networks*. MSc thesis, University of Bath, UK, 2002.

- [Morris et Lecar, 1981] C. Morris et H. Lecar, Voltage oscillations in the barnacle giant muscle fiber, *Biophysical Journal*, 35:193–213, 1981.
- [Morrison et al., 2007] A. Morrison, A. Aertsen et M. Diesmann. Spike timing-dependent plasticity in balanced random networks. *Neural Computation*, 19:1437–1467, 2007.
- [Mouroutis et al. 1998] T. Mouroutis, S.J. Roberts et A.A. Bharath. Robust cell nuclei segmentation using statistical modeling. *BioImaging*, 6: 79–91, 1998.
- [Murashov, 2004] D. Murashov. Two-level method for segmentation of cytological images using active contour model. In *Proceeding of the 7th Int. Conference on Pattern Recognition and Image Analysis, (PRIA'2004)*, St. Petersburg, Russian Federation, 3:814–817, October 18-23, 2004.
- [Nagumo et al., 1962] J. Nagumo, S. Arimoto et S. Yoshizawa. An active pulse transmission line simulating nerve axon. In *Proceedings of the Institute of Radio Engineers*, 50:2061–2070, 1962.
- [Natschläger et Ruf, 1998] T. Natschläger et B. Ruf. Spatial and temporal pattern analysis via spiking neurons. *Network: Computation in Neural Systems*, 9(3):319–332, 1998.
- [Natschläger, 1999] T. Natschläger. *Efficient computation in networks of spiking neurons - simulations and theory*. PhD thesis, Institute for Theoretical Computer Science, Graz University of Technology, Austria, 1999.
- [Natschläger et al., 2002] T. Natschläger, W. Maass et H. Markram. The liquid computer, a novel strategy for real-time computing on time series. *Special issue on Foundations of Information Processing of TELEMATIK*, 8(1):32–36, 2002.
- [Nelson et al., 1993] P. G. Nelson, R. D. Fields, C. Yu et Y. Liu. Synapse elimination from the mouse neuromuscular junction in vitro: A non-Hebbian activity-dependent process. *Journal of Neurobiology*, 24(11):1517-1530, 1993.
- [Ourdighi et Benyettou, 2010] A. Ourdighi et A. Benyettou. An adaptive time-delay neural network training using parallel genetic algorithm in time series prediction and classification. *Journal of applied sciences*, 10(18):2115-2120, 2010.
- [Pajares, 2006] G. Pajares. A Hopfield Neural Network for Image Change Detection. *IEEE Transactions on Neural Networks*, 17(5):1250-1264, 2006.
- [Paugam-Moisy, 2006] H. Paugam-Moisy. *Spiking neuron networks a survey*. Idiap-rr-11-2006, IDIAP, 2006.
- [Paugam-Moisy et Bohte, 2009] H. Paugam-Moisy et S. M. Bohte. *Computing with spiking neuron networks*. In J. Kok, T. Heskes (Eds) *Handbook of natural computing*, Springer Verlag, Heidelberg, 2009

