

N° d'ordre : 1728

50376
1996
119

THESE

présentée à

L'UNIVERSITE DES SCIENCES ET TECHNOLOGIES DE LILLE

pour obtenir le titre de

DOCTEUR DE L'UNIVERSITE

*Spécialité : **Electronique***

par

Abdelaziz YACOUBI

**CONTRIBUTION DE L'APPROCHE NEUROMIMETIQUE
A L'ANALYSE DU MOUVEMENT DANS UNE SEQUENCE D'IMAGES**

Soutenue le 05 Avril 1996 devant la commission d'examen composée de:

MM.

J. G. Postaire

Président

H. Vu-Thien

Rapporteur

M. Benjelloun

Rapporteur

V. Devlaminck

Co-Directeur de Thèse

N. Kerkeni

Examineur

J. P. Dubus

Examineur



Remerciements

Je tiens à exprimer ici ma reconnaissance aux membres du jury:

Monsieur J. G. Postaire, Professeur au Centre d'Automatique de Lille, qui m'a fait l'honneur de présider le jury de cette thèse,

Messieurs H. Vu-Thien, Professeur au Conservatoire National des Arts et Métiers de Paris et M. Benjelloun, Docteur d'état à l'Institut National des Télécommunications d'IVRY qui ont accepté d'être rapporteurs de ce travail,

Monsieur N. Kerkeni, Maître de Conférences à l'Université de Valenciennes, pour sa participation à ce jury et pour l'intérêt qu'il porte à ces travaux qui s'intègrent aux activités du GT1 du GRAISYHM,

Monsieur le Professeur J. P. Dubus, Directeur du Laboratoire de Mesures Automatiques, pour avoir bien voulu m'accueillir dans son laboratoire pour effectuer le travail exposé dans cette thèse. Je le remercie vivement pour le soutien moral que son enthousiasme m'a apporté et pour la confiance qu'il m'a accordée. Je lui témoigne ici tout mon respect,

Monsieur V. Devlaminck, Maître de Conférences, qui est à l'origine de ce travail, pour le temps aussi précieux qu'il a réussi à libérer pour étudier ma thèse, ainsi que pour son soutien scientifique permanent et ses remarques à la fois judicieuses, constructives et précieuses.

Je n'oublierai pas mon ami L. Lelaurin, pour sa perspicacité et son désir de mieux approfondir les méthodes connexionnistes. Son scepticisme m'a souvent poussé à rechercher encore et toujours de nouvelles idées à mettre en oeuvre, pour améliorer et valider les résultats.

J'adresse mes remerciements au personnel du Crous de Lille, en particulier "Les services de la vie étudiante", à leur tête Mr. Vanhaecke, pour les inoubliables quatre dernières années passées avec eux dans le cadre de l'allocation d'une bourse de coopération Algéro-Française qui m'a permis de réaliser ce travail.

Un petit mot pour Régine qui, lorsqu'elle n'est pas plongée dans son ordinateur ou dans sa machine à écrire, cherche toujours à rendre service avec diligence et amabilité. Pour Brigitte, je lui dirai de continuer d'être toujours fidèle au poste. Enfin, un grand merci à toute l'équipe du L.M.A avec qui j'ai partagé de merveilleux moments.

*"Le commencement de toutes les sciences,
c'est l'étonnement de ce que les choses
sont ce qu'elles sont."*

Aristote, Métaphysique.

TABLE DES MATIERES

CHAPITRE I. INTRODUCTION GENERALE

I.1. LE PROBLEME : L'ANALYSE DE SCENES DYNAMIQUES	10
I.2. LES OUTILS : LES RESEAUX DE NEURONES	11
I.3. L'APPLICATION : L'ESTIMATION DU MOUVEMENT DANS DES SEQUENCES D'IMAGES	17

CHAPITRE II. LES MODELES CONNEXIONNISTES

II.1. NEURONE FORMEL ET NOTIONS D'APPRENTISSAGE	21
II.1.1. Le neurone formel	21
II.1.1.1. <i>Le neurone formel binaire</i>	23
II.1.1.2. <i>Le neurone formel à valeurs réelles</i>	24
II.1.1.1. <i>Le neurone formel probabiliste</i>	24
II.1.2. Le choix de la fonction d'activation	25
II.1.3. Conclusion.....	26
II.2. NOTIONS D'APPRENTISSAGE	27
II.2.1. La règle de Hebb	27
II.2.2. La règle de Widrow & Hoff	28
II.3. RESEAUX NEUROMIMETIQUES.....	28
II.3.1. Réseaux multicouches	29
II.3.1.1. <i>Le perceptron</i>	29

II.3.1.2. Réseaux à plusieurs couches	31
II.3.2. Réseaux entièrement connectés	31
II.4. RESEAU A APPRENTISSAGE SUPERVISE	33
II.4.1. La rétro-propagation du gradient	33
II.4.2. L'algorithme de rétro-propagation de gradient.....	33
II.5. RESEAU A APPRENTISSAGE NON SUPERVISE	34
II.5.1. Cartes auto-organisatrices de Kohonen	34
II.5.2. Réseau de Grossberg/Carpenter	36
II.6. LES MODELES BIOLOGIQUES.....	37
II.7. LES MODELES STOCHASTIQUES	37
CHAPITRE III. LE MODELE DE HOPFIELD	
III.1. INTRODUCTION.....	40
III.2. DEFINITION	41
III.3. DYNAMIQUE DU RESEAU	43
III.4. ENERGIE ET STABILITE	46
III.5. RESEAUX NUMERIQUES DE HOPFIELD	48
III.5.1. Définition de l'énergie en mode asynchrone	48
III.5.2. Définition de l'énergie en mode synchrone	49
III.6. LES RESEAUX ANALOGIQUES DE HOPFIELD	50
III.6.1. Equation dynamique du réseau.....	50
III.6.2. Définition de l'énergie	51
III.7. LES RESEAUX DE NEURONES STOCHASTIQUES: LA MACHINE DE BOLTZMANN	52

III.7.1. Mode asynchrone	53
III.7.2. Mode synchrone	54
III.8. APPLICATIONS	54
III.8.1. Mémoire associative	54
III.8.2. Résolution des problèmes d'optimisation	56
III.9. CONCLUSION	57

CHAPITRE IV. LA PROBLEMATIQUE DE L'ESTIMATION DU MOUVEMENT

IV.1. ESTIMATION DU MOUVEMENT DANS UNE SEQUENCE D'IMAGES ...	60
IV.1.1. Haut niveau (méthodes structurelles)	63
IV.1.2. Niveau intermédiaire	63
IV.1.3. Bas niveau (méthodes orientées mouvement).....	64
IV.2. METHODES D'ESTIMATION DU CHAMP DES DEPLACEMENTS.....	65
IV.2.1. Introduction	65
IV.2.2. Les méthodes différentielles	66
<i>IV.2.2.1. Principe</i>	<i>68</i>
<i>IV.2.2.2. Conclusion sur les approches différentielles</i>	<i>73</i>
IV.2.3. Les méthodes par transformée	73
<i>IV.2.3.1. La transformée de Hough.....</i>	<i>74</i>
<i>IV.2.3.2. L'approche fréquentielle</i>	<i>75</i>
<i>IV.2.3.3. Les transformées spatio-temporelles.....</i>	<i>75</i>
<i>IV.2.3.4. Conclusion sur les approches par transformées.....</i>	<i>75</i>
IV.2.4. Méthodes de mise en correspondance	76
<i>IV.2.4.1. Conclusion sur les approches de mise en correspondance.....</i>	<i>77</i>
IV.2.5. Méthodes probabilistes	78
IV.3. CONCLUSION.....	78

CHAPITRE V. LES RESEAUX DE NEURONES POUR L'ESTIMATION DU MOUVEMENT

V.1. RESEAUX DE NEURONES ET MOUVEMENTS	81
V.2. LE MODELE DE ZHOU ET CHELLAPPA.....	84
V.2.1. Présentation de la méthode	84
V.2.2. Architecture du réseau de neurones	86
V.2.3. Calcul du "flux optique"	89
V.2.4. Conclusion	90
V.3. REFORMULATION DU PROBLEME	90
V.3.1. Calcul de la fonction d'énergie	91
V.3.2. La phase d'identification.....	92
V.3.3. Discussion sur la variation d'énergie ΔE	93
V.3.4. Résultats expérimentaux et commentaires	94
V.4. CONCLUSION	100

CHAPITRE VI. UN RESEAU GENERIQUE

VI.1. POSITION DU PROBLEME.....	104
VI.2. L'ARCHITECTURE DU RESEAU PROPOSE	105
VI.3. LA DYNAMIQUE.....	110
VI.4. LA FONCTION D'ENERGIE	112
VI.5. LA PHASE D'IDENTIFICATION.....	114
VI.6. EVOLUTION DE L'ENERGIE AU COURS DE LA DYNAMIQUE	118
VI.7. PRISE EN COMPTE DES DISCONTINUITES	121

VI.8. STRATEGIES DE PARCOURS	123
VI.9. RESULTATS EXPERIMENTAUX	124
VI.9.1. Séquence d'images de synthèse	124
<i>VI.9.1.1. Cas d'une translation pure</i>	<i>124</i>
<i>VI.6.1.2. Cas d'une rotation.....</i>	<i>127</i>
VI.9.2. Séquence d'images réelles	131
VI.9.3. Analyse des parcours types	133
VI.9.4. Comparaison avec d'autres approches.....	140
VI.10. CONCLUSION.....	145
 CHAPITRE VII. CONCLUSION GENERALE	
VII.1 EN CONCLUSION.....	148
ANNEXE A: PRINCIPAUX RESEAUX DE NEURONES	151
ANNEXE B: PROPRIETES DE CONVERGENCE DU RESEAU DE HOPFIELD....	153
ANNEXES C: PUBLICATIONS DANS LE CADRE DE LA THESE	158
REFERENCES BIBLIOGRAPHIQUES	160

CHAPITRE I

INTRODUCTION GENERALE

CHAPITRE I

INTRODUCTION GENERALE

Le travail que nous présentons porte sur l'utilisation des réseaux de neurones formels dans le cadre de l'analyse de scènes dynamiques. L'étude réside sur la simulation d'un réseau de neurones discret basé sur le modèle de Hopfield destiné à la minimisation de fonctionnelles. L'application étant l'estimation du mouvement dans une séquence d'images monoculaires. Dans ce chapitre, nous allons tout d'abord présenter le problème, les outils que nous avons mis en oeuvre, puis l'application particulière que nous avons choisie pour valider notre approche.

I.1 LE PROBLEME : L'ANALYSE DE SCENES DYNAMIQUES

Le traitement de séquences d'images a connu ces dernières années un essor considérable, dû essentiellement au développement de la technologie, et plus particulièrement dans le domaine des mémoires (capacités, temps d'accès), point sensible à l'implantation de tout algorithme d'analyse spatio-temporelle. L'objectif recherché de l'analyse de scènes dynamiques est de permettre l'extraction de l'information concernant directement l'environnement 3D dans lequel évolue le capteur. Cet environnement étant perçu par le calculateur comme une scène 2D.

Les séquences d'images peuvent être, soit une série d'images issues d'un même capteur à des instants différents, soit un ensemble d'images d'une même scène prises simultanément par des capteurs complémentaires. Le déplacement des objets dans le monde réel se traduit par un déplacement apparent sur l'image provenant d'une caméra. Celle-ci pouvant être fixe ou mobile par rapport à la scène. La qualité de l'estimation de ces déplacements réels à partir des représentations des objets contenus dans une séquence d'images est cruciale. Elle dépend en grande partie des étapes situées en amont de la chaîne de traitement.

Parmi les applications de l'analyse de scènes dynamiques, on peut citer:

- Le codage d'images TV;
- Le suivi d'objets ou de cibles;
- Le domaine biomédical;
- L'imagerie satellitaire;
- La robotique et la navigation autonome.

I.2 LES OUTILS : LES RESEAUX DE NEURONES

Le traitement d'images semble être un des domaines d'application privilégié des réseaux connexionnistes. En plus de l'attrait que peut apporter sur le plan intellectuel l'étude des problèmes posés par les réseaux de neurones, la diversité de ses applications potentielles constitue un facteur non négligeable dans la motivation des recherches réalisées dans ce domaine. Les applications sont fort nombreuses. On peut distinguer parmi celles-ci:

- La reconnaissance des formes visuelles: c'est le domaine d'application le plus vaste et le plus diversifié. On distingue quatre axes:

- * La reconnaissance des formes bruitées: (Imagerie médicale, Robotique, secteurs militaires);
- * La reconnaissance des écritures typographiques et manuscrites;
- * Le contrôle de qualité et la recherche de défauts;
- * L'identification d'images à des fins de sécurité (reconnaissance d'empreintes digitales, identification de rétine de l'oeil, reconnaissance d'un visage).

- La reconnaissance de signaux dynamiques:
 - * Le traitement du signal médical;
 - * La surveillance et le contrôle de processus industriels;
 - * La reconnaissance de la voix et de la parole.

- Le codage et la compression de l'information:
 - * Les télécommunications (satellites, TV haute définition, TV3D...);
 - * Le secteur militaire (encryptage/décryptage de messages);
 - * La banque (contrôle d'accès, cartes à puces).

- La classification.

- L'optimisation et la résolution de problèmes.
 - * L'ordonnancement des prises de vues des photos satellites SPOT;
 - * L'optimisation de l'implantation de lignes de production;
 - * L'optimisation des tracés de circuits sur des puces VLSI.

- La modélisation de phénomènes complexes.

La multitude et la variété des applications entraînent des méthodologies algorithmiques et des architectures matérielles très diversifiées en fonction des différents objectifs [LIP - 87] [WAS - 89] [KOH - 88a] [HEC - 87] [PAO - 90].

Le domaine des réseaux de neurones n'est pas nouveau. Ainsi le physiologiste Sigmund Exner proposait déjà en 1894 un modèle neuronal de la détection du mouvement par l'oeil de la mouche. Cependant le concept de ce qui est appelé réseaux neuromimétiques est apparu avec les travaux de Clark et Farley, qui les premiers, ont simulé sur un ordinateur numérique un système de "neurones" rebouclé comportant des connexions dynamiquement variables [KAM - 90]. Ces réseaux sont constitués d'un grand nombre de processeurs élémentaires opérant en parallèle appelés donc "neurones" ou "neurones formels", connectés à un certain nombre d'autres processeurs. La nature de la fonction à assurer par ces réseaux est principalement définie par la topologie des connexions.

On peut, de façon très schématique, distinguer trois grandes approches principales de la neuromimétique:

- Celles qui s'appuient sur l'étude du fonctionnement réel du cerveau et les propriétés des systèmes nerveux biologiques;
- Celles qui sont particulièrement orientées vers des études de type psychologique. Des modèles plus complexes permettant de décrire le processus de perception et de cognition sont imaginés;
- Enfin, celles qui exploitent ou reproduisent les connaissances biologiques mais dont le but est d'aborder ou de réaliser des tâches technologiques bien précises. La modélisation des neurones est en général plus simple que dans les cas précédents puisqu'elle ne nécessite pas une prise en compte complète de la complexité biologique.

Selon Parrochia [PAR - 94], "c'est donc moins la parenté avec le cerveau ou la plausibilité biologique du processus de reconnaissance qui est recherchée que la performance pour elle-même. Dès lors, le concept de "réseau de neurones" s'autonomise, devient un outil à part entière qui, même s'il peut parfois présenter des analogies avec le fonctionnement du cerveau (Kohonen), est développé en fait pour lui-même, selon des algorithmes qu'on essaie d'optimiser".

Les premières investigations sur le concept de "réseau de neurones" ont vu le jour au cours des années 40 dans les premières tentatives de modélisation du cerveau menées par deux chercheurs, le neurophysiologiste Warren McCulloch et le mathématicien Walter Pitts [McC - 43]. En s'inspirant de leurs propres travaux sur les neurones biologiques, ces auteurs ont proposé le premier modèle mathématique d'un neurone formel. Un des premiers résultats de leurs travaux a été de montrer que l'on peut modéliser n'importe quelle fonction logique avec de tels réseaux. La fonction seuil utilisée comme fonction d'activation a été adaptée dans plusieurs réseaux neuroniques. On citera à titre d'exemple le réseau récurrent de Amari [AMA - 71] [AMA - 77], le réseau discret de Hopfield [HOP - 82] ou encore les mémoires associatives bidirectionnelles [KOS - 88].

Un peu plus tard, Donald Hebb [HEB - 49] expliquait la façon dont un réseau s'organise (les règles d'apprentissage) selon un postulat basé sur des observations biologiques: l'activation répétée de deux neurones tend à renforcer leur force de connexion. Ce résultat est connu sous le nom de *loi de Hebb*.

La première machine adaptative ayant remporté un certain succès est le "perceptron". Les travaux réalisés sur le système oculaire sont à l'origine de ce premier réseau de neurones formels réalisé par Rosenblatt [ROS - 57]. L'idée était de traduire l'image dans l'architecture d'un réseau de petites cellules susceptibles de prendre différentes valeurs.

Widrow et Hoff [WID - 59] ont développé en 1959 les adaptateurs linéaires ADALINE et MADALINE (Multiple ADAPtative LINear Element). Ces derniers pouvaient apprendre plus rapidement que le perceptron. La différence essentielle entre le perceptron et ce type de réseau se situe au niveau de sa règle d'apprentissage. Pour le perceptron, la méthode minimise l'erreur simple alors que pour MADALINE la méthode minimise l'erreur quadratique, d'où la *règle de Widrow-Hoff*. Cette règle ou algorithme des moindres carrés fut mis au point en 1974 [LEM - 88] [LIP - 87].

Plus tard, Minsky et Papert [MIN - 69] démontrèrent théoriquement les limites de ces réseaux (complexité d'un réseau de très grande taille, problème de séparabilité

linéaire,...). Les recherches se restreignirent alors durant les années soixante dix au profit d'une autre technique "l'intelligence artificielle" qui se développait rapidement et avec succès. Toutefois, certains travaux sur le sujet ont été publiés par quelques rares auteurs. Ils ont principalement concerné les réseaux de neurones formels totalement connectés. Parmi ces auteurs, Amari [AMA - 71] [AMA - 72] étudie les propriétés de tels réseaux pour leur utilisation en tant que mémoires associatives. Il propose l'idée de stocker des informations sous forme d'états stables du système, ou de séquences d'états. La règle de calcul des poids est la règle de Hebb. Fukushima [FUK - 75] propose le cognitron, un réseau de neurones multicouches pour la reconnaissance d'images, malheureusement la règle d'apprentissage proposée est non supervisée, ce qui rend le système difficilement utilisable. Ce modèle a été amélioré un peu plus tard par l'apparition d'une structure multicouche qui déboucha sur le néocognitron [FUK - 84]. Little [LIT - 74] pour sa part, propose l'utilisation des méthodes de la physique statistique pour l'étude des réseaux de neurones. Il présente un modèle de réseau neuronal composé de neurones interconnectés ayant deux états possibles, +1 ou -1. Il remarque qu'il est possible d'identifier la configuration d'une rangée de spins du modèle d'Ising à deux dimensions ($n \times n$) avec l'état d'un réseau neuronal de taille n .

L'ouvrage de Duda et Hart [DUD - 73] constitue une synthèse très complète des premiers travaux sur les réseaux de neurones. Les travaux de la fin des années 70 sont regroupés dans [HIN - 81], en particulier le modèle "Brain-State-in-the-Box" de Anderson, les réseaux sémantiques distribués de Hinton et les cartes topologiques de Kohonen. Les travaux de Kohonen sont regroupés dans [KOH - 84].

L'approche neuromimétique connaît une évolution significative à partir du début des années quatre-vingts. Des progrès remarquables ont été réalisés grâce d'abord, à l'apparition de machines parallèles puissantes rendant les aspects opérationnels assez attrayants, à l'utilisation de concepts et méthodes importés d'autres disciplines - en particulier la physique statistique des systèmes magnétiques, et grâce notamment aux travaux de Fogelmann-Soulié, Werbos, Parker, de Rumelhart, Hinton et Williams [RUM - 86a] sur les réseaux multicouches et l'algorithme de rétropropagation du gradient [LEC - 85], [RUM - 86b], [BAR - 89], à ceux de Hopfield [HOP - 82]

[HOP - 84] [HOP - 85], de Kohonen sur les mémoires associatives et les cartes auto organisatrices [KOH - 84] [KOH - 88b] [KOH - 89], sur les méthodes LVQ (= Learning Vector Quantization) [KOH - 90] ou encore aux travaux de Grossberg et Carpenter [GRO - 88] sur la théorie de la résonance adaptative. Ackley, Hinton et Sejnowski [ACK - 85] introduisent la machine de Boltzmann, le premier modèle à lever les limitations du perceptron de façon satisfaisante et dont la dynamique et l'apprentissage sont probabilistes. C'est sans doute Kohonen [KOH - 84] qui a conçu le modèle dont l'esprit est le plus proche de la manière dont fonctionne le cortex visuel. On pourra aussi citer les BAM (= Bidirectional Associative Memories) de Kosko [KOS - 87] qui prolongent le modèle auto-associatif de Hopfield en mémoire hétéro-associatives, les réseaux à apprentissage compétitif [RUM - 86c] ou encore ceux à apprentissage renforcé [BAR - 85] ainsi que les réseaux récurrents et leurs dynamiques étudiées par Pineda [PIN - 87].

Signalons enfin que la tendance actuelle consiste à construire et à étudier des réseaux modulaires, en composant ou cascasant plusieurs modèles distincts. Des expériences récentes montrent sur des cas particuliers qu'un tel chaînage permet souvent d'accroître les performances. Il devient toutefois difficile de contrôler l'apprentissage dans des réseaux de ce type.

Cette liste ne prétend pas être exhaustive. Une illustration de l'évolution historique des principaux systèmes de réseaux de neurones est présentée à titre indicatif dans *l'annexe A*.

Il est à noter enfin l'évolution de la technologie VLSI qui ne cesse d'apporter un autre intérêt de recherche à ce domaine [CHO - 91] [SAT - 92] [SHE - 92]. En effet, des réseaux neuronaux électroniques (silicium) permettront non seulement d'atteindre les ordres de grandeur nécessaires pour traiter en temps réel certaines applications, mais également pour une utilisation des techniques neuronales sur des systèmes spécialisés où l'encombrement est limité [MEA - 89].

On recense aujourd'hui un nombre important d'ouvrages et d'articles traitant de la neuromimétique. Bien qu'il devient de plus en plus délicat de prétendre couvrir un

domaine aussi vaste et en rapide évolution, on peut néanmoins citer [RUM - 86] [AMI - 89] [PER - 90] [LEC - 87] [BUR - 88] [ROS - 59] [AND - 88] [HEC - 90] [DAV - 90] [LIP - 87] [KAM - 90] [FRE - 92] [WEI - 89] [ZHO - 92] [HER - 94] [JOD - 94a] [JOD - 94b] [ABD - 94].

I.3 L'APPLICATION : L'ESTIMATION DU MOUVEMENT DANS DES SEQUENCES D'IMAGES

L'estimation du mouvement est un problème très complexe. Bien qu'il s'agit d'une tâche dont les organismes vivants les plus simples s'acquittent sans effort apparent, les ordinateurs traditionnels semblent mal adaptés. C'est sans doute pour cette raison que les premières tentatives d'application des réseaux de neurones ont porté sur ce domaine. Il s'avère important de pouvoir définir précisément l'apport que peuvent avoir les réseaux de neurones dans les diverses tâches qui entrent en jeu dans un système d'analyse de mouvements.

Chez la plupart des êtres vivants munis d'un système visuel, l'analyse et l'interprétation de l'environnement passe par une analyse poussée du mouvement perçu. Des études physiologiques et neurophysiologiques ont montré l'importance du "flux optique" dans la perception de l'espace par les systèmes visuels des êtres vivants. Il a été en effet prouvé que de tels êtres disposent dans leur cortex de récepteurs spécifiques sensibles à la détection du mouvement.

Il existe cependant de nombreuses méthodes spécialisées pour l'estimation du mouvement. On peut décrire une méthode selon le niveau d'abstraction qu'elle utilise afin de représenter l'information [CAB - 92]. On distingue ainsi le "bas niveau" qui fait appel à des techniques d'analyse des variations locales de l'intensité de l'image, le "haut niveau" où l'information contenue dans l'image est décrite en termes d'objets, et le "niveau intermédiaire" assurant la liaison entre ces deux représentations. Les méthodes basées sur l'analyse du "flux optique" qui consistent en l'étude du champ du vecteur vitesse apparente sur les images, appartiennent à ce dernier niveau.

L'approche connexionniste en s'inspirant des modèles neurobiologiques apporte des solutions à certains problèmes de calcul du "flux optique". En effet, les réseaux de neurones se présentent comme des dispositifs caractérisés par un haut degré de parallélisme où le traitement de l'information est largement distribué au travers de toute la structure d'où une plus grande robustesse vis-à-vis d'un éventuel mauvais fonctionnement d'un petit nombre d'éléments [KER - 92].

Le travail que nous proposons dans ce mémoire porte l'utilisation des réseaux de neurones pour la minimisation de fonctionnelles avec comme application l'estimation du mouvement dans le cadre de l'analyse de scènes dynamiques. Plus précisément, l'idée de base est de fonder cette estimation sur le champ des vitesses apparentes dans l'image. L'utilisation d'un tel champ de vitesse permet de disposer d'un champ de mesures dense donc d'informations très riches.

Le chapitre II décrit d'une part, les notions fondamentales indispensables à la définition d'un neurone formel et ses variantes, et présente d'autre part, les principaux types de réseaux constitués de tels neurones parmi les plus connus ainsi que leurs algorithmes d'apprentissage associés.

Les chapitres III et IV situent le contexte. Ainsi le chapitre III porte sur le modèle de Hopfield. Nous étudierons plus particulièrement sa dynamique et nous montrons l'importance de ce type de réseau totalement interconnecté dans la résolution des problèmes d'optimisation.

Le chapitre IV, lui, est dédié au problème de l'estimation du mouvement dans une séquence d'images. A l'aide d'une étude bibliographique, nous rappellerons l'état des diverses techniques classiques existantes sur l'analyse du mouvement 2D, puis nous exposerons les principales approches d'estimation du "flux optique".

Dans le chapitre V, on replace le problème de l'estimation du mouvement dans l'optique connexionniste. Une première méthode est présentée. Elle est inspirée d'une approche neuronale proposée antérieurement [ZHO - 88]. Le réseau de neurones développé est basé sur des considérations physiologiques et anatomiques du système

visuel. L'étude a été motivée par les travaux réalisés sur l'organisation structurelle en "colonnes d'orientation" des cellules du cortex visuel.

Le chapitre VI expose une autre méthode neuronale pour l'estimation du "flux optique" dans une séquence d'images. Le principe de l'approche consiste dans une première étape à coder le problème en termes d'états qui sont les variables discrètes représentant les deux composantes du vecteur vitesse. Ces dernières peuvent être alors déduites par une sommation des variables des états des neurones du réseau associé. Dans une seconde phase, la construction du réseau de neurones discret qui résout ce problème revient à déterminer les paramètres - poids synaptiques des interconnexions et entrées externes - qui lient les neurones de telle sorte que les points fixes soient les solutions recherchées. Nous introduisons une fonction d'énergie qui caractérise un état du réseau et nous montrons que les minima locaux de cette fonction ainsi formée et qui contient toute l'information nécessaire, correspondent à des stables du réseau quand celui-ci suit la règle d'évolution définie par Hopfield. Ce modèle est basé sur une structure rebouclée et fonctionne en exploitant les propriétés de convergence vers des points ou des cycles stables.

Enfin, nous comparons les résultats avec ceux fournis par les méthodes classiques d'estimation du mouvement de façon à valider les solutions obtenues par le réseau de neurones proposé.

CHAPITRE II

LES MODELES CONNEXIONNISTES

CHAPITRE II

LES MODELES CONNEXIONNISTES

Différents types de réseaux de neurones ont déjà fait l'objet de recherches. Nous rappellerons dans ce chapitre les notions élémentaires sur les modèles de neurones puis nous exposerons très rapidement les règles d'apprentissage qui furent alors proposées pour permettre à ces réseaux de s'auto-adapter. Nous tenterons ensuite de faire une synthèse des différentes approches existantes. Enfin, nous nous intéresserons aux développements des réseaux entièrement interconnectés. L'étude portera par la suite sur le réseau de Hopfield (et ses variantes) utilisé comme résolveur de problèmes d'optimisation.

II.1. NEURONE FORMEL ET NOTIONS D'APPRENTISSAGE

II.1.1. Le neurone formel

Appelé aussi classifieur linéaire en reconnaissance de formes, automate à seuil ou élément linéaire à seuil en théorie des automates, c'est l'élément de base des réseaux connexionnistes. Il est apparu pour la première fois dans un article de McCulloch et Pitts [McC - 43]. Ces deux auteurs ont en effet proposé le premier modèle

mathématique du neurone biologique. Dans ce modèle, simplifié à l'extrême, chaque neurone calcule périodiquement son degré d'activation. Si ce dernier dépasse un seuil prédéfini, le neurone se met ou reste dans l'état actif. Si par contre, l'activation est en dessous de ce seuil, il devient inactif. Un des premiers résultats de leurs travaux a été de montrer que l'on peut modéliser n'importe quelle fonction logique avec de tels réseaux.

On retrouve dès lors différents types de modèles de neurones formels. La structure générale reste la même et peut être représentée comme l'illustre la *figure II.1*.

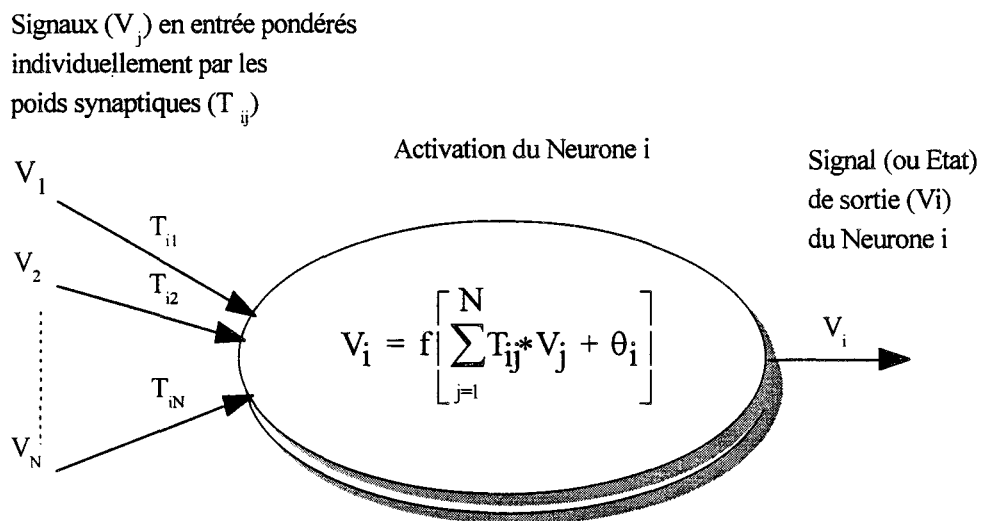


Figure II. 1 : *Le neurone formel.*

Les différents choix de modèles de neurones formels se caractérisent par le type de la fonction d'activation employée. On distingue ainsi trois grands modèles de neurones:

- le neurone formel binaire;
- le neurone formel réel;
- le neurone formel probabiliste.

II.1.1.1. Le neurone formel binaire

C'est un élément de calcul possédant N entrées booléennes ou réelles (V_1, V_2, \dots, V_N) et une seule sortie booléenne. Il est modélisé par deux opérateurs:

- Un opérateur de sommation qui élabore un "potentiel" égal à la somme pondérée des entrées du neurone;
- Un opérateur qui calcule l'état de la sortie du neurone en fonction de son potentiel.

La fonction de changement d'état (ou fonction de transition) du neurone formel consiste à effectuer la somme pondérée de ses entrées et à comparer ce résultat à un seuil θ . Si le potentiel du neurone u_i est positif, l'état V_i de l'automate devient actif et égal à +1; dans le cas contraire, il est inactif et égal à 0:

$$V_i = 1 \quad \text{si} \quad \left(u_i = \sum_{j=1}^{j=N} T_{ij} \cdot V_j + \theta_i \right) > 0 \quad (\text{II.1})$$

$$V_i = 0 \quad \text{sinon}$$

La fonction d'activation utilisée est en général la fonction de Heaviside (*figure II.2.a*).

Il est possible, dans certaines situations, de trouver la valeur -1 plutôt que 0 pour représenter l'inactivité du neurone. La fonction de Heaviside est alors remplacée par la fonction Signe (*figure II.2.b*). Naturellement, d'un point de vue général, ces deux notations sont équivalentes puisqu'il est possible de passer du domaine $\{0, +1\}$ au domaine $\{-1, +1\}$ et vice-versa par un simple changement de variable. Ce changement se traduira par une modification de la valeur du seuil θ_i du neurone [KAM - 90].

Il est à noter que, dans certaines écritures, le seuil est considéré comme un poids particulier, $T_{i0} = \theta_i$, dont son entrée correspondante V_0 vaudrait toujours +1. L'équation du potentiel devient alors:

$$u_i = \sum_{j=0}^{j=N} T_{ij} \cdot V_j \quad (\text{II.2})$$

II.1.1.2. Le neurone formel à valeurs réelles

L'automate à seuil est intéressant lorsqu'il s'agit de traiter des discriminations. Du fait des deux seuls états possibles, on s'est vite aperçu de ses limites dans des applications plus complexes où l'emploi de neurones binaires devient inadéquat. Le modèle de neurones se généralise donc vers des sorties à valeurs réelles. Son évolution se situe alors dans l'utilisation de fonctions d'activation continues à seuils. Ces dernières sont généralement choisies croissantes et impaires (mais ceci ne constitue pas une restriction de principe). On trouve, en particulier, les fonctions linéaires à seuils (*figure II.3*) et les fonctions sigmoïde (*figure II.4*) [MOU - 92] [LEM - 88].

II.1.1.3. Le neurone formel probabiliste

Ce neurone formel particulier utilise une fonction d'activation stochastique (généralement une sigmoïde). L'état Y prend les valeurs $[0,1]$ selon une probabilité:

$$P(Y_i = 1) = \frac{1}{1 + \exp\left(-\frac{\Delta E_i}{T}\right)} \quad P(Y=0) = 1 - P(Y=1) \quad (\text{II.3})$$

où T est un paramètre appelé "Température" et ΔE_i représente la différence d'énergie entre les états 1 et 0 du neurone i .

Le comportement de cet "automate probabiliste" est proche de celui des particules en physique statistique [WEI - 89].

Le neurone formel probabiliste est utilisé principalement dans les machines de Boltzmann [HIN - 84] et de Cauchy [SZU - 87].

II.1.2. Le choix de la fonction d'activation

Le choix de la fonction d'activation (ou de transition) F est a priori laissé au concepteur de l'application neuronale. Cette fonction prend différentes formes, suivant l'utilisation du réseau, et selon la nature continue ou discrète, voire binaire de l'état de sortie du neurone. Il est néanmoins souhaitable de respecter les contraintes suivantes:

- La fonction est non linéaire: En effet, un réseau entièrement linéaire peut toujours se ramener à un réseau équivalent sans couche intermédiaire. Il suffit pour cela de faire le produit des matrices de transformation entre les différentes couches. On perd donc tout l'avantage d'une structure multicouches;
- La fonction est strictement croissante: Il est en effet souhaitable que la dérivée ne s'annule pas afin d'éviter un blocage de l'apprentissage (gradient nul);
- La fonction est saturante: Cette condition correspond aux observations biologiques, et de plus, la présence de saturations procure au réseau une forte robustesse par rapport au bruit. Enfin, la présence de saturations permet au neurone de fonctionner comme un calculateur binaire quand ceci est nécessaire.

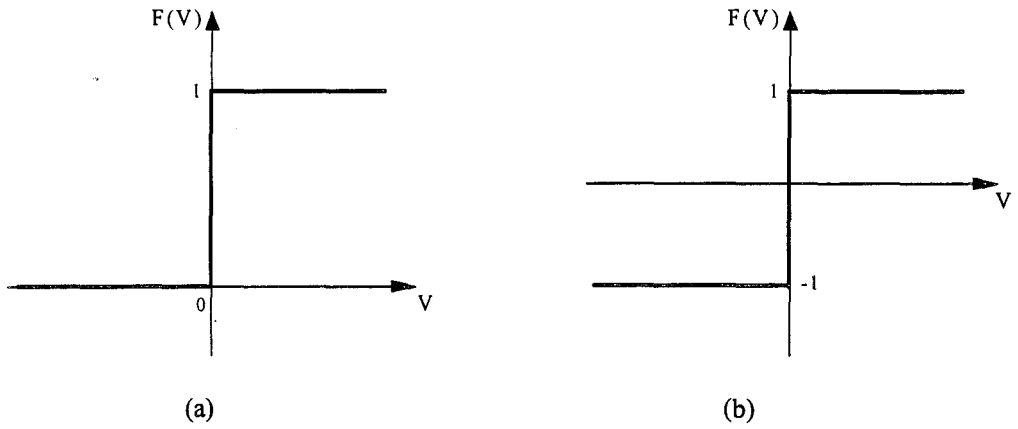


Figure II.2 : Fonctions d'activation binaires.

(a) Fonction de Heaviside.

(b) Fonction Signe.

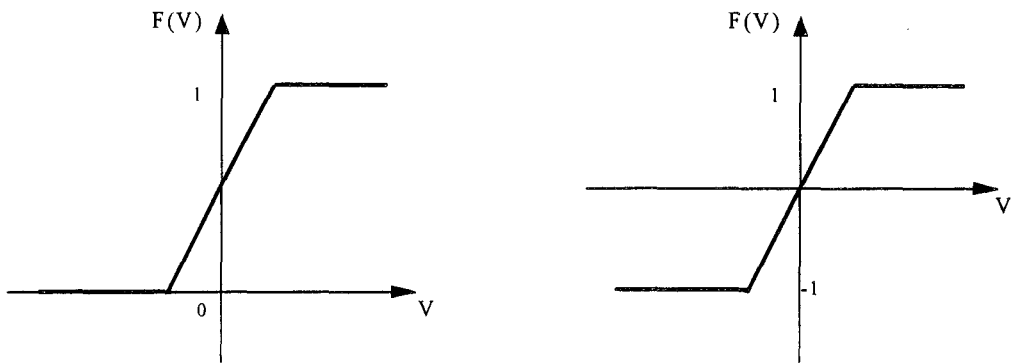


Figure II.3 : Fonctions linéaires à seuils.

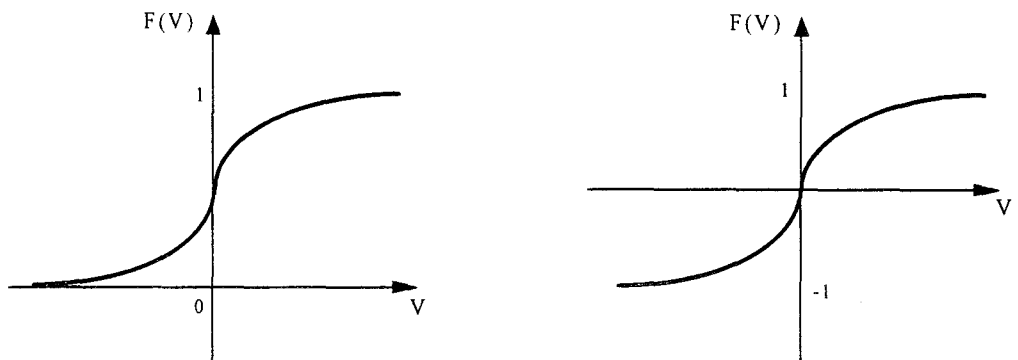


Figure II.4 : Fonctions sigmoïdes.

II.1.3. Conclusion

Il résulte de la présentation précédente qu'un réseau de neurones est défini généralement par 5 éléments:

- La nature des entrées qui peuvent être binaires ($\{0,1\}$ ou $\{-1,1\}$) ou bien réelles;
- La nature de la fonction d'entrée totale (ou potentiel) qui définit le prétraitement sur les entrées. Il s'agit le plus souvent d'une somme pondérée. La fonction sommation peut être du type linéaire, affine ou polynomiale;
- La nature de la fonction d'activation du neurone qui définit son état interne en fonction de son entrée totale. Les fonctions d'allure sigmoïde sont très utilisées car elles modélisent facilement les phénomènes continus. Les fonctions à seuil binaire, à seuil linéaire ou encore stochastiques peuvent être également utilisées;
- La nature de la fonction de sortie qui calcule la sortie du neurone en fonction de son état d'activation;
- Enfin la nature de la sortie du neurone qui servira d'entrée aux neurones qui sont connectés.