- [Penrose et al., 2003] R. Penrose, S. Hawking et M. Lachèze-Rey. *La Nature de l'espace et du temps*. Traduction F. Balibar, Gallimard (Eds), 2003.
- [Perkell et Bullock, 1968] D. H. Perkel et P. H. Bullock. Neural coding. *Neuroscience Research Program*, 3:221-248, 1968.
- [Petersen et al., 2001] R. S. Petersen, S. Panzeri, et M. E. Diamond. Population coding of stimulus location in rat somatosensory cortex. *Neuron*, 32:503-14, 2001.
- [Pervej et al., 2008] R. Pervej, N. Akter, A. R. Chowdhury et Md. A. Bhuiyan. On image recognition using hopfield net and ellipse fitting. *International Journal of Computer Science and Network Security*, 8(5):148-153, 2008.
- [Petrou et al., 1998] M. Petrou, L. Shafarenko et J. Kittler. Histogram-based segmentation in a perceptually uniform color space. *IEEE Transactions on Image Processing*, 7:1354-1358, 1998.
- [Ponulak et kasiński, 2006] F. Ponulak et A. kasiński. Comparison of supervised learning methods for spike time coding in spiking neural networks. *International Journal of Applied Mathematics and Computer Science*, 16(1):101-113, 2006.
- [Ponulak et kasiński, 2010] F. Ponulak et A. kasiński. Supervised Learning in Spiking Neural Networks with ReSuMe: Sequence Learning, Classification, and Spike Shifting. *Neural Computation*, 22(2):467-510, 2010.
- [Pouget et al., 2003] A. Pouget, P. Dayan et R. S. Zemel. Inference and computation with population codes. *Annual Review of Neuroscience*, 26:381-410, 2003.
- [Privitera et Morasso, 1994] C. M. Privitera et P. Morasso. The analysis of continuous temporal sequences by a map of sequential leaky integrators. In *Proceedings of the International Conference on Artificial Neural Networks (ICANN'1994)*, Sorrento, Italy: 3127-3130, 1994.
- [Psarrou et al., 1995] A. Psarrou, S. Gong, H. Buxton. Modelling spatio-temporal trajectories and face signatures on partially recurrent neural networks. *Neural Networks*, 5: 2226-2231, 1995.
- [Rieke et al., 1996] F. Rieke, D. K. Warland, R. R. de Ruyter van Steveninck et W. Bialek. *Spikes: exploring the neural code*. MIT Press, Cambridge, MA, 1996.
- [Risinger et Kaikhah, 2008] L. Risinger et K. Kaikhah. Motion detection and object tracking with discrete leaky integrate-and-fire neurons. *Applied Intelligence*, 29(3): 248-262, 2008.
- [Rochel, 2004] O. Rochel. *Une approche événementielle pour la modélisation et la simulation de réseaux de neurones impulsifs*. Thèse de doctorat, université Henri Poincaré, Nancy I, France, 2004.

- [Rosenblatt, 1958] F. Rosenblatt. The perceptron : a probabilistic model for information storage and organization in the brain. *Psychological Review*, 65:386-408, 1958.
- [Rovelli, 2006] C . Rovelli.. *Qu'est-ce que le temps ? Qu'est-ce que l'espace ?*. Essai (Eds), 2006.
- [Rubin et al., 2001] J. Rubin, D. D. Lee et H. Sompolinsky. Equilibrium properties of temporally asymmetric Hebbian plasticity. *Physical Review Letters*, 86(2):364-367, 2001.
- [Ruf et Schmitt, 1997] B. Ruf et M. Schmitt. Learning temporally encoded patterns in networks of spiking neurons. *Neural Processing Letters*, 5(1):9-18, 1997.
- [Ruf et Schmitt, 1998] B. Ruf et M. Schmitt. Self-organisation of spiking neurons using action potential timing. *IEEE Transactions on Neural Networks*, 9(3):575-578, 1998.
- [Rumelhart et al., 1986a] D. E. Rumelhart, G. Hinton et R. Williams. Learning representations by back-propagating errors. *Nature*, 323:533-536, 1986.
- [Rumelhart et McClelland, 1986b] D.E. Rumelhart et J. L. McClelland. *Parallel Distributed Processing, Exploration in the Micro-Structure of Cognition*. Bradford MIT Press, 1986.
- [Salhi et al., 2009] M. S. Salhi, N. Arous et N. Ellouze. Principal temporal extensions of SOM: Overview. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, 2(4):61-84, 2009.
- [Sato et al., 2003] S. Sato, T. Umezaki, et K. Inagaki. A recurrent neural network approach to rear vehicle detection which considered state dependency. *Journal of Systemics, Cybernetics and Informatics*, 1(4):72-77, 2003.
- [Schrauwen et Campenhout, 2004] B. Schrauwen et J. V. Campenhout. Improving spikeprop: enhancements to an error-backpropagation rule for spiking neural networks. In *Proceedings of the 15th ProRISC Workshop*, 2004.
- [Schrauwen et Van Campenhout, 2006] B. Schrauwen et J. Van Campenhout. Backpropagation for population-temporal coded spiking neural networks. In *Proceedings of the International Joint Conference on Neural Networks (IJCNN'2006)*, Vancouver, Canada: 3463-3470, July 16-21, 2006.
- [Schrauwen et al., 2007] B. Schrauwen, D. Verstraeten et J. Van Campenhout. An overview of reservoir computing: theory, applications and implementations. In *Proceedings of the European Symposium on Artificial Neural Networks (ESANN'2007)*, Bruges, Belgium: 471-482, April 25-27, 2007.
- [Schuster et Paliwal, 1997] M. Schuster et K. K. Paliwal. Bidirectional recurrent neural networks. *IEEE Transactions on Signal Processing*, 45(11): 2673-2681, 1997.

- [Senn et al., 2002] W. Senn, M. Schneider et B. Ruf. Activity-dependent development of axonal and dendritic delays, or, why synaptic transmission should be unreliable. *Neural Computation*, 14(3): 583–619, 2002.
- [Siroh et Miikkulainen, 1997] J. Siroh et R. Miikkulainen. Topographic receptive fields and patterned lateral interaction in a self-organizing model of the primary visual cortex. *Neural Computation*, 9(3): 577–594, 1997.
- [Shadlen et Newsome, 1994] M. N. Shadlen et W. T. Newsome. Noise, neural codes and cortical organization. *Current Opinion in Neurobiology*, 4: 569–579, 1994.
- [Song et al., 2000] S. Song, K. D. Miller et L. F. Abbott. Competitive hebbian learning through spike-timing dependent synaptic plasticity. *Nature Neuroscience*, 3: 919–926, 2000.
- [Song et Abbott, 2001] S. Song et L. F. Abbott. Cortical development and remapping through spike timing-dependent plasticity. *Neuron*, 32: 339–350, 2001.
- [Stagge et Sendhoff, 1997] P. Stagge et B. Sendhoff. An extended Elman net for modeling time series. In *Proceedings of the International Conference on Artificial Neural Networks, LNCS 1327*, Springer-Verlag : 427–432, 1997.
- [Stratton et Wiles, 2007] P. Stratton et J. Wiles. *Why spiking neurons*. Technical Report TS-2007001, University of Queensland, Australia, 2007.
- [Strickert et Hammer, 2005] M. Strickert et B. Hammer. Merge SOM for temporal data. *Neurocomputing*, 64: 39–71, 2005.
- [Sulehria et Zhang, 2007] H. K. Sulehria et Y. Zhang. Hopfield neural networks - a survey. In *Proceedings of the 6th International Conference on Artificial Intelligence, Knowledge Engineering and Data Bases (WSEAS'2007)*, Corfu Island, Greece: 125–130, February 16–19, 2007.
- [Sulehria et Zhang, 2008] H. K. Sulehria et Y. Zhang. Study on the capacity of Hopfield neural networks. *Information Technology Journal*, 6(4): 684–688, 2008
- [Sun et al., 1995] Y. Sun, J. G. Li et S. Y. Yu. Improvement on performance of modified hopfield neural network for image restoration. *IEEE Transactions on Image Processing*, 4(5): 688–692, 1995.
- [Sun, 2000] Y. Sun. Hopfield neural network based algorithms for image restoration and reconstruction. *IEEE Transactions on Signal Processing*, 48(7), 2000.
- [Szilas, 1997] N. Szilas. *Les réseaux de neurones récurrents : une revue critique*. Rapport de recherche IMAG RR 972-I, 1997.
- [Swadlow, 1988] H. A. Swadlow. Efferent neurons and suspected interneurons in binocular visual cortex of the awake rabbit: Receptive fields and binocular properties. *Journal of Neurophysiology*, 59(4): 1162–1187, 1988.