II.2. NOTIONS D'APPRENTISSAGE

II.2.1. La règle de Hebb

La principale règle d'apprentissage a été déduite de la loi de Hebb [HEB - 49]. Bien que celle-ci fut élaborée pour les réseaux biologiques; les réseaux de neurones formels doivent également s'y conformer. Selon Hebb, l'apprentissage est lié à une modification fonctionnelle dans les synapses: des activités neuroniques corrélées augmentent

l'efficacité du lien synaptique. Si $V = [V_1, V_2, \dots, V_n]$ est un vecteur prototype, les poids sont donc modifiés selon:

$$\delta T_{ij} = \alpha \cdot V_i \cdot V_j, \quad i, j=1, \dots, n \quad \text{soit} \quad \delta T = \alpha \cdot V \cdot V^T \quad (\text{II.4})$$

L'efficacité synaptique entre les neurones i et j est notée T_{ij} . V_i est l'entrée du neurone i , V_j est l'activité du neurone j , α est le "gain d'apprentissage".

Ainsi:

- (a) si deux neurones sont actifs en même temps, alors leur liaison est renforcée.
- (b) si les deux neurones sont au repos, alors leur poids reste inchangé.
- (c) si la cellule excitatrice ou la cellule réceptive est activée et l'autre non, alors le poids prend une valeur plus faible.

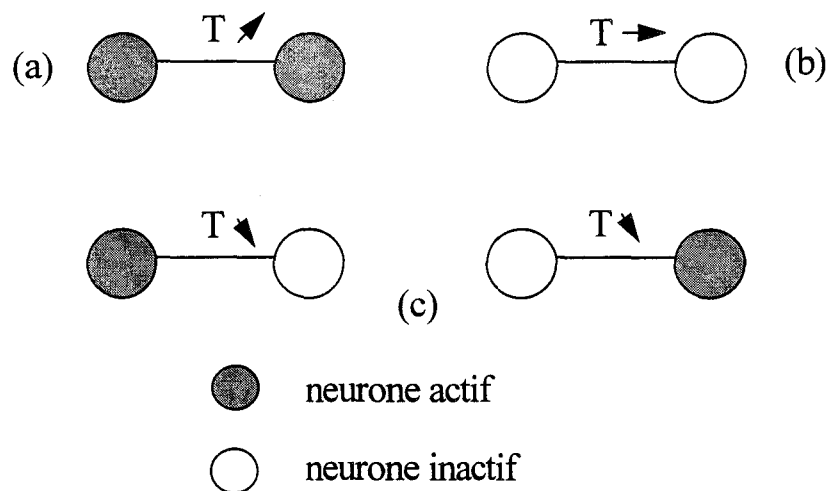


Figure II.5 : La règle de Hebb.

La règle de Hebb fut la première et reste la principale règle d'apprentissage utilisée dans les réseaux de neurones.

II.2.2. La règle de Widrow & Hoff

En plus de la règle de Hebb, il existe de nombreuses méthodes d'apprentissage des poids d'un neurone formel qui consistent généralement à minimiser une fonction de coût. Parmi celles-ci, on peut citer la procédure de Widrow & Hoff [WID - 60] dans laquelle la fonction de transfert à seuil binaire est remplacée par une fonction à seuil linéaire par morceaux. La modification des poids ne dépend alors que des entrées pondérées non seuillées.

II.3. RESEAUX NEUROMIMETIQUES

Un réseau neuromimétique (ou de neurones formels) est un ensemble d'unités de calcul interconnectées, inspiré des structures et de certains effets neurophysiologiques relatifs au fonctionnement des neurones biologiques et du cortex cérébral. Les principales caractéristiques qui sont retenues pour un fonctionnement satisfaisant de tels réseaux sont les suivantes:

- le choix du neurone formel;
- l'architecture du réseau;
- l'évolution ou la règle d'apprentissage.

Afin de présenter les différents types de réseaux, il est souhaitable d'en faire une classification. En effet, on peut distinguer parmi un florilège de méthodes plusieurs classes de réseaux qu'on peut regrouper soit en fonction de leur architecture soit sur le critère de la fonction réalisée par le réseau.

II.3.1. Réseaux multicouches

II.3.1.1. Le perceptron

Le perceptron peut être considéré comme le premier modèle des réseaux de neurones artificiels. Il fut réalisé par F. Rosenblatt [ROS - 57] [ROS - 62]. Fondé sur un principe similaire à celui de McCulloch et Pitts, son but est d'associer des configurations (des formes ou des stimuli) présentées en entrée à des réponses. Il s'appuie sur le fonctionnement du système oculaire et comprend dans son modèle général trois éléments principaux (*figure II.6*):

- une rétine constituée de cellules où s'inscrit le stimulus;
- une unité d'association qui extrait les caractéristiques des stimuli reçus par la rétine;
- une unité de décision qui fournit la réponse du réseau aux stimuli de la rétine.

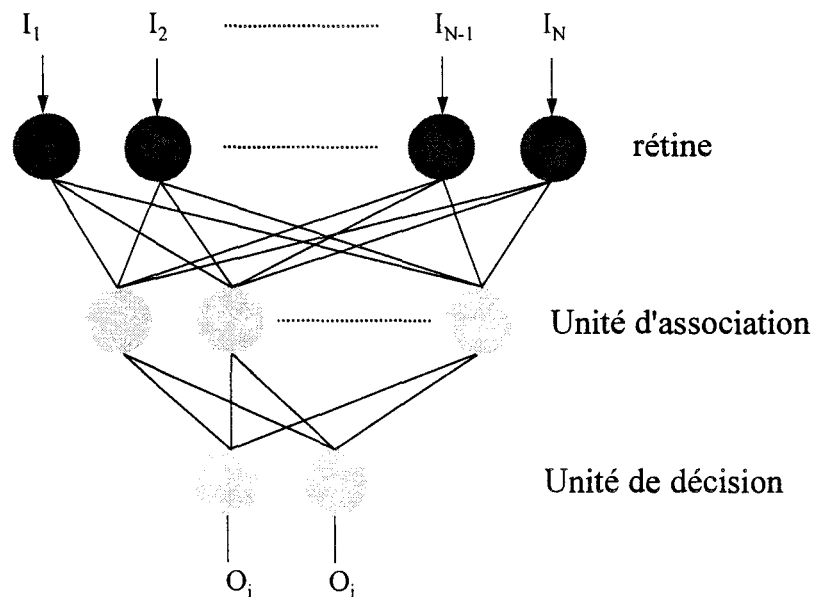


Figure II.6 : Architecture du perceptron: la configuration d'activité dans la couche d'entrée est le résultat d'un codage réalisé par une "rétine".

Les connexions entre la rétine et l'unité d'association ne sont pas pondérées. Seuls, les poids synaptiques entre l'unité d'association et l'unité de décision sont modifiés lors de l'apprentissage. Le perceptron peut être donc considéré comme un réseau à deux couches, une d'entrées et une de sorties, les neurones à l'intérieur d'une même couche n'étant pas connectés.

Cependant, il convient de noter les limites du perceptron qui sont de nature diverses, comme par exemple:

- la complexité d'un réseau de très grande taille;
- le fait qu'un perceptron ne puisse reconnaître si une forme est connexe ou non;
- le problème de la séparabilité linéaire (limite toutefois levée grâce à l'utilisation de couches supplémentaires dans le réseau).

II.3.1.2. Réseaux à plusieurs couches

Ce type de réseaux (figure II.7) présente une couche d'entrée recevant les informations extérieures et une couche de sortie renvoyant les résultats du traitement. Entre les deux, se situent plusieurs couches cachées qui effectuent ce traitement. Les neurones utilisés sont de même type dans une même couche. Cependant, ceux de couches différentes peuvent être de types différents.

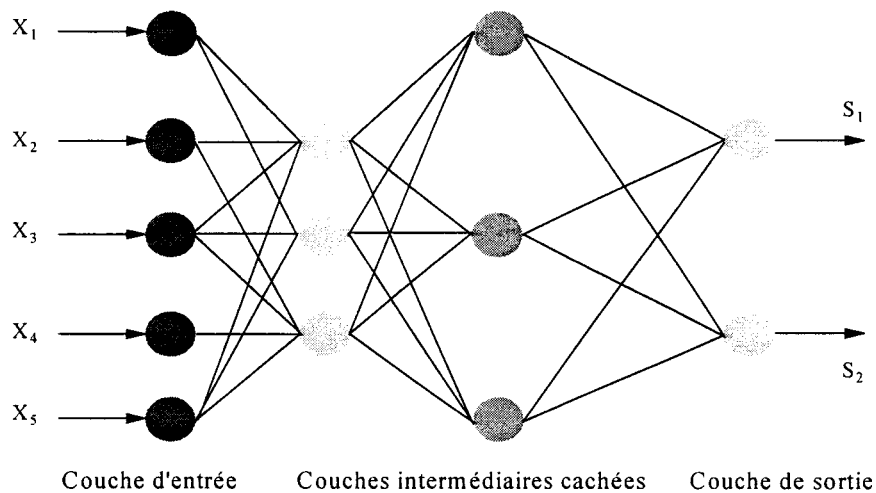


Figure II.7 : Réseau à couches.

L'information à traiter est appliquée au réseau sous la forme d'un vecteur d'entrée puis propagée vers la couche de sortie suivant le même sens de déplacement. Le résultat apparaît lui aussi sous forme de vecteur. Les connexions entre chaque neurone de couches différentes peuvent être totales ou partielles. Ainsi un neurone peut être connecté soit à certains, soit à tous ceux de la couche suivante. Cependant, les neurones d'une même couche ne sont pas interconnectés.

Ce type de réseau est aussi appelé *réseau feedward*.

II.3.2. Réseaux entièrement connectés

Face aux problèmes de structuration et de définition des réseaux multicouches, certains chercheurs se sont orientés vers un autre type d'architecture: les réseaux entièrement connectés (*figure II.8*).

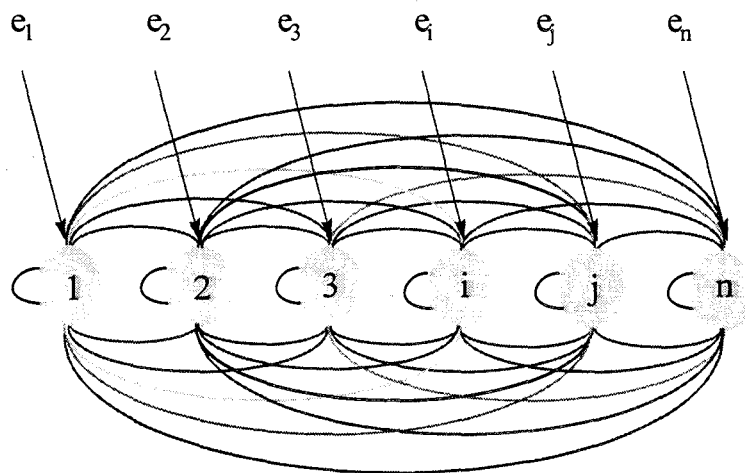


Figure II.8: Réseau totalement connecté.

L'information est appliquée sur chaque neurone, sous forme de vecteurs. Ceux-ci communiquent avec tous les autres et peuvent aussi communiquer avec eux-mêmes. Ce sont les sorties de ces neurones qui déterminent les composantes du vecteur de sortie.

Ce type de réseau est aussi appelé *réseau feed-back*.

A partir de ces architectures, un certain nombre de modèles mettant en jeu des fonctionnements et des règles d'apprentissages différents ont été développés. Nous

nous contenterons de décrire rapidement quelques réseaux en donnant en quelques mots leurs grandes caractéristiques et éventuellement leur domaine d'application.

Nous regrouperons les modèles de la manière suivante:

- Les réseaux à apprentissage supervisé pour lesquels nous présenterons l'algorithme le plus utilisé, la rétro-propagation du gradient;
- Les réseaux à apprentissage non supervisé, avec en particulier les cartes auto-organisatrices de Kohonen et le modèle ART de Grossberg et Carpenter;
- Les modèles biologiques;
- Les modèles stochastiques.

II.4. RESEAU A APPRENTISSAGE SUPERVISE

On soumet au réseau un modèle en entrée et l'on compare alors l'erreur obtenue entre la sortie obtenue et la sortie désirée. L'apprentissage supervisé consiste à ajuster les poids de chaque connexion, à l'aide d'un algorithme itératif, jusqu'à ce que l'erreur soit nulle ou minimale. Le réseau doit être stable lorsque l'apprentissage est réalisé. Il faut donc s'assurer que l'erreur minimale entraîne la stabilité du réseau (perceptron, mono ou multicouches).

II.4.1. La rétro-propagation du gradient

La méthode de rétro-propagation du gradient est l'une des mieux connues. Avant la phase d'apprentissage, les poids synaptiques ont des valeurs aléatoires et les réponses fournies par le réseau à la présentation d'une configuration d'entrée sont quelconques. Durant la phase d'apprentissage, la procédure de rétro-propagation consiste à modifier de façon itérative les poids synaptiques afin d'établir une correspondance entre la présentation d'une entrée et la sortie désirée. Pour chaque entrée, le réseau calcule par propagation l'état de sa couche de sortie. La différence

entre l'état de la couche de sortie et l'état désiré est rétro-propagée de façon à modifier les poids des connexions.

II.4.2. L'algorithme de rétro-propagation de gradient

Il a été introduit simultanément et indépendamment par Y. Le Cun [LEC - 85], D. Rumelhart [RUM - 86] et D. Parker [PAR - 85]. Cette technique utilise une procédure de descente de gradient travaillant sur une erreur quadratique. C'est donc une généralisation de la règle de Widrow & Hoff à une fonction de coût non linéaire. L'algorithme consiste à calculer les dérivées partielles de l'erreur en sortie par rapport à tous les poids du réseau puis à appliquer une procédure de gradient pour minimiser l'erreur en fonction de tous les poids synaptiques.

En résolvant le problème du calcul de ces poids pour les couches "cachées" des réseaux multicouches, l'algorithme a redonné une nouvelle vigueur aux recherches sur les réseaux neuromimétiques, ce qui a permis l'apparition de nouveaux modèles.

Cet algorithme est utilisé pour des applications variées. On peut citer les travaux de [SEJ - 87] en reconnaissance de la parole ou ceux de [LON - 90] [THI - 90] en analyse d'images.

II.5. RESEAU A APPRENTISSAGE NON SUPERVISE

Basé sur des propriétés statistiques des exemples, le système essaye de s'organiser lui-même afin de diviser les exemples en différentes classes (le modèle de Kohonen). Les procédures de modification des poids synaptiques sont généralement inspirées de la règle de Hebb, et lorsque l'algorithme converge, les poids synaptiques atteignent des valeurs stables. Du fait que le système agit tout seul sans avoir d'influence extérieure on l'appelle parfois apprentissage par auto-organisation.

II.5.1. Cartes auto-organisatrices de Kohonen

L'algorithme de T. Kohonen a été initialement conçu en 1982 par Kohonen [KOH - 82] [KOH - 84] pour modéliser l'auto-organisation des liaisons neuronales

entre les cellules du cortex visuel et les cellules de la rétine sous l'effet de stimuli successifs.

La carte d'auto-organisation de Kohonen est un réseau à deux couches qui peut organiser une carte topologique à partir de points de départ aléatoires. Le réseau combine une couche d'entrée avec une couche compétitive de neurones entraînés de façon non supervisée. Cette dernière se présente comme une grille à deux dimensions. Toutes les interconnexions vont de la couche d'entrée vers la seconde couche. Chaque neurone de la couche d'entrée est connecté à tous les neurones de la couche compétitive. La *figure II.9* représente une telle structure de base.

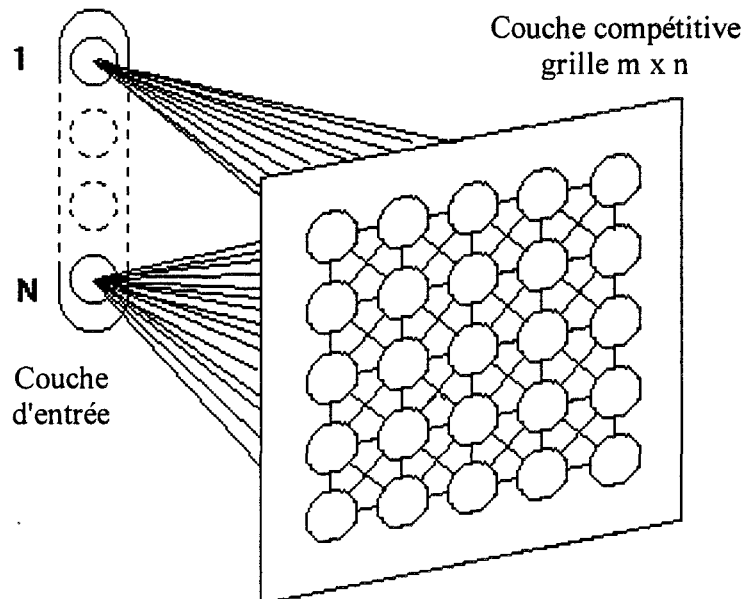


Figure II.9 : Carte auto-organisatrice de Kohonen.

Quand une entrée est présentée, chaque neurone de la première couche prend la valeur correspondant à son entrée dans le vecteur. Les neurones de la seconde couche calculent la somme des entrées et déterminent le neurone gagnant.

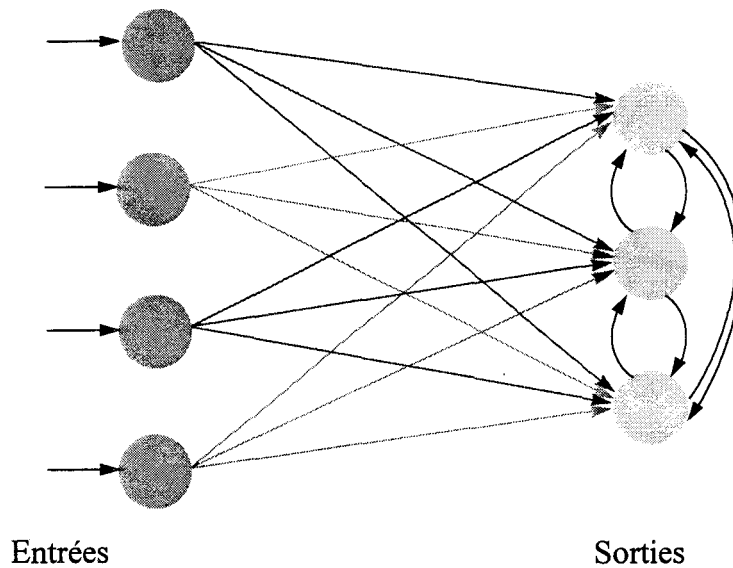
Dans le réseau de Kohonen, chaque connexion a un poids. Au départ, les poids des connexions sont fixés aléatoirement. Ces valeurs sont ensuite mises à jour durant la phase d'apprentissage du réseau.

Cet algorithme a donné lieu à de nombreuses applications. Celles-ci touchent des domaines aussi variés que le traitement des images [NAS - 88], l'analyse de la parole et d'une façon plus générale l'analyse de données [ULT - 90]. On peut citer, par exemple, l'analyse d'une grande quantité de données technologiques pour une ligne de fabrication de VLSI [TRY - 89], la création d'un spectre topographique pour l'analyse de la parole [KOH - 88], ou encore le contrôle d'un bras de robot dans un espace pourvu d'obstacles [GRA - 88]. Parmi les exploitations possibles du réseau de Kohonen, citons également des applications à la résolution de problèmes d'optimisation parmi lesquels on trouve les travaux sur le problème du voyageur de commerce [ANG - 88] [FOR - 88].

II.5.2. Réseau de Grossberg/Carpenter

Les systèmes ART (Adaptative Resonance Theory) [CAR - 89] [GRO - 87] tentent de modéliser par des réseaux de neurones des processus à la fois psychologiques et biologiques. Ce sont des modèles possédant différents modules (module d'attention et module d'orientation) dont les composants ont un comportement obéissant à des systèmes d'équations différentielles généralement non linéaires.

L'apprentissage dans ce type de modèle se développe lorsqu'il y a résonance entre entrées et sorties. En effet, si le réseau a déjà appris à reconnaître un vecteur d'entrée alors un état de résonance est atteint rapidement quand l'entrée est présentée et il y a alors renforcement de la forme mémorisée. En revanche, si la forme n'est pas reconnue, le système recherche dans les formes apprises si aucune mise en correspondance ne réussit. L'obtention de la résonance permet alors l'apprentissage de la nouvelle forme sans effacer les anciennes données.

Figure II.10 : *Le réseau de Grossberg.*

L'intérêt du modèle de Grossberg/Carpenter repose essentiellement sur sa faculté à apprendre sans exemple. Il faut noter aussi que les résultats sont assez mauvais lorsqu'il s'agit de reconnaître des formes bruitées.

II.6. LES MODELES BIOLOGIQUES

Ils sont directement inspirés, dans leur conception, de considérations neurobiologiques. On peut citer les réseaux auto-adaptatifs de Linsker [LIN - 86a] [LIN - 86b] [LIN - 86c]. Ce dernier simule la formation et l'agencement des cellules d'orientation dans le cortex visuel des mammifères. Certains auteurs proposent des approches inspirées de la structure histologique et biologique du cortex cérébral [BAL - 88] [ALE - 89]. D'autres préconisent une étude précise de l'organisation anatomique et physiologique du cerveau humain avant de fournir des modèles de représentation adéquats [SPI - 88].

II.7. LES MODELES STOCHASTIQUES

Il s'agit surtout de la machine de Boltzmann et des techniques de recuit simulé [KIR - 83] [ACK - 85] [AND - 88]. Ces modèles permettent de résoudre des problèmes non linéairement séparables. L'approche par le recuit simulé, proposée par Kirkpatrick, Gelatt et Vecchi en 1983, utilise un algorithme mis au point en physique statistique: l'algorithme de Metropolis [MET - 53]. Il permet, lorsqu'un système est à l'équilibre thermique, et que l'on baisse la température (d'où le terme de recuit), de n'explorer que des configurations voisines de l'énergie minimale.

La machine de Boltzmann est de type connexionniste, inspirée d'un des modèles issus de la physique statistique qui est le réseau de Hopfield. Dans sa forme la plus simple, cette machine est un réseau numérique de Hopfield dont le comportement est dicté par un algorithme de recuit simulé. Elle sera présentée dans le chapitre III.

Les méthodes d'analyse de réseaux neuromimétiques issues de la physique statistique s'avèrent intéressantes pour des problèmes d'optimisation. Un autre aspect à considérer est celui des réalisations matérielles de tels réseaux à l'aide de circuits analogiques, donc pour le calcul hautement parallèle.

CHAPITRE III

LE MODELE DE HOPFIELD

CHAPITRE III

LE MODELE DE HOPFIELD

III.1 INTRODUCTION

Parmi les multiples configurations de réseaux et méthodes d'optimisation, le choix du modèle pour notre étude s'est porté sur le réseau de type Hopfield. En effet, les propriétés de ce type de réseau permettent d'envisager leur application à une grande classe de problèmes de traitement d'images formalisés en termes de minimisation de fonctionnelles. L'estimation du "flux optique" en étant un exemple typique. De tels réseaux peuvent être utilisés soit comme mémoires associatives soit comme solveurs de problèmes d'optimisation [HOP - 85] [HOP - 86]. Leur modèle est identique à celui d'Amari [AMA - 74] caractérisé par le rebouclage des sorties des neurones sur les entrées (réseau récuratif).

De tels réseaux peuvent être caractérisés par leurs capacités de traitement non linéaire de l'information, leur structure massivement parallèle et leur faculté d'apprentissage qui joue un rôle prépondérant dans de nombreux cas.

Ces réseaux sont basés sur le fait que, dans le cerveau humain, l'information circule et qu'une décision est prise par étapes successives. Selon Hopfield, le système nerveux recherche des états stables, attracteurs dans son espace d'états. Les états voisins tendent à se rapprocher d'un état stable, ce qui autorise la correction des erreurs, et la capacité à compléter des informations manquantes. C'est également une mémoire adressable par son contenu: une forme mémorisée est retrouvée par une stabilisation du réseau.

Le modèle que nous allons décrire suscite, depuis les travaux de Little en 1974 [LIT - 74], un grand engouement dans la communauté scientifique en raison de ce qui suit:

- ce modèle est proche d'un autre modèle, celui des "verres de spins", utilisé par les physiciens pour décrire les propriétés magnétiques des alliages dilués;
- possibilité d'utiliser des méthodes de la mécanique statistique pour déterminer les attracteurs;
- possibilité de générer un réseau dont les attracteurs sont déterminés à l'avance.

III.2 DEFINITION

En 1982, Hopfield [HOP - 82] étudie les propriétés d'un ensemble de neurones, du type McCulloch et Pitts, pouvant se trouver dans deux états, excité ou non excité, et ayant un seuil de fonctionnement. Chaque neurone est connecté à tous les autres neurones et joue à la fois le rôle de neurone d'entrée (au temps t) et de sortie ($t + 1$). Comme on peut le constater sur la *figure III.1*, il s'agit donc d'un système rebouclé.

Dans ce réseau, chaque neurone reçoit synchroniquement ou d'une manière aléatoire et asynchrone les entrées $\sum T_{i,j} v_j$ de tous les autres neurones et une entrée externe I_i :

$$u_i = \sum_j^N T_{i,j} v_j + I_i, \quad v_i = g(u_i) \quad (\text{III.1})$$

où u_i représente l'état interne (ou potentiel) du neurone i . La connexion de ce dernier vers le neurone j est caractérisée par un coefficient de couplage ou poids synaptique $T_{i,j}$. Ce couplage correspond à l'importance de l'action du neurone i sur le neurone j . v_j est la sortie (ou état) du neurone j et I_i représente l'action de l'extérieur sur le neurone i . $g(x)$ étant une fonction d'activation du neurone.

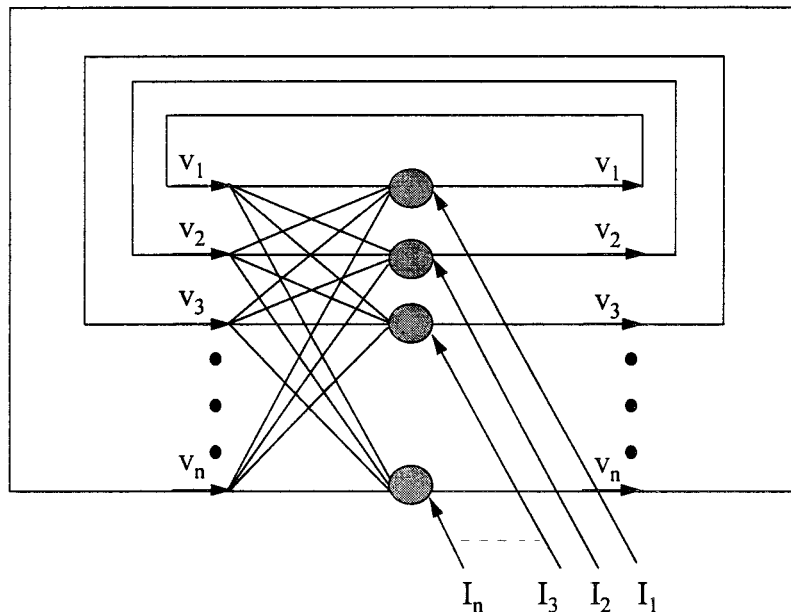


Figure III.1 : Architecture du réseau de Hopfield discret.

Hopfield a imaginé deux types de réseaux entièrement interconnectés:

- **Les réseaux numériques ou binaires:** La sortie des neurones est binaire (± 1 ou 0-1);
- **Les réseaux analogiques:** La sortie des neurones peut prendre toutes les valeurs comprises dans l'intervalle $[0,1]$.

Plus récemment, **des réseaux de neurones stochastiques**, à base de recuit simulé, ont été présentés. Ces réseaux tentent d'effectuer la synthèse entre les réseaux analogiques de Hopfield, de convergence rapide mais locale et le recuit simulé, de convergence lente mais globale. Ces réseaux sont basés sur l'analyse de processus stochastiques.

III.3 DYNAMIQUE DU RESEAU

Les systèmes rebouclés ont été étudiés par les physiciens et il n'est pas surprenant de constater que le modèle de Hopfield s'est développé grâce à l'intérêt important qu'ils y ont manifesté, essentiellement à cause de l'analogie formelle avec des réseaux de particules orientées en interaction: les verres de spins [AMI - 85] [CAS - 92] [HIN - 84].

La plupart des réseaux connexionnistes qui existent actuellement utilisent des dynamiques discrètes qui fournissent une alternative intéressante au modèle continu utilisant des systèmes différentiels [FOG - 87].

Dans le modèle de Hopfield, la dynamique du réseau est décrite par l'équation classique du neurone de McCulloch et Pitts:

$$u_i(t) = \sum_{j=1}^N T_{i,j}(t) v_j(t) \quad \text{et} \quad v_i(t) = \text{signe}(u_i(t)) \quad (\text{III.2})$$

Hopfield a montré que la dynamique d'un système rebouclé peut conduire à un état stable sous certaines conditions [HOP - 82]:

- Les interconnexions entre les neurones doivent être symétriques (la matrice des connexions T est symétrique);
- Les éléments diagonaux de cette matrice doivent être non négatifs.

La démonstration est basée sur le fait qu'on peut associer une fonction d'énergie à un système dynamique, et que tout changement d'état des éléments du réseau doit conduire à une décroissance de cette fonction d'énergie. De plus cette fonction doit admettre un minimum ce qui implique que la décroissance est finie.

La fonction d'énergie s'exprime comme suit:

$$E = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N T_{i,j} v_i v_j \quad \text{avec } T_{ii} \geq 0 \quad (\text{III.3})$$

N représente le nombre de composantes du vecteur V , c'est-à-dire le nombre de neurones du réseau.

L'évolution du réseau se fait soit en mode synchrone, c'est à dire que toutes les variables d'état sont modifiées simultanément, mode de fonctionnement semblable au schéma itératif de Jacobi pour la solution des systèmes linéaires [GOL - 90]; soit en mode asynchrone où ici le vecteur d'état est mis à jour élément par élément en tenant compte des nouvelles valeurs des composantes dès qu'elles sont disponibles. Dans cette optique, plusieurs variantes sont encore possibles, mais la plus importante est le *mode asynchrone séquentiel*, où on calcule les nouvelles valeurs des éléments dans l'ordre où ils apparaissent dans le vecteur d'état. Ce mode de fonctionnement rappelle le schéma itératif de Gauss-Seidel en algèbre linéaire [GOL - 90].

Considérons que la mise à jour du système est de type asynchrone aléatoire. A chaque instant t , on tire un neurone au hasard et on le met à jour. Au cours de la dynamique, l'énergie évolue comme suit: lorsqu'un neurone i change d'état au cours d'un pas d'itération, la variation de la fonction d'énergie vaut:

$$\Delta E = \sum_{j=1}^N T_{i,j} v_j v_i + \sum_{j=1}^N T_{j,i} v_j v_i \quad \text{avec } T_{ii} \geq 0 \quad (\text{III.4})$$

Cette expression peut également être écrite en utilisant un indice temporel qui met en évidence les instants auxquels sont lues les valeurs des neurones du réseau et en exploitant le fait que la matrice T est supposée symétrique ($T_{ij} = T_{ji}$).

$$\Delta E = 2 v_i(t+1) v_i(t) \text{ car } \sum_{j=1}^N T_{ij} v_i v_j = v_i(t+1) v_i(t) \quad (\text{III.5})$$

La variation de l'énergie engendrée par la variation de l'état d'un seul neurone du réseau est toujours négative. En effet, cette variation n'est rien que le produit de l'état du neurone i aux instants t et $(t+dt)$. Or, si nous considérons que les états des neurones ne valent que $+1$ ou -1 , le produit est toujours négatif s'il y a changement d'état. Ainsi, lorsqu'il y a changement d'état, la variation de l'énergie est toujours négative. Si il n'y a pas de changement d'état, ΔE reste bien sûr nul. C'est une fonction monotone décroissante dont la valeur changera jusqu'à atteindre un minimum au moins local et restera stable après.

Il reste à déterminer la borne de cette décroissance, c'est-à-dire savoir si la fonction d'énergie est bornée par valeurs inférieures. La réponse affirmative à cette question est due au fait que le nombre de neurones du réseau est fini, et que les valeurs des neurones ne peuvent être que $+1$ ou -1 . La fonction d'énergie définie est une double somme sur tous les neurones du réseau est elle même finie. Par conséquent, au bout d'un certain nombre d'itérations, une énergie minimum est atteinte et constitue l'accès à un attracteur du réseau. La *figure III.2* donne une représentation imagée du processus de décroissance de l'énergie du réseau dans l'espace des états.

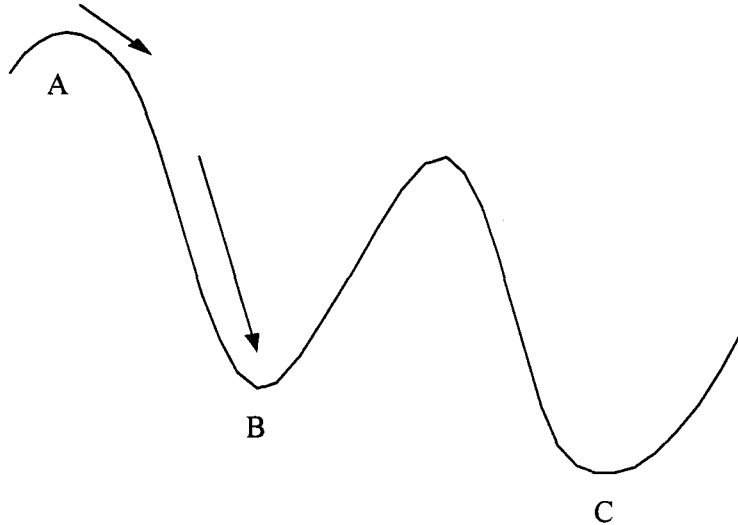


Figure III.2 : Le problème des minima locaux: étant donné la solution A, l'algorithme donnera B et C comme minima locaux; la solution optimale étant C.

Il est important de noter qu'un même réseau peut admettre plusieurs minima qui sont accessibles à partir de configurations initiales différentes et qui peuvent avoir des énergies associées différentes.

Nous venons de voir comment la dynamique d'un réseau de neurones peut nous permettre, sous certaines conditions de converger vers des états stables. L'approche qui sera retenue impose de déterminer une "bonne" fonction d'énergie qui place les solutions du problème comme des minima locaux de cette fonction.

III.4 ENERGIE ET STABILITE

La propriété qu'a le réseau de "rechercher" un minimum d'énergie le rend apte à résoudre des problèmes d'optimisation où l'objectif est de minimiser une fonction de coût. Nous rappelons qu'un problème d'optimisation sous contrainte s'exprime de la façon suivante:

$$\text{minimiser } f(\bar{x}) \text{ sous la contrainte } g(\bar{x}) = 0$$

où \bar{x} est vecteur d'état des variables du problème. Ce vecteur représente l'espace des solutions du problème. $g(\bar{x}) = 0$ est une équation scalaire qui décrit un sous-espace de l'espace des états. Durant le processus de convergence de l'algorithme, l'état doit être attiré vers le sous-espace $g(\bar{x}) = 0$ et y rechercher la plus petite valeur de $f(\bar{x})$. Beaucoup de méthodes ont été développées dans le passé: elles utilisent les multiplicateurs de Lagrange ou les relations de Kuhn et Tucker pour caractériser les solutions éventuelles. Un panorama est donné dans [CIA - 88].

Hopfield a montré que les réseaux récurrents étaient bien adaptés pour résoudre les problèmes d'optimisation combinatoires. Il définit une application, appelée *fonction énergie* du réseau, de l'hypercube C^N dans R qui à tout vecteur d'état \vec{V} associe:

$$\begin{aligned}
 E(\vec{V}) &= -\frac{1}{2} \vec{V}^T \cdot T \cdot \vec{V} + \vec{V}^T \cdot \vec{I} \\
 &= -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N T_{i,j} \cdot v_i \cdot v_j + \sum_{i=1}^N I_i \cdot v_i
 \end{aligned} \tag{III.6}$$

On définit ainsi un *paysage énergétique* discret à N dimensions et 2^N points. La fonction est décroissante le long de toute la trajectoire si la matrice des poids synaptiques T vérifie les propriétés précédemment citées [HOP - 82] [HOP - 84]. Quand T est d'une forme quelconque, la décroissance de $E(\vec{V})$ n'est pas prouvée théoriquement, bien que constatée le plus souvent.

Par analogie avec les résultats d'Hopfield, la première étape, face à un problème d'optimisation, consiste donc à trouver une représentation de la fonction objectif à minimiser et des contraintes sous la forme d'une fonctionnelle simple, de préférence quadratique. Une telle fonctionnelle, appelée énergie du problème, est issue généralement de la théorie de régularisation. Elle revêt la forme suivante:

$$E = E_{\text{objectif}} + \sum_{\text{contraintes } c} \lambda_c \cdot E_c$$

$$= -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N T_{ij} \cdot v_i \cdot v_j + \sum_{i=1}^N I_i \cdot v_i \quad (\text{III.7})$$

Le problème est donc ramené à un problème de minimisation non contrainte. Les coefficients λ_c sont des réels positifs pondérant l'importance relative des termes. Des méthodes de détermination automatique des λ_c existent [TIK - 76] mais elles concernent surtout les problèmes dans lesquels il est facile d'estimer la différence entre la valeur théorique et la valeur courante de l'estimation de la solution. De telles méthodes sont généralement très coûteuses à mettre en oeuvre et sont peu applicables en présence du bruit.

Néanmoins, des fonctions énergie non quadratiques pourraient être définies [STY - 90] [SZU - 86] [SZU - 87]. On note que les poids synaptiques n'évoluent pas durant le processus de convergence du réseau mais sont complètement spécifiés par l'énergie à minimiser [HOP - 85] [HOP - 86]. Une solution du problème d'optimisation sera un sommet de l'hypercube C^N recherchée sous la forme d'un vecteur d'état du réseau. Mais le paysage énergétique (i.e. espace des solutions) peut posséder beaucoup de minima locaux du fait de sa discrétisation.

Il est possible de définir l'évolution de l'énergie au cours de la dynamique pour les deux types de réseau entièrement interconnectés imaginés par Hopfield.

III.5 RESEAUX NUMERIQUES DE HOPFIELD

III.5.1 Définition de l'énergie en mode asynchrone

On considère un réseau fonctionnant en mode asynchrone. Par conséquent, les N neurones changent d'état les uns après les autres dans un ordre aléatoire. On suppose aussi que la matrice des poids synaptiques est symétrique ($T_{ij} = T_{ji}$). En notation matricielle, un état V est un état stable si on a $V = f_s (T \cdot V + I)$. La fonction d'énergie est définie par:

$$E = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N T_{ij} v_i(t) v_j(t) - \sum_{i=1}^N I_i v_i(t) \quad (\text{III.8})$$

Si le neurone i change d'état, la variation correspondante de la fonction d'énergie est donnée par:

$$\Delta E = -\Delta V_i \cdot u_i - \frac{1}{2} T_{ii} \cdot \Delta v_i^2 \quad (\text{III.9})$$

- Si $T_{ii} \geq 0$, la fonctionnelle E est monotone décroissante et l'état final sera un état stable. E est une fonction de Lyapunov pour le réseau.
- Si $T_{ii} < 0$, la fonctionnelle E ne sera plus monotone décroissante et la convergence du réseau n'est plus assurée. E n'est plus une fonction de Lyapunov.

III.5.2 Définition de l'énergie en mode synchrone

Si le réseau fonctionne en mode synchrone, la variation d'énergie correspondant à un changement d'état de tous les neurones est:

$$E = - \sum_{i=1}^N \Delta v_i \cdot u_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N T_{ij} \cdot \Delta v_i \Delta v_j \quad (\text{III.10})$$

Le premier terme de l'expression est toujours négatif, mais le second peut être de signe quelconque. On doit donc modifier la définition de l'énergie et introduire une nouvelle fonction [GOL - 86]:

$$L = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N T_{ij} v_i(t) v_j(t+1) - \frac{1}{2} \sum_{i=1}^N I_i (v_i(t) + v_i(t+1)) \quad (\text{III.11})$$

On peut alors démontrer que cette fonctionnelle est monotone décroissante et en conclure que le réseau convergera vers un état stable ou un cycle de longueur inférieure ou égale à 2 qui minimise L [BRU - 87].

Un inconvénient d'un réseau fonctionnant en mode synchrone est l'obligation d'avoir $T_{ii} \geq 0$. Or, pour résoudre un problème d'optimisation, on calcule les coefficients du réseau à partir de la forme quadratique à optimiser par:

$$T_{ij} = -\frac{\partial^2 E}{\partial v_i \partial v_j} \quad (\text{III.12})$$

Si celui-ci est négatif, il est nécessaire d'introduire dans l'énergie de Lyapunov un terme supplémentaire de contrainte de manière à rendre positifs les coefficients T_{ij} , ce qui en général diminue la qualité des solutions.

Dans le cas d'un fonctionnement asynchrone ou synchrone, l'évolution d'un réseau de Hopfield binaire, dont la matrice des poids synaptiques est symétrique, est restreinte aux coins de l'hypercube unité de dimension N , ce qui implique que ces réseaux ont une probabilité importante de tomber dans un minimum local ou dans un cycle de longueur 2 et que leur évolution est très fortement tributaire de l'état initial [GOL - 85].

III.6 LES RESEAUX ANALOGIQUES DE HOPFIELD

Les neurones de ces réseaux fonctionnent de manière continue dans l'espace des états: la sortie des neurones peut prendre n'importe quelle valeur entre 0 et 1 et le fonctionnement est continu dans le temps. La rapidité de calcul des réseaux analogiques est donc supérieure à celle des réseaux binaires, car les neurones changent leur état de manière continue et simultanée. Pour le neurone i , on a:

$$v_i = f_i(u_i) \quad \text{avec} \quad f_i(x) = \frac{1}{1 + \exp(-x / T_i)} \quad (\text{III.13})$$

T est la température du neurone i .

III.6.1 Equation dynamique du réseau

En supposant que chaque neurone a la même constante de temps τ , l'équation décrivant l'évolution temporelle du neurone i est donnée par l'équation différentielle [HOP - 85]:

$$\frac{\partial u_i}{\partial t} = -\frac{u_i}{\tau} + \sum_{j=1}^N T_{i,j} \cdot v_j - I_i \quad (\text{III.14})$$

où $\tau = RC$ est la constante de temps généralement fixée à 1.

Pour un vecteur donnant l'état interne du réseau à $t = 0$, le système d'équation donné ci-dessus décrit complètement l'évolution du réseau: une convergence à partir d'un état initial donné est déterministe, mais, à partir d'états initiaux différents, on peut obtenir plusieurs solutions, en raison de l'existence de minima locaux.

III.6.2 Définition de l'énergie

A partir du critère général de convergence de Lyapunov, on peut associer au réseau la fonction d'énergie suivante:

$$E = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N T_{ij} v_i(t) v_j(t) - \sum_{i=1}^N I_i v_i(t) + \sum_{i=1}^N \frac{1}{\tau} \int_0^{v_i(t)} f^{-1}(v) dv \quad (\text{III.15})$$

Si la fonction de transfert f est monotone croissante et la matrice des poids synaptiques symétrique, on peut montrer que $\frac{\partial E}{\partial t} \leq 0$ (voir annexe B).

Lorsque la température de la fonction sigmoïde est faible (cas limite d'une fonction à grand gain), les états stables du réseau sont les minima locaux de la quantité:

$$E = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N T_{ij} v_i(t) v_j(t) - \sum_{i=1}^N I_i v_i(t) \quad (\text{III.16})$$

Les deux types de réseaux décrits précédemment conduisent à une convergence locale. Les réseaux stochastiques ont été développés afin de remédier à ce problème. Ils assurent une convergence globale du processus, mais celle-ci reste très lente.

III.7 LES RESEAUX DE NEURONES STOCHASTIQUES: LA MACHINE DE BOLTZMANN

Ces réseaux ont été développés à partir des travaux de Kirkpatrick [KIR - 82]. L'analogie est ici faite avec la physique statistique: pour un système de particules, la probabilité de le trouver dans un état donné est proportionnelle au facteur de Boltzmann $\exp(-E/T)$ où E représente l'énergie de l'état et T la température du système. Comme d'autre part, l'approche neuronale des problèmes d'optimisation ne peut conduire qu'à un minimum local, il est tentant de relier le facteur de Boltzmann à l'énergie d'un réseau de Hopfield telle qu'elle a pu être définie ci-dessus, afin d'améliorer la probabilité de convergence vers le minimum global.

La machine de Boltzmann [ACK - 85] [HIN - 84] qui se fonde sur une architecture semblable à celle des réseaux de Hopfield, introduit 2 nouvelles notions:

- l'introduction d'une technique d'optimisation importante, le "recuit simulé".
- l'emploi d'une technique d'apprentissage à deux temps.

Dans sa forme la plus simple, la machine de Boltzmann est un réseau numérique de Hopfield, en ce sens qu'elle impose les mêmes contraintes de connectivité. De plus, son comportement est dicté par un algorithme de recuit simulé [KOR - 89]. Pour un réseau de Hopfield déterministe, l'évolution se produit toujours avec diminution de l'énergie de Lyapunov: pour cette nouvelle machine, on va donc permettre à un neurone un changement d'état autorisant une augmentation d'énergie.

Comme pour les réseaux déterministes, deux modes de fonctionnement existent:

- Mode asynchrone: à chaque itération, un seul neurone change d'état;
- Mode synchrone: à chaque itération, tous les neurones changent d'état.

III.7.1 Mode asynchrone

Pour un réseau numérique symétrique, la variation d'énergie correspondant au changement d'état du neurone i est égale à $\Delta E_{B_i} = -\Delta v_i \cdot u_i$ si $T_{ii} = 0$. Pour une machine de Boltzmann de poids synaptiques T_{ij} et de seuils I_i , le changement d'état ne s'effectue plus suivant le signe de l'état interne u_i , mais par tirage au sort avec des probabilités proportionnelles au facteur de Boltzmann $\exp(-\frac{\Delta E_{B_i}}{T})$.

En normalisant, on obtient les probabilités de transition suivantes:

$$p(v_i(t+1) = 1) = \frac{1}{1 + \exp(-u_i / T)} \quad (\text{III.17})$$

et

$$p(v_i(t+1) = 0) = \frac{\exp(-u_i / T)}{1 + \exp(-u_i / T)} \quad (\text{III.18})$$

que l'on peut aussi écrire:

$$p(\Delta v_i) = \frac{1}{1 + \exp(\Delta E_{B_i} / T)} \quad (\text{III.19})$$

Il est intéressant d'observer que la valeur moyenne prise par un neurone correspond à la valeur de la sortie d'un neurone dans un réseau de Hopfield analogique. Pour compléter l'algorithme, il reste à choisir une loi décroissante de température pour s'assurer de la convergence. En début de convergence, à des températures très élevées, chaque neurone évolue de façon quasi aléatoire,

$$p(v_i = 1) = p(v_i = 0) \approx \frac{1}{2} \quad (\text{III.20})$$

ensuite, lorsque la température est plus faible, c'est l'état interne qui détermine la sortie d'un neurone; enfin, pour des températures proches de 0, le comportement aléatoire est supprimé et le réseau se comporte comme un réseau déterministe.

III.7.2 Mode synchrone

Une machine de Boltzmann synchrone est une machine dans laquelle tous les neurones (ou une partie) examinent à chaque itération la probabilité de changement de leur état de sortie. Le comportement asymptotique d'une telle machine est plus complexe à analyser que celui d'une machine asynchrone, car un réseau de Hopfield déterministe synchrone peut converger vers un cycle de longueur 2.

Bien que ce type de réseaux présente l'avantage d'une convergence globale, cette dernière reste très lente et il devient difficile de maîtriser le processus en cours d'évolution. De plus, de tels réseaux sont difficilement réalisables à grande échelle, ce qui en limite les possibilités d'utilisation.

Une version "déterministe" de la machine de Boltzmann a été proposée par Hinton [HIN - 89]. Ici, l'activation bivaluée des neurones a été remplacée par des variables continues représentant la probabilité moyenne que le neurone soit activé. Il a été démontré que la technique d'apprentissage associée effectue une descente de gradient sur la même surface d'erreur que la machine originale. Cette nouvelle version permet de réduire le temps d'apprentissage, tout en conservant les avantages de la machine de Boltzmann.

III.8 APPLICATIONS

III.8.1 Mémoire associative

Les réseaux totalement connectés ont des propriétés de mémoire associative qui ressemblent de manière frappante aux propriétés de la mémoire biologique. Ces réseaux peuvent en effet associer une version déformée d'une information à un original préalablement mémorisé, et ainsi "reconnaître" l'information déformée.

La première phase consiste à initialiser les poids synaptiques qui revient en fait à mémoriser les M modèles. Lorsqu'on présente au réseau une forme, il exécute plusieurs itérations jusqu'à ce que deux états consécutifs des neurones soient identiques. La forme présentée s'assimilera donc au résultat obtenu autrement dit à la classe du modèle préalablement appris. Durant la phase d'apprentissage, l'ensemble des prototypes à mémoriser est présenté au réseau qui calcule le poids des connexions pour arriver à une stabilisation du réseau. Ce calcul peut être effectué de différentes façons, soit en suivant une règle de type Widrow-Hoff, ou de type Hebb.

Toutefois, les limites suivantes apparaissent:

- La capacité d'apprentissage de la règle de Hebb est très réduite. D'après [AMI - 85], en supposant le nombre N de neurones très grand, le nombre maximal d'états correctement retrouvés est inférieur à $0.138 N$. Si on dépasse $0.138 N$ qui est la capacité limite, c'est-à-dire si on veut "surcharger" la mémoire, tous les prototypes précédemment appris sont perdus.
- L'information ne peut pas être stockée brutalement dans la mémoire. Il faut prétraiter les données afin de les orthogonaliser, par exemple ou de les rendre au pire linéairement indépendantes afin d'approcher au plus près les capacités théoriques des réseaux.
- Il faut éviter de classer une configuration de manière erronée. Ce comportement est à comparer à la non reconnaissance qui consiste à classer la configuration présentée dans un état "poubelle".

Une application de ce réseau en tant que mémoire associative est donnée par Kohonen [KOH - 88] sur la reconnaissance de visages, comme le montre la *figure III.3*.

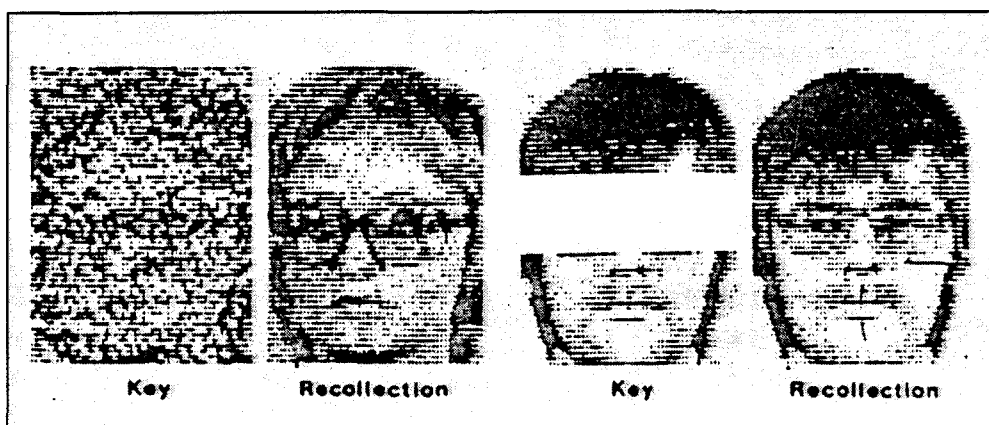


Figure III.3 : Extrait de "Self Organization and Associative Memory",
T. Kohonen, Springer Verlag 1988.

Le réseau utilisé pour cet exemple est bâti avec 3024 neurones. Il y a donc 9144576 interconnexions. L'auteur a codé 100 configurations de référence dans ce réseau. Dans le premier exemple, l'entrée du réseau est une configuration bruitée, le bruit étant dérivé d'une distribution uniforme. Dans le second exemple, c'est un prototype dans lequel on a enlevé 25% de l'information qui est présentée. Dans les deux exemples, le système associe à l'entrée proposée le représentant correct de la classe.

III.8.2 Résolution des Problèmes d'optimisation

Hopfield et Tank furent les premiers à utiliser les réseaux neuromimétiques pour résoudre des problèmes d'optimisation. Le problème le plus connu est celui du voyageur de commerce dans lequel il s'agit de relier n villes entre elles en minimisant le chemin parcouru [HOP - 86]. Pour cela, Hopfield et Thank ont codé le problème avec un réseau cartésien de $(N \times N)$ neurones, chaque ligne étant associée à une ville, et chaque colonne au numéro d'ordre de la ville dans le chemin. Ainsi, un chemin est caractérisé dans le réseau par l'activité de N neurones parmi les N^2 : un seul neurone actif par ligne et par colonne.

Ce problème d'optimisation est de type NP-complet, le temps nécessaire à la résolution croît exponentiellement avec le nombre de villes. L'approche retenue

impose de déterminer une "bonne" fonction d'énergie qui place les solutions du problème comme des minima locaux de cette fonction. Ainsi, le réseau évolue vers une configuration stable, constituant une solution optimisée du problème. La démonstration de ce résultat est faite en utilisant une fonction de Lyapunov. Appliqué à 30 villes, le problème est résolu en quelques millièmes de secondes alors que cela demande plus d'une seconde sur un gros ordinateur. Toutefois, dans le cas d'un nombre de villes beaucoup plus important, l'utilisation de réseaux neuromimétiques devient beaucoup moins évidente du fait de l'explosion combinatoire du nombre de calculs à effectuer.

Dans ce domaine d'application, les contraintes du problème sont exprimées par la connectivité du réseau, et par des limites supplémentaires imposées à sa dynamique. L'état initial du réseau correspond aux données du problème, et l'état final du réseau à la solution qu'il propose. Ici aussi, plusieurs chercheurs ont effectué des travaux dans ce contexte, comme par exemple Tagliarini [TAG - 87], Amatur [AMA - 92], Park [PAR - 89], Matsui [MAT - 95], Cooper [COO - 95].

Cependant, un des problèmes avec les réseaux de Hopfield lorsqu'ils sont utilisés pour résoudre des problèmes d'optimisation découle de leur tendance à rester bloqués dans des minima locaux. Afin de s'affranchir de cette difficulté, des techniques semblables au recuit simulé [WEI - 89] [DAV - 90] sont utilisées. En effet, si l'on souhaite que l'on arrive au minimum global, lors de la minimisation et ne pas rester bloqué dans un minimum local, l'utilisation des machines de Boltzmann munies de fonctions de transfert stochastiques en employant la méthode du recuit simulé s'avèrent nécessaires pour résoudre ce type de problème.

III.9 CONCLUSION

Nous avons vu qu'un réseau de neurones récurrent est un système dynamique et que son évolution peut être agencée de plusieurs façons. Le réseau reçoit des entrées et propage son activation pendant un nombre fixe de cycle d'activation. L'activation initiale des neurones du réseau est fixée par l'utilisateur. On laisse ensuite le réseau évoluer jusqu'à un état stable. Cet état est considéré comme la réponse du réseau.

Deux principales "contraintes" du modèle de Hopfield sont ses liaisons symétriques et la connectivité totale. Des modèles du type Hopfield mais avec des connexions partielles (diluées) et non symétriques ont été étudiés par Amit [AMI - 87] et Sompolinsky [SOM - 86].

Les travaux sur les réseaux de Hopfield constituent une source très fructueuse en connexionnisme. Ces travaux touchent plusieurs disciplines, de la psychologie à l'analyse théorique même des réseaux neuromimétiques, en passant par d'autres sciences cognitives. De leur côté, les réseaux récurrents de type Hopfield ont prouvé leur efficacité dans la résolution de problèmes d'optimisation. De plus, leur architecture a permis des réalisations électroniques [GRA - 89] et optiques [HSU - 90] [FAR - 89] performantes.

De tels réseaux sont bien adaptés à l'étude des séquences d'images dynamiques. Par exemple, Zhou et Chellappa [ZHO - 92] ont bien montré le lien entre les réseaux neuronaux récurrents dont les neurones utilisent des fonctions d'activation non linéaire et l'analyse du mouvement dans une séquence d'images.

Toutes ces applications nécessitent un codage du problème sous la forme d'un réseau de neurones artificiels décrit au moyen d'une forme quadratique, appelée *énergie du problème*, et dont le minimum correspond à une solution optimale du problème. Une fois cette énergie définie, il est nécessaire de choisir la méthodologie neuronale à mettre en oeuvre pour sa minimisation.

CHAPITRE IV

LA PROBLEMATIQUE DE L'ESTIMATION DU MOUVEMENT

CHAPITRE IV

LA PROBLEMATIQUE DE

L'ESTIMATION DU MOUVEMENT

Afin de limiter notre étude dans le contexte très vaste des méthodes d'analyse de mouvements à partir d'une séquence d'images, nous nous bornerons à présenter les grands axes des méthodes générales d'estimation du mouvement 2D s'appuyant sur un dispositif monoculaire. Par contre, nous présenterons de façon plus approfondie les techniques d'estimation du champ de vecteur vitesse apparente ("flux optique").

IV.1. ESTIMATION DU MOUVEMENT DANS UNE SEQUENCE D'IMAGES

L'intérêt de l'analyse de mouvement dans une séquence d'images a été principalement suscité par de grands domaines de recherches [BOU - 88]:

- ✓ La transmission de séquences d'images où l'estimation du mouvement permet d'extraire la redondance spatio-temporelle présente dans la suite des valeurs d'intensité d'une séquence d'images;

- ✓ L'analyse de scènes dynamiques qui consiste à caractériser le mouvement 3D des objets de la scène et du capteur à partir de l'information de mouvement fournie par la séquence d'images.

On peut citer quelques exemples d'applications dans les domaines suivants:

- La Météorologie: Estimation du mouvement de masses nuageuses à partir d'images satellite.
- La Biologie Humaine, Animale, Microscopique
 - * Etude cinématique des organes humains;
 - * Analyse du mouvement du coeur;
 - * Analyse des articulations des animaux;
 - * Trajectométrie automatique de micro-organisme.
- Les Transports
 - * Surveillance de trafic;
 - * Guidage de robots mobiles.
- L'agriculture: Cueillette automatique de fruits.
- Le Domaine militaire
 - * Suivi de cibles;
 - * Guidage de sous-marins.
- La Robotique
 - * Guidage de robots.
- Les Télécommunications
 - * Transmissions;
 - * Codage et compression de données en télévision.

Les premières études menées dans le cadre de l'analyse de séquences d'images ont porté sur la détection des variations temporelles de la fonction intensité lumineuse, leurs statistiques et leurs corrélations avec des primitives spatiales. Au début, l'analyse des séquences d'images a été considérée comme une extension de l'analyse d'images statiques. Le principe consistait à reconnaître les objets de la scène sur chaque image prise de façon séparée, puis à partir de leurs positions, calculer leur mouvement. Très vite des méthodologies différentes de celles de l'analyse statique ont enrichi la littérature. Ces méthodes se concentrent plus sur le mouvement lui-même que sur la détermination explicite des objets.

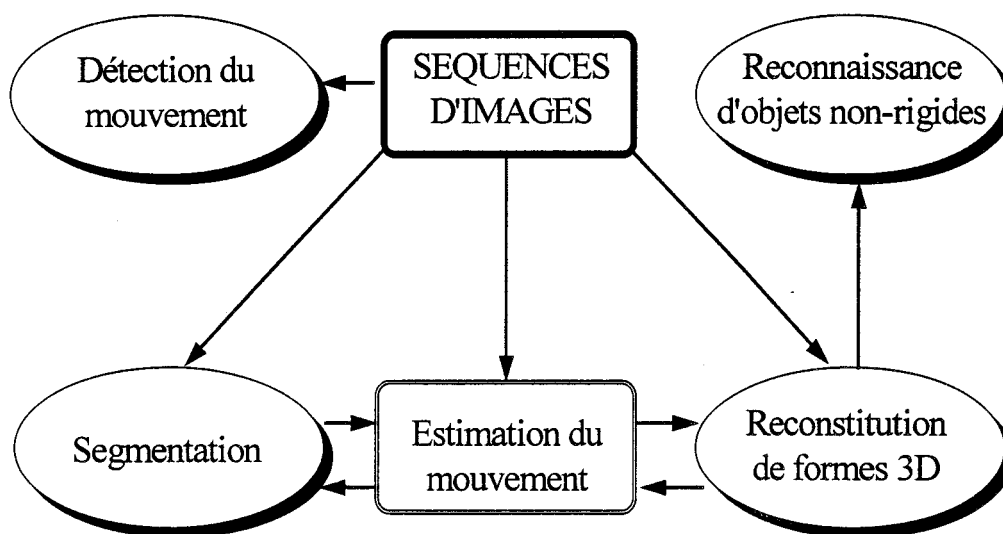


Figure IV.1 : Principaux domaines de l'analyse de séquences d'images

Les phases importantes de l'analyse du mouvement peuvent être résumées comme suit:

- Détection du mouvement;
- Segmentation par contours spatio-temporels ou en régions de mouvement apparent différent;
- Estimation du mouvement (du champ des vitesses);

- Interprétation du mouvement (quantitative, qualitative).

Ces modules peuvent être utilisés de façon autonome ou en combinaison avec d'autres. Tous ne sont pas détaillés ici.

Il est possible de classer les techniques d'analyse de mouvement dans les séquences d'images suivant les niveaux d'interprétation des données auxquelles elles font appel [CAB - 92].

IV.1.1. Haut niveau (méthodes structurelles)

Les méthodes structurelles sont basées sur l'étude de l'évolution temporelle de certaines caractéristiques suffisamment "riches" et robustes d'objets présents dans différentes images. L'extraction de celles-ci est donc un préalable indispensable à la mise en correspondance proprement dite.

La recherche des éléments structurels peut se faire soit sur l'image initiale, soit sur une image transformée (image de contours, transformée de Hough). La fiabilité de la mise en correspondance est directement liée au choix de ces éléments structurels qui peuvent être, par exemple, des:

- pixels;
- éléments de contours (segments de droite, arc de cercle, ...);
- contours fermés;
- régions de l'image.

IV.1.2. Niveau intermédiaire

Des traitements effectués sur la séquence d'images pouvant combiner des caractéristiques locales (par exemple l'évolution des niveaux de gris dans une région)

et des caractéristiques globales (évolution de l'illumination de la scène observée) permettent de segmenter l'image en zones homogènes vis à vis du mouvement.

Le problème principal de la méthode est l'estimation de la carte de vitesses en suivant le déplacement de chaque point de l'image. L'interprétation de cette carte de vitesse permet de retrouver le mouvement ou la structure 3D de l'objet [ADI - 85].

L'analyse du "flux optique", qui consiste en l'étude du champ du vecteur vitesse apparente sur les images, appartient à cette catégorie.

IV.1.3. Bas niveau (méthodes orientées mouvement)

Des traitements locaux, mettant en jeu tant le voisinage spatial d'un pixel que son voisinage temporel (pixels situés à la même position dans les autres images) fournissent une nouvelle image dans laquelle les pixels affectés par le mouvement se différencient simplement des pixels des zones stationnaires.

Leur principe consiste souvent en l'élimination des zones de l'image restant immobiles durant la séquence. Ces méthodes sont presque toujours basées sur l'exploitation des différences entre images.

La mise en oeuvre de ces méthodes est simple avec cependant une difficulté de segmenter globalement l'image différence quand les objets mobiles présents dans une séquence d'images ne sont pas constitués de zones de niveau de gris homogène.

Dans de nombreux cas, l'étude du mouvement se restreint à estimer dans un premier temps un paramètre particulier, en général le vecteur vitesse ou le vecteur déplacement. Il est possible de définir une classification des méthodes d'estimation du champ des déplacements qui rejoint la classification présentée précédemment.

IV.2. METHODES D'ESTIMATION DU CHAMP DES DEPLACEMENTS

IV.2.1. Introduction

Conformément à l'application qui nous intéresse, seul le "flux optique", ou champ de vitesse apparente à deux dimensions (*figure IV.1*), sera considéré dans la suite.

Nous présentons dans ce qui suit un panorama des techniques continues d'estimation spatio-temporelles d'un tel champ de vitesse. On peut en effet, classer les méthodes permettant l'estimation du champ des déplacements en 4 grandes classes. Ce sont les méthodes:

- Différentielles;
- Par transformées;
- Par mise en correspondance;
- Probabilistes.

Certaines de ces techniques font appel à plusieurs niveaux de la classification établie dans le paragraphe IV.1. Les méthodes de mise en correspondance, par exemple, font coopérer le niveau bas et le niveau intermédiaire ou encore les niveaux intermédiaire et haut, selon les types de primitives extraites de l'image.

Il est à noter enfin qu'une distinction peut être faite entre ces méthodes. Cette différence est liée au caractère *déterministe* ou *probabiliste* de ces approches [HEI - 93] [LAL - 89].

IV.2.2. Les méthodes différentielles

Parmi les méthodes continues d'estimation spatio-temporelle du champ de vitesse permettant d'effectuer la segmentation ou l'interprétation d'une scène $3D+t$, on distingue l'approche du "flux (ou flot) optique". Cette méthode se situe entre les techniques de différences et les techniques d'appariement (niveau intermédiaire). Le "flux optique" peut être considéré comme la projection du champ de vitesse 3D, qui est le mouvement réel de chaque point de la scène, sur le plan image du capteur. Il s'agit d'associer à chaque pixel d'une image de la séquence un vecteur déplacement permettant de le situer dans l'image suivante. Le champ de vecteur ainsi obtenu est appelé champ des déplacements. En divisant chacun de ces vecteurs par l'intervalle de temps séparant les acquisitions d'images successives, on obtient un nouveau champ de vecteur représentant cette fois la vitesse de chaque point.

Le "flux optique" fait d'abord appel à un traitement local pour le calcul de gradients d'intensités spatiaux et temporels; ces gradients serviront ensuite à estimer les vitesses à l'aide d'une minimisation globale d'une fonctionnelle sur l'ensemble de l'image.

Il est intéressant de remarquer le lien existant entre le champ du "flux optique" et celui des vitesses. En effet, ces deux derniers ne sont identiques que dans le cas d'un mouvement parallèle au plan image. Dans certains cas, le "flux optique" estimé peut ne pas correspondre au champ de vitesse réel, c'est-à-dire à la réalité physique. Afin d'écartier toute confusion entre le "flux optique" et le champ de vitesse, nous donnons deux exemples:

1. Imaginons une scène parfaitement fixe dans laquelle une boule de billard, réfléchissante ou non, est animée d'une rotation pure autour d'un de ses axes. Dans cette image:

- le "flux optique" est nul,
- le champ de vitesse n'est pas nul.

2. Soit une scène parfaitement fixe dans laquelle se trouve un prisme réfléchissant. Lorsque la source d'éclairage se déplace:

- le "flux optique" n'est pas nul,
- le champ de vitesse est nul.

On attribue généralement la paternité de la méthode à Horn et Schunck [HOR - 81]. De fait, c'est là qu'on retrouve le premier algorithme à caractère itératif pour le calcul d'un champ de vitesses dense (c'est-à-dire un champ où un vecteur vitesse est attribué à chacun des pixels de l'image) à partir de l'équation du "flux optique". Une contrainte de luminance est introduite: elle suppose qu'un point d'intensité lumineuse donnée dans une image conserve son intensité lumineuse dans l'image suivante. En introduisant une contrainte supplémentaire de régularisation sur le champ de vitesse, un champ de vitesse unique est évalué dans les séquences d'images 2D. Toutefois, plusieurs des idées de base auraient déjà été retracées dans les travaux de Limb et Murphy [LIM - 74], Cafforio et Rocca [CAF - 76], Fennema et Thompson [FEN - 79], Netravali et Robbins [NET - 79]. Par ailleurs, il est intéressant de noter que l'expression même "flux optique" avait déjà été introduite par Gibson en 1959 lors d'une étude de la modélisation des mouvements perçus par l'œil [HAY - 66].

Les techniques basées sur les gradients d'intensités d'images en niveaux de gris se sont avérées efficaces pour l'estimation de champs denses de vecteurs déplacement dans une séquence d'images. L'équation qui en est à la base est celle de la *contrainte du mouvement apparent (ECMA)* qui décrit par une relation linéaire le lien entre les gradients d'intensités et les composantes du champ de vitesse.

Ces méthodes se distinguent entre elles par [BOU - 88]:

- la façon dont s'exprime l'équation de contrainte du "flux optique";
- le choix de la procédure de minimisation;

premier ordre ce qui exclut les zones de masquages ou de mises à découvert d'objets par d'autres.

Le développement de Taylor permet d'écrire la relation (IV.1) :

$$f(x+dx, y+dy, t+dt) = f(x, y, t) + \frac{\partial f(x, y, t)}{\partial x} dx + \frac{\partial f(x, y, t)}{\partial y} dy + \frac{\partial f(x, y, t)}{\partial t} dt + \varepsilon$$

ε représente les termes d'ordres supérieurs. Nous ferons l'hypothèse que ces termes d'ordres supérieurs à un sont négligeables.

Si l'on admet que l'illumination de la scène ne change pas significativement entre les instants t et $t+dt$ séparant la prise de deux images successives, alors la fonction de niveaux de gris $f(x,y,t)$ au point de coordonnées (x,y) de l'image vérifie:

$$f(x+dx, y+dy, t+dt) = f(x,y,t) \Rightarrow df(x,y,t) = 0 \quad (\text{IV.2})$$

dx et dy sont alors solutions de:

$$\frac{\partial f(x, y, t)}{\partial x} \cdot dx + \frac{\partial f(x, y, t)}{\partial y} \cdot dy + \frac{\partial f(x, y, t)}{\partial t} \cdot dt = 0 \quad (\text{IV.3})$$

ou encore en divisant les deux termes de l'équation précédente par dt , on obtient:

$$\frac{dx}{dt} \cdot \frac{\partial f(x, y, t)}{\partial x} + \frac{dy}{dt} \cdot \frac{\partial f(x, y, t)}{\partial y} + \frac{\partial f(x, y, t)}{\partial t} = 0 \quad (\text{IV.4})$$

Les termes $\frac{dx}{dt}$ et $\frac{dy}{dt}$, représentent les composantes selon les directions x et y du

vecteur vitesse ω à estimer. Si l'on note $\omega_x = \frac{dx}{dt}$ et $\omega_y = \frac{dy}{dt}$ les composantes du

vecteur vitesse dans le plan image, on obtient la relation de l'E.C.M.A définie par Fennema et Thompson [FEN - 79] ainsi que par Horn et Schunck [HOR - 81]:

- les contraintes additionnelles prises en compte.

IV.2.2.1 Principe

Il est connu aujourd'hui que les variations temporelles de niveaux de gris dans une séquence d'images permettent, sous certaines hypothèses, de "remonter" à des estimations des paramètres du mouvement - principalement les composantes du vecteur vitesse - dans le cas d'analyse tridimensionnelle (3D) ou de leur projection dans un plan - celui de la caméra par exemple - dans le cas bidimensionnel (2D). Le calcul de ces estimations repose généralement sur un certain nombre d'hypothèses quant à la nature qualitative du mouvement. Ce dernier est pris par exemple comme un processus continu qui se manifeste par des changements réguliers au regard de la cadence de succession des images, ainsi que sur la nature des objets en déplacement - rigides - non rigides - par exemple. A partir d'hypothèses de ce type, il est possible de définir l'équation de contrainte du mouvement apparent (E.C.M.A).

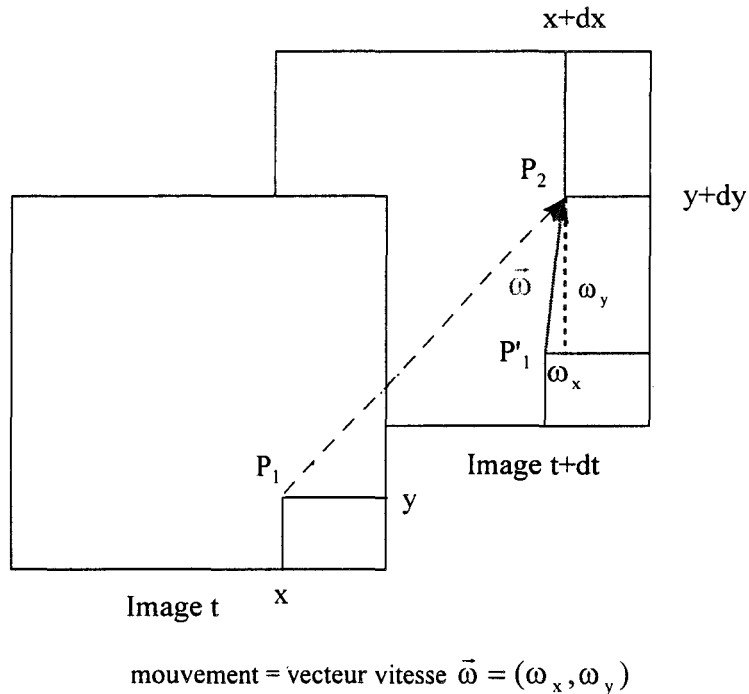


Figure IV.2 : Vitesse d'un point.

Soit la projection d'une scène réelle sur un plan image à deux dimensions et soit $f(x,y,t)$ son intensité lumineuse. On suppose que f est différentiable au moins jusqu'au

$$f_t + f_x \omega_x + f_y \omega_y = 0 \quad \text{avec} \quad f_i = \frac{\partial f}{\partial i} \quad (\text{IV.5})$$

$$\text{que l'on peut noter vectoriellement par } \nabla \vec{f} \cdot \vec{\omega} = -f_t \quad (\text{IV.6})$$

où $\nabla \vec{f} = (f_x, f_y)$ représente le gradient spatial de la fonction d'intensité, et $f_x = \frac{\partial f}{\partial x}$, $f_y = \frac{\partial f}{\partial y}$ et $f_t = \frac{\partial f}{\partial t}$ représentent les dérivées partielles de la fonction f par les variables indiquées en indice.

On remarque par la relation IV.6 que seule la composante du vecteur vitesse parallèle au gradient d'intensité est directement mesurable. Son amplitude est donc:

$$\frac{-f_t}{\sqrt{(f_x)^2 + (f_y)^2}} \quad (\text{IV.7})$$

L'ensemble des solutions décrit une droite (figure IV.3) :

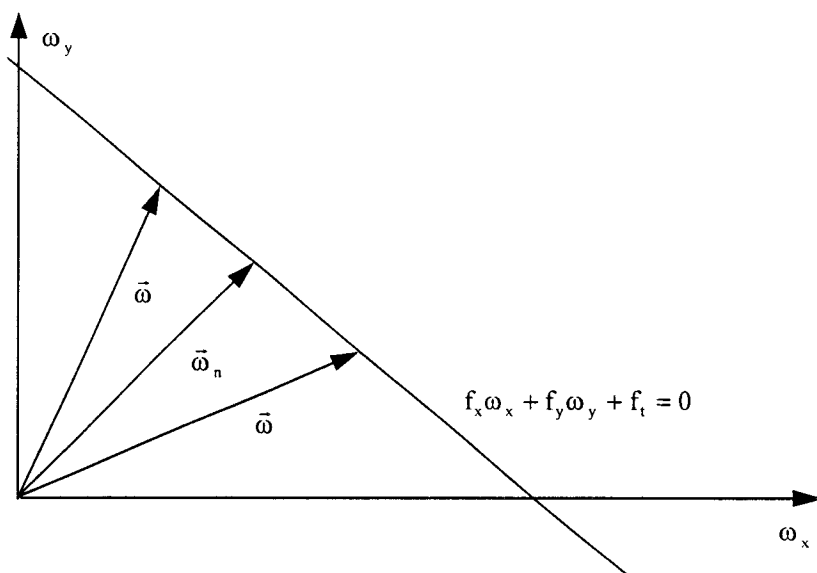


Figure IV.3: Droite définie par l'ECMA.

Dans la pratique toutefois, il faut noter que seuls les gradients spatio-temporels (f_i) sont observables et que l'on ne connaît pas la loi d'évolution de $\omega_x(t)$ et $\omega_y(t)$. Il

résulte de ceci que l'équation IV.6 ne donne pas véritablement les vecteurs vitesses mais les variations de position des structures de niveaux de gris.

L'ensemble des vecteurs $\bar{\omega}$ obtenus avec l'équation IV.6 est donc appelé "flux optique". On constate néanmoins que dans la très grande majorité des cas, il est possible d'assimiler le "flux optique" avec les vecteurs vitesses. Cette équation est cependant insuffisante puisqu'elle fournit un système sous-déterminé. De plus, la nature éventuellement bruitée des images fait qu'il n'est pas garanti d'obtenir une solution stable pour cette équation, ce qui fait dire que la détermination du vecteur vitesse est un problème mal posé au sens de Hadamard. On peut résumer les autres problèmes posés par l'ECMA en quatre principaux points suivants [SCH - 86] [THO - 87] [VER - 86] [BOU - 89]:

- Phénomènes d'ombrage;
- Propriétés de la fonction intensité:
 - * fortes textures: non respect de l'hypothèse de linéarité locale de la fonction intensité;
 - * zones uniformes: gradients spatiaux nuls, pas suffisamment d'informations.
- Comportement aux frontières:
 - * contours d'intensité: dans les zones de contraste fort et abrupt de l'intensité lumineuse, cette intensité ne présente pas un profil linéaire, même localement;
 - * contours d'occlusion: à position spatiale fixe, on change d'objet d'un instant au suivant; la mesure du gradient temporel de l'intensité lumineuse ne peut plus être analytiquement reliée aux variations de l'intensité.
- Restrictions sur le mouvement:

- * grands déplacements: problème des occlusions; non recouvrement des objets d'un instant à l'autre.

Ces difficultés et restrictions ont amené certains auteurs à développer des approches reposant sur l'utilisation d'une ou plusieurs contraintes de lissage (optimisation globale, régularisation) que doivent vérifier les dérivées partielles spatio-temporelles de l'intensité de l'image. Le problème de la recherche de ω peut être alors formulé comme celui de la minimisation d'une fonctionnelle comportant un terme de contrainte lié à la nature du problème (ici l'E.C.M.A) et un terme de régularisation lié en général à des hypothèses de lissage de la solution. Dans ce cadre, deux classes d'approches ont été abordées pour pallier nombre de problèmes que présente l'utilisation de l'ECMA dans la mesure du mouvement apparent. Il s'agit de l'approche de type multirésolution [ENK - 86] où il est possible d'accéder à une mesure du mouvement fiable pour des amplitudes de déplacement grandes et celle de type markovienne [KON - 90] [HEI - 91] qui permet d'intégrer plusieurs sources de mesure du mouvement et de propager des informations de mouvement dans les zones où l'ECMA n'est pas opérationnelle.

On peut alors distinguer diverses approches pour estimer le champ des vitesses. Elles se classent en deux grands schémas [BOU - 88] [NAG - 87]:

1. **Optimisation locale:** utilisation de l'ECMA sans contrainte de lissage [MIT - 87] [MIL - 89].
2. **Optimisation globale, régularisation:** introduction d'une contrainte de lissage à l'équation fondamentale. Dans cet optique, il existe deux techniques fondamentales pour mener l'estimation:
 - Méthodes de régularisation déterministes [HOR - 81] [SCH - 86] [NAG - 87] [ENK - 86] [TER - 86]: minimisation d'une fonctionnelle d'énergie comportant un terme relatif aux variations spatio-temporelles de l'intensité, et l'autre favorisant la continuité du champ de vecteurs.

- Méthodes d'estimation stochastique: introduction de contraintes à travers de modèles statistiques appropriés (estimation d'un processus stochastique 2D [ROU - 87]), modélisation markovienne avec prise en compte des discontinuités de mouvement [KON - 90] [HEI - 91].

IV.2.2.2 Conclusion sur les approches différentielles

La méthode de détermination du "flux optique" basée sur des paramètres de continuité impose des déplacements très faibles dans la scène. Le calcul est souvent réalisé par des algorithmes utilisant des estimations itératives de dérivées partielles spatiales et temporelles des signaux.

La limitation principale des méthodes de mise en évidence du mouvement basées sur l'évaluation du "flux optique" est leur sensibilité aux bruits de mesure et aux variations pouvant apparaître dans les conditions d'acquisition. Ceci rendra difficile l'obtention d'une bonne estimation des dérivées partielles, spatiales et temporelles d'ordre de plus en plus élevé. Certaines approches ont été proposées afin de diminuer la sensibilité au bruit des méthodes utilisant le gradient spatio-temporel [AIS - 89], mais elles entraînent souvent une perte de précision concernant les mouvements détectés.

Cependant, ces méthodes fournissent des solutions très élégantes théoriquement dans le domaine de l'analyse du mouvement apparent ou de l'estimation des structures 3D. En effet, le "flux optique" est très riche en informations et peut ainsi permettre l'obtention d'une multitude de caractéristiques concernant les objets observés.

IV.2.3. Les méthodes par transformée

Comme méthode, on peut citer:

- La transformée de Hough [SAM - 84] [LI - 86];

- Les transformées fréquentielles, ex: transformée de Fourier [FLE - 85] [SAG - 89] [MIT - 92];
- Les transformées spatio-temporelles, exemple: fenêtre de Gabor [HEE - 88] [EAG - 89] ou transformée en ondelettes [DAU - 88] [MAL - 89].