- [[Tai et Huang, 2009](#)] S. Tai et H. Huang. Facial expression recognition in video sequences advances. In *Proceedings of the 6th International Symposium on Neural Networks (ISNN'2009)*, Wuhan, China, 5553: 1026-1033, May 26-29, 2009.
- [[Tan et De Silva, 2003](#)] C. H. Tan et L. De Silva. Human activities recognition by head movement using partial recurrent neural network. In *Proceedings of the Visual communications and image processing Conference*, Lugano, Italie, 5150: 2007-2014, July 8-11, 2003.
- [[Thorne, 2009](#)] K. S. Thorne. *Trous noirs et distorsions du temps: L'héritage sulfureux d'Einstein*. Traduction A. Bouquet et J. Kaplan, Flammarion (Eds), 2009.
- [[Thorpe, 1990](#)] S. J. Thorpe. *A highly efficient coding scheme for neural networks*. In *Parallel processing in neural systems and computers*, R. Eckmiller, G. Hartmann and G. Hauske (Eds.), Elsevier, 91-94, 1990.
- [[Tiño et Mills, 2005](#)] P. Tiño et A. J. Mills. Learning beyond finite memory in recurrent networks of spiking neurons. In L. Wang, K. Chen, and Y. Ong, editors, *Advances in Natural Computation – (ICNC'2005), Lecture Notes in Computer Science*, 3611: 666-675, Springer-Verlag, 2005.
- [[VanRossum et al., 2000](#)] M. C. W. VanRossum, G. -Q. Bi et G. G. Turrigiano. Stable hebbian learning from spike timing-dependent plasticity. *Journal of Neuroscience*, 20(23): 8812-8821, 2000.
- [[VanRullen et al., 1998](#)] R. VanRullen, J. Gautrais, A. Delorme et S. J. Thorpe. Face processing using one spike per neurone. *Biosystems*, 48: 229-239, 1998.
- [[VanRullen et al., 2005](#)] R. VanRullen, R. Guyonneau et S. J. Thorpe. Spike times make sense. *Trends in Neurosciences*, 28(1): 1-4, 2005.
- [[Varsta et al., 1997](#)] M. Varsta, J. Millán et J. Heikkonen. A recurrent self-organizing map for temporal sequence processing. In *Proceedings of the International Conference on Artificial Neural Networks (ICANN'1997)*, Lausanne, France, LNCS 1327: 421-426, October 8-10, 1997.
- [[Vaucher, 1993](#)] G. Vaucher. Un modèle de neurone artificiel conçu pour l'apprentissage non supervisé de séquences d'événements asynchrones. *VALGO*, 1: 66-107, 1993.
- [[Vesanto, 1997](#)] J. Vesanto. Using the SOM and local models in time-series prediction. In *Proceedings of the Workshop on self-Organizing Maps*, Espoo, Finland: 209-214, June 4-6, 1997.
- [[Victor et Purpura, 1996](#)] J. D. Victor et K. P. Purpura. Nature and precision of temporal coding in visual cortex: a metric-space analysis. *Journal of Neurophysiology*, 76: 1310-26, 1996.