IV.2.3.1 La transformation de Hough

La recherche des différents mouvements présents dans une séquence d'images peut se faire par transformation de Hough du champ des déplacements. En effet, cette transformation est généralement bien adaptée au problème de segmentation spatio-temporelle, posé lors de la détermination du "flux optique", puisqu'elle permet la mise en correspondance d'objets ayant subi au cours du temps des déformations géométriques (translations, rotations, homothéties, etc).

Le principe de la transformée de Hough consiste à passer d'un *espace image* à un autre appelé *espace des paramètres*. Ce dernier est subdivisé en une grille d'*accumulateurs*, chacun de ces derniers correspondant à une combinaison de paramètres. La détection des points d'accumulation dans l'espace de Hough peut se réaliser en employant des techniques de classification automatique non supervisée [ORO - 81], ou par mise en évidence de formes par utilisation du principe des nuées dynamiques (DID - 82).

La transformée de Hough est cependant robuste au bruit et elle est insensible aux occlusions partielles. Le calcul des paramètres de mouvement ne paraît possible que:

- pour des mouvements très simples (translation ou rotation);
- si l'on connaît déjà une ou plusieurs valeurs de ces paramètres;
- dans le cas d'images simples aux contours très droits et bien définis.

Le principal inconvénient de la méthode réside dans le fait que la complexité croît exponentiellement avec la dimension de l'espace des paramètres ou lorsqu'on désire augmenter le nombre d'accumulateurs dans le but d'améliorer la résolution.

IV.2.3.2 L'approche fréquentielle

Cette méthode se base sur les propriétés de la transformée de Fourier. En analysant la phase de la transformée de Fourier de chaque image de la séquence étudiée, il est alors possible de déterminer un terme de déplacement [HUA - 81].

Il faut remarquer que ces méthodes ne sont utilisables que pour des cas très simples de translations ou de rotations pures sur un fond uniforme. Bien que prometteuses, de telles méthodes demandent des calculs très lourds. C'est pourquoi, leur emploi s'avère très limité.

IV.2.3.3 Les transformées spatio-temporelles

On peut citer principalement la méthode de Wigner-Ville [DEH - 94], ainsi que des approches multirésolution par transformation en ondelettes. Cette dernière représentation est obtenue par la projection récursive de l'image de départ sur des bases orthogonales particulières appelées ondelettes. Ces bases sont obtenues par translation et dilatation d'une fonction unique appelée ondelette mère.

Les avantages de la représentation en ondelette sont nombreux:

- Représentation complète;
- Existence d'ondelettes orthogonales;
- A l'inverse de la transformée de Fourier, elle possède la propriété de localisation spatiale et fréquentielle.

IV.2.3.4 Conclusion sur les approches par transformée

Les méthodes par transformée se distinguent en général par de lourds calculs, ce qui réduit considérablement le champ d'application potentiel. Cependant, on peut penser que l'introduction d'informations complémentaires peut permettre des estimations de qualité. L'intégration de la multirésolution par exemple, présente l'avantage de pouvoir estimer les mouvements de grandes amplitudes bien que le problème du passage d'un niveau à un autre n'est pas résolu d'une manière satisfaisante. Toutefois, elle peut améliorer les estimations et réduire les temps de calculs.

IV.2.4. Méthodes de mise en correspondance

Une autre catégorie de détermination du champ de déplacement est représentée par les méthodes dites de mise en correspondance [MOU - 90] [KAO - 93] [BEN - 92]. La mise en correspondance implique une segmentation statique des images successives. L'information mouvement est alors déduite des relations entre les primitives extraites. Les principaux aspects qui conditionnent l'efficacité et la fiabilité de mise en correspondance sont:

- Le choix des primitives: certains types de primitives sont plus aisés à appairer. De plus, ce choix détermine la représentation finale de la scène;
- La fiabilité de leur extraction;
- Les conditions d'appariement: des critères et contraintes doivent être satisfaits lors de cet appariement.

Les caractéristiques ou primitives à appairer peuvent être des pixels, des contours fermés, des segments de droite, ou des régions de l'image. Il n'existe pas de méthode générale pour réaliser la mise en correspondance. Cependant, il est utile de citer les plus usuelles telles que les méthodes de relaxation, celles utilisant le principe de prédiction et vérification d'hypothèses ou encore celles basées sur l'appariement de segments par graphes d'adjacence.

Qu'ils s'agissent de séquences d'images monoculaires ou de séquences d'images binoculaires, la plupart des algorithmes de mise en correspondance sont transposables d'un type de séquence à l'autre. Dans le premier cas, les mises en correspondance se font exclusivement entre les images successives de la séquence (appariement temporel). Dans le second cas, elles se font entre chaque paire d'images droite - gauche acquises aux mêmes instants [AYA - 89]. Le champ de vecteurs mettant en correspondance les primitives des deux images stéréoscopiques est aussi nommé "champ de disparités".

On distingue deux types de méthodes de mise en correspondance. Ce sont les méthodes de mise en correspondance bas niveau et haut niveau. Dans le premier type d'approche, le principe de mesure de ressemblance entre deux ensembles de données, appelé aussi corrélation, est utilisé. Les mesures de similarité ou de corrélation sont généralement basées sur le calcul d'une distance. Cette technique est souvent utilisée en codage d'images où le problème de la précision de l'estimation est secondaire. On connaît actuellement encore des développements par l'utilisation de blocs de formes ("Block Matching") adaptées aux contours de l'image.

Le deuxième type d'approche s'appuie sur des caractéristiques des objets étudiés. L'extraction de celles-ci est donc un préalable indispensable à la mise en correspondance. Parmi les approches de mise en correspondance haut niveau, on peut citer l'approche prédiction-vérification d'hypothèses [MOU - 90] [SDI - 94] et l'approche classification [GIV - 89].

IV.2.4.1 Conclusion sur les approches de mise en correspondance

L'avantage de telles méthodes réside dans une plus grande précision dans l'estimation du mouvement quand les objets possèdent des caractéristiques facilement repérables dans l'image. Un autre avantage est qu'elles peuvent supporter des changements relativement importants (translations, rotation,...). Cependant, ces méthodes fournissent des champs de déplacement épars car l'information de mouvement n'est mesurée qu'en certains endroits de l'image. La détermination du "flux optique" est moins sensible au bruit de mesure, mais le volume de calcul est très grand.

Le problème lié à leur emploi est le choix difficile de l'extracteur qui impose en général une connaissance a priori de la scène.

IV.2.5. Méthodes probabilistes

Ces méthodes constituent une alternative aux méthodes déterministes. On peut distinguer l'approche Bayésienne qui consiste à aborder le problème de l'extraction de l'information de mouvement [HEI - 89] à partir d'une séquence d'images en utilisant des méthodes basées sur le critère du maximum de probabilité *a posteriori* de Bayes. Une autre approche aux problèmes de l'estimation de mouvement consiste à une modélisation en termes de champs de Markov [HEI - 93]. La théorie de ces champs fait appel aux notions de "sites" et de "voisinage".

Ces méthodes s'affranchissent de l'hypothèse très restrictive de la constance de la luminance, et peuvent traiter des mouvements beaucoup plus complexes que la translation ou des rotations pures. Les discontinuités de mouvement sont aussi mieux traités. Cependant, un des problèmes rencontrés dans la seconde approche est lié à la détermination des paramètres définissant les distributions des champs de Markov.

IV.3. CONCLUSION

L'étude menée n'est pas exhaustive en raison du nombre de publications parues dans ce domaine. Toutefois, nous remarquons qu'aucune de ces approches ne s'avère utilisable en toute situation. Le choix d'une méthode particulière ne peut se faire qu'en analysant les spécificités de l'application envisagée. Il est à noter aussi que la classification que nous avons adoptée n'est pas unique et qu'il est difficile de classer certaines méthodes dans l'une ou l'autre des approches présentées.

Comme on vient de le voir, le problème d'estimation du mouvement par les méthodes différentielles et plus particulièrement la détermination du "flux optique" est un problème "mal posé". Ce qui est aussi le cas pour d'autres problèmes principaux de la vision [POG - 85]. Il s'avère donc nécessaire afin de transformer de

tels problèmes et de restreindre l'espace de recherche des solutions d'introduire dans les fonctionnelles à minimiser des connaissances a priori sur la nature de l'estimation. De telles fonctionnelles sont généralement issues de la théorie de la régularisation au sens de Tikhonov [TIK - 76]. Les contraintes additionnelles introduites consistent à minimiser une erreur de préférence quadratique. Le problème se ramène donc à un problème d'optimisation qui peut être abordé par le biais de différentes méthodes. L'un des premiers algorithmes, sinon le premier, à caractère itératif pour le calcul d'un champ dense de vitesses c'est-à-dire un champ où un vecteur vitesse est attribué à chacun des points de l'image considérée, a été développé par Horn et Schunck [HOR - 81]. Par la suite, d'autres méthodologies de minimisation dédiées à l'estimation du champ de vitesse ont été développées. On peut citer par exemple l'approche présentée dans [HEI - 91] qui permet la prise en compte de la contrainte de lissage tout en respectant les frontières de mouvement grâce à l'utilisation de la modélisation markovienne, ou encore d'autres approches utilisant la multirésolution.

Depuis quelques années, ce genre de problème d'estimation du mouvement, vu comme un cas particulier de l'étude des systèmes complexes, a pu être abordé par le biais des réseaux de neurones récurrents de type Hopfield [HOP - 82]. Ces derniers permettent de résoudre divers problèmes d'optimisation.

L'étude proposée dans le chapitre suivant porte, dans un premier temps, sur une présentation de quelques méthodes d'estimation du mouvement par l'utilisation des réseaux de neurones, et plus particulièrement sur la description d'une approche proposée par Zhou et Chellappa [ZHO - 88] pour le calcul du "flux optique" à l'aide d'un réseau de Hopfield. Dans une deuxième étape, et à la lumière des problèmes que cette méthode pose, nous proposerons une approche dérivée de celle-ci. L'analyse d'une telle méthode constituera le point de départ vers la mise en oeuvre d'un nouveau réseau destiné au calcul du "flux optique".

CHAPITRE V

LES RESEAUX DE NEURONES POUR L'ESTIMATION DU MOUVEMENT

CHAPITRE V

LES RESEAUX DE NEURONES ET L'ESTIMATION DU MOUVEMENT

V.1 RESEAUX DE NEURONES ET MOUVEMENTS

Comme dans un certain nombre de domaines d'application en traitement d'images, il se dégage deux axes de recherches pour l'estimation du mouvement par utilisation des réseaux de neurones. Le premier exploite les connaissances physiologiques et anatomiques du système visuel. Cet axe est ainsi orienté vers des modèles complexes permettant une description du processus de perception du mouvement. Le deuxième axe ne prend pas en compte la complexité biologique. Son but est de résoudre des problèmes de calculs des vecteurs vitesses à l'aide de techniques neuronales. Nos travaux se situent dans ce dernier axe.

Parmi les travaux utilisant une approche de type biologique, on peut citer Grzywacz et Yuille [GRZ - 87], Grzywacz et Poggio [GRZ - 90], Poggio et Torre [POG - 88], Lappe et Rauschecker [LAP - 93] ou encore Daugman [DAU - 88] qui

introduit dans sa méthode un théorème fondamental, le "Spectral Coplanarity Theorem", qui montre comment l'approche coopération / compétition dans un réseau de type "Winner-Take-All" permet aux sorties des filtres neurophysiologiques spatio-temporels d'être combinés pour donner une information de vitesse indépendante de la forme des objets en mouvement ou de leurs contours. Le concept classique d'un champ réceptif neuronal est généralisé au concept de *champ de vitesse réceptif*, dans lequel les entrées sont ajoutées le long d'un plan incliné dans un espace 3D de filtres linéaires qui sont organisés (comme le sont les neurones du cortex strié) suivant l'orientation, la fréquence spatiale et la fréquence temporelle. La coopération dans, et la compétition à travers, les plans inclinés de filtres neuronaux spatio-temporaux, identifient la vitesse et la direction de mouvement locale de l'image.

Ces approches sont en général destinées à modéliser le système visuel humain et comme celui-ci, elles fournissent plus d'informations sur la nature du mouvement que sur la mesure des paramètres caractérisant ce mouvement.

L'utilisation des réseaux de neurones en tant qu'outil de mesure des paramètres du mouvement devient significative à partir des travaux de Zhou et Chellappa [ZHO - 88]. Même si ceux-ci utilisent également des notions tirées du modèle visuel humain, leur objectif n'est plus qualitatif mais bien quantitatif.

Un certain nombre de travaux avaient été réalisés précédemment à ceux de Zhou mais leur objectif était plus une validation architecturale des modèles neuroniques de type analogique ou numérique. Ainsi, un réseau de neurones proposé par Hopfield et Tank [HOP - 85] est utilisé dans [KOC - 87] pour calculer le "flux optique". Pour éviter que la contrainte de lissage n'affecte les forts gradients en vitesse, Koch a proposé d'incorporer un processus linéique sous forme d'un terme additionnel dans l'équation du "flux optique" [KOC - 87]. L'équation résultante est non convexe, ce qui engendre des termes cubiques voire d'ordre supérieur. Plutôt que d'utiliser un algorithme de recuit simulé très coûteux en temps de calcul, un algorithme déterministe basé sur un réseau mixte analogique-numérique est employé afin d'obtenir une solution sous-optimale.

Une deuxième version de réseau de neurones artificiels, reprenant des données physiologiques et anatomiques dans le système visuel pour la perception du mouvement, a été présentée dans [ZHO - 90]. Cette perception est réalisée par l'évolution des neurones en utilisant un système de mise à jour parallèle. Deux algorithmes sont alors présentés pour le calcul du "flux optique" dans une séquence d'images monoculaires. Le premier intègre les informations de toutes les images simultanément en les incluant dans les entrées externes du réseau alors que le second qui est récursif utilise une méthode des moindres carrés pour la mise à jour des entrées externes du réseau.

Des travaux se situant dans le même axe ont été réalisés par Samy [SAM - 90], Ogmen et Gagne [OGM - 90], Tsao et al [TSA - 91], Shibata et Fukuda [SHI - 94].

Enfin un certain nombre de circuits ont été développés à partir des concepts proposés. On peut citer par exemple Lee *et al* [LEE - 93] qui proposent, en se basant sur les travaux de Zhou et Chellappa [ZHO - 88] [ZHO - 90], un système de neuroprocesseurs VLSI pour la détection du mouvement. Il s'agit d'un réseau neuronal compétitif localement connecté. L'information de mouvement à partir des données d'une séquence d'images peut être déterminée à travers un tableau de multiprocesseurs à deux dimensions dans lequel chaque élément de traitement représente un neuroprocesseur analogique. Le transfert des données locales entre les neuroprocesseurs est réalisé en utilisant un système d'interconnexion analogique point à point. Ainsi, un neuroprocesseur VLSI comprenant 25 neurones et 25×27 synapses est capable d'estimer le mouvement de chaque pixel avec 25 vitesses différentes. On peut également citer les réalisations proposées par Hutchinson [HUT - 88] et Koch [KOC - 88].

Nous décrivons plus particulièrement dans la suite de ce chapitre la méthode proposée par Zhou et Chellappa [ZHO - 88] pour le calcul du "flux optique" à l'aide d'un réseau de Hopfield. A partir de l'analyse de cette méthode et des problèmes qu'elle pose, nous proposerons dans un premier temps une approche dérivée de celle-ci. Ceci constituera le point de départ vers une nouvelle approche basée cette fois non

plus sur des techniques de "classification" de vecteurs de déplacement mais sur un principe de construction de ces vecteurs.

V.2 LE MODELE DE ZHOU ET CHELLAPPA

Sous l'hypothèse de rigidité locale et de contraintes de lissage, Zhou propose un réseau de neurones discret afin d'implanter une procédure de calcul basée sur les valeurs estimées des intensités et de leurs courbures principales. Puisque les valeurs d'intensité et leurs courbures principales sont invariantes en rotation, cette méthode peut détecter des objets aussi bien en translation qu'en rotation.

Au début, une procédure d'estimation des valeurs d'intensité et des principales courbures a été mise en oeuvre. En effet, dans le but de réduire les effets du bruit et de l'erreur de quantification, une technique polynomiale appropriée à deux variables est utilisée pour estimer les valeurs d'intensité et de leurs principales courbures. Pour cela, les auteurs utilisent les polynômes de Chebyshev. En minimisant une erreur quadratique, ils obtiennent les coefficients de ces polynômes. Une fonction d'intensité lissée peut être obtenue si des polynômes d'ordre faible et de larges fenêtres sont utilisées. Les principales courbures au point (i,j) sont définies comme des valeurs maximale et minimale des courbures normales de la fonction d'intensité au même point (i,j) . Une importante propriété de ces courbures est leur invariance en rotation, ce qui permet la détection des objets en rotation.

V.2.1 Présentation de la méthode

La formulation du réseau de neurones a été motivée par les recherches faites par Hubel et Wiesel [HUB - 65] [HUB - 77] sur l'organisation en "colonnes d'orientation" des cellules du cortex visuel. Les études entreprises par ces deux chercheurs chez les chats et les singes révèlent une organisation structurelle en couches du cortex visuel.

De nombreuses observations du cortex cérébral prouvent en effet que les cellules sont organisées en unités fonctionnelles et anatomiques appelées "colonnes

corticales" en raison du caractère ascendant et descendant de la transmission des signaux à l'intérieur de ces unités. Ces colonnes, particulièrement bien connues dans le cortex visuel primaire, grâce aux travaux des deux chercheurs cités précédemment [HUB - 77] réagissant de manière préférentielle aux informations provenant de l'oeil correspondant. En effet, lorsque l'on se déplace dans une direction privilégiée sur le cortex, on remarque que les colonnes présentent périodiquement une sensibilité préférentielle à l'oeil droit puis à l'oeil gauche: ce sont les colonnes de dominance oculaire. Lorsque l'on se déplace orthogonalement à ces zones, on remarque d'autres colonnes qui, elles, sont sensibles à l'orientation d'un stimulus dans le champ visuel: l'angle de l'orientation préférentielle croît régulièrement à mesure que l'on se déplace d'une colonne à la suivante. En outre, les cellules de deux colonnes voisines selon ces zones sont sensibles à des orientations voisines. On dispose ainsi d'une continuité spatiale de la réponse corticale à des stimulations visuelles voisines.

Ces propriétés jouent probablement des rôles très importants en conservant les sélectivités d'axe et de direction et en forçant les champs réceptifs voisins à avoir la même préférence directionnelle.

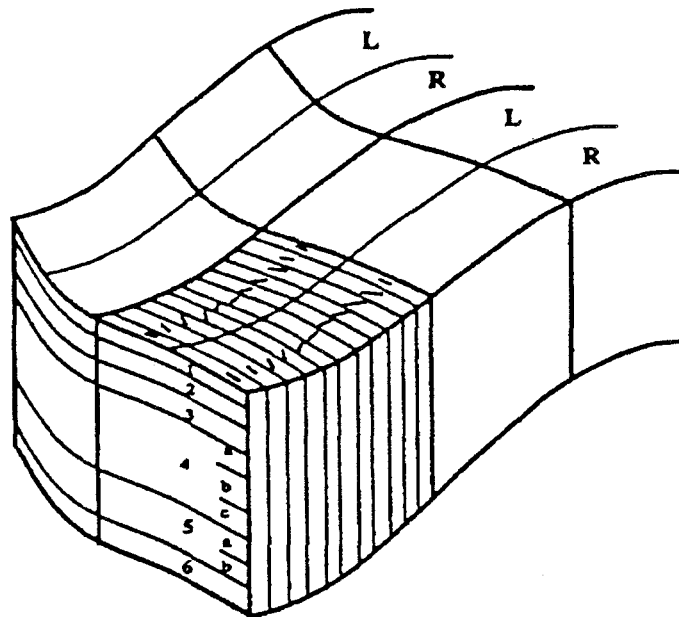


Figure V.1 : Représentation schématique des hypercolonnes dans le cortex visuel (aire 17)

L et R représentent respectivement l'oeil gauche et l'oeil droit.

(Extrait de [ZHO - 90] figure 1, ©1990 IEEE.)

La colonne d'orientation est formée par un ensemble de cellules sélectives à une orientation spécifique de 0° à 180° . A l'intérieur d'une hypercolonne, les orientations des cellules sont disposées régulièrement; lorsqu'on étudie la sélectivité des cellules dans une hypercolonne, un changement progressif des orientations est observé, ainsi qu'une alternance de réponses préférentielles pour l'oeil droit et pour l'oeil gauche.

V.2.2 Architecture du réseau de neurones

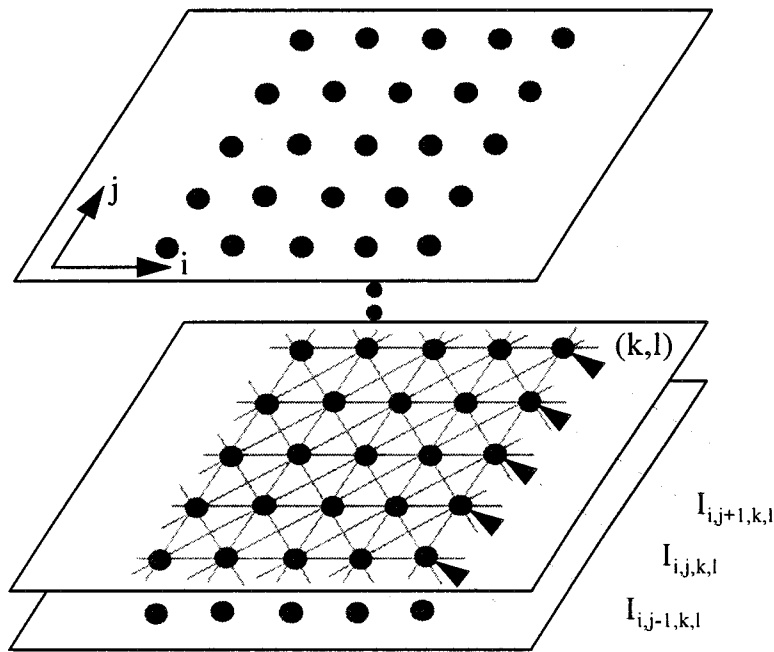
Le modèle est constitué de plusieurs "modules de déplacement". Chaque module du réseau est constitué d'un ensemble de neurones répartis uniformément sur un plan cartésien 2D. On peut représenter le module par une grille de dimension identique à la taille de l'image. Chaque module contient ainsi $(N \times M)$ neurones si $(N \times M)$ représente la taille de l'image (*ligne x colonne*). A l'intérieur d'un même module, chaque neurone a sa sortie branchée sur une de ses entrées et est également interconnecté avec d'autres neurones dans un voisinage de taille $(I \times I)$.

Si k et l sont les deux composantes du vecteur vitesse, alors $(2D_k+1)(2D_l+1)$ modules de neurones sont utilisés pour représenter le "flux optique". D_k et D_l étant respectivement les valeurs maximum des composantes du vecteur vitesse dans les directions verticale 'k' et horizontale 'l'. Si le déplacement maximum est D alors $(2D+1)^2$ neurones mutuellement exclusifs sont requis pour chaque pixel et un nombre total de $(N \times M) \times (2D+1)^2$ neurones est nécessaire pour cette image de taille $(N \times M)$. Puisque deux objets ne peuvent pas occuper la même position en même temps, seule une valeur de vitesse peut être attribuée à chaque pixel. Par conséquent, dans chaque hypercolonne, seul un neurone est dans un état actif. La valeur de la vitesse peut être déterminée selon sa sélectivité de direction.

Comme les champs réceptifs voisins sont forcés pour avoir la même direction préférentielle, nous supposons que les neurones avec une sélectivité identique dans les hypercolonnes voisines tendent à influencer d'autres neurones à travers les entrées qu'ils reçoivent comme le montre la *figure V.2*.

Pour chaque pixel, seul le neurone qui a l'excitation maximum parmi tous les neurones des autres modules est actif tandis que les autres restent inactifs. Quand le neurone au point (i, j) dans le $(k, l)^{\text{ième}}$ module est dans l'état actif, cela signifie que les vitesses dans les directions k et l au point (i, j) sont respectivement $k\omega$ et $l\omega$.

Les figures V.2.a et V.2.b représentent la manière dont les neurones sont interconnectés. A chaque position (i, j) , on associe une hypercolonne de $(2D_k+1)(2D_l+1)$ neurones mutuellement exclusifs. Chacun de ces neurones représente un vecteur de déplacement possible du pixel de coordonnée (i, j) . Le vecteur retenu correspond au neurone actif dans l'hypercolonne.



(a)

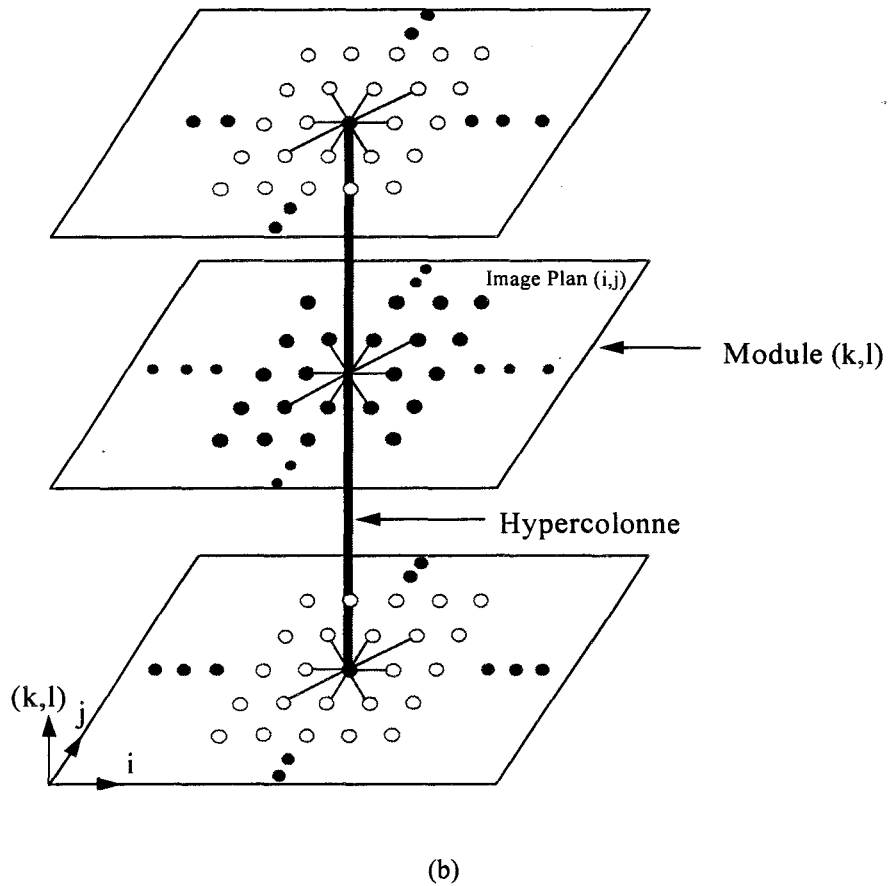


Figure V.2 : Architecture du réseau

(a) Les neurones sont disposés dans des plans vitesses.

(b) Chaque hypercolonne représente la sélectivité de la vitesse pour un point de l'image.

Ce réseau de neurones est décrit comme suit:

- $V = \{v_{i,j,k,l} \mid 1 \leq i \leq N, 1 \leq j \leq M, -D_k \leq k \leq D_k, -D_l \leq l \leq +D_l\}$ est un ensemble d'états binaires du réseau avec $v_{i,j,k,l}$ (1 pour actif, 0 pour inactif) indiquant l'état du (i,j,k,l) ième neurone situé au point (i,j) dans le (k,l) ième module;
- $T_{i,j,k,l;m,n,p,q}$ est le poids synaptique de l'interconnexion entre le neurone (i,j,k,l) et le neurone (m,n,p,q) ;
- $I_{i,j,k,l}$ est l'entrée externe du neurone (i,j,k,l) .

A chaque étape, chaque neurone (i,j,k,l) reçoit de manière synchrone les entrées de tous les autres neurones ainsi que de lui-même et une entrée externe:

$$u_{i,j,k,l} = \sum_{m=1}^N \sum_{n=1}^M \sum_{p=-D_k}^{D_k} \sum_{q=-D_l}^{D_l} T_{i,j,k,l;m,n,p,q} v_{m,n,p,q} + I_{i,j,k,l} \quad (V.1)$$

Chaque $u_{i,j,k,l}$ est réinjecté aux neurones correspondants d'après la fonction d'activation:

$$v_{i,j,k,l} = g(u_{i,j,k,l}) \quad (V.2)$$

où g est une fonction non linéaire d'activation maximum définie comme suit:

$$g(x_{i,j,k,l}) = \begin{cases} 1 & \text{si } x_{i,j,k,l} = \max(x_{i,j,p,q}; -D_k \leq p \leq +D_k, -D_l \leq q \leq +D_l) \\ 0 & \text{ailleurs} \end{cases} \quad (V.3)$$

Cette fonction est utilisée pour assurer que seul un neurone dont l'excitation est maximale est actif dans une colonne tandis que les autres $(2K+1)(2L+1)-1$ neurones sont inactifs.

V.2.3 Calcul du "flux optique"

En plus des courbures principales, deux autres contraintes sont introduites dans la fonctionnelle à minimiser. Celle-ci est donc composée d'une somme de trois termes: le premier terme représente les principales courbures des deux images, le second est une contrainte de lissage de la solution et le troisième terme est le processus de lignes pour la détection des discontinuités. Des constantes déterminent l'importance relative des trois termes. Le problème dans les approches basées sur les réseaux de neurones comme dans les problèmes de minimisation de fonctionnelles est le difficile choix des valeurs de ces constantes. Les paramètres (poids synaptiques et entrées externes) du réseau neuronal sont obtenus après identification de cette fonctionnelle avec la fonction d'énergie du réseau de neurones. Enfin, le processus de mise à jour, lors de la convergence, est implanté en parallèle.

V.2.4 Conclusion

La méthode a été testée sur différents types d'images [ZHO - 88]. Bien qu'elle donne des résultats satisfaisants, elle est néanmoins très lourde en temps de calcul. La difficulté principale de sa mise en oeuvre n'est pas tant le réseau de neurones lui-même que les opérations nécessaires pour alimenter les entrées de ce réseau: estimation de l'interpolation de la fonction intensité de l'image, calcul des moments, etc ... On peut même penser que la qualité de la solution initiale présentée au réseau est telle que son rôle est quasiment secondaire dans le processus complet d'estimation des déplacements.

A partir de cette étude, il nous a semblé intéressant de voir le comportement d'une telle approche mais d'une manière beaucoup plus simplifiée. A cet effet, l'idée est de proposer une méthode plus simple, basée uniquement sur les caractéristiques de l'image. Ni les courbures principales, ni le processus de lignes pour la détection des discontinuités ne sont mis en évidence. Des critères simples sont introduits.

V.3 REFORMULATION DU PROBLEME

La même architecture du réseau neuronal vue précédemment est reprise ici, mais la dynamique assurant la convergence du réseau est redéfinie par une nouvelle fonctionnelle d'erreur à minimiser. Cette fonctionnelle ne tient compte que de la seule contrainte de régularité [YAC - 95].

Le réseau étant rebouclé, il est donc caractérisé par l'évolution de l'état de ses neurones au cours du temps. Cette évolution peut permettre, sous certaines conditions, de converger vers un état stable.

V.3.1 Calcul de la fonction d'énergie

A ce système dynamique, on peut associer une énergie dont la décroissance est une condition de stabilité du réseau. On remarquera qu'un même réseau peut admettre *plusieurs minima* qui sont accessibles à partir de configurations initiales différentes et qui peuvent avoir des énergies associées différentes. Les paramètres, poids

synaptiques des interconnexions et entrées externes, du modèle neuronal précédent peuvent être remis dans une équation caractérisant l'énergie du réseau suivant le calcul de Hopfield et Tank [HOP - 85]. Dans le cas de notre réseau, la fonction d'énergie prend la forme suivante:

$$E = - \sum_{i=1}^N \sum_{j=1}^M \sum_{k=-D_k}^{D_k} \sum_{l=-D_l}^{D_l} \left[\frac{1}{2} \sum_{m=1}^N \sum_{n=1}^M \sum_{p=-D_k}^{D_k} \sum_{q=-D_l}^{D_l} T_{i,j,k,l;m,n,p,q} V_{i,j,k,l} V_{m,n,p,q} + I_{i,j,k,l} V_{i,j,k,l} \right] \quad (V.4)$$

Nous choisissons alors comme fonctionnelle à minimiser afin de détecter le "flux optique", la relation suivante (V.5):

$$E = \sum_{i=1}^N \sum_{j=1}^M \sum_{k=-D_k}^{D_k} \sum_{l=-D_l}^{D_l} \left[\left\{ A(g_1(i, j) - g_2(i \oplus k, j \oplus l))^2 v_{i,j,k,l} \right\} + \left\{ \frac{B}{2} \sum_{s \in S} (v_{i,j,k,l} - v_{(i,j) \oplus s, k, l})^2 \right\} \right]$$

où $\{g_1(i, j)\}$ et $\{g_2(i \oplus k, j \oplus l)\}$ sont les valeurs d'intensité de la première et deuxième image. S est une série d'index excluant (0,0) pour tous les voisins dans une fenêtre ($I \times I$) centrée au point (i, j), et le symbole \oplus indique que:

$$f_{a \oplus b} = \begin{cases} f_{a \oplus b} & \text{si } 0 \leq a + b \leq N, M \\ 0 & \text{ailleurs} \end{cases} \quad (V.6)$$

Le premier terme de cette fonction d'énergie (V.5) recherche les valeurs de déplacements tels que les points des deux images successives soient minimisés sous forme d'une erreur quadratique. L'introduction du second terme présente l'avantage d'induire un lissage du champ estimé et donc de minimiser l'influence du bruit. Dans cette expression, les coefficients réels et positifs A et B permettent de pondérer l'influence relative de chacune des contraintes. Le rôle de ces deux coefficients est donc important puisqu'à différentes valeurs de A et B correspondront des solutions différentes.

V.3.2 La phase d'identification

La construction du réseau de neurones qui résout le problème revient donc à déterminer les connexions $T_{i,j,k,l;m,n,p,q}$ qui lient les neurones (i,j,k,l) et (m,n,p,q) de telle sorte que les points stables du réseau soient les solutions recherchées. Il suffit, en prenant $\Gamma = 5$ d'identifier terme à terme les deux expressions (V.4) et (V.5).

Nous obtenons alors:

$$T_{i,j,k,l;m,n,p,q} = -48B\delta_{i,m}\delta_{j,n}\delta_{k,p}\delta_{l,q} + 2B\sum_{s \in S} \delta_{(i,j),(m,n) \oplus s} \delta_{k,p}\delta_{l,q} \quad (V.7)$$

$$I_{i,j,k,l} = -A(g_1(i,j) - g_2(i \oplus k, j \oplus l))^2 \quad (V.8)$$

où $\delta_{a,b}$ est le symbole de Kronecker. $\delta_{a,b} = 1$ si et seulement $a = b$ et $\delta_{a,b} = 0$ dans les cas contraires. On remarque que la matrice des poids synaptiques est symétrique, ce qui augure de la convergence du réseau.

L'état initial des neurones est défini par:

$$v_{i,j,k,l} = \begin{cases} 1 & \text{si } I_{i,j,k,l} = \max_{p,q} (x_{i,j,p,q}) \\ 0 & \text{ailleurs} \end{cases} \quad (V.9)$$

S'il existe deux entrées externes maximales au point (i,j) , alors seul le neurone correspondant à la plus petite vitesse est initialement défini à 1 et l'autre est mis à 0. Quand le réseau atteint un état stable, le "flux optique" est déterminé par l'état du neurone.

V.3.3 Discussion sur la variation d'énergie ΔE

Pour deux neurones (i,j,k_1,l_1) et (i,j,k_2,l_2) changeant d'état à chaque itération, les variations de leur état $\Delta v_{i,j,k_1,l_1}$ et $\Delta v_{i,j,k_2,l_2}$ sont définies comme suit:

$$\Delta v_{i,j,k_1,l_1} = v_{i,j,k_1,l_1}^{\text{nouveau}} - v_{i,j,k_1,l_1}^{\text{ancien}} \quad (\text{V.10})$$

$$\Delta v_{i,j,k_2,l_2} = v_{i,j,k_2,l_2}^{\text{nouveau}} - v_{i,j,k_2,l_2}^{\text{ancien}} \quad (\text{V.11})$$

La variation d'énergie correspondante est notée:

$$\Delta E = E^{\text{nouveau}} - E^{\text{ancien}} \quad (\text{V.12})$$

En considérant la fonction d'énergie établie en (V.5), l'expression de ΔE peut être écrite sous la forme suivante:

$$\begin{aligned} \Delta E = & - \left(\sum_{m=1}^N \sum_{n=1}^M \sum_{p=-D_k}^{D_k} \sum_{q=-D_l}^{D_l} T_{i,j,k_1,l_1;m,n,p,q} v_{m,n,p,q} + I_{i,j,k_1,l_1} \right) \Delta v_{i,j,k_1,l_1} \\ & - \left(\sum_{m=1}^N \sum_{n=1}^M \sum_{p=-D_k}^{D_k} \sum_{q=-D_l}^{D_l} T_{i,j,k_2,l_2;m,n,p,q} v_{m,n,p,q} + I_{i,j,k_2,l_2} \right) \Delta v_{i,j,k_2,l_2} \\ & - \frac{1}{2} T_{i,j,k_1,l_1;i,j,k_1,l_1} (\Delta v_{i,j,k_1,l_1})^2 - \frac{1}{2} T_{i,j,k_2,l_2;i,j,k_2,l_2} (\Delta v_{i,j,k_2,l_2})^2 \\ & - T_{i,j,k_1,l_1;i,j,k_2,l_2} (\Delta v_{i,j,k_1,l_1} v_{i,j,k_2,l_2}^{\text{nouveau}} + \Delta v_{i,j,k_2,l_2} v_{i,j,k_1,l_1}^{\text{nouveau}}) \end{aligned} \quad (\text{V.13})$$

Lorsque $v_{i,j,k_1,l_1}^{\text{ancien}} = 0$, $v_{i,j,k_2,l_2}^{\text{ancien}} = 1$

et $u_{i,j,k_2,l_2} < u_{i,j,k_1,l_1}$

on a $v_{i,j,k_1,l_1}^{\text{nouveau}} = 1$, $v_{i,j,k_2,l_2}^{\text{nouveau}} = 0$

et $\Delta E = (u_{i,j,k_2,l_2} - u_{i,j,k_1,l_1}) - \frac{1}{2} (T_{i,j,k_1,l_1;i,j,k_1,l_1} + T_{i,j,k_2,l_2;i,j,k_2,l_2})$ (V.14)

Le premier terme $(u_{i,j,k_2,l_2} - u_{i,j,k_1,l_1})$ est négatif. Le second, c'est-à-dire

$-\frac{1}{2} (T_{i,j,k_1,l_1;i,j,k_1,l_1} + T_{i,j,k_2,l_2;i,j,k_2,l_2})$ vaut 48B. Il est positif. Si le premier terme est très

inférieur par rapport au second, une règle de décision déterministe est appliquée pour la mise à jour des états des neurones afin d'assurer la convergence du réseau. Celui-ci converge ainsi vers un minimum local. Nous pourrions aussi envisager l'utilisation d'une règle de décision stochastique pour réaliser cette convergence. Cependant, ce choix impliquerait des temps de calculs importants.

V.3.4 Résultats expérimentaux et commentaires

Cette méthode a été validée sur plusieurs variétés d'images. Trois exemples sont présentés ici. La séquence d'images de synthèse (séquence 1), présentée en *figure V.1* est constituée d'un carré noir en mouvement sur un fond fixe blanc. Le carré se translatant avec un déplacement de 1 pixel vers le coin bas gauche de l'image. L'amplitude des déplacements est choisie faible pour que le calcul des dérivées partielles soit fondé.

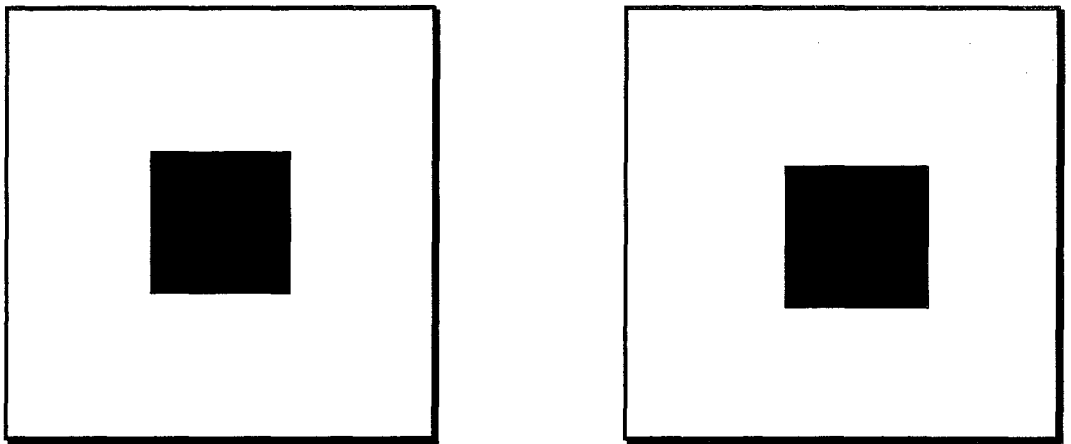


Figure V.1 : Séquence de synthèse n°1

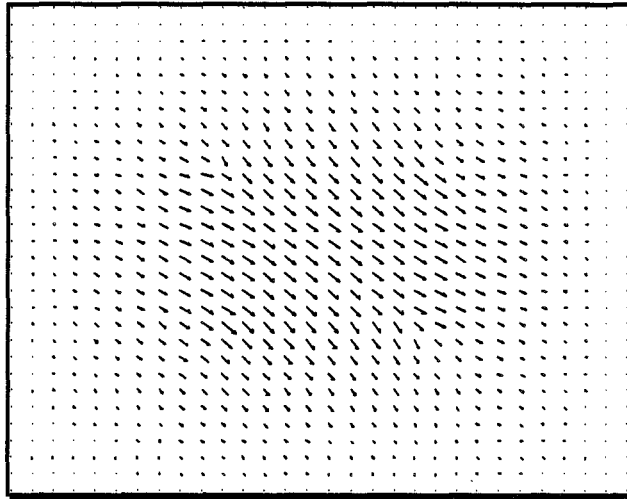


Figure V.2 : Champ de vecteurs généré

La séquence n°2 présentée en *figure V.3* représente une voiture se déplaçant en translation oblique.

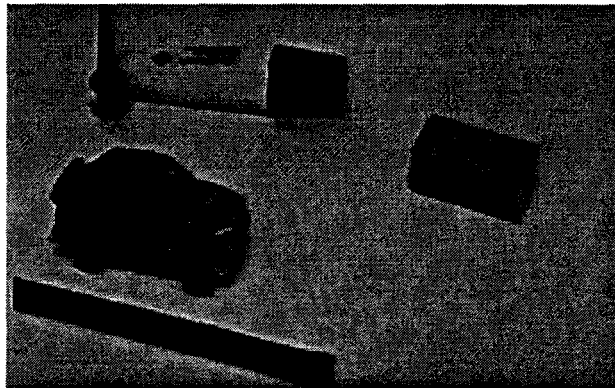


Figure V.3 : Séquence n°2.

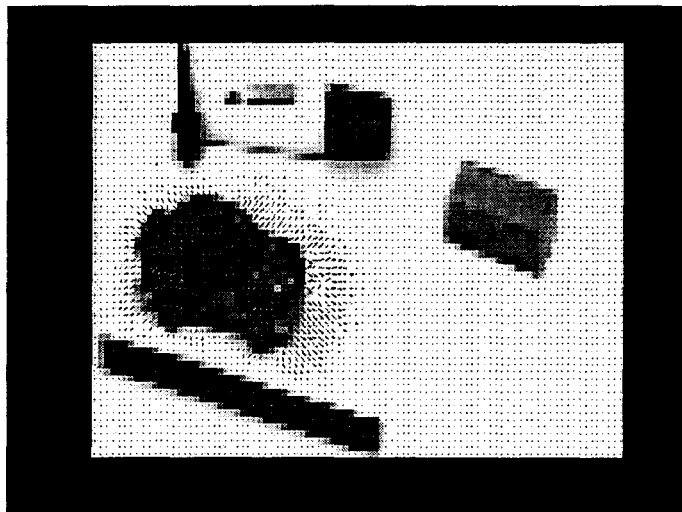


Figure V.4 : Champ de vecteurs généré.

Comme le montre la figure, les résultats obtenus sur cet exemple sont satisfaisants. En effet, les zones de mouvements sont correctement estimées tout au long de la séquence.

La séquence n°3 présentée en *figure V.5* représente des images échocardiographiques. Ces dernières sont obtenues à l'aide d'un échographe VINGMED 750 en incidence apicale selon 3 plans sécants à 60 degrés à la même position dans le cycle cardiaque mais ne correspondant pas nécessairement au même cycle. Des séquences de quelques secondes pour chaque plan sont disponibles sur bande magnétique et la sélection des vues se fait par synchronisation des images des séquences à partir de l'ECG enregistré en même temps que l'échographie. Pour pouvoir représenter le mouvement, on définit une série de N (de l'ordre de 15 à 20) triplets de coupes apicales pris aux différents instants du cycle cardiaque.

Les résultats sont présentés sur un échantillon d'images extraites de la séquence décrite précédemment. Les *figures V.5* et *V.7.a* présentent l'image initiale sur laquelle seules ont été reportées, à titre illustratif, les fenêtres sur lesquelles a été estimé le mouvement. Ces fenêtres ont pour taille 30×30 . Les *figures V.6* et *V.7.b* montrent l'estimation locale des mouvements dans la zone sélectionnée pour deux images successives de la séquence prises à des instants différents. Les valeurs des paramètres utilisés sont respectivement $A = 1$, $B = 50$ et $A = 1$, $B = 200$ avec $K = L = 5$. La

convergence a été obtenue après 8 itérations dans le premier test et une quinzaine dans le second.



Figure V.5 : Séquence n°2, Images échocardiographiques
Image n°1 avec sélection d'une zone d'intérêt.

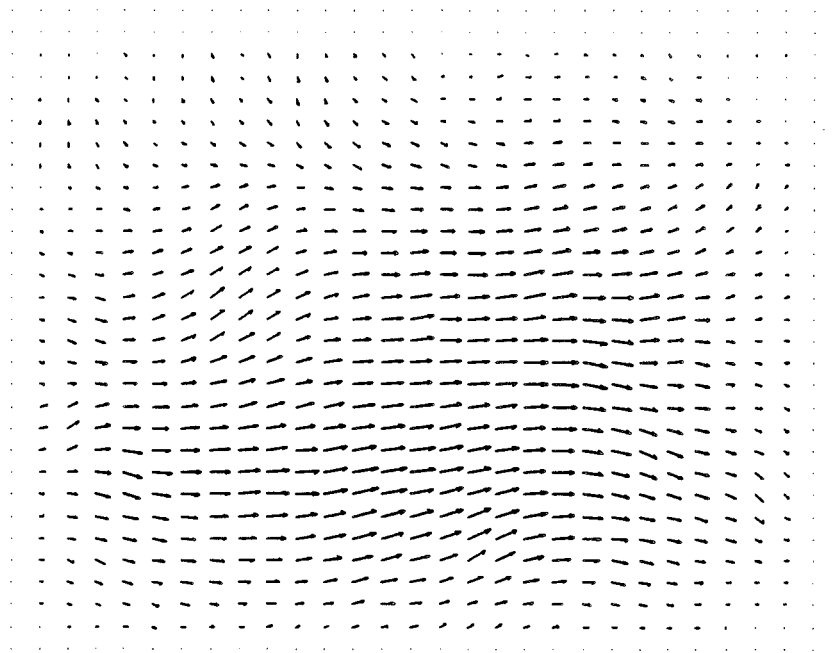
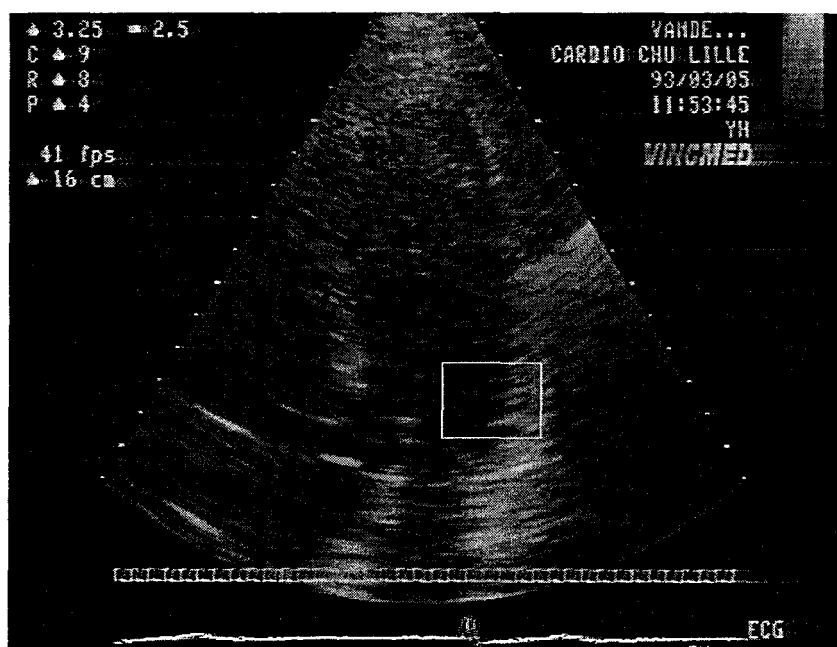
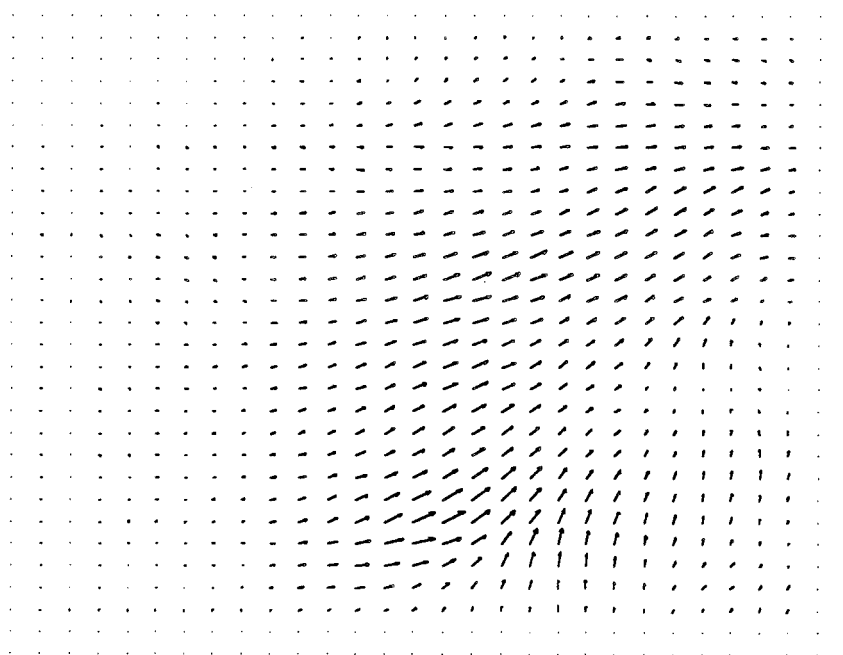


Figure V.6 : Estimation locale des mouvements



(a)



(b)

Figure V.7 : Séquence n°2, Images échocardiographiques

(a) Image n°1 avec sélection d'une autre zone d'intérêt, (b) Estimation locale des mouvements.

Nous constatons à l'étude de l'ensemble des estimations de mouvements obtenues que la cartographie des zones homogènes de mouvements est très proche de celle des régions texturées présentes dans l'image. Ceci met bien en évidence le rôle du lissage grâce au deuxième terme présent dans la fonction énergétique minimisée dont l'influence vis-à-vis du terme caractérisant la "ressemblance" est paramétrée par le rapport entre les coefficients A et B.

Il faut noter toutefois, les deux obstacles importants responsables des difficultés d'évaluation du "flux optique" dans ce type de séquences médicales:

- la mauvaise qualité des images échocardiographiques (fortement perturbées par la présence du bruit) qui nécessite donc un coefficient B élevé;
- la complexité des mouvements du ventricule gauche (mouvements composés de translations, rotations, dilatations et contractions) que l'on perçoit surtout par leur projection dans le plan de l'image dans le cas du 2D.

Choix des paramètres :

Le choix des paramètres A et B est un point important de la méthode. Si le choix de la valeur même de A et B mérite attention, il est aussi essentiel de regarder s'il est préférable de définir leurs valeurs en fonction de la taille des zones analysées, ou bien définir des valeurs constantes quelque soit les zones considérées.

Le champ de déplacement obtenu sera différent suivant l'influence de l'une ou l'autre des contraintes supplémentaires. Si A est grand et B est petit, alors c'est l'influence de la contrainte de lissage qui est faible. Dans le cas où A est petit et B est grand, c'est l'influence de la contrainte de l'E.C.M.A qui est peu influente.

Le choix des paramètres A et B est crucial. Il importe donc de leur donner une valeur convenable.



V.4 CONCLUSION

La méthode, appliquée à l'estimation de mouvements sur des séquences d'images de synthèse représentant des objets subissant des mouvements simples connus a priori, donne des résultats satisfaisants. Cette méthode permet de montrer d'autre part l'influence importante du choix des paramètres de réglage sur la forme du champ de vitesse calculé.

Appliquée aux images échocardiographiques, cette méthode permet d'obtenir, dans des zones où l'on décèle visuellement des mouvements, une information objective sur ceux-ci. Par ailleurs, la cartographie obtenue pourrait constituer un outil de diagnostic précieux pour le clinicien dans l'étude de la contractilité cardiaque et lui permettre de confronter son appréciation subjective des mouvements à une mesure objective et précieuse.

Les expérimentations sur des images représentant des scènes réelles et bruitées ont montré que la mise en oeuvre de l'algorithme est difficile. Ces images sont en effet très complexes. Des problèmes sont constatés lorsque l'on a affaire à des déplacements apparents de grandes amplitudes ou lorsqu'il y a des changements de luminosité en cours de séquence. Un autre problème important est le choix empirique des divers paramètres (paramètres de régularisation, d'initialisation,...).

A la lumière de cette étude, nous pouvons conclure que la méthode simplifiée décrite ci-dessus donne des solutions, certes, vraisemblablement moins bonnes que celles obtenues par l'approche proposée par Zhou et Chellappa, mais toutefois de qualité suffisante pour l'obtention d'un champ de vecteurs vitesse cohérent.

Il convient de noter que cette technique présente un certain nombre de limites. En effet, ne se baser que sur les différences d'images est trop simpliste. Il faut donc injecter dans le réseau des connaissances a priori plus fines sur le problème qui permettent de trouver une représentation de la fonction à minimiser dont le minimum correspond à une solution optimale du réseau. Si l'on ne veut pas uniquement "déplacer" le problème en déportant les difficultés sur un calcul préalable à

l'utilisation du réseau, il apparaît nécessaire de construire un réseau sur un principe différent de celui de Zhou.

L'analyse faite du réseau de Zhou et Chellappa montre clairement que ce système repose sur un principe de classification des vecteurs. En effet, il identifie le module où se trouve le neurone actif dans l'hypercolonne et lui assigne ainsi la valeur du vecteur vitesse. En d'autres termes, la minimisation de la fonctionnelle est établie sous forme de "classification" parmi les solutions prédéfinies. Ceci explique en grande partie les difficultés que le système rencontre si les "critères" de classification ne sont pas robustes. Par ailleurs, le code de sortie prend la forme d'un ensemble de vecteurs où seule une composante est active. Cette technique demande donc énormément de neurones pour ne retenir à la fin qu'un seul neurone actif par hypercolonne. Le nombre de neurones croît considérablement avec l'augmentation de la dynamique des vecteurs vitesse. Enfin, on peut remarquer que compte tenu de sa formulation, la prise en compte d'informations comme la notion de dérivée des variables est difficilement possible. Or de nombreuses méthodes comme nous l'avons vu, reposent sur des techniques différentielles.

Notre objectif dans le chapitre suivant est donc de développer un système neuronal permettant de s'affranchir des difficultés rencontrées précédemment. L'idée est donc de trouver un autre mode de codage qui peut mettre en évidence, parmi les informations disponibles, celles qui favorisent la résolution du problème. Pour cela, le modèle de type classification est abandonné au profit d'un modèle de "calcul de la solution". La méthode adoptée consiste à transcrire la variable recherchée sous forme d'une sommation de neurones. Ceci est réalisé par la redéfinition des deux composantes ω_x et ω_y à la position (i,j) . Chaque composante du vecteur vitesse sera alors déduite par une somme des états des neurones du sous réseau associé.

CHAPITRE VI

UN RESEAU GNERIQUE

CHAPITRE VI

UN RESEAU GENERIQUE

La méthode que nous adoptons consiste dans un premier temps à coder le problème en termes d'états qui sont les variables discrètes. Nous proposons un modèle neuronal composé de deux réseaux totalement connectés afin de représenter les deux composantes ω_x et ω_y du vecteur vitesse ω . Chacune de ces composantes correspond à un réseau. Le modèle consiste donc en $(2 * L * M * N)$ neurones mutuellement interconnectés, où $(L * M)$ représente la taille de l'image et N le nombre total des neurones dans un réseau. L'image peut être décrite par un ensemble fini de vecteur vitesse $\{\omega(i, j), \text{ où } 1 \leq i, j \leq L, M\}$ avec $\omega(i, j)$ un nombre entier indiquant le vecteur vitesse au point (i, j) . Soit $V = \{v_{(i,j)k}, \text{ où } 1 \leq (i, j) \leq L \times M, 1 \leq k \leq 2N\}$ un ensemble d'états binaires du réseau de neurones avec $v_{(i,j)k}$ indiquant l'état du $(i, j, k)^{\text{ème}}$ neurone situé au point (i, j) dans le réseau de composante ω_x ou ω_y . Il est alors possible de représenter les deux composantes du vecteur vitesse et les états des neurones. En effet, ω_x et ω_y peuvent être déduites par une sommation des variables des états des neurones du sous-réseau associé:

$$\omega_x = \sum_{h=1}^N v_{(i,j)h} \quad \text{et} \quad \omega_y = \sum_{p=1}^N v_{(i,j)p} \quad (\text{VI.1})$$

$v_{(i,j)h}$ et $v_{(i,j)p}$ représentent respectivement l'état des neurones h et p correspondant au sous-réseau de la composante ω_x (ω_y) au point (i,j) .

VI.1 POSITION DU PROBLEME

Afin de valider l'adaptabilité du problème d'estimation du mouvement à une solution par les réseaux de neurones, il nous faut:

1. définir l'architecture de réseau adaptée où un ensemble de neurones formels est connecté à un ensemble d'autres neurones par des poids synaptiques. Le choix de cette architecture se base sur:
 - la nature de l'application;
 - la nature des données;
 - les considérations sur les performances.

2. définir la dynamique assurant la convergence du réseau:
 - la fonction d'activation;
 - le mode de changement d'états: asynchrone, synchrone;
 - l'évolution du réseau: déterministe, stochastique (l'évolution se terminant toujours sur un état stable, minimum local ou global de la fonction d'énergie).

3. définir l'énergie à minimiser:

- choix de la fonction sur les états du réseau dont la minimisation amène à résoudre le problème.

4. procéder à la phase d'identification avec la fonctionnelle à résoudre:

- identifier la fonction d'énergie établie par Hopfield à la fonction sur les états afin d'obtenir les valeurs des entrées externes et des poids synaptiques.

Il s'agit donc d'injecter dans le réseau une connaissance a priori sur le problème, en espérant que la solution, au sens du minimum d'une fonction de coût, soit proche de notre solution initiale.

VI.2 L'ARCHITECTURE DU RESEAU PROPOSE

Nous réalisons un réseau de neurones conforme aux idées que nous venons de développer [YAC - 95] [LEL - 95]. Le système se compose de deux réseaux binaires identiques de type Hopfield représentant les deux composantes du vecteur vitesse $\vec{\omega} (\omega_x, \omega_y)$. Chacun de ces deux réseaux est à son tour composé de deux sous-réseaux également identiques qui tiennent compte des déplacements positifs et (ou) négatifs des objets en mouvement dans la séquence d'images. Ainsi, si L et M représentent les dimensions de l'image et N le nombre de neurones, nous aurons alors au total $(4 \times N)$ neurones dans chaque point de l'image et $(4 \times L \times M \times N)$ neurones dans l'image entière (*Figure VI.1*).

La *figure VI.2* montre la connectivité des quatre sous-réseaux représentant les deux composantes du vecteur vitesse. Le réseau est alors constitué par un assemblage de motifs de base. L'ensemble présente une grande régularité, ce qui permet une implémentation matérielle par simple répétition du même motif. Le réseau obtenu est caractérisé d'une part, par les nombreuses interconnexions agissant entre les neurones dans chaque sous-réseau et d'autre part, par celles agissant entre les différents quatre sous-réseaux. A chacune de ces connexions est associé un poids qui détermine

l'influence réciproque de deux neurones connectés. Nous verrons un peu plus loin que ces poids synaptiques jouent un rôle prépondérant dans la convergence du réseau vers des états stables. La *figure VI.5* donne un exemple de l'architecture du réseau de neurones dans un point (i,j) . Les interconnexions entre les neurones dans ce point sont totales. Les liaisons entre ces neurones et ceux se trouvant dans les sous réseaux aux points voisins ne sont pas figurées.

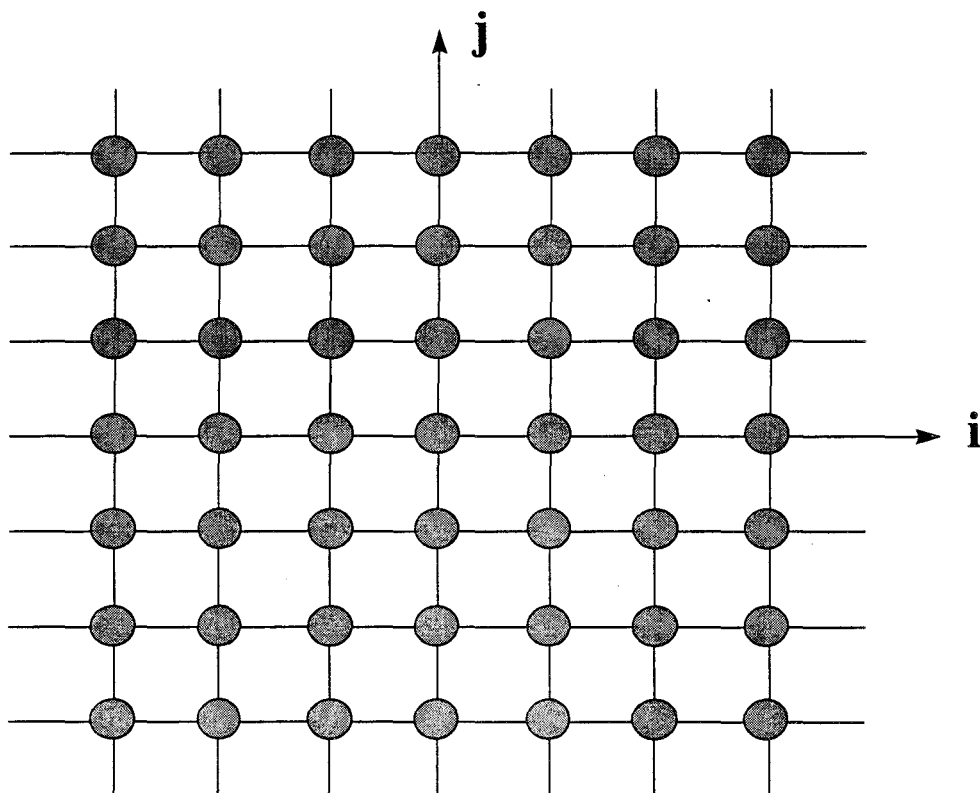


Figure VI.1 : Les 4 sous réseaux sont distribués uniformément sur une image.

Chaque cercle représente 4 sous réseaux.

Chacune des deux composantes (ω_x , ω_y) à la position (i,j) est déduite par une somme des états des neurones du sous réseau associé, autrement dit en tenant compte des composantes à valeurs positives et négatives, nous avons:

$$\omega_x^+ = k \sum_{h^+=1}^N v_{i,j,h^+} \quad \omega_x^- = k \sum_{h^-=1}^N v_{i,j,h^-}$$

$$\omega_{(i,j)x} = \omega_x^+ - \omega_x^- \quad (\text{VI.2})$$

$$\omega_y^+ = k \sum_{p^+=1}^N v_{i,j,p^+} \quad \omega_y^- = k \sum_{p^-=1}^N v_{i,j,p^-}$$

$$\omega_{(i,j)y} = \omega_y^+ - \omega_y^- \quad (\text{VI.3})$$

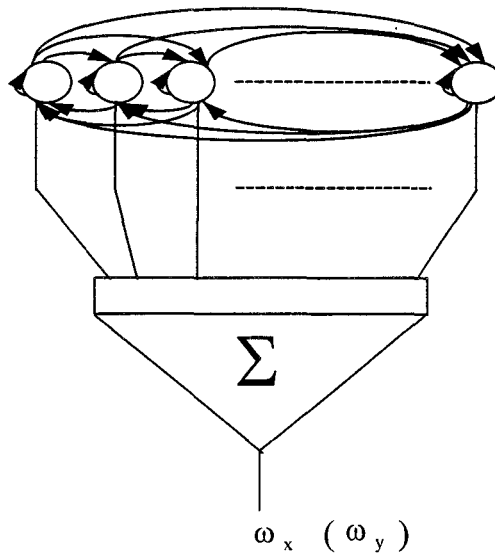


Figure VI.3 : Sous réseau neuronal.

v_{i,j,h^+} , v_{i,j,h^-} (v_{i,j,p^+} , v_{i,j,p^-}) représentent respectivement l'état du neurone h^+ , h^- (p^+ , p^-) correspondant au sous réseau de la composante ω_x (ω_y) au point (i,j).

Le facteur k est un réel positif dont le réglage dépend de la séquence d'images à traiter. Il est introduit afin d'avoir une échelle de quantification dans la dynamique de la norme des vecteurs vitesse. Il est ainsi lié au nombre total des neurones actifs, étant donné que le produit $|k \times N|$ représente la valeur maximale détectable de la vitesse. Notons que la méthode du calcul des gradients utilisée impose a priori de faibles déplacements des objets présents dans l'image.

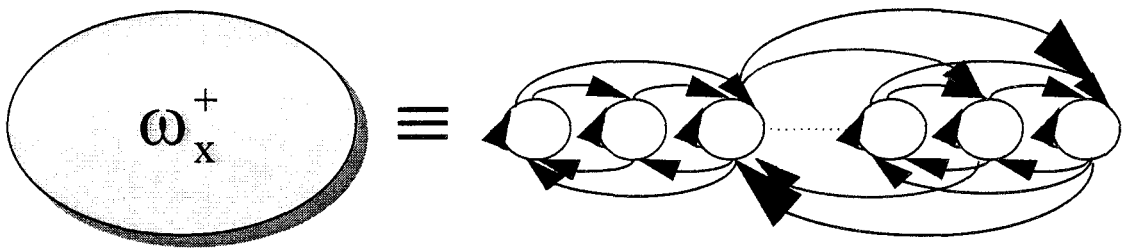


Figure VI.4 : Exemple de sous réseau.

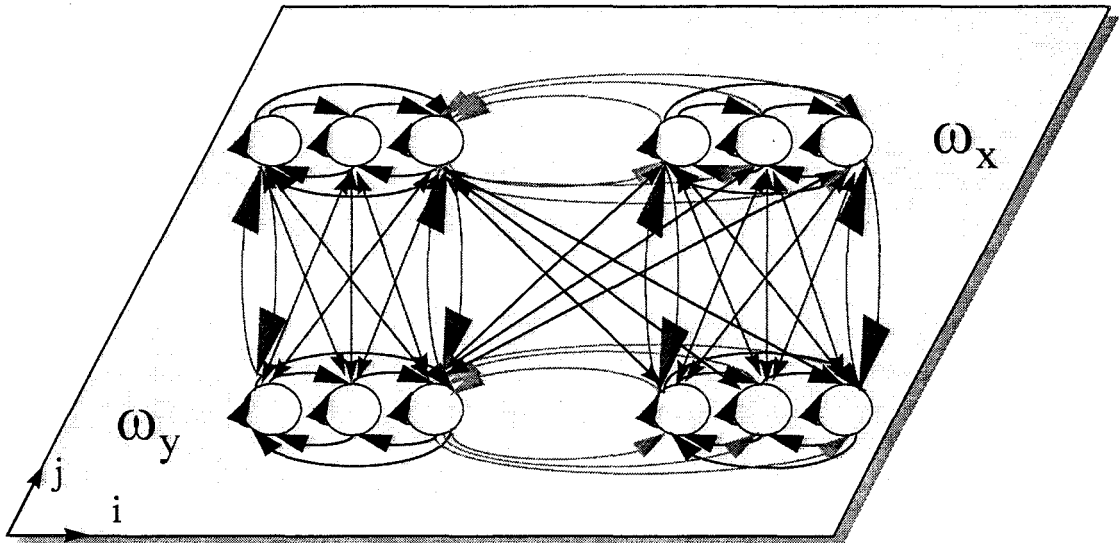


Figure VI.5 : Exemple de liaisons dans un point (i,j) pour $N=3$.
Les interconnexions vers les points voisins ne sont pas figurées.

VI.3 LA DYNAMIQUE

La dynamique du réseau est entièrement définie par la nature des connexions, la fonction d'activation des neurones et enfin par la définition du mode d'itération. La structure des connexions présente un grand intérêt vis-à-vis de la convergence du réseau. Dans notre cas, cette convergence est assurée par les poids de connexions symétriques et une mise à jour asynchrone dans laquelle le vecteur d'état est actualisé élément par élément en tenant compte des nouvelles valeurs des composantes dès qu'elles sont disponibles.

Pour chacun des 4 sous-réseaux, le potentiel $u_{i,j}$ du neurone $(i, j, .)$ est défini par les relations suivantes:

➤ **réseau de composante x :**

$$\begin{aligned}
 u_{i,j,h^+} = & \sum_{m=1}^L \sum_{n=1}^M \sum_{k^+=1}^N T_{i,j,h^+;m,n,k^+} V_{m,n,k^+} + \sum_{m=1}^L \sum_{n=1}^M \sum_{h^-=1}^N T_{i,j,h^+;m,n,h^-} V_{m,n,h^-} \\
 & + \sum_{m=1}^L \sum_{n=1}^M \sum_{p^+=1}^N T_{i,j,h^+;m,n,p^+} V_{m,n,p^+} + \sum_{m=1}^L \sum_{n=1}^M \sum_{p^-=1}^N T_{i,j,h^+;m,n,p^-} V_{m,n,p^-} + I_{i,j,h^+}
 \end{aligned} \tag{VI.4}$$

$$\begin{aligned}
 u_{i,j,h^-} = & \sum_{m=1}^L \sum_{n=1}^M \sum_{k^-=1}^N T_{i,j,h^-;m,n,k^-} V_{m,n,k^-} + \sum_{m=1}^L \sum_{n=1}^M \sum_{h^+=1}^N T_{i,j,h^-;m,n,h^+} V_{m,n,h^+} \\
 & + \sum_{m=1}^L \sum_{n=1}^M \sum_{p^-=1}^N T_{i,j,h^-;m,n,p^-} V_{m,n,p^-} + \sum_{m=1}^L \sum_{n=1}^M \sum_{p^+=1}^N T_{i,j,h^-;m,n,p^+} V_{m,n,p^+} + I_{i,j,h^-}
 \end{aligned} \tag{VI.5}$$

➤ **réseau de composante y :**

$$\begin{aligned}
 u_{i,j,p^+} = & \sum_{m=1}^L \sum_{n=1}^M \sum_{q^+=1}^N T_{i,j,p^+;m,n,q^+} V_{m,n,q^+} + \sum_{m=1}^L \sum_{n=1}^M \sum_{p^-=1}^N T_{i,j,p^+;m,n,p^-} V_{m,n,p^-} \\
 & + \sum_{m=1}^L \sum_{n=1}^M \sum_{h^+=1}^N T_{i,j,p^+;m,n,h^+} V_{m,n,h^+} + \sum_{m=1}^L \sum_{n=1}^M \sum_{h^-=1}^N T_{i,j,p^+;m,n,h^-} V_{m,n,h^-} + I_{i,j,p^+}
 \end{aligned} \tag{VI.6}$$

$$\begin{aligned}
u_{i,j,p^-} = & \sum_{m=1}^L \sum_{n=1}^M \sum_{q=1}^N T_{i,j,p^-;m,n,q^-} v_{m,n,q^-} + \sum_{m=1}^L \sum_{n=1}^M \sum_{p^+=1}^N T_{i,j,p^-;m,n,p^+} v_{m,n,p^+} \\
& + \sum_{m=1}^L \sum_{n=1}^M \sum_{h=1}^N T_{i,j,p^-;m,n,h^-} v_{m,n,h^-} + \sum_{m=1}^L \sum_{n=1}^M \sum_{h^+=1}^N T_{i,j,p^-;m,n,h^+} v_{m,n,h^+} + I_{i,j,p^-}
\end{aligned} \tag{VI.7}$$

où $T_{i,j,h^+;m,n,k^+}$ et $T_{i,j,h^+;m,n,h^-}$ représentent les interconnexions synaptiques à l'intérieur du réseau de composante (ω_x). $T_{i,j,h^+;m,n,p^+}$ et $T_{i,j,h^+;m,n,p^-}$ celles liant le neurone (h^+) à la position (i,j) aux autres neurones du réseau de composante (ω_y). I_{i,j,h^+} étant l'entrée externe du neurone(h^+).

Ces différentes interconnexions doivent représenter les contraintes réelles du problème et fournir une matrice de connexions pour laquelle le changement d'état considéré ne mène le réseau qu'à des points stables, ce qui implique que:

- la matrice de connexions T est symétrique;
- les éléments diagonaux doivent être non négatifs.

Cette hypothèse s'explique par le fait que la règle de Hebb décrite au *chapitre II* est la plus usuelle pour fixer les poids synaptiques et qu'elle produit spontanément une matrice symétrique. Cette symétrie garantit des propriétés importantes [KAM - 90].

L'état $v_{i,j,\alpha}$ de chaque neurone est défini par une fonction d'activation:

$$v_{i,j,h^+} = g(u_{i,j,h^+}) \quad v_{i,j,h^-} = g(u_{i,j,h^-}) \quad v_{i,j,p^+} = g(u_{i,j,p^+}) \quad v_{i,j,p^-} = g(u_{i,j,p^-}) \tag{VI.8}$$

où $g(x)$ est la fonction de Heaviside dont le rôle est prépondérant dans le comportement du réseau. Cette fonction est définie comme suit:

$$g(x) = \begin{cases} 1 & \text{si } x > 0 \\ 0 & \text{si } x < 0 \end{cases} \quad (\text{VI.9})$$

VI.4 LA FONCTION D'ENERGIE

Compte tenu de la représentation adoptée, le codage du problème sous la forme d'une énergie de Hopfield est directe. Le réseau de Hopfield associé au problème de l'estimation du mouvement est la transcription neuronale de la représentation vectorielle définie précédemment.

Les paramètres, poids synaptiques des interconnexions et entrées externes, du modèle neuronal précédent peuvent être remis dans une équation caractérisant l'énergie du réseau suivant le calcul de Hopfield et Thank [HOP - 85]. Dans le cas de notre réseau, La fonctionnelle d'énergie adaptée au problème peut alors s'écrire comme la somme de quatre termes:

$$E = \sum_{i,j} (E_x^+ + E_x^- + E_y^+ + E_y^-) \quad (\text{VI.10})$$

Les quantités E_x^+ , E_x^- , E_y^+ et E_y^- sont déterminées respectivement pour les quatre sous-réseaux représentant les composantes du vecteur vitesse en tenant compte de leurs éventuelles valeurs positives et négatives. Pour la simplicité de l'écriture, nous ne donnons que l'expression du premier terme E_x^+ :

$$E_x^+ = -\frac{1}{2} \sum_{m,n,h^+} \left[\sum_{k^+} T_{i,j,h^+;m,n,k^+} v_{m,n,k^+} v_{i,j,h^+} + \sum_{h^-} T_{i,j,h^+;m,n,h^-} v_{m,n,h^-} v_{i,j,h^+} \right. \\ \left. + \sum_{p^+} T_{i,j,h^+;m,n,p^+} v_{m,n,p^+} v_{i,j,h^+} + \sum_{p^-} T_{i,j,h^+;m,n,p^-} v_{m,n,p^-} v_{i,j,h^+} \right] \\ - \sum_{h^+} I_{i,j,h^+} v_{i,j,h^+} \quad (\text{VI.11})$$

avec: $i = 1,2,\dots,L$; $j = 1,2,\dots,M$; $m = 1,2,\dots,L$; $n = 1,2,\dots,M$; $h^+ = 1,2,\dots,N$;
 $h^- = 1,2,\dots,N$; $p^+ = 1,2,\dots,N$; $p^- = 1,2,\dots,N$.

Le coefficient $T_{i,j., m,n.}$ représente le poids synaptique entre les neurones d'indices $(i,j,.)$ et $(m,n,.)$.

Les trois autres termes de E se forment de la même manière, en remplaçant bien sûr les indices correspondant au sous-réseau donné.

Il s'agit maintenant d'exprimer les contraintes du problème c'est-à-dire construire une fonction d'énergie de manière à les décrire assez parfaitement. Sous l'hypothèse que les variations temporelles des niveaux de gris en un point de l'espace ne sont dues qu'au mouvement ou à la déformation, nous choisissons comme fonctionnelle à minimiser afin d'estimer le champ de vecteur vitesse apparente, la fonctionnelle typique d'estimation du "flux optique":

$$E = \gamma^2 \left\{ \sum_{i=1}^L \sum_{j=1}^M (f_{(i,j)x} \omega_{(i,j)x} + f_{(i,j)y} \omega_{(i,j)y} + f_{(i,j)t})^2 + \alpha^2 \sum_{i=1}^L \sum_{j=1}^M \left[\left(\frac{\partial \omega_{(i,j)x}}{\partial x} \right)^2 + \left(\frac{\partial \omega_{(i,j)x}}{\partial y} \right)^2 + \left(\frac{\partial \omega_{(i,j)y}}{\partial x} \right)^2 + \left(\frac{\partial \omega_{(i,j)y}}{\partial y} \right)^2 \right] \right\} \quad (\text{VI.12})$$

où $f_{(i,j)x}$, $f_{(i,j)y}$ et $f_{(i,j)t}$ représentent les gradients spatio-temporels de la fonction

intensité avec $f_{(i,j)\alpha} = \left(\frac{\partial f(x, y, t)}{\partial \alpha} \right)_{(i,j)}$.

$\left(\frac{\partial \omega_{(i,j)x}}{\partial x} \right)$, $\left(\frac{\partial \omega_{(i,j)x}}{\partial y} \right)$, $\left(\frac{\partial \omega_{(i,j)y}}{\partial x} \right)$ et $\left(\frac{\partial \omega_{(i,j)y}}{\partial y} \right)$ sont les dérivées partielles de ω_x et de ω_y par rapport aux coordonnées spatiales x et y .

Le premier terme de cette fonctionnelle (VI. 12) est déduit de l'équation de contrainte du mouvement apparent. Le second terme correspond à une hypothèse additionnelle sur la faible variation du champ de vecteurs vitesses (contrainte de lissage de la solution). L'introduction de cette dernière présente le double avantage d'induire un lissage du champ estimé et de préserver une estimation cohérente dans les zones où les observations locales de mouvement ne sont plus significatives.

α^2 est un paramètre à ajuster expérimentalement afin de régler le degré de lissage souhaité. Une grande valeur de α^2 provoque un fort lissage alors qu'une très faible valeur n'induit que fort peu de contraintes sur la nature de l'image. Il est important de donner une valeur convenable à α^2 . Comme l'architecture du réseau est établie en fonction des deux composantes du vecteur vitesse, la fonctionnelle à minimiser doit être construite de façon à les décrire parfaitement. A cet effet, une quantité γ^2 est introduite afin de satisfaire l'homogénéité entre les signaux d'entrée et de sortie d'un neurone. Elle est inversement proportionnelle au carré de gradient.

VI.5 LA PHASE D'IDENTIFICATION

La fonction d'énergie ayant été choisie, il reste donc à minimiser la fonctionnelle. La construction du réseau de neurones qui résout le problème revient donc à déterminer les poids synaptiques $T_{i,j,..,m,n..}$ qui lient les neurones $(i,j,..)$ et $(m,n,..)$ de telle sorte que les points stables du réseau soient les solutions recherchées. Le point fixe trouvé doit correspondre à la réponse apportée par le réseau pour estimer le "flux optique" lié à la configuration de départ. Il suffit d'identifier terme à terme les deux expressions (VI.10) et (VI.12) pour déterminer les valeurs des interconnexions propres à chacun des quatre sous-réseaux ainsi que celles les liant et les valeurs de leur entrée externe.