- [Voegtlin, 2002] T. Voegtlin. Recursive self-organizing maps. *Neural Networks*, 15: 979–991, 2002.
- [Waibel et Lang, 1989] G. Waibel et K. Lang. Phoneme recognition using time delay neural network. *IEEE Transaction in Acoustics, Speech and Signal Processing*, 37(3): 328–339, 1989.
- [Wang et al., 2010] Z. Wang, Y. Ma, F. Cheng et L. Yang. Review of pulse-coupled neural networks. *Image and Vision Computing*, 28: 5–13, 2010.
- [Wang et Buzsaki, 1996] X. Wang et G. Buzsaki. Gamma oscillation by synaptic inhibition in a hippocampal interneuronal network model. *Journal of Neuroscience*, 16(20) : 6402–6413, 1996.
- [Wang et Sun, 2010] X. Wang et Y. Sun. A color- and texture-based image segmentation algorithm. *Machine Graphics & Vision*, 19(1): 3–18, 2010.
- [Wang et al., 2011] X-Y. Wang, T. Wang et J. Bu. Color image segmentation using pixel wise support vector machine classification. *Pattern Recognition*, 44:777-787, 2011.
- [Werbos, 1990] P. J. Werbos. Backpropagation through time: what is does and how to do it. In *Proceedings of the IEEE*, 78(10): 1550-1560, October 6, 1990.
- [Wiemer, 2003] J. C. Wiemer. The time-organized map (TOM) algorithm: extending the self-organizing map (SOM) to spatiotemporal signals. *Neural networks*, 15(5): 1143-1171, 2003.
- [Williams et Zipser, 1989] R. Williams et D. Zipser. A learning algorithm for continually running fully recurrent neural networks. *Neural Computation*, 1(3): 270–280, 1989.
- [Wohler et Anlauf, 1999a] C. Wohler et J. K. Anlauf. A Time delay neural network algorithm for estimating image-pattern shape and motion. *Image and Vision Computing*, 17(3-4): 281-294, 1999.
- [Wohler et Anlauf, 1999b] C. Wohler et J. K. Anlauf. An adaptable time-delay neural-network algorithm for image sequence analysis. *Neural Networks*, 10(6): 1531-1536, 1999.
- [Wohler et al., 2000] C. Wohler, U. Kresel et J. K. Anlauf. Pedestrian recognition by classification of image sequences - global approaches vs. local spatio-temporal processing. In *Proceedings of the 15th International Conference on Pattern Recognition (ICPR'2000)*, 2: 2540, 2000.
- [Wu et al., 2000] H.S. Wu, J. Barba et J. Gil. Iterative thresholding for segmentation of cells from noisy images. *Journal of Microsc*, 197: 296–304, 2000.

- [Xiang Wu et al., 2005] Q.X.Wu, T.M. McGinnity, L.P. Maguire, A. Belatreche et B. Glackin. Adaptive co-ordinate transformation based on spike timing-dependent plasticity learning paradigm. *Lecture Notes in Computer Science*, 3610 :420-429, 2005.
- [Xiang Wu et al., 2006] Q.X. Wu, T.M. McGinnity, L.P. Maguire, B. Glackin et A. Belatreche. Learning mechanism in networks of spiking neurons. *Studies in Computational Intelligence, Springer-Verlag*, 35 :171-197, 2006.
- [Xiang Wu et al., 2007] Q.X. Wu, M. McGinnity, L. Maguire, A. Belatreche et B. Glackin. Edge Detection Based on Spiking Neural Network Model, *Lecture Notes in Computer Science*, 4682:26-34, 2007.
- [Xiang Wu et al., 2008] Q. Xiang Wu, T.M. McGinnity, L. Maguire, J. Cai et G.D. Valderrama-Gonzalez. Motion Detection Using Spiking Neural Network Model. In *Proceedings of the 4th international conference on Intelligent Computing: Advanced Intelligent Computing Theories and Applications - with Aspects of Artificial Intelligence (ICIC'2008)*, Shanghai, China, LNCS 5227 : 76-83, September 15-18, 2008.
- [Xin et Embrechts, 2001] J. Xin et M. J. Embrechts. Supervised learning with spiking neuron networks. In *Proceedings of the IEEE International Joint Conference on Neural Networks, (IJCNN'2001)*, Washington: 1772-1777, July 15-19, 2001.
- [Yang et al., 1997] R. Yang, S. Ambellouis et F. Cabestaing. Détection d'obstacles dans une séquence d'images de rétrovision par analyse du mouvement. In *Proceedings du seizième colloque GRETSI*, Grenoble : 913-916, 15-19 septembre, 1997.
- [Yedjour et al., 2011] H. Yedjour, B. Meftah, D. Yedjour et A. Benyettou. Combining spiking neural network with Hausdorff distance matching for object tracking. *Asian Journal of Applied Sciences*, 4(1): 63-71, 2011.
- [Yilmaz et al, 2006] A. Yilmaz, O. Javed et M. Shah. Object Tracking: A Survey. *ACM Computing Surveys*, 38(4): 1-45, 2006.
- [Young et al., 1997] S. S. Young, P. D. Scott et N.M. Nasrabadi. Object recognition using multi-layer Hopfield neural network. *IEEE Transactions on Image Processing*, 6(3): 357-372, 1997.
- [Zehraoui et Bennani, 2004] F. Zehraoui et Y. Bennani. M-SOM: matricial self organizing map for sequences clustering and classification. *Proceedings of the IEEE International Joint Conference on Neural Network (IJCNN'04)*, Budapest, Hungary, July 25-29, 2004.
- [Zou et Bhanu, 2005] X. Zou et B. Bhanu. Tracking Humans using Multi-modal Fusion. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'2005)*, San Diego, CA, USA, 3, June 20-25, 2005.
- [Zucker, 1976] S. Zucker. Region growing: childhood and adolescence. *Computer Graphics and Image Processing*, 5: 382-399, 1976.