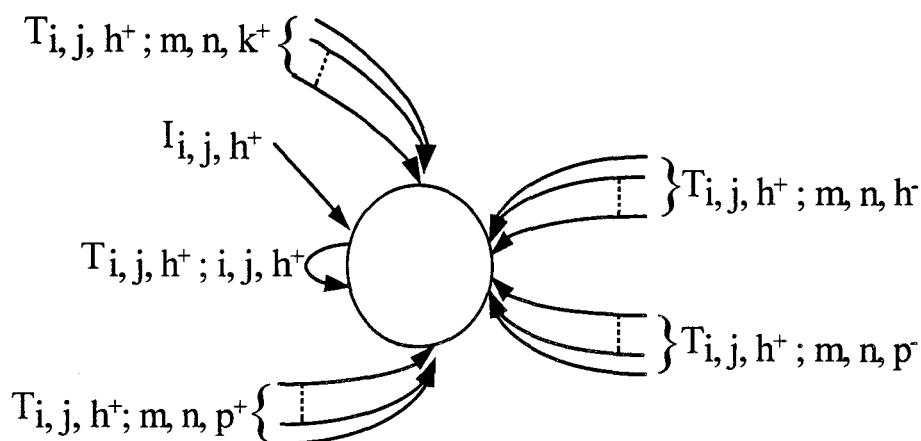


Figure VI.6 : Exemple de connexions arrivant sur un neurone.

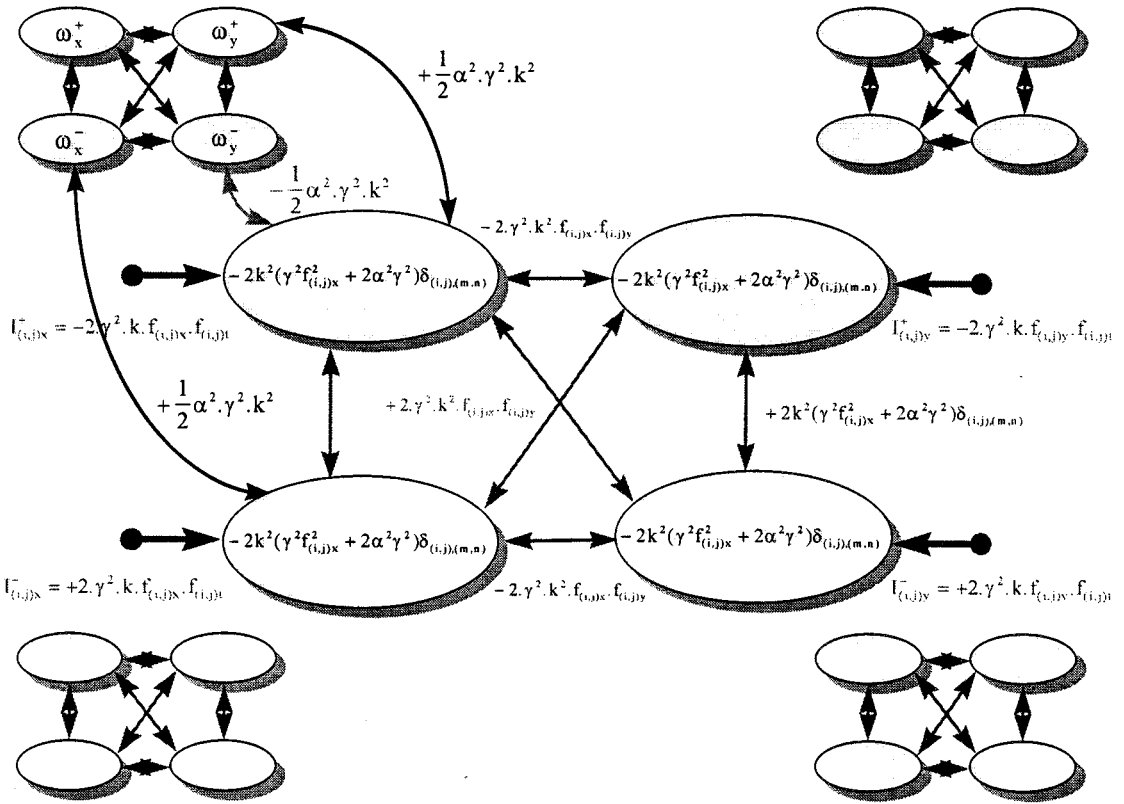


Figure VI.7 : Quelques liaisons synaptiques dans le réseau au point (i,j) , et entre celui-ci et les autres réseaux situés aux points voisins.

➤ Pour le réseau de composante ω_x

- les interconnexions entre neurones dans les sous-réseaux représentant les composantes ω_x^+ et ω_x^- :

$$T_{i,j,h^+,m,n,k^+} = -2 \cdot k^2 \cdot \left[\left(\gamma^2 \cdot f_{(i,j)x}^2 + 2\alpha^2 \cdot \gamma^2 \right) \cdot \delta_{(i,j),(m,n)} - \frac{1}{2} \alpha^2 \cdot \gamma^2 \delta_{(i,j),(m,n) \oplus s} \right] \tag{VI.13}$$

$$T_{i,j,h^+,m,n,k^+} = T_{i,j,h^-,m,n,k^-}$$

- les interconnexions entre neurones liant les sous-réseaux représentant les composantes ω_x^+ et ω_x^- :

$$T_{i,j,h^+;m,n,h^-} = + 2.k^2 \cdot \left[\left(\gamma^2 \cdot f_{(i,j)x}^2 + 2\alpha^2 \cdot \gamma^2 \right) \cdot \delta_{(i,j),(m,n)} - \frac{1}{2} \alpha^2 \cdot \gamma^2 \delta_{(i,j),(m,n)\Phi_s} \right] \quad (VI.14)$$

$$T_{i,j,h^+;m,n,k^-} = T_{i,j,h^-;m,n,k^+}$$

- les entrées externes :

$$I_{i,j,h^+} = - 2\gamma^2 \cdot k \cdot f_{(i,j)x} \cdot f_{(i,j)t} \quad \text{et} \quad I_{i,j,h^-} = -I_{i,j,h^+} \quad (VI.15)$$

➤ Pour le réseau de composante ω_y

- les interconnexions entre neurones dans les sous-réseaux représentant les composantes ω_y^+ et ω_y^- :

$$T_{i,j,p^+;m,n,q^+} = - 2.k^2 \cdot \left[\left(\gamma^2 \cdot f_{(i,j)y}^2 + 2\alpha^2 \cdot \gamma^2 \right) \cdot \delta_{(i,j),(m,n)} - \frac{1}{2} \alpha^2 \cdot \gamma^2 \delta_{(i,j),(m,n)\Phi_s} \right] \quad (VI.16)$$

$$T_{i,j,p^+;m,n,q^+} = T_{i,j,p^-;m,n,q^-}$$

- les interconnexions entre neurones liant les sous-réseaux représentant les composantes ω_y^+ et ω_y^- :

$$T_{i,j,p^+;m,n,p^-} = + 2.k^2 \cdot \left[\left(\gamma^2 \cdot f_{(i,j)y}^2 + 2\alpha^2 \cdot \gamma^2 \right) \cdot \delta_{(i,j),(m,n)} - \frac{1}{2} \alpha^2 \cdot \gamma^2 \delta_{(i,j),(m,n)\Phi_s} \right] \quad (VI.17)$$

$$T_{i,j,p^+;m,n,q^-} = T_{i,j,p^-;m,n,q^+}$$

- les entrées externes :

$$I_{i,j,p^+} = -2\gamma^2 \cdot k \cdot f_{(i,j)y} \cdot f_{(i,j)t} \quad \text{et} \quad I_{i,j,p^-} = -I_{i,j,p^+} \quad (\text{VI.18})$$

➤ **les interconnexions entre neurones liant les réseaux de composantes ω_x et ω_y dans un point (i,j) :**

$$T_{i,j,h^+;m,n,p^+} = T_{i,j,h^-;m,n,p^-} = -2 \cdot \gamma^2 \cdot k^2 \cdot f_{(i,j)x} \cdot f_{(i,j)y} \cdot \delta_{(i,j),(m,n)} \quad (\text{VI.19})$$

$$T_{i,j,h^+;m,n,p^-} = T_{i,j,h^-;m,n,p^+} = +2 \cdot \gamma^2 \cdot k^2 \cdot f_{(i,j)x} \cdot f_{(i,j)y} \cdot \delta_{(i,j),(m,n)} \quad (\text{VI.20})$$

$$T_{i,j,p^+;m,n,h^+} = T_{i,j,p^-;m,n,h^-} = -2 \cdot \gamma^2 \cdot k^2 \cdot f_{(i,j)x} \cdot f_{(i,j)y} \cdot \delta_{(i,j),(m,n)} \quad (\text{VI.21})$$

$$T_{i,j,p^+;m,n,h^-} = T_{i,j,p^-;m,n,h^+} = +2 \cdot \gamma^2 \cdot k^2 \cdot f_{(i,j)x} \cdot f_{(i,j)y} \cdot \delta_{(i,j),(m,n)} \quad (\text{VI.22})$$

où $\gamma^2 = \left(\frac{\sum_{i,j} (f_{(i,j)x}^2 + f_{(i,j)y}^2)}{L \times M} \right)^{-1}$, $\delta_{a,b}$ est le symbole de Kronecker, et

$S = \{(-1,-1), (-1,1), (1,-1), (1,1)\}$ est une série d'index excluant (0,0). Le symbole \oplus indique:

$$f_{a \oplus b} = \begin{cases} f_{a+b} & \text{si } 0 \leq a + b \leq L, M \\ 0 & \text{ailleurs} \end{cases} \quad (\text{VI.23})$$

L'utilisation du réseau consiste à forcer l'état initial des neurones et à laisser évoluer la structure bouclée vers un état d'équilibre qui fournit le résultat du traitement neuronal. Il est intéressant de partir avec une initialisation assez proche de la solution, ce qui permettra d'atteindre un minimum local correspondant à une solution cohérente avec le problème traité.

Nous remarquons que la matrice des poids synaptiques est symétrique, ce qui augure de la convergence du réseau. Il est à noter par ailleurs que ces poids

synaptiques n'évoluent pas durant le processus de convergence du réseau car ils sont complètement spécifiés par l'énergie à minimiser.

VI.6 EVOLUTION DE L'ENERGIE AU COURS DE LA DYNAMIQUE

Afin de faciliter la compréhension, nous raisonnons sur un sous-réseau au lieu de tout le réseau complet pour montrer comment la dynamique du réseau au cours des itérations converge vers un minimum local. Nous considérons le sous-réseau de composante à valeurs positives ω_x^+ .

Soit la variation d'état du neurone h^+ au point (i,j) :

$$\Delta v_{i,j,h^+} = v_{i,j,h^+}^{\text{nouveau}} - v_{i,j,h^+}^{\text{ancien}} \quad (\text{VI.24})$$

et soit la variation d'énergie associée qui se déduit de l'équation de mise à jour:

$$\Delta E = E^{\text{nouveau}} - E^{\text{ancien}} \quad (\text{VI.25})$$

Nous supposons qu'à chaque itération, seul un neurone change son état de 1 vers 0 ou de 0 vers 1. Comme la fonction d'activation est une fonction Heaviside, la variation d'énergie ΔE associée à une variation d'état $\Delta v_{i,j,h^+}$ est donnée par:

$$\Delta E = - \left(\sum_{h^+}^N T_{i,j,h^+,i,j,h^+} + I_{i,j,h^+} \right) \Delta v_{i,j,h^+} - \frac{1}{2} T_{i,j,h^+,i,j,h^+} (\Delta v_{i,j,h^+})^2 \quad (\text{VI.26})$$

$$\Delta E = - \left(u_{i,j,h^+} \Delta v_{i,j,h^+} + \frac{1}{2} T_{i,j,h^+,i,j,h^+} \Delta v_{i,j,h^+}^2 \right) \quad (\text{VI.27})$$

- Quand $u_{i,j,h^+} \gg 0$, v_{i,j,h^+} change son état de 0 (état ancien) vers 1 (état nouveau).

Ce changement vaut $\Delta v_{i,j,h^+} = 1$. La variation d'énergie est alors:

$$\Delta E = -u_{i,j,h^+} - \frac{1}{2} T_{i,j,h^+,i,j,h^+} \quad (\text{VI.28})$$

Si $T_{i,j,h^+,i,j,h^+} < -2 u_{i,j,h^+}$, alors $\Delta E > 0$ (VI.29)

- De même manière, lorsque $u_{i,j,h^+} \ll 0$, v_{i,j,h^+} change son état de 1 (état ancien) vers 0 (état nouveau) et le changement d'état vaut $\Delta v_{i,j,h^+} = -1$. La variation d'énergie est alors:

$$\Delta E = +u_{i,j,h^+} - \frac{1}{2} T_{i,j,h^+,i,j,h^+} \quad (\text{VI.30})$$

Si $T_{i,j,h^+,i,j,h^+} < +2 u_{i,j,h^+}$, alors $\Delta E > 0$ (VI.31)

Enfin lorsque $T_{i,j,h^+,i,j,h^+} < -2 |u_{i,j,h^+}|$, nous avons $\Delta E > 0$ qui signifie que la variation d'énergie n'est pas toujours négative et que la fonction d'énergie E ne sera plus monotone décroissante. Le réseau devient instable et par conséquent sa convergence n'est plus garantie. Afin d'éliminer ce problème, nous adoptons le fonctionnement suivant:

L'algorithme de fonctionnement :

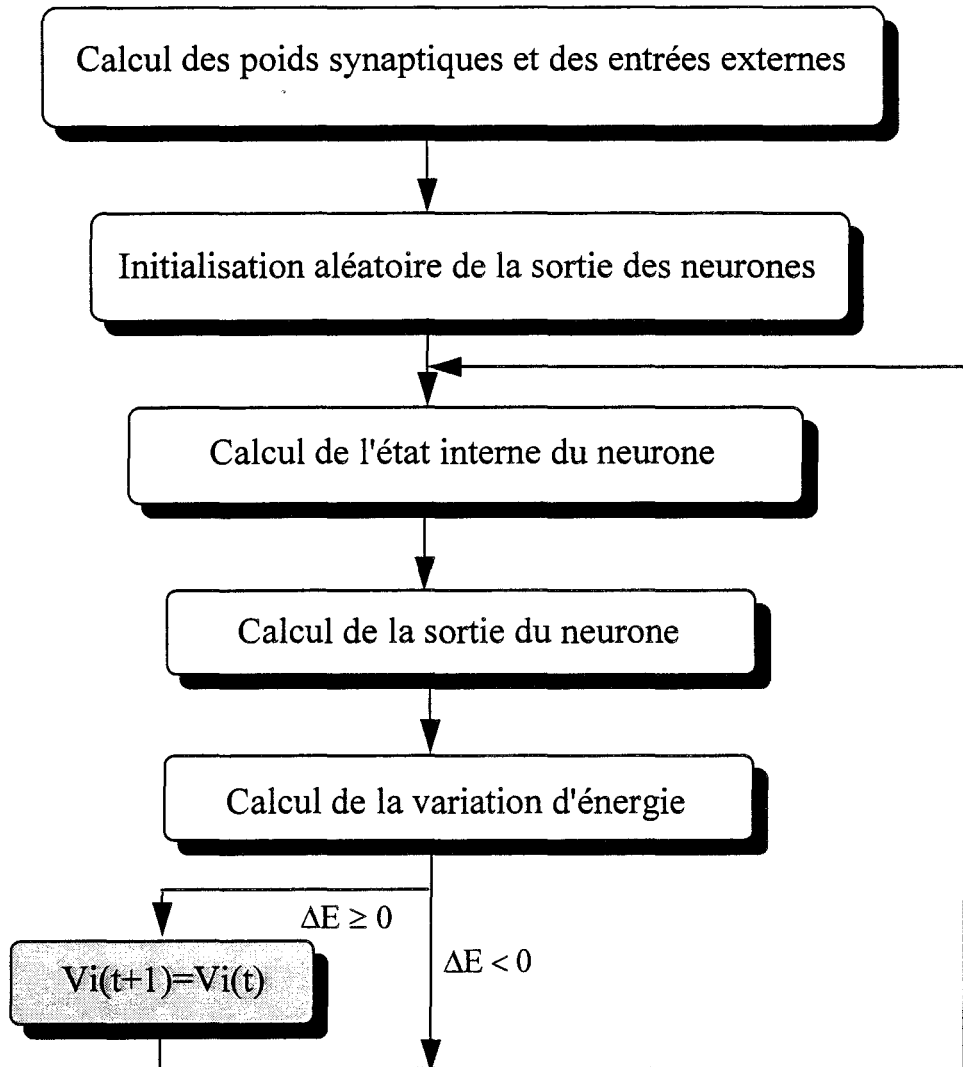


Figure VI.8 : Organigramme de l'algorithme.

Il s'agit d'un algorithme itératif dont le but est d'optimiser la fonction d'énergie définie précédemment. Pour atteindre cet objectif, l'algorithme travaille sur l'ensemble des points de l'image. Chaque point étant constitué d'un ensemble de neurones. La dynamique du réseau est entièrement définie par la donnée des paramètres, poids synaptiques symétriques et entrées externes, celle de la fonction d'activation des neurones et enfin par le mode d'itération qui est asynchrone et aléatoire. Les neurones changent d'état les uns après les autres et dans un ordre aléatoire. Le vecteur d'état est ainsi actualisé élément par élément en tenant compte des nouvelles valeurs des composants dès qu'elles sont disponibles. Cet algorithme de minimisation est parfaitement déterministe. En effet, le réseau passe une deuxième fois par la configuration précédemment atteinte. A chaque itération, l'énergie peut

rester fixe lorsqu'il n'y a pas de changement d'état ou bien elle décroît. Le critère de terminaison est donc basé sur le principe suivant: la recherche est terminée si au bout d'un certain temps, une énergie minimum est atteinte. Dans ce cas, la mise à jour de chaque neurone i conserve l'état dudit neurone; la configuration atteinte est par conséquent un état stationnaire qui est la réponse au problème.

La convergence et la stabilité du processus itératif dépendent donc des données telles que les paramètres, poids synaptiques et entrées externes, ou encore du choix de l'initialisation qui joue aussi un rôle important dans l'évolution du processus.

Il convient de noter que cet algorithme ne réalise qu'une descente dans un puits d'énergie (minimum local) déterminé par une configuration initiale convenable. Il ne comporte pas la caractéristique essentielle du recuit simulé qui est de visiter plusieurs puits d'énergie avant de se stabiliser dans l'un d'eux, mais présente l'avantage si le système n'est pas trop loin de la solution lors de l'initialisation, de converger beaucoup plus rapidement vers une solution "acceptable".

Enfin, il est important de remarquer que contrairement aux réseaux à couches, il n'y a pas d'apprentissage car les poids synaptiques sont calculés avant le fonctionnement du réseau.

VI.7 PRISE EN COMPTE DES DISCONTINUITES

L'approche présentée dans ce chapitre ne tient pas compte des discontinuités du champ de déplacement. Or l'un des problèmes importants liés à l'estimation du mouvement est de savoir détecter et gérer les discontinuités pouvant apparaître dans le champ de vitesses. Nous pouvons résoudre ce problème en introduisant par exemple dans la fonctionnelle d'énergie (VI.12) une troisième contrainte:

$$\begin{aligned}
E = \gamma^2 & \left\{ \sum_{i=1}^L \sum_{j=1}^M (f_{(i,j)x} \omega_{(i,j)x} + f_{(i,j)y} \omega_{(i,j)y} + f_{(i,j)t})^2 \right. \\
& + \alpha^2 \sum_{i=1}^L \sum_{j=1}^M \left[\left(\frac{\partial \omega_{(i,j)x}}{\partial x} \right)^2 + \left(\frac{\partial \omega_{(i,j)x}}{\partial y} \right)^2 + \left(\frac{\partial \omega_{(i,j)y}}{\partial x} \right)^2 + \left(\frac{\partial \omega_{(i,j)y}}{\partial y} \right)^2 \right] \left. \right\} \quad (\text{VI.32}) \\
& + \frac{C}{2} \gamma^2 \left\{ \left[(\omega_{(i,j)x} - \omega_{(i,j+1)x})^2 + (\omega_{(i,j)y} - \omega_{(i,j+1)y})^2 \right] [1 - L_{i,j}^v] \right. \\
& \left. + \left[(\omega_{(i,j)x} - \omega_{(i+1,j)x})^2 + (\omega_{(i,j)y} - \omega_{(i+1,j)y})^2 \right] [1 - L_{i,j}^h] \right\}
\end{aligned}$$

$$\text{où } L_{i,j}^v = 1 \text{ si } \|\bar{\omega}_{i,j} - \bar{\omega}_{i,j+1}\|^2 \neq 0 \text{ et } L_{i,j}^h = 1 \text{ si } \|\bar{\omega}_{i,j} - \bar{\omega}_{i+1,j}\|^2 \neq 0 \quad (\text{VI.33})$$

C est un réel positif à ajuster. Le choix des trois paramètres α , γ et C s'avère important, en particulier pour une bonne estimation des discontinuités. Les paramètres doivent être ajustés manuellement sur chaque exemple traité.

Il s'agit en fait d'un processus de lignes permettant de tenir compte des ruptures locales du champ de mouvement. Le système de processus de lignes consiste à introduire des lignes verticale et horizontale L^v , L^h . Chaque ligne peut prendre deux états:

- 1 pour l'état actif. Les pixels voisins sont indépendants, il y a présence de discontinuité.
- 0 pour l'état inactif. Le pixel est lié avec son voisin (forte interaction locale).

Le principe de détection des ruptures utilise le lien existant entre l'évolution temporelle et la distribution spatiale du champ de vecteurs vitesse, au voisinage d'une frontière de mouvement. Ce même principe peut conférer au réseau une bonne prise en compte des discontinuités du champ de vecteurs vitesse et peut ainsi améliorer les résultats de l'estimation.

D'un point de vue de l'architecture neuronale, cette prise en compte des discontinuités est faite sans grande modification de la structure du réseau. Le processus de ligne vient se superposer comme un deuxième réseau sur le réseau de base [GEM - 88] [HUT - 88] [DEV - 95].

VI.8 STRATEGIES DE PARCOURS

Une autre particularité associée au réseau et à la méthode que nous proposons réside dans la possibilité de parcourir le réseau de différentes manières. Ceci est rendu possible par l'utilisation d'un facteur de quantification variable entre les différentes itérations. Le principe est le suivant: au début, la relaxation est effectuée sur un réseau comportant très peu de neurones (faible résolution, mais rapidité de convergence). Puis aux étapes suivantes, les neurones des réseaux (de meilleure résolution) héritent des activations de leur neurone parent. Ces réseaux sont de nouveau amenés à se relaxer dans un état stable, et ainsi de suite jusqu'à atteindre un degré de résolution satisfaisant (*figure VI.9*). Le champ vectoriel de déplacement est ainsi estimé sur les basses résolutions, puis raffiné avec les résolutions supérieures.

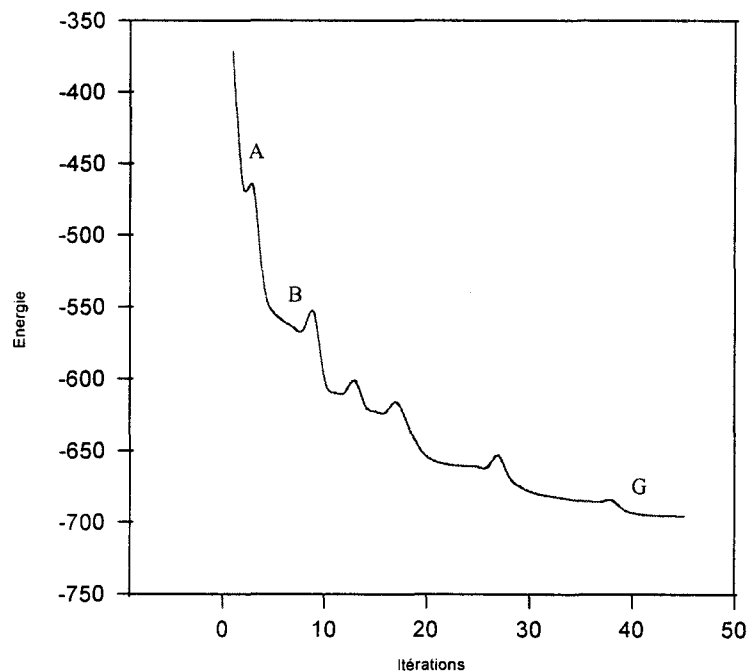


Figure VI.9 : A partir d'une solution initiale (A), des solutions successivement meilleures peuvent être obtenues. La solution finale (G) est un minimum de E.

Cette approche présente deux avantages principaux. Le premier est de fournir éventuellement des vecteurs vitesses suivant un schéma de type "progressif". La progression de la solution est faite en minimisant l'énergie à chaque étape. La solution finale est celle qui possède la plus petite valeur d'énergie. Le second avantage permet, comme nous le verrons plus tard, une accélération de la vitesse de convergence du réseau d'où une optimisation des temps de calcul et de l'espace mémoire utilisé.

VI.9 RESULTATS EXPERIMENTAUX

Les performances de l'algorithme présenté sont évaluées sur des images de synthèse. Des résultats sur des images réelles sont également donnés afin de montrer les avantages et les limites de notre approche.

VI.9.1 Séquence d'images de synthèse

Nous supposons dans un premier temps que les mouvements sont des translations simples et des rotations.

VI.9.1.1 Cas d'une translation pure

Nous reprenons la séquence d'images de synthèse (*figure V.1*) présentée dans le chapitre précédent. Le carré (10 x 10) se translatant vers le coin bas gauche de l'image ($\omega_x = 1$, $\omega_y = 1$). Nous étudions le comportement de l'algorithme dans les deux approches de convergence du réseau, directe et suivant le type de parcours adopté.

L'analyse se fait sur les points suivants:

- la garantie de convergence en un point stable et le temps de convergence;
- la qualité des champs de vecteurs vitesse obtenus;

- la simplicité d'utilisation (lié au nombre de paramètres à ajuster pour un bon fonctionnement de la méthode)

Dans un premier temps, nous utilisons la structure de convergence directe et nous étudions l'impact du nombre de neurones sur la qualité des résultats obtenus. Nous fixons un nombre N de neurones donné et nous laissons le système évoluer jusqu'à l'obtention d'une valeur minimale de l'énergie finale. Ainsi, des tests lorsque $N = 5$, $N = 20$, $N = 100$ et $N = 200$ sont réalisés. Le paramètre de régularisation étant le même ($\alpha = 100$) pour tous les tests.

Nous comparons la carte des vecteurs vitesse estimés pour ces différentes configurations. Les champs de vecteurs (*figure VI.10*) sont obtenus respectivement après 6, 20, 47 et 57 itérations. Nous remarquons, lorsque $N = 5$, que le champ n'est pas dense et présente des valeurs aberrantes sur les bords. Le résultat peut ne plus être significatif à cause du déplacement quelquefois très important du minimum local. Par contre, les champs de vecteurs obtenus pour les autres valeurs de N sont de plus en plus denses et cohérents au fur et à mesure que le nombre de neurones augmente ($N = 20$, $N = 100$, puis $N = 200$). La régularité de la solution augmente avec l'ordre du nombre de neurones comme cela était prévisible.

L'analyse des courbes de la *figure VI.11* conduit à observer que la valeur minimale de l'énergie est d'autant plus faible que le nombre de neurones augmente. De plus, ces courbes présentent deux zones distinctes de variation; dans un premier temps, l'énergie décroît fortement puis la variation d'énergie diminue. La zone de faible variation correspond à une très faible amélioration de la qualité de la solution. Il est donc possible d'asservir la durée du calcul (arrêt des itérations) sur le gradient de la courbe d'énergie. Le critère de terminaison préalable des itérations étant lié à un seuil de ce gradient.

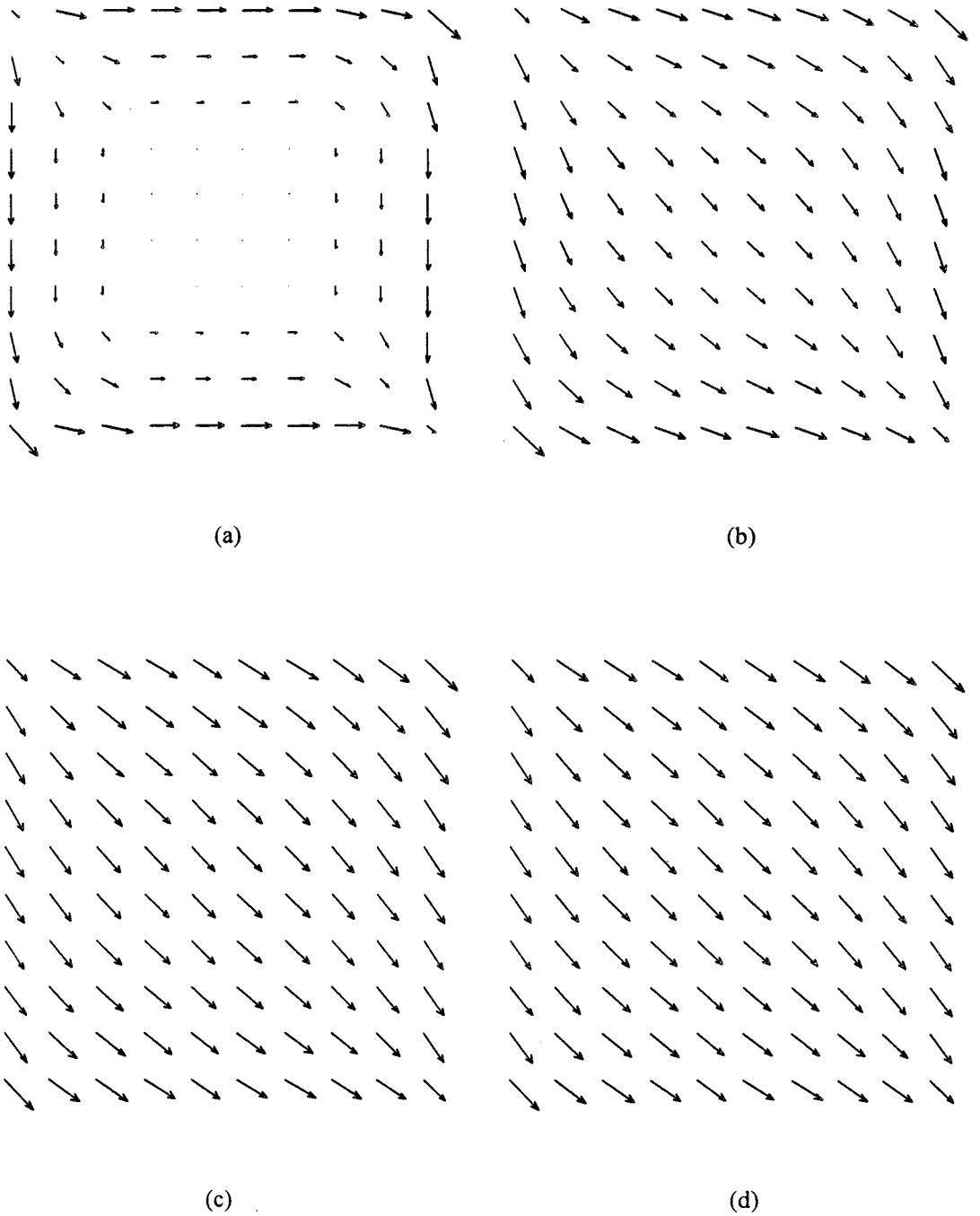


Figure VI.10 : Champs de vitesses générés par l'approche directe

(a) $N=5$, (b) $N=20$, (c) $N=100$, (d) $N=200$.

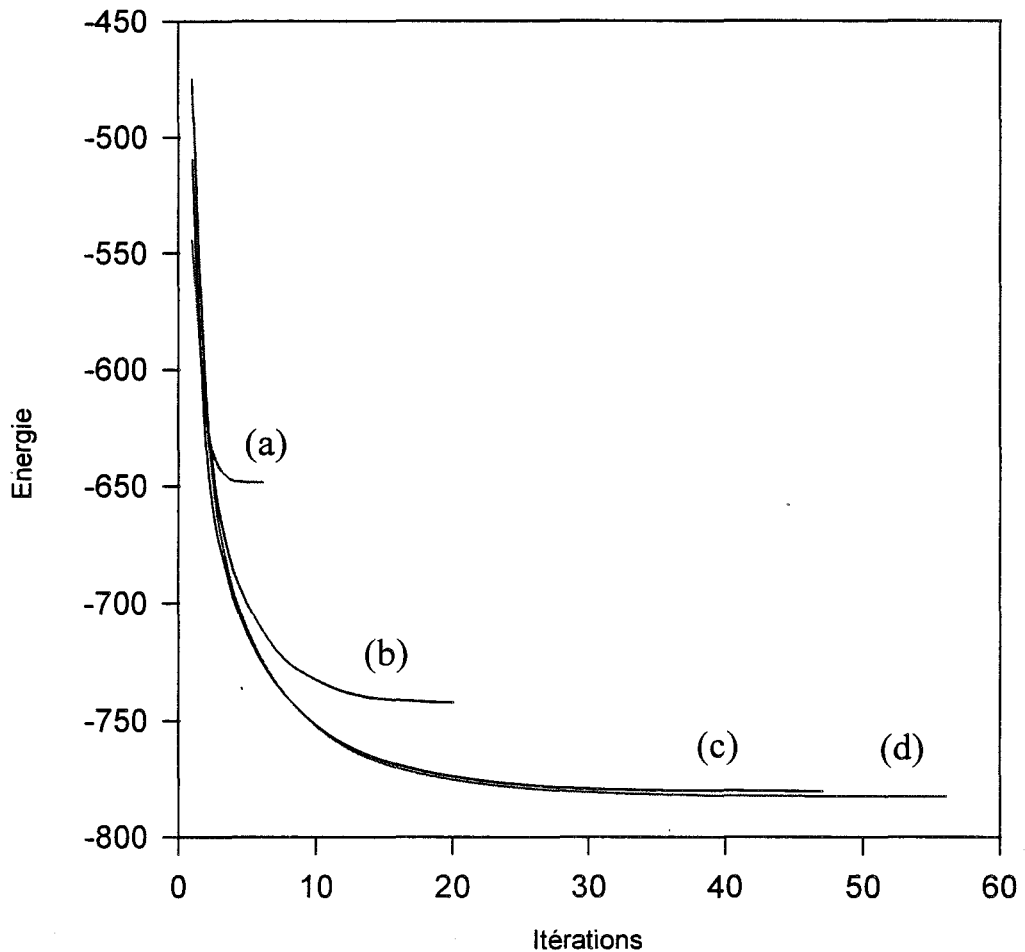


Figure VI.11 : Décroissance de l'énergie en fonction du nombre d'itérations dans le cas de l'approche directe : (a) $N=5$, (b) $N=20$, (c) $N=100$, (d) $N=200$.

VI.9.1.2 Cas d'une rotation

De façon analogue, nous avons testé le système dans le cas de champs plus complexes. Le champ rotationnel est le résultat d'une rotation autour de l'axe optique de la caméra. Un exemple sur les résultats obtenus sur la même séquence "carré" munie d'une rotation centrée de 2° sont présentés en *figure VI.12*

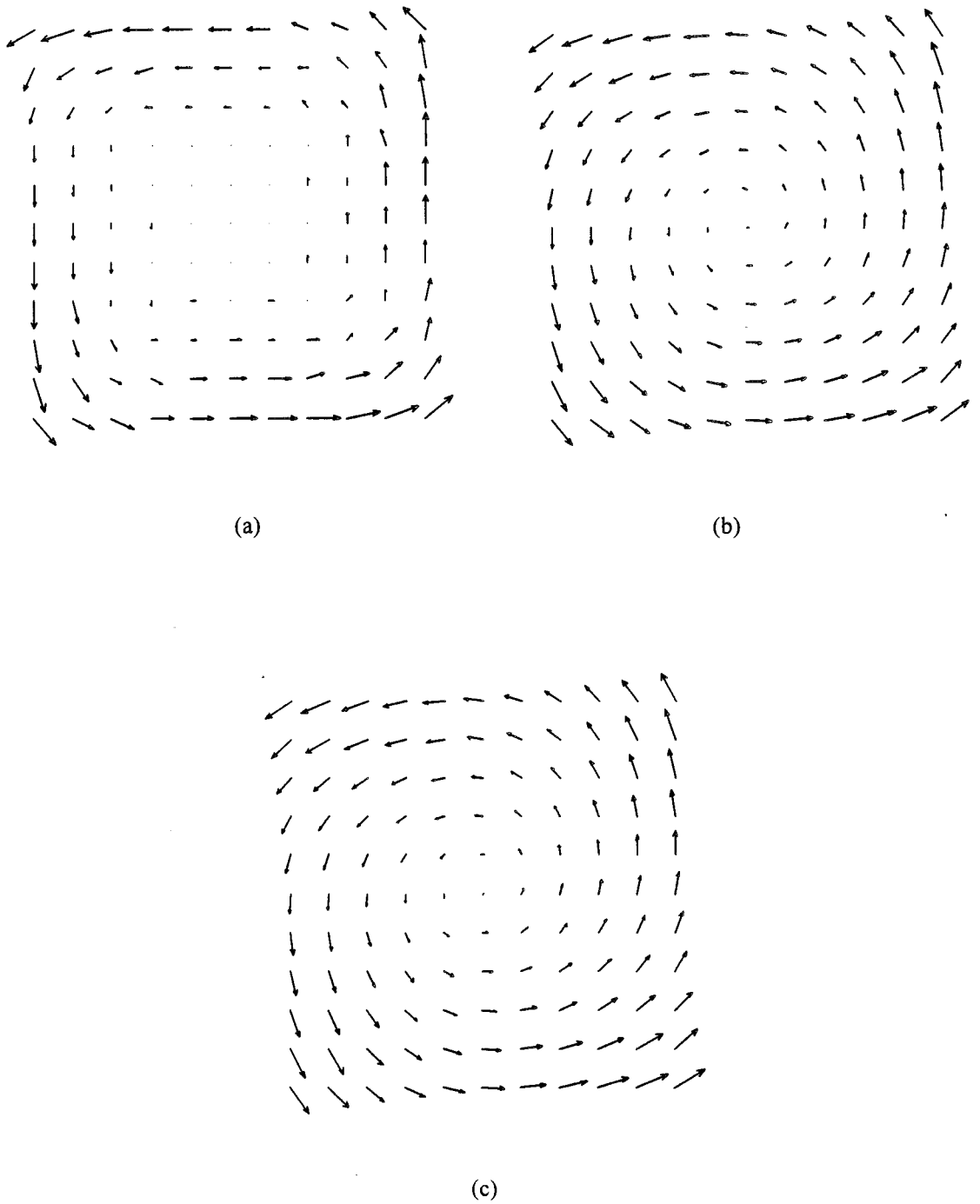


Figure VI.12 : Champs de vecteurs vitesse obtenu par:
(a) et (b) le réseau neuronal lorsque $N=20$ et $N=200$
(c) la méthode de Horn et Schunck.

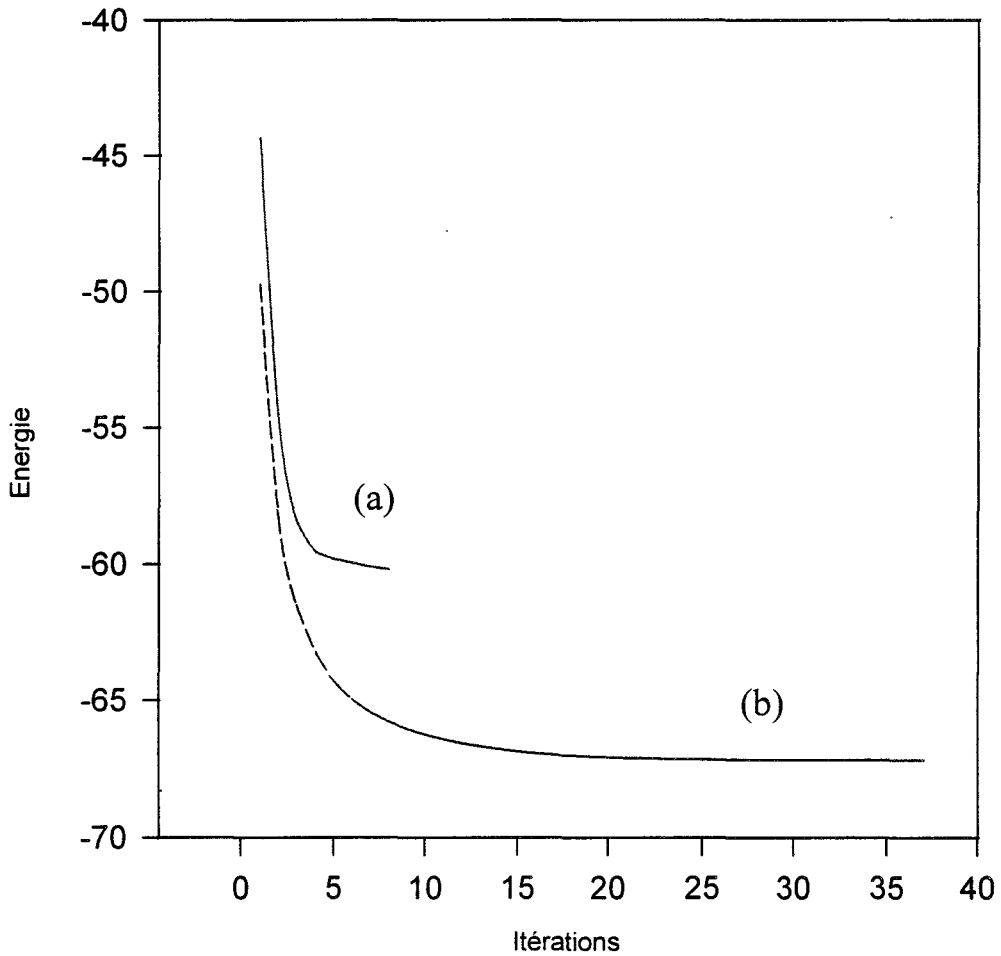


Figure VI.13 : Décroissance de l'énergie en fonction du nombre d'itérations dans le cas de l'approche directe : (a) $N=20$, (b) $N=200$.

Nous remarquons dans ce cas aussi que l'énergie la plus basse est obtenue, comme dans le cas de la translation, pour des valeurs de N élevés.

A titre de comparaison, le résultat obtenu avec l'algorithme de Horn et Schunck est également donné (figure VI.12.c).

Enfin, il est intéressant d'exprimer la géométrie d'autres types de champs de vitesse. En effet, l'algorithme est capable de générer des champs de vecteurs apparents plus complexes comme le champ divergent ou encore les champs hyperboliques.

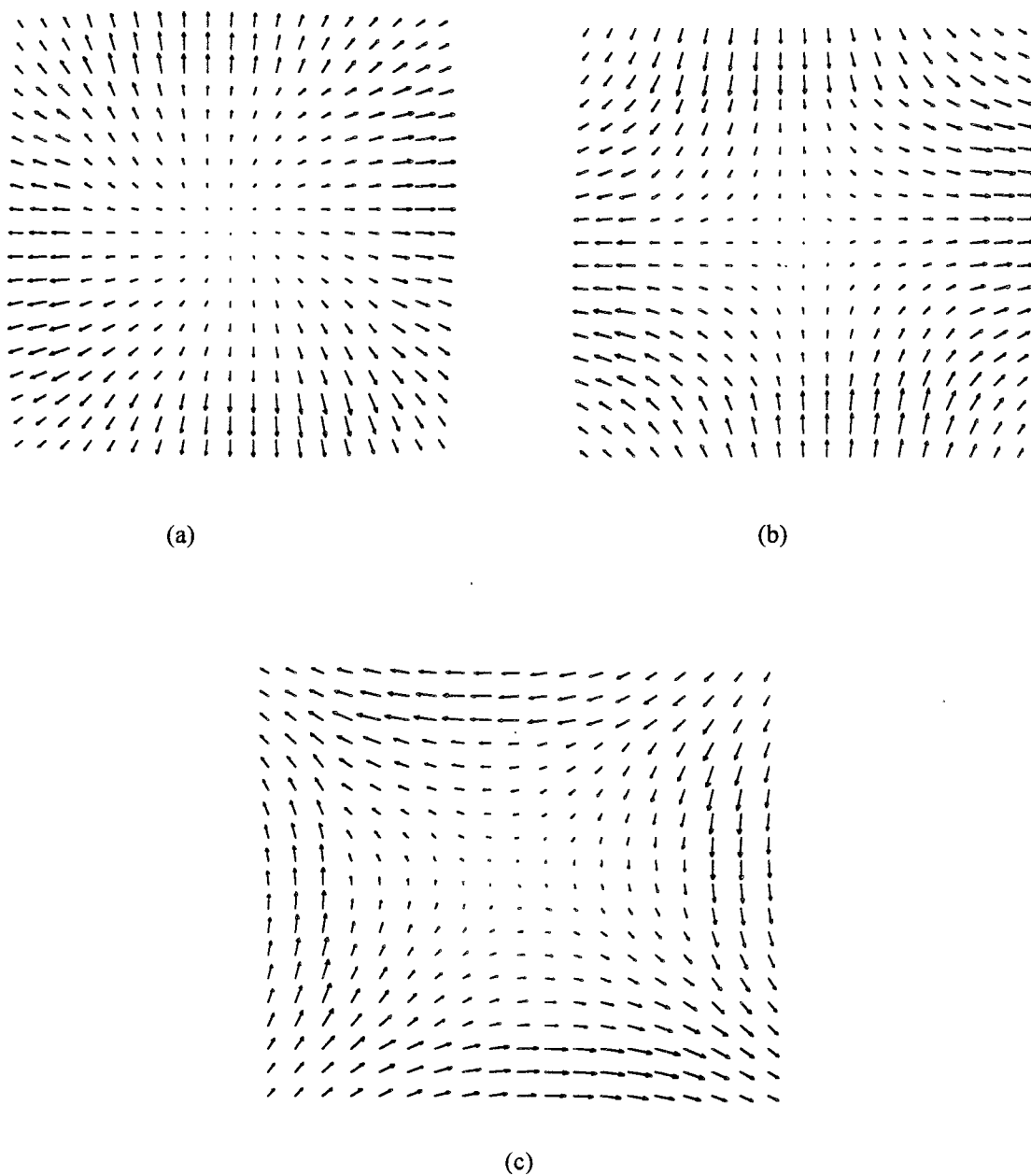


Figure VI.14: (a) champ divergent, (b) et (c) champs hyperboliques.

Les trois champs représentés sur la *figure VI.14* sont obtenus pour $N = 100$ et $\alpha = 100$. Le champ de vitesse divergent est le résultat d'un mouvement axial, c'est à dire le long de l'axe optique de la caméra. Les termes hyperboliques traduisent des mouvements plus complexes.

VI.9.2 Séquence d'images réelles

La première séquence testée est celle de la voiture se déplaçant en translation oblique vue dans le chapitre précédent. Les paramètres utilisés pour ce test sont : $N = 20$ et $\alpha = 100$.

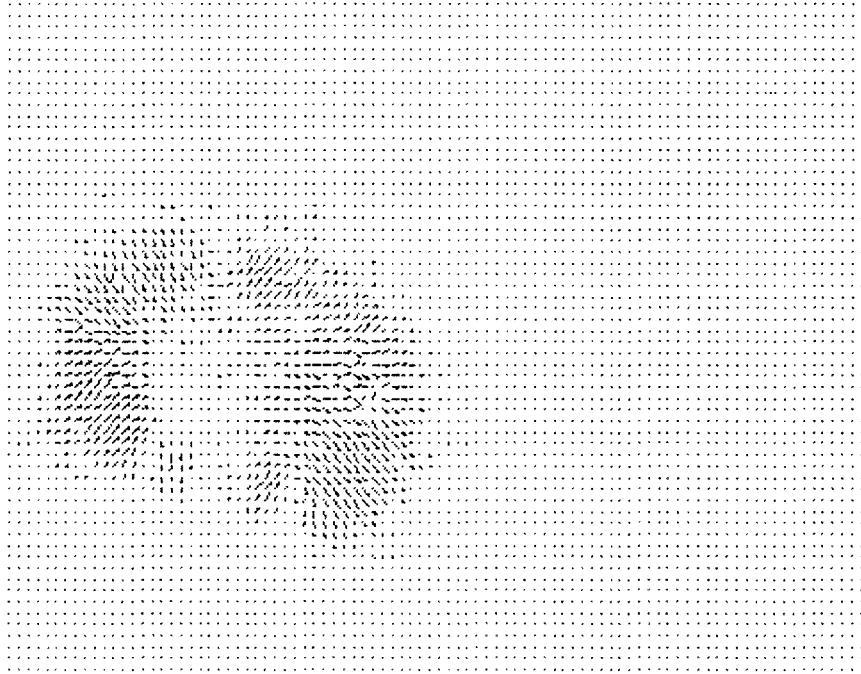


Figure VI.15 : *Champ de vecteurs vitesse.*

Cette figure représente un champ des vitesses tout à fait cohérent avec le mouvement réel des objets de la scène. On y constate en effet que le fond présente un mouvement apparent de translation vers l'avant, ce qui correspond bien à son mouvement apparent réel. Quant à la voiture, elle-aussi présente un mouvement cohérent, la zone correspondant à l'avant du véhicule ayant un mouvement apparent plus important que la zone correspondant à l'arrière du véhicule.

L'autre séquence présentée en *figure VI.16* est une séquence représentant une personne se déplaçant légèrement vers la droite. Les champs de vitesse sont estimés lorsque $N = 20$ et $\alpha = 100$, puis pour $N = 100$ et $\alpha = 100$.

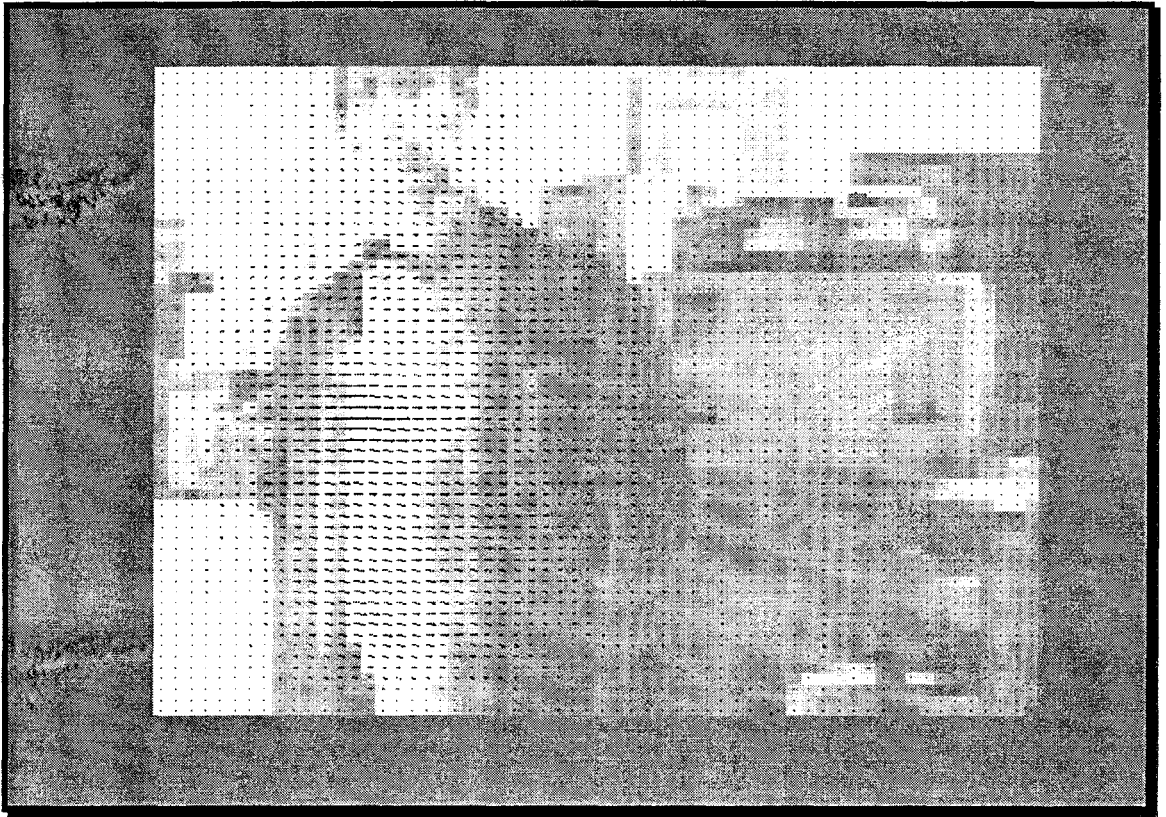


Figure VI.16 : Image originale avec superposition des vecteurs vitesse obtenus pour $N=100$.

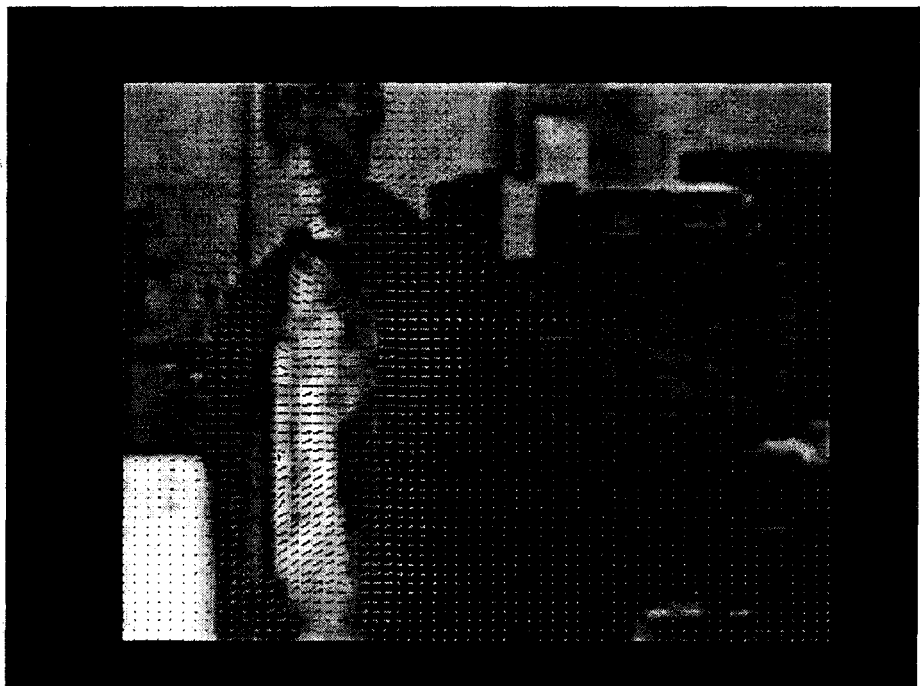


Figure VI.17 : Image originale avec superposition des vecteurs vitesse obtenus pour $N=20$.

Nous constatons que les résultats obtenus sur cet exemple sont satisfaisants. En effet, les zones de mouvement sur la totalité de l'image sont correctement reconstituées.

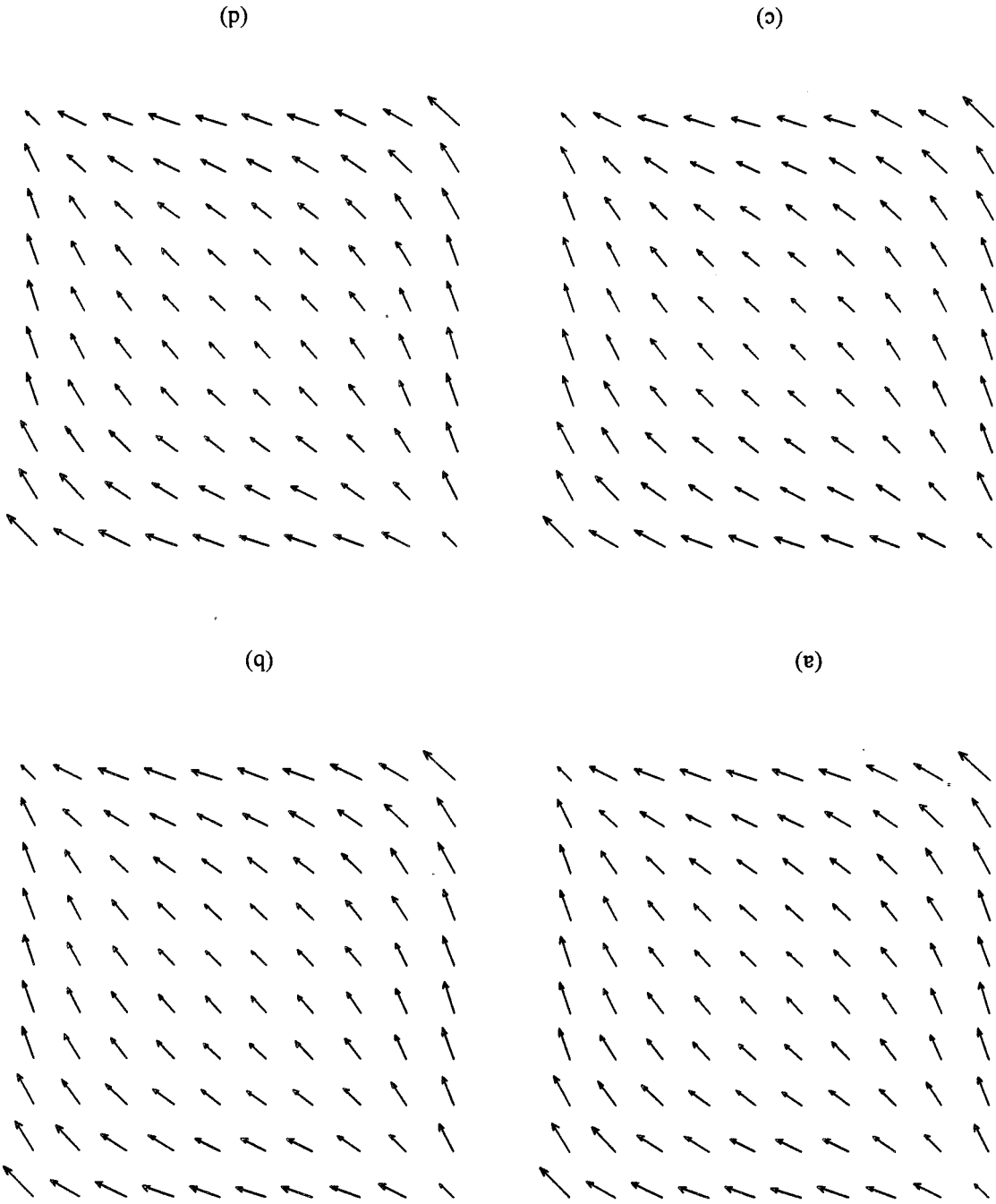
La méthode permet donc d'obtenir un champ de vitesse relativement régulier et de bonne cohérence temporelle. Nous remarquons aussi que pour un nombre de neurones important, la fiabilité du champ est encore meilleure.

VI.9.3 Analyse des parcours types

Des tests sur l'évolution du réseau pour différentes stratégies de parcours au sein de cette structure sont réalisés. Nous avons expérimenté le cas du parcours des quantifications grossières aux plus fines. Les parcours retenus sont le parcours direct ($N = 20$), un parcours de type exponentiel ($N = 3, 7, 20$), un parcours de type logarithmique ($N = 2, 17, 20$), et un parcours de type linéaire ($N = 1, 10, 20$). Les *figures VI.18* et *VI.19* illustrent les performances de tels parcours sur les champs de vecteurs vitesse, ainsi que sur la décroissance de l'énergie. Les champs de déplacement sont obtenus respectivement après 20, 26, 35 et 25 itérations. Comme nous le remarquons sur la *figure VI.18*, les champs obtenus sont sensiblement similaires.

Le réseau évolue à partir d'une stratégie de parcours donné vers les configurations d'énergie les plus basses. Sur les courbes illustrées en *figure VI.18*, nous constatons l'existence de plusieurs minima qui correspondent en fait aux différents parcours. Le "bon" minimum local est celui dont l'énergie est la plus basse. Parmi les exemples que nous avons pris, le parcours ($N = 1, 10, 20$) donne la plus faible énergie.

Figure VI.18 : Champs de vitesses g n r s par l'approche pyramidale pour les configurations suivantes : (a) $N = 20$, (b) $N = 3, 7, 20$; (c) $N = 2, 17, 20$; (d) $N = 1, 10, 20$.



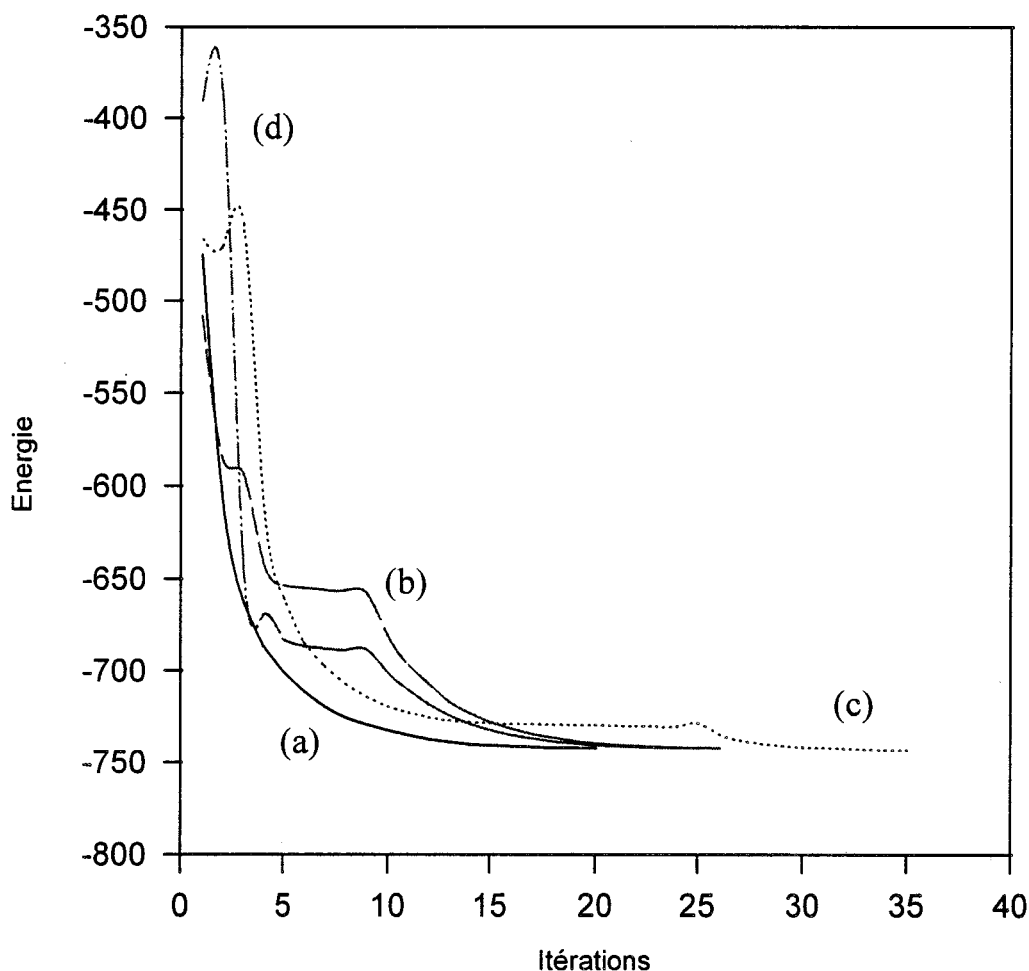


Figure VI.19 : *Décroissance de l'énergie en fonction du nombre d'itérations dans le cas de l'approche pyramidale pour les configurations :*
 (a) $N=20$, (b) $N=3, 7, 20$; (c) $N=2, 17, 20$; (d) $N=1, 10, 20$.

Des tests sur d'autres types de parcours ont été effectués. A titre d'exemple, nous avons retenu les parcours de type exponentiel ($N = 3, 7, 20$), logarithmique ($N = 2, 9, 13, 15, 17, 18, 19, 20$) et linéaire ($N = 1, 2, 3, \dots, 19, 20$). Différents minima correspondant aux différents parcours sont obtenus, mais parmi les exemples pris, le parcours ($N = 1, 2, 3, \dots, 19, 20$) donne la plus faible énergie.

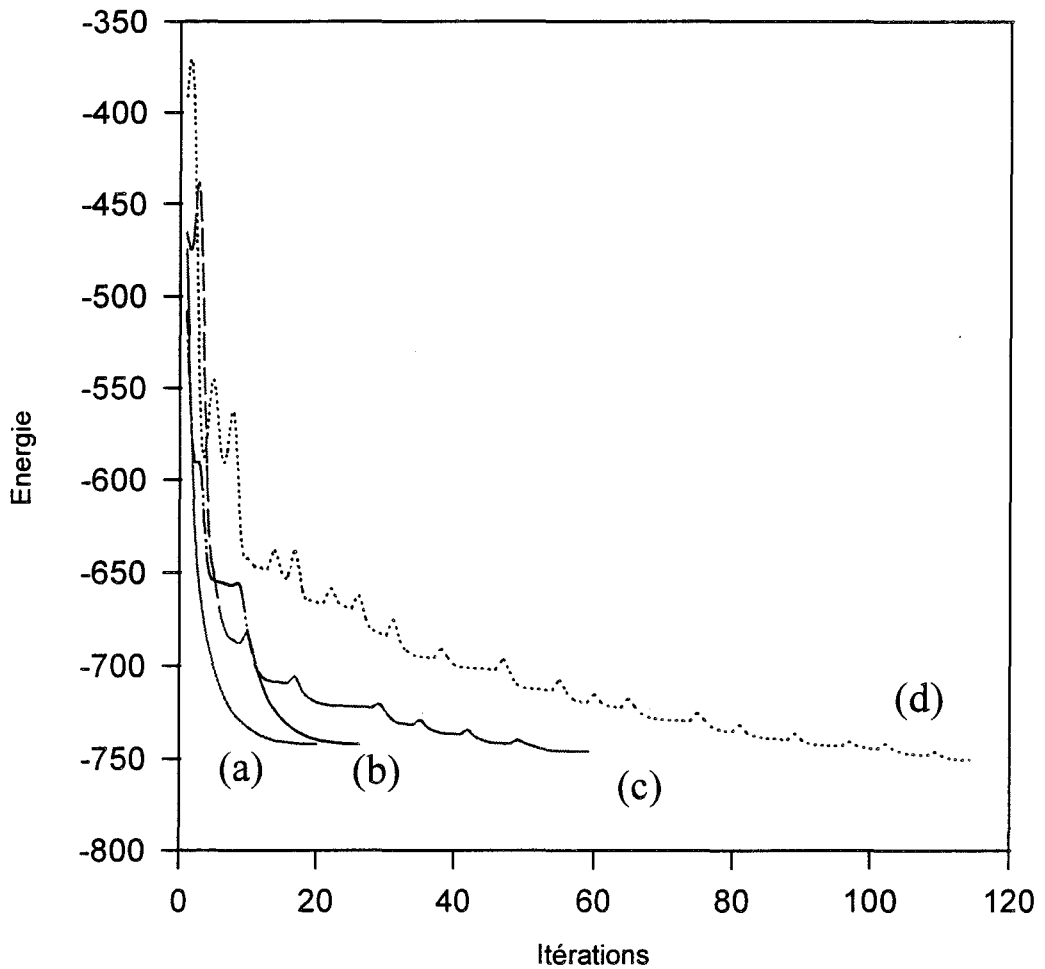


Figure VI.20 : Décroissance de l'énergie en fonction du nombre d'itérations dans le cas de l'approche pyramidale pour les configurations :
 (a) $N=20$, (b) $N=3, 7, 20$; (c) $N=2, 9, 13, 15, 17, 18, 19, 20$; (d) $N=1, 2, 3, \dots, 19, 20$.

Nous désirons maintenant évaluer l'activité du réseau. Nous avons vu que pour différentes configurations, nous obtenons différentes solutions. Or, nous savons que la connectivité du réseau est complète. L'état de chaque neurone dépend par conséquent de tous les autres, y compris de lui-même. Pour spécifier le nombre de modifications d'état (changements de neurones), il suffit de tenir compte des successeurs de chacun des états du réseau. Nous pouvons alors évaluer l'activité du réseau par le nombre de neurones changés par chaque parcours au cours de la convergence.

Dans le cas du parcours direct, le réseau converge continûment vers la valeur finale de l'énergie et le nombre de neurones changés est dans ce cas le plus important. Ce qui n'est pas le cas pour des configurations à différents parcours. En effet, nous remarquons que dans le cas des autres types de parcours, le nombre de neurones modifiés est dans tous les cas plus faible que dans celui du parcours direct. L'activité du réseau est donc globalement plus faible.

Configuration	Nombre de neurones modifiés
N=20	4500
N=3, 7, 20	3314
N=2, 17, 20	4057
N=1, 10, 20	3307

Tableau VI.1: *Nombre de neurones changés.*

On peut également constater que plus le nombre de neurones actifs est élevé, meilleure est la qualité de la solution et corrélativement plus basse est l'énergie finale.

Cependant, d'un point de vue pratique, il est à remarquer que les temps de calculs nécessaires à chaque type de parcours pour atteindre sa valeur minimale d'énergie, peuvent être notablement différents dans la mesure où une itération pour chaque type de parcours ne correspond pas aux tests et à l'évaluation du potentiel du même nombre de neurones. Ainsi, dans l'exemple traité, si on prend comme unité de temps, le temps nécessaire à l'évaluation d'un neurone, les temps nécessaires effectifs de convergence des quatre parcours traités (N = 20; N = 3, 7, 20; N = 2, 17, 20; et N = 1, 10, 20) sont donnés par:

Configuration	Temps effectif de convergence
N=20	400 unités / pixel
N=3, 7, 20	378 unités / pixel
N=2, 17, 20	559 unités / pixel
N=1, 10, 20	371 unités / pixel

Tableau VI.2: Temps effectif de convergence.

Détermination automatique de k

Une autre expérimentation consiste à fixer le nombre de neurones et faire varier le pas de quantification k . Ce dernier étant introduit pour obtenir un échantillon de quantification dans la dynamique des vecteurs vitesse. k est un réel positif à ajuster en fonction de l'image à traiter. Il est lié au nombre total de neurones, étant donné que le produit $|k \times N|$ spécifie la dynamique des vecteurs déplacement que le réseau est capable de traiter.

Nous avons réalisé une série de tests dans le cas du parcours direct pour différentes valeurs du nombre N de neurones ($N = 10$, $N = 20$, $N = 30$). Pour chaque valeur de N fixée, nous faisons varier k . Le paramètre de régularisation α ne change pas et reste égal à 100. Pour chaque N fixé, on donne en *figure VI.21* les différentes valeurs d'énergie minimale obtenues lorsque k varie. Nous constatons que ces courbes présentent toutes un minimum pour une valeur de k différente lorsque N varie. Toutefois, il apparaît que dans les trois cas traités, les positions des minima de ces courbes se produisent pour une valeur constante du produit $|k \times N|$. Dans l'exemple présent, ceci se produit dans les trois cas pour une valeur du produit $(k \times N)$ sensiblement égale à 0.875.

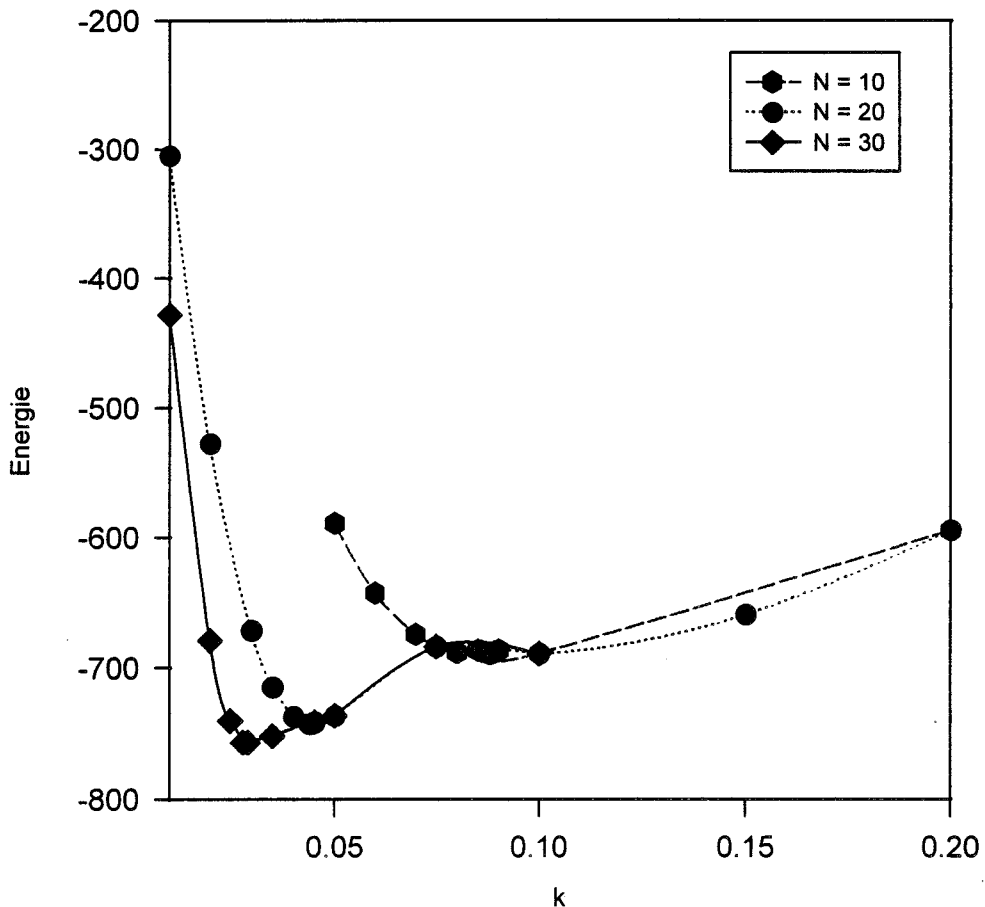


Figure VI.21 : Evolution de l'énergie en fonction de k

(a) $N=10$, (b) $N=20$, (c) $N=30$.

Le système semble donc capable quelque soit N , lorsque k varie, de retrouver automatiquement la dynamique des vecteurs vitesse. Le résultat obtenu est donc caractéristique de la séquence d'images traitée et non de la configuration (k, N) retenue. Il résulte de ces tests que l'introduction de la quantification ne perturbe pas le fonctionnement du réseau. Il est donc possible d'automatiser le calcul de k . On fixe N et on fait évoluer k jusqu'à trouver la valeur minimale de la courbe. Cet ajustement automatique permet de s'affranchir d'une série manuelle d'essais pour trouver la valeur convenable. Toutefois, la valeur k trouvée reste une fonction du paramètre α de régularisation.

Nous remarquons comme précédemment que le "bon" minimum local est atteint pour une valeur de N de plus en plus élevée. En effet, plus le nombre de neurones

augmente, plus la solution s'affine et l'énergie finale diminue. Le pas de quantification est de plus en plus petit et dans le voisinage du minimum local, la recherche de la solution est très sensible à de petites variations des paramètres utilisés.

VI.9.4 Comparaison avec d'autres approches

Comme cela a été montré précédemment, il est possible de faire évoluer les solutions obtenues par le réseau en fonction du mode de parcours retenu. Partant de cette constatation, il paraît intéressant d'évaluer les avantages que présentent les différents modes de parcours vis à vis du mode de parcours direct. Nous testons cette approche dans le cas particulier des images bruitées. Pour ce faire, nous comparons les résultats obtenus par notre réseau en fonction du type de parcours retenu lorsque le rapport signal sur bruit se dégrade dans les images. Les résultats obtenus sont par ailleurs comparés avec ceux obtenus par l'algorithme classique de Horn et Schunck [HOR - 81]. Les critères de comparaison retenus sont les suivants :

- la qualité des champs de vecteurs générés;
- la précision en amplitude et en phase des vecteurs par rapport à la solution théorique.

Afin de mener à bien cette comparaison, nous avons à chaque fois utilisé les mêmes images de synthèse et les mêmes valeurs des paramètres à ajuster. La *figure VI.22* donne le résultat obtenu pour les composantes ω_x et ω_y dans le cas de la translation pure de valeurs ($\omega_x = 1$, $\omega_y = 1$) du carré noir sur fond blanc par la méthode classique de Horn et Schunck. Nous pouvons alors le comparer avec notre approche lorsque $N = 200$. Nous constatons que les deux champs de vecteurs vitesse obtenus sont très proches l'un de l'autre et ne diffèrent que par quelques valeurs d'amplitude différente. La régularité de la solution, pour le cas de notre système, augmente avec le nombre de neurones alloué au réseau comme cela est prévisible.

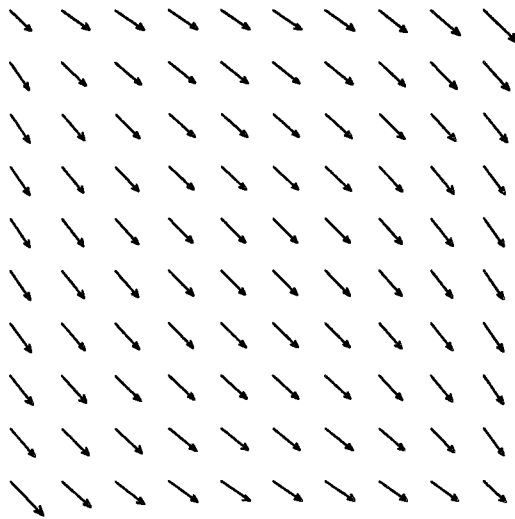


Figure VI.22 : *Champ de vecteurs vitesse obtenu par la méthode de Horn et Schunck.*

Le carré est ensuite bruité par un bruit additif de distribution normale, pour différentes valeurs de rapport signal sur bruit (S/B). Une série d'expérimentations est ainsi réalisée pour les différentes valeurs de variance suivantes: 3, 6, 10, 15, 20, 30 et 40. L'évaluation des résultats se fait en terme d'erreurs moyennes d'amplitude et de direction des vecteurs vitesses, calculées sur toute la surface du carré noir.

Nous comparons la carte des vecteurs déplacements obtenus dans le cas de la méthode de Horn et Schunck et dans le cas de notre méthode en utilisant le même paramètre de régularisation avec le champ de vitesse théorique. Deux calculs d'erreurs sont alors effectués en chaque point: l'erreur d'amplitude obtenue en calculant pour chaque vecteur le pourcentage d'erreur de la norme du vecteur vitesse estimé par rapport à celle du vecteur vitesse théorique et l'erreur de direction donnée par l'angle (en degré) qui sépare la direction du vecteur vitesse calculé de celle du vecteur théorique.

Nous constatons que quelque soit la méthode utilisée, l'erreur moyenne de direction obtenue est faible lorsque les images sont non bruitées (de l'ordre de 4° dans Horn et Schunck et de 6° lorsque $N = 100$ dans le cas de notre réseau). Les erreurs d'amplitude et de direction obtenues sont plus fortes lorsqu'elles sont

calculées par le réseau neuronal avec $N = 20$. Ces erreurs diminuent notablement quand elles sont calculées pour $N = 100$. L'erreur d'amplitude est beaucoup plus sensible aux valeurs des paramètres. Lorsque les images sont bruitées, elle varie de 14% jusqu'à 40% dans le cas de Horn et Schunck et de 17% jusqu'à 30% dans notre réseau lorsque $N = 100$. Ces valeurs sont sensiblement constantes au fur et à mesure qu'augmente le bruit d'une manière non brutale. Nous remarquons que ces erreurs d'amplitude augmentent rapidement et qu'elles sont surtout sensibles au paramètre N . En effet, plus on alloue au réseau un nombre important de neurones, plus leurs valeurs deviennent de plus en plus faibles.

Le tableau des valeurs d'erreurs d'amplitude et de direction obtenues par le système lorsque est $N = 20$ et $N = 100$, ainsi que par la méthode de Horn et Schunck, est présenté ci-dessous pour le cas des images bruitées et non bruitées.

erreur		Image Sans bruit $\sigma^2 = 0$	Image bruitée						
			$\sigma^2 = 3$	$\sigma^2 = 6$	$\sigma^2 = 10$	$\sigma^2 = 15$	$\sigma^2 = 20$	$\sigma^2 = 30$	$\sigma^2 = 40$
Horn et Schunck	d'amplitude (%)	14,1819	14,2948	14,7362	16,7134	18,2498	28,0293	26,2292	39,1994
	de direction (degré)	4,4381	4,5337	4,6211	4,9713	5,6398	7,4466	7,8214	12,0515
Réseau neuronal $N=100$	d'amplitude (%)	17,7604	16,7401	16,4321	17,0890	17,5804	22,5895	22,0848	30,5359
	de direction (degré)	6,0243	5,7302	5,6407	5,7649	5,8115	6,7928	6,7611	9,2109
Réseau neuronal $N=20$	d'amplitude (%)	35,0264	30,1601	29,7207	32,0053	32,1939	32,2118	32,5386	38,8353
	de direction (degré)	12,4727	10,4624	10,4975	11,5399	10,7356	12,0277	11,3553	13,5560

Tableau VI.3 : Erreur moyenne d'amplitude et de direction des vecteurs vitesse, calculée sur toute la surface du carré noir pour une translation par la méthode de Horn et Schunck et par le réseau neuronal lorsque $N = 100$ et $N = 20$.