Résumé

Les réseaux de neurones classiques peuvent offrir des solutions très intéressantes dans des applications de reconnaissance des formes ou approximation de fonctions mais ne peuvent en aucun cas être appliqués sur des données où le temps joue un rôle déterminant dans la résolution du problème.

Comment ces réseaux de neurones statiques peuvent être adaptés pour être dynamiques ? Quelles sont les architectures de réseaux temporels qui existent en littérature ? Comment est mené l'apprentissage pour la prise en compte de la dimension temporelle des données ? Nous essaierons de répondre à ces questions à travers cette thèse consacrée à la représentation du temps dans les réseaux de neurones en correspondance avec le traitement d'images animées. Nos intérêts de recherche concernent principalement deux domaines complémentaires, à savoir, d'une part, l'étude et la synthèse des réseaux de neurones temporels et d'autre part, la mise en application des réseaux de neurones impulsionsnels en traitement d'images animées.

Mots clés : Réseaux de neurones temporels, Réseaux de neurones impulsionsnels, Apprentissage, Codage, Traitement d'images animées, Détection de contours, Segmentation, Suivi d'objets.

ملخص

الشبكات العصبية التقليدية يمكن أن توفر حلاً مثيرة جداً للاهتمام في تطبيقات التعرف على نمط أو وظيفة تقريب ولكن لا يمكن بأي حال من الأحوال أن تطبق على البيانات حيث الوقت يلعب دوراً حاسماً في حل المشكلة.

كيف يمكن تكييف هذه الشبكات العصبية الثابتة لتكون ديناميكية؟ ما هي تصميمات الشبكات الزمنية الموجودة؟ كيف يتم التعلم الذي يؤدي إلى مراعاة البعد الزمني للبيانات؟ سنحاول الإجابة على هذه الأسئلة من خلال هذه المذكرة من خلال تمثيل الوقت في الشبكات العصبية مع تطبيقاتها في معالجة الصور. اهتماماتنا البحثية تمحورت أساساً في مجالين متكاملين، وهما، أولاً، دراسة وتحليل الشبكات العصبية الزمنية وثانياً، تنفيذ الشبكات العصبية الاندفاعية في معالجة الصور المتحركة.

كلمات مفتاحية : الشبكات العصبية الزمنية، الشبكات العصبية الاندفاعية، التعلم، والترميز، معالجة الصور، الكشف عن الحافة، تجزئة الصور، تتبع الأشياء.

Abstract

Conventional neural networks can provide very interesting solutions in applications of pattern recognition or function approximation but can in no way be applied to data where time plays a decisive role in resolving the problem.

How these static neural networks can be adapted to be dynamic? What are the temporal network architectures that exist in literature? How is learning led to the consideration of the temporal dimension of data? We try to answer these questions through this thesis on the representation of time in neural networks in correspondence with image processing.

Our research interests mainly concern two complementary areas, namely, first, the study and synthesis of temporal neural networks and secondly, the implementation of spiking neural networks in image processing.

Keywords: Temporal neural networks, spiking neural networks, Learning, Coding, Image processing, Edge detection, Segmentation, Object tracking.