Les valeurs ainsi obtenues sont sensiblement très proches de celles obtenues par la méthode de Horn et Schunck (tableau VI.3). Comme on le constate, les erreurs moyennes de direction et d'amplitude restent d'autant plus fortes que le nombre de neurones est faible (cas: $N = 20$). Ces erreurs sont divisées pratiquement par 2 lorsque le réseau fonctionne avec $N = 100$.

Il convient cependant de noter que, pour une valeur de N donnée, l'erreur d'amplitude reste à peu près constante jusqu'à ce que le bruit devienne suffisamment important pour être prépondérant vis-à-vis de l'erreur due à la quantification. C'est à dire au fait que le nombre de neurones n'est pas infini. Par exemple dans le cas où $N = 100$, l'erreur d'amplitude est identique tant que la variance du bruit est inférieure à 15. La valeur de cette erreur correspondant effectivement à l'erreur d'amplitude lorsque $\sigma^2 = 15$ dans le cas de Horn et Schunck. Au delà de $\sigma^2 = 15$, cette erreur augmente mais moins vite que dans le cas de Horn et Schunck.

On peut dire que tant que N est faible, l'erreur est due à la faiblesse de sa valeur qui ne permet pas de prendre en compte la solution. On voit apparaître d'autant plus vite l'influence du bruit dans l'image que le nombre de neurones est élevé.

Des tests réalisés sur les différents types de parcours, exponentiel ($N = 5, 20, 100$), logarithmique ($N = 10, 40, 65, 80, 90, 95, 98, 100$) et linéaire ($N = 10, 20, \dots, 100$) sur les mêmes images bruitées montrent que l'erreur moyenne d'amplitude de direction obtenue reste faible, et très peu sensible aux valeurs de N et de bruits introduits. Cette erreur varie de 5° à 9° sur la surface du carré étudiée.

Configuration	Erreur	$\sigma^2 = 0$	$\sigma^2 = 6$	$\sigma^2 = 15$	$\sigma^2 = 40$
Exponentielle	d'amplitude (%)	17,7604	16,4321	18,1825	30,6017
	de direction (degré)	6,0243	5,6407	6,2398	9,2482
Logarithmique	d'amplitude (%)	16,6627	15,6493	16,7593	29,8642
	de direction (degré)	5,6999	5,3952	5,6464	8,9742
Linéaire	d'amplitude (%)	17,3839	16,0582	17,3884	29,9679
	de direction (degré)	5,9640	5,4799	5,7130	8,9977

Tableau VI.4 : erreur moyenne d'amplitude et de direction des vecteurs vitesse, calculée sur toute la surface du carré noir pour une translation par le réseau neuronal pour les trois types de parcours.

En ce qui concerne les champs de déplacement, nous avons obtenu dans le cas de notre méthode, pour le même bruit introduit, un champ de vecteurs vitesse plus régulier que celui de Horn et Schunck. Les figures VI.23.a et VI.23.b donnent les champs de vitesse générés par l'approche de Horn et Schunck après 85 itérations et par le réseau neuronal après 44 itérations sur les images bruitées de même variance ($\sigma^2 = 40$). L'introduction du bruit ne modifie donc pas fondamentalement l'évolution de l'énergie même si le minimum local peut changer de place. Le déplacement du minimum local peut être quelquefois très important surtout lorsque le nombre de neurones est faible.

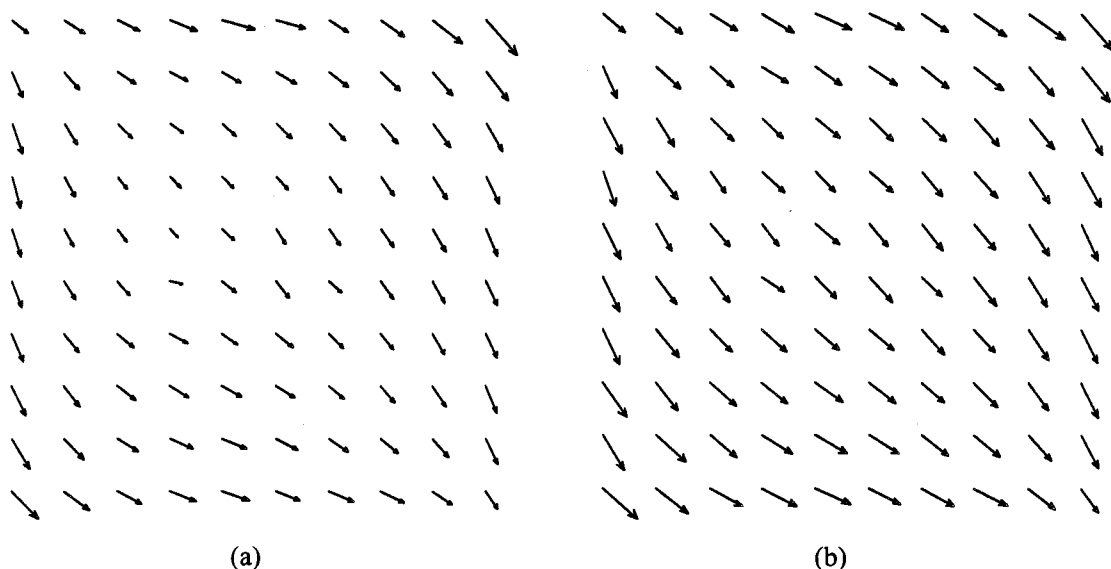


Figure VI.23 : *Champs de vitesse apparents obtenus sur des images bruitées de variance 40*
 (a) *par la méthode de Horn et Schunck*
 (b) *par le réseau neuronal.*

VI.10 CONCLUSION

Les résultats que nous avons observé sur les séquences d'images de synthèse ou réelles ont permis de valider et de montrer l'intérêt de la technique neuronale mise en oeuvre dans le cadre de l'estimation du mouvement. En outre, nous avons tenu à souligner l'importance d'une bonne connaissance, à la fois du problème que l'on désire inculquer au réseau, et des capacités du réseau lui-même.

Par ailleurs, l'utilisation des différents modes de parcours permettent de diminuer la complexité du problème d'estimation du mouvement sur deux plans:

- allégement des coûts calculatoires;
- moindre sensibilité aux minima locaux.

Enfin, d'un point de vue implantation du réseau sous forme d'un circuit VLSI, il est à noter que la structure proposée est nettement moins coûteuse que celle basée sur le modèle de classification proposé par Zhou et Chellappa. Pour une image de taille (LxM), le réseau de Zhou nécessite un nombre de neurones égal à $(2N+1)^2 \times (LxM)$ pour estimer les vecteurs vitesse dont les composantes ont une dynamique de -N à N par pas entier. Dans notre cas, avec la même configuration, le réseau ne nécessite que $(4xNxLxM)$ neurones. Le rapport est donc de $\frac{(2N+1)^2}{4N} = N + 1 + \frac{1}{4N}$ en notre faveur.

A titre d'exemple, le circuit réalisé par Lee [LEE - 93] et basé sur le concept de classification de Zhou, permet d'estimer des vecteurs avec une valeur de $N = 5$. Une réalisation à partir du réseau que nous proposons nécessiterait 6 fois moins de neurones, soit pour un circuit traitant une image de taille (512 x 512) pixels, environ 26 millions de neurones de moins que dans le circuit réalisé par Lee.

Un autre avantage concerne les problèmes qui peuvent surgir dans le cas d'un mauvais fonctionnement des neurones. En effet, dans la mesure où la méthode décrite au chapitre V s'apparente à une classification; si un neurone ne fonctionne plus, on perd alors complètement la valeur de vitesse correspondante. Ceci est dû au traitement de l'information qui n'est pas largement distribuée au travers de toute la structure. L'approche présentée dans ce chapitre est plus robuste au défaut architectural, puisqu'elle permet au système d'avoir différentes combinaisons pour réaliser la même solution. En effet, si l'on a N neurones par composante de vitesse et que la composante a une valeur de k, le nombre de possibilité pour obtenir k avec les N neurones est C_N^k . Si p neurones avec $(N-p > k)$ ne fonctionnent plus, il reste au système C_{N-p}^k combinaisons possibles pour réaliser la solution.

CHAPITRE VII

CONCLUSION GENERALE

CHAPITRE VII

CONCLUSION GENERALE

VII.1 EN CONCLUSION

Dans ce mémoire, nous avons présenté une étude portant sur l'utilisation des réseaux de neurones pour la minimisation de fonctionnelles avec comme application l'estimation du mouvement dans le cadre de l'analyse de scènes dynamiques. Le réseau de neurones proposé est de type Hopfield. Il est complètement connecté, et les connexions sont symétriques. Les neurones sont discrets et répondent selon une fonction d'activation.

Dans un premier temps, nous avons analysé une approche proposée par Zhou et Chellappa [ZHO - 88] qui permet de calculer le "flux optique" à l'aide d'un réseau de Hopfield. Cette étude nous a conduit à développer d'abord une méthode dérivée de celle-ci, ensuite à proposer un système neuronal qui permet de s'affranchir des difficultés rencontrées dans les précédentes approches. Nous avons donc proposé une méthodologie qui s'articule sur deux points:

- Codage du problème d'estimation des vecteurs vitesse sous forme d'un réseau de Hopfield et établissement de "l'énergie du problème" dont le minimum correspond à une solution optimale de ce problème;
- Mise en oeuvre d'une approche neuronale pour la minimisation de cette énergie.

Des tests sur le comportement du réseau neuronal ont été effectués. Des résultats sur des images de synthèse et réelles et des comparaisons avec des approches classiques sont présentés afin de montrer les avantages et les limites de notre approche.

Le réseau de neurones développé donne des résultats de qualité équivalente à ceux obtenus par les méthodes déjà existantes. Il présente cependant un visage séduisant pour différentes raisons. Parmi ces dernières, son architecture qui est moins complexe et moins consommatrice en neurones que celle proposée par Zhou et Chellappa. Son implémentation sur un circuit numérique ou sur un circuit mixte analogique-numérique, bien que non étudiée dans cette thèse, est donc tout à fait réalisable dans la mesure où une architecture basée sur le concept de classification de Zhou et Chellappa a déjà été réalisée. Cette architecture présente une plus grande robustesse vis-à-vis d'un éventuel dysfonctionnement des neurones qui la constituent.

Par ailleurs, le concept de construction de la grandeur à estimer fait que le modèle proposé n'est pas tributaire de l'approche du mouvement retenu, ni même de l'application considérée. En effet, il semble évident que cette méthode peut être généralisée à tous les types de problèmes de minimisation de fonctionnelles dans lesquelles la variable à estimer apparaît de manière explicite. C'est par exemple le cas de la restauration ou de la stéréovision. Il suffit pour cela de modifier la fonctionnelle d'énergie à minimiser et de calculer les deux paramètres importants du réseau qui sont les poids synaptiques et les entrées externes par identification de celle-ci avec l'énergie du réseau.

Cet aspect de la méthode utilisée pour construire le réseau fait qu'il est tout à fait envisageable dans le contexte de l'analyse du mouvement de prendre en compte avec

le réseau, le phénomène des discontinuités de mouvements sous forme d'un processus linéique par exemple, afin d'assurer une qualité et une précision d'estimation meilleures que celles qui sont pour le moment accessibles.

La généralisation au contexte tridimensionnel ne pose pas de difficultés supplémentaires, toujours en raison de la méthodologie adoptée.

Une autre évolution possible concerne cette fois non plus la fonctionnelle implantée mais le modèle neuronal utilisé. Il semble ainsi envisageable d'utiliser un modèle de type Hopfield "étendu" dans lequel la fonction d'énergie est donnée par une expression d'ordre supérieur des sorties des neurones. En effet, cette fonction d'énergie converge plus facilement vers un minimum global et évite de tomber dans des minima locaux. Ce type de modèle étendu constitue alors un outil efficace pour résoudre des problèmes d'optimisation plus complexes.

Le travail relaté dans ce mémoire, ainsi que ses prolongements, montre donc combien cette approche d'utilisation de réseaux neuronaux pour la minimisation de fonctionnelles constitue une alternative intéressante aux problèmes d'optimisation en traitement d'images.

Annexe A

Principaux réseaux de neurones

Réseau	Année	Inventeur / Concepteur	Principales Applications	Avantages	Désavantages
Adaline / Madaline	1960	B. Widrow	Filtrage du signal adaptatif	Rapide, simple à implémenter et réalisation VLSI.	Relation linéaire entre entrée et sortie. Séparabilité linéaire.
Adaptive Resonance Theory (ART)	1983	G. Carpenter & S. Grossberg	Reconnaissance de formes	Apte à apprendre de nouveaux modèles et former de nouvelles catégories en plus de celles déjà apprises.	Inadéquate si modèle trop bruité.
Brain-state- In-a-Box (BSB)	1977	J. Anderson	Mémoire auto-associative	Amélioration sensible des performances / réseau d'Hopfield.	Analyse partielle en termes de performance et de potentiel.
Bidirectional Associative Memory (BAM)	1987	B. Kosko	Mémoire hétéro-associative	Architecture et apprentissage simples. Stabilité prouvée.	Faible capacité mémoire et faible exactitude de rappel.

Machine de Boltzmann	1984	G. Hinton, T. Sejnowski, D. Ackley	Reconnaissance de formes (images, sonar, radar), Optimisation.	Représentation optimale des caractéristiques du modèle en atteignant le minimum de la fonction d'énergie.	Temps d'apprentissage très long pour la machine de Boltzmann, réduit dans le cas de la machine de Cauchy.
et Machine de Cauchy	1986	H. Szu			

Perceptron / Perceptron multicouches	1958 1974-1986	F. Rosenblatt P. Werbos, D. Parker, D. Rumelhart	filtrage, reconnaissance de formes, classification, compression de données, segmentation, ...	Réalise de nombreuses opérations. Classifieur très étudié et employé avec beaucoup de succès.	Temps d'apprentissage important.
Hopfield / Tank	1982-1984	J. Hopfield	Mémoire associative, Optimisation.	Concept simple, stabilité dynamique prouvée, facile à implémenter en VLSI.	Faible capacité mémoire, Possibilité d'états poubelles.
Learning Vector Quantization	1981	T. Kohonen	Mémoires auto-associatives, Compression de données	Auto-organisation de représentations de vecteurs.	Apprentissage lent
Cognitron/ Néocognitron	1975-1982	K. Fukushima	Reconnaissance de caractères manuscrits et autres figures.	Insensible aux phénomènes d'échelles, de rotation et de translation.	Nécessite plusieurs unités et couches de traitement, Structure complexe.
Radial Basis Function	1987-1988	Multiplés chercheurs	Classification.	Emploi de gaussienne sur couche cachée pour meilleure classification.	Méconnu.
Boundary Contour System (B.C.S)	1985	S. Grossberg & E. Mingola	Traitement des images bas niveau.	Approche excellente pour la segmentation.	Architecture à couches complexes.

Annexe B

Propriétés de convergence du réseau de Hopfield

La fonction d'activation d'un neurone proposé par McCulloch et Pitts est donnée par:

$$V_i = f(U_i) = \begin{cases} 1 & \text{si } U_i \geq 0 \\ 0 & \text{ailleurs} \end{cases} \quad (1)$$

Hopfield et Thank [HOP - 85] ont utilisé une fonction Sigmoidé. Celle-ci est une fonction différentiable, continue et non décroissante:

$$V_i = g(U_i) = \frac{1}{2} (1 + \tanh(\lambda U_i)) \quad (2)$$

Le paramètre λ représente un gain constant qui change la raideur de la courbe de la sigmoïde. La variation de U_i est donnée par l'équation dynamique du i ème neurone. Cette équation est définie dans le modèle de Hopfield comme suit:

$$\frac{dU_i}{dt} = -\frac{U_i}{\tau} - \frac{\partial E}{\partial V_i} \quad (3)$$

Le terme $-\frac{U_i}{\tau}$ perturbe la convergence du réseau car il augmente l'énergie E sous certaines conditions. Bien que certains chercheurs attribuent une grande valeur à τ afin d'éliminer les effets indésirables, il est théoriquement suggéré d'éliminer ce terme de l'équation (4).

Afin de garantir la convergence vers un minimum local, la matrice des poids synaptiques T est symétrique avec les éléments diagonaux nuls. La fonction d'énergie de Lyapunov E est donnée par:

$$\begin{aligned} E(t) &= -\frac{1}{2} \sum_i \sum_j T_{i,j} V_i(t) V_j(t) - \sum_i I_i V_i(t) + \frac{1}{\lambda} \sum_i \frac{1}{R_i} \int_0^{V_i(t)} g^{-1}(V(t)) dV(t) \\ &= E_D(t) + \frac{1}{\lambda} \sum_i \frac{1}{R_i} \int_0^{V_i(t)} g^{-1}(V(t)) dV(t) \end{aligned} \quad (4)$$

où $E_D(t) = E_D(V_1(t), V_2(t), \dots, V_n(t))$. L'équation dynamique pour le i ème neurone est donnée par:

$$\frac{dU_i(t)}{dt} = \sum_j T_{i,j} V_j(t) + I_i - \frac{U_i}{\tau} = \frac{\partial E_D(t)}{\partial V_i} - \frac{U_i(t)}{\tau} \quad (5)$$

où τ est une constante de temps. Pour un problème donné, la fonction d'énergie E est construite par l'introduction de contraintes supplémentaires. On notera que cette fonction n'est pas unique pour le problème. Elle détermine les poids synaptiques T_{ij} entre les neurones et l'entrée externe I_i dans (2). Hopfield a démontré que cette fonction basée sur l'équation (2) est une fonction de Lyapunov. En effet, pour le dernier terme:

$$\frac{1}{\lambda} \sum_i \frac{1}{R_i} \int_0^{V_i(t)} g^{-1}(V(t)) dV(t)$$

de l'équation (1), la cartographie devient très difficile et la solution "minimum local" est déplacée par le paramètre λ . Il est supposé qu'un gain très élevé est utilisé pour simplifier cette cartographie. Dans le cas où $\lambda \rightarrow \infty$, La fonction de Lyapunov devient la fonction E_D en dehors du terme de (1), et l'équation dynamique du ième neurone est donnée par $\frac{dU_i(t)}{dt} = -\frac{\partial E_D(t)}{\partial V_i(t)}$. Quand la fonction E_D est utilisée, le terme $-\frac{U_i(t)}{\tau}$ doit être exclu de l'équation dynamique. Si le gain n'est pas infini, la fonction de Lyapunov doit inclure le terme:

$$\frac{1}{\lambda} \sum_i \frac{1}{R_i} \int_0^{v_i(t)} g^{-1}(V(t)) dV(t) \quad (6)$$

dans (4).

Le théorème 1 montre que le terme $-\frac{U_i(t)}{\tau}$ augmente l'énergie E sous certaines conditions. En d'autre terme, E_D n'est pas une fonction de Lyapunov du réseau.

Théorème 1: Le terme $-\frac{U_i(t)}{\tau}$ dans l'équation dynamique du ième neurone augmente l'énergie $E_D(t)$ lorsque:

$$\left| \sum_i \frac{dV_i(t)}{dt} \frac{U_i(t)}{\tau} \right| > \left| \sum_i \left(\frac{dV_i(t)}{dU_i(t)} \right) \left(\frac{dU_i(t)}{dt} \right)^2 \right| \quad (7)$$

et si l'une des conditions ($U_i(t) > 0$ et $\frac{dV_i(t)}{dt} < 0$) ou ($U_i(t) < 0$ et $\frac{dV_i(t)}{dt} > 0$) est satisfaite.

Démonstration:

Soit la dérivée de l'énergie E_D au temps t:

$$\begin{aligned} \frac{dE_D(t)}{dt} &= \sum_i \frac{dV_i(t)}{dt} \frac{\partial E_D(t)}{\partial V_i(t)} \\ &= \sum_i \frac{dV_i(t)}{dt} \left(-\frac{U_i(t)}{\tau} - \frac{dU_i(t)}{dt} \right) \end{aligned} \quad (8)$$

où $\frac{\partial E_D(t)}{\partial V_i(t)}$ est remplacé par $\left(-\frac{U_i(t)}{\tau} - \frac{dU_i(t)}{dt} \right)$

$$\begin{aligned} &= -\sum_i \frac{dV_i(t)}{dt} \frac{U_i(t)}{\tau} - \sum_i \frac{dV_i(t)}{dt} \frac{dU_i(t)}{dt} \\ &= -\sum_i \frac{dV_i(t)}{dt} \frac{U_i(t)}{\tau} - \sum_i \left(\frac{dU_i(t)}{dt} \frac{dV_i(t)}{dU_i(t)} \right) \left(\frac{dU_i(t)}{dt} \right) \\ &= -\sum_i \frac{dV_i(t)}{dt} \frac{U_i(t)}{\tau} - \sum_i \left(\frac{dV_i(t)}{dU_i(t)} \right) \left(\frac{dU_i(t)}{dt} \right)^2 \end{aligned} \quad (9)$$

Le premier terme $-\sum_i \frac{dV_i(t)}{dt} \frac{U_i(t)}{\tau}$ peut être positif, négatif ou nul. Le second terme

$-\sum_i \left(\frac{dV_i(t)}{dU_i(t)} \right) \left(\frac{dU_i(t)}{dt} \right)^2$ est toujours négatif ou nul car la sortie $V_i(t) = f(U_i(t))$ est

une fonction non décroissante. La condition suivante est vraie:

$$= -\sum_i \frac{dV_i(t)}{dt} \frac{U_i(t)}{\tau} - \sum_i \left(\frac{dV_i(t)}{dU_i(t)} \right) \left(\frac{dU_i(t)}{dt} \right)^2 > 0 \quad (10)$$

lorsque

$$\left| \sum_i \frac{dV_i(t)}{dt} \frac{U_i(t)}{\tau} \right| > \left| \sum_i \left(\frac{dV_i(t)}{dU_i(t)} \right) \left(\frac{dU_i(t)}{dt} \right)^2 \right|$$

et lorsque l'une des conditions suivantes est satisfaite: ($U_i(t) > 0$ et $\frac{dV_i(t)}{dt} < 0$) ou ($U_i(t) < 0$ et $\frac{dV_i(t)}{dt} > 0$). Sous les mêmes conditions, la dérivée de $E_D(t)$ doit être positive: $\frac{dE_D(t)}{dt} > 0$. Par conséquent, le terme $-\frac{U_i(t)}{\tau}$ augmente l'énergie sous les mêmes conditions, qui signifie que $E_D(t)$ n'est pas une fonction de Lyapunov du réseau.

Le lemme suivant est introduit pour prouver que le réseau converge vers un minimum local ou vers une solution toute proche de celui-ci.

Lemme 1: $\frac{dE}{dt} \leq 0$ est satisfaite sous deux conditions, $\frac{dU_i}{dt} = -\frac{\partial E}{\partial V_i}$ et $V_i = f(U_i)$

où $f(U_i)$ est une fonction non décroissante.

Preuve:

$$\frac{dE}{dt} = \sum_i \frac{dU_i}{dt} \frac{dV_i}{dU_i} \frac{\partial E}{\partial V_i} = -\sum_i \left(\frac{dU_i}{dt} \right)^2 \frac{dV_i}{dU_i} \quad (11)$$

où $\frac{\partial E}{\partial V_i}$ est remplacé par $\frac{dU_i}{dt}$ (condition 1) ≤ 0 quand $\frac{dV_i}{dU_i} > 0$ (condition 2).

Le lemme 1 montre que l'état du réseau recherche finalement un état d'équilibre.

Annexe C

Publications dans le cadre de la thèse

- [YAC-94] **A. Yacoubi, L. Lelaurin, V. Devlaminck, J.P. Dubus.** "Contribution des réseaux neuromimétiques à l'estimation du mouvement des cavités cardiaques". *Actes du Septième Forum Jeunes Chercheurs en Biologique et Médical*, Lyon, France, pp. 6 - 7, 26 - 27 Mai 1994.
- [YAC-94] **A. Yacoubi, L. Lelaurin, V. Devlaminck, J.P. Dubus.** "Motion estimation using a neural network". *Proceedings of the International Conference on Electrical & Electronic Engineering (ICEEE-94)*, University of Sciences and Technology of Oran, Algeria, Vol. 1-2, pp. 306 - 312, October 1 - 3, 1994.
- [YAC-95a] **A. Yacoubi, L. Lelaurin, V. Devlaminck, J. P. Dubus.** "Contribution des réseaux neuromimétiques à l'estimation du mouvement des cavités cardiaques". *Innovation et Technologie en Biologie et Médecine (ITBM)*, France, Vol. 16, n° 4, pp. 441-453, 1995.
- [YAC-95b] **A. Yacoubi, L. Lelaurin, V. Devlaminck, J.P. Dubus.** "Optimization by neural networks: Application for optical flow estimation in echocardiographic images". *Proceedings of the 17th Annual International Conference IEEE Engineering in Medicine and Biology Society (EMBS'95)*, Montréal, Québec, Canada, September 20 - 23, 1995.
- [LEL-94a] **L. Lelaurin, A. Yacoubi, V. Devlaminck, J.P. Dubus.** "Contribution d'un réseau de neurones à l'estimation des paramètres du mouvement dans une séquence d'images". *Colloque sur le neuromimétisme, AIDRI*, Lyon, France, pp. 271 - 275, 20 - 23 Juin 1994.

- [LEL-94b] **L. Lelaurin, A. Yacoubi, V. Devlaminck, J.P. Dubus.** "A neural network for motion estimation". *Proceedings of the 2nd Japan - France Congress of Mechatronics (International)*, Takamatsu, Kagawa, Japan, Vol. 2, pp. 549 - 552, 01 - 03 Nov. 1994.
- [LEL-95] **L. Lelaurin, A. Yacoubi, V. Devlaminck, J.P. Dubus.** "Motion estimation by using a neuromimetic approach". *Proceedings of the IEEE International Conference on Neural Networks (ICNN'95)*, Perth, Western Australia, Vol. 6, pp. 2917-2921, 27 Nov - 01 Dec. 1995.
- [DEV-95a] **V. Devlaminck, L. Lelaurin, A. Yacoubi, J.P. Dubus.** "Estimation of discontinuous optical flow using a neural network for echocardiographic images". *Medical Imaging 1995, Proceedings of the SPIE - Image Processing*, San Diego, California, USA, pp. 210-211, (paper 2434-51), 26 Feb - 02 Mars 1995.
- [DEV-95b] **V. Devlaminck, L. Lelaurin, A. Yacoubi, J.P. Dubus.** "Stress measure and non rigide motion estimation from density images using a neural network". *Accepté. IEEE Computation Engineering in Systems Applications (CESA'96)* , IMACS, Lille, France, 9-12 July 1996.

REFERENCES BIBLIOGRAPHIQUES

Chapitre I

- [ABD-94] H. Abdi. "Les réseaux de neurones". *Ed. Presses Universitaires de Grenoble*, 264 pages, 1994.
- [ACK-85] D. H. Ackley, G. E. Hinton and T. J. A. Sejnowski. "A Learning algorithm for Boltzmann machines". *Cognitive Science*, Vol. 9, pp. 147-169, 1985.
- [AMA-71] S. Amari. "Characteristics of randomly connected threshold-element networks and network systems". *Proceedings IEEE*, Vol. 59, pp. 35-47, Jan.1971.
- [AMA-72] S. Amari. "Learning patterns and pattern sequences by self-organizing nets of threshold elements". *IEEE Transaction on Computer*, Vol. C-21, pp. 1197-1206, Nov.1972.
- [AMA-77] S. Amari. "Neural theory of association and concept formation". *Biological Cybernetics*, Vol. 26, pp. 175-185, 1977.
- [AMI-89] D. J. Amit. "Modeling brain function, the world of attractor dynamics". *Cambridge University Press*, 1989.
- [AND-88] J. A. Anderson & Rosenfeld. "Neurocomputing: Foundations of research". *MIT Press*, Cambridge MA, 1988.
- [BAR-89] H. B. Barlow. "Unsupervised learning". *Neural Computation*, Vol. 1, pp. 295-311, 1989.
- [BAR-85] A. G. Barto. "Learning by statistical cooperation of self-interested neuron-like computing elements", *Human Neurobiology*, Vol. 4, pp. 229-256, 1985.
- [BUR-88] Y. Burnod., "An adaptive neural network: the cerebral cortex". *Collection Biologie théorique*, Ed. Masson, 1988.

- [CAB-92] F. Cabestaing. "Détection des contours en mouvement dans une séquence d'images. Conception et réalisation d'un processeur câblé temps-réel". *Thèse de Doctorat de Lille I*, 1992.
- [CHO-91] J. Choi, B. J. Sheu, and S. M. Gowda. "Analog VLSI neural network implementations of hardware annealing and winner-take-all functions". *Presented at the 34 th Midwest Symp. Circuits and Syst. (invited paper)*, Monterey, CA, May 1991.
- [DAV-90] E. Davalo et P. Naim., "Des réseaux de neurones". Ed. Eyrolles, 1990.
- [DUD-73] R. O. Duda et P. E. Hart., "Pattern classification and scene analysis". *Wiley, New York*, 1973.
- [FRE-92] J. A. Freeman & D. M. Skapura. "Neural networks: Algorithms, applications and programming techniques". *Addison-Wesley*, 1992.
- [FUK-75] K. Fukushima. "Cognitron: a self-organizing multilayered neural network". *Biological Cybernetics*, n° 20, pp. 121-136, 1975.
- [FUK-84] K. Fukushima. "A hierarchical neural network model for associative memory". *Biological Cybernetics*, n° 50, pp. 105-113, 1984.
- [GRO-88] S. Grossberg . "Neural network and natural intelligence". *MIT Press*, 1988.
- [HEB-49] D. O. Hebb. " The organization of behaviour". *Wiley, New York*, 1949.
- [HEC-87] R. Hecht-Nielsen. "Counterpropagation networks". *Applied Optics* 26 (23), pp. 4979-4984, 1987.
- [HEC-90] R. Hecht-Nielsen. "Neurocomputing". *HNC, Inc. and University of California, San Diego, Addison-Wesley Publishing Compagny*, 1990.
- [HER-94] J. Herault et C. Jutten., "Les réseaux neuronaux et traitement du signal". *INPG, Université Joseph Fourier de Grenoble, Ed. Hermes*, 320 pages, 1994.
- [HIN-81] G. E. Hinton, J. A. Anderson (eds.). "Parallel models of associative memory". *Hillsdale, Erlbaum*, 1981.
- [HOP-82] J.J. Hopfield. "Neural networks and physical systems with emergent collective computational abilities". *Proc. Natl. Acad. Sci. USA*, Vol. 79, pp. 2554-2558, 1982.
- [HOP-84] J.J. Hopfield. " Neurons with graded response have collective computational properties lie those of two-states neurons." *Proc. Natl. Acad. Sci. USA*, Vol. 81, pp. 3088-3092, 1984.

- [HOP-85] J.J. Hopfield., D.W. Tank. "Neural computation of decisions in optimisation problems". *Biological Cybernetic*, Vol. 52, pp. 141-152, 1985.
- [JOD-94a] J. F. Jodouin., "Les réseaux de neurones. Principes et définitions". Université Paris-Sud, Ed. Hermes, 128 pages, 1994.
- [JOD-94b] J. F. Jodouin., "Les réseaux de neuromimétiques". *Université Paris-Sud, Ed. Hermes*, 256 pages, 1994.
- [KAM-90] Y. Kamp, M. Hasler., "Réseaux de neurones récursifs pour mémoires associatives". *Ed. Presses polytechniques et Universitaires Romandes, Collection électricité*, 1990.
- [KER-92] P. Kerlizin et P. Réfregier., "Robustesse des réseaux neuronaux multicouches: analyse théorique et application à la compression d'images". *Neuro-Nîmes*, pp. 479-490, 1992.
- [KOH-84] T. Kohonen. "Self organization and associative memory". *Springer-Verlag*, Berlin, 1984.
- [KOH-88a] T. Kohonen. "The 'neural' phonetic typewriter", *IEEE Computer*, Vol. 21, n° 3, pp. 11-22, 1988.
- [KOH-88b] T. Kohonen. "Self-organization and associative memory". *Second Edition, Springer-Verlag*, Berlin, 1988.
- [KOH-89] T. Kohonen. "Self organisation and associative memory" *Springer-Verlag*, 1989.
- [KOH-90] T. Kohonen. "The self-organizing map", *Proceedings of the IEEE*, Vol. 78, n° 9, pp. 1464-1479, 1990.
- [KOS-87] B. Kosko. "Adaptive bidirectional associative memories", *Applied Optics*, Vol. 26, n° 23, pp. 4947-4960, 1987.
- [KOS-88] B. Kosko. "Bidirectional associative memories". *IEEE Trans. Systems, Man, and Cybernetics*, Vol. SMC-18, n° 1, pp. 42-60, 1988.
- [LEC-85] Y. Le Cun. "A learning scheme for asymmetric threshold network". *In Cognitiva 85, CESTA-AFCET Ed.*, Vol. 2, pp. 599-604, 1985.
- [LEC-87] Y. Le Cun. "Modèles connexionnistes de l'apprentissage". *Thèse de Doctorat de l'Université de Paris VI*, 1987.
- [LEM-88] Lemberg. "Les réseaux neuronaux artificiels". *Langages & Systèmes, Info. P.C.*, Vol. 71, 1988.
- [LIP-87] R. P. Lippmann. "An introduction to computing with neural nets". *IEEE ASSP magazine*, pp. 4-23, Avril 1987.

- [LIT-74] Little W.A. "The existence of persistent states in the brain". *Math. Biosci.*, 19, pp. 101-120, 1974.
- [McC-43] W.S. McCulloch and W. Pitts. "A logical calculus of the ideas immanent in nervous activity". *Bulletin of Mathematical Biophysics*, Vol. 5, pp. 115-133, 1943.
- [MEA-89] C. Mead. "Analog VLSI and neural systems". Addison-Wesley, 1989.
- [MIN-69] M. Minsky and S. Papert. "Perceptrons". *MIT Press, Cambridge*, 1969.
- [PAO-90] Y. Pao. "Adaptive pattern recognition and neural networks". *Case Western Reserve University Addison-Wesley Publishing Company, Inc.*, 1990.
- [PAR-94] D. Parrochia. "Quelques remarques historiques et philosophiques au sujet du concept de réseau de neurones de Mc Culloch et Pitts à Kohonen". *Colloque sur le neuromimétisme, AIDRI, Lyon*, pp.35-39, 20-23 Juin, 1994.
- [PER-90] J-C. Perez. "La révolution des ordinateurs neuronaux". *Ed. Hermes, Paris*, 1990.
- [PIN-87] F. J. Pineda. "Generalization of back-propagation to recurrent neural networks", *Physical Review Letters*, Vol. 59, n° 19, pp. 2229-2232, 1987.
- [ROS-57] F. Rosenblatt. "The perceptron: a perceiving and recognising automaton". *Project PARA, Cornell Aeronautical Lab. Report*, n° 85-460-1, 1957.
- [ROS-59] F. Rosenblatt. "Two theorems of statistical separability in the perceptron". *In Mechanisation of Thought Processes: Proceedings of a symposium Held at the National Physical Laboratory, HM Stationery Office, London*, pp. 421-456, November, 1959.
- [RUM-86] D. E. Rumelhart, J. L. McClelland and the PDP research group. "Parallel Distributed Processing - Exploration in the Microstructure of Cognition". Vol. I, II and III. *A Bradford book, MIT Press, Cambridge (MA)*, 1986.
- [RUM-86a] D. E. Rumelhart, G. Hinton, and R. J. Williams. "Learning internal representations by error propagation" *In Parallel Distributed Processing - Exploration in the Microstructure of Cognition, D. E. Rumelhart & J. C. McClelland (Eds), MIT Press Cambridge, USA*, Vol. 1, pp. 318-362, 1986.

- [RUM-86b] D. E. Rumelhart and D. A. McClelland. "PDP models and general issues in cognitive science". In *Parallel Distributed Processing*, D. E. Rumelhart & J. C. McClelland (Eds), MIT Press Cambridge, USA), Vol. I, 1986.
- [RUM-86c] D. E. Rumelhart, D. Zisper. "Feature discovery by competitive learning", in: *Parallel Distributed Processing*, MIT Press, Vol.I, pp. 151-193, 1986.
- [SAT-92] S. Satanarayama, Y. P. Tsividis, and H.P. Graf. "A reconfigurable VLSI neural network". *IEEE J. Solid-State Circuits*, Vol. 27, pp. 67-81, Jan. 1992.
- [SHE-92] B. J. Sheu, J. Choi, and C.-F. Chang. "An analog neural network processor for self-organizing mapping". *Tech. Dig. IEEE Int. Solid-State Circuits Conf.* San Francisco, CA, pp. 136-137, 266, Feb. 1992.
- [WAS-89] P. D. Wasserman. "Neural computing theory and practise". V.N.R, New York, 1989.
- [WEI-89] G. Weisbuch. "Dynamique des systèmes complexes. Une introduction aux réseaux d'automates". *Savoirs Actuels, InterEditions/Editions du CNRS*, 1989.
- [WID-59] B. Widrow. "Adaptive sampled-data system-a-statistical theory of adaptation". In *Institute of Radio Engineers, Western Electronic Show and Convention, Convention Record*, Vol. Part 4, pp. 74-85, 1959.
- [ZHO-88] Y. T. Zhou, R. Chellappa. "Computation of optical flow using a neural network". *IEEE Proc. Int. Conf. Neural Networks*, San Diego, CA, Vol. 2, pp. 71-78, 1988.
- [ZHO-92] Y. T. Zhou, R. Chellappa. "Artificial neural network for computer vision". *Springer-Verlag*, Berlin, 1992.

Chapitre II

- [ACK-85] D. Ackley, G. Hinton and T. Sejnowski. "A learning algorithm for Boltzmann machines". *Cognitive Science*, Vol. 9, pp. 147-169, 1985.
- [AND-88] J. R. Anderson. "A mean field computational model for PDP". *Proceedings of the 1988 Connectionist Summer School*, pp. 217-223, Morgan-Kaufman, 1988.
- [ANG-88] B. Angéniol, & al. "Self-organizing feature maps and the travelling salesman problem". *Neural Networks*, Vol. 4, n° 4, pp. 289-293, 1988.
- [ALE-89] F. Alexandre, Y. Burnod, F. Guyot, J. P. Haton. "La colonne corticale: une nouvelle unité de base pour les réseaux multicouches." - In *Actes CIRILLES - 3ème Université d'été, "Réseaux connexionnistes en informatique, méthodes et applications"*, 2 au 7 Juillet 1989, ed. par Dussauchoy A., Guivarch M., Laurini R. - pp. XII.1 à XII.16.
- [BAL-88] D. H. Ballard. "Cortical Connections and Parallel Processing: Structure and Function." - In "Vision, Brain, and Cooperative Computation." Ed. by M.A. Arbib and A.R. Hanson, a Bradford book, the MIT Press, Cambridge, 2nd pr., pp. 563-621, 1988.
- [CAR-89] G.A. Carpenter, S. Grossberg, and C. Meharian. "Invariant recognition of cluttered scenes by a self-organizing ART architecture: CORT-X Boundary Segmentation". *Neural Networks*, Vol. 2, pp. 169-181, 1989.
- [FOR-88] J. C. Fort. "Solving a combinatorial problem via self-organizing process: an application of the Kohonen algorithm to the travelling salesman problem. In *Proc. IEEE ICNN*, Vol. I, pp. 69-76, 1988.
- [GRA-88] D. H. Graf and W. R. Lalonde. "A neural controller for collision-free movement of general robot manipulators". In *Proc. IEEE ICNN*, Vol. I, pp. 77-84, 1989.
- [GRO-87] S. Grossberg. "The adaptive brain". Ed. Vol. I and II, Amsterdam: Elsevier/North-Holland, 1987.
- [HEB-49] D. O. Hebb. "The organization of behaviour". *Wiley, New York*, 1949.
- [HIN-84] G. Hinton, T. Sejnowsky, D. H. Ackley. "Boltzmann machines: constraint satisfaction networks that learn". *Technical report Carnegie Mellon Univ*, 1984.

- [KAM-90] Y. Kamp, M. Hasler., "Réseaux de neurones récurrents pour mémoires associatives". *Ed. Presses polytechniques et Universitaires Romandes, Collection électricité*, 1990.
- [KIR-83] S. Kirkpatrick, C. D. Gelatt et M. P. Vecchi. "Optimization by simulated annealing". *Science*, n° 220, pp. 671-680, 1983.
- [KOH-82] T. Kohonen. "Analysis of a simple self-organizing process". *Biological Cybernetics*, Vol. 44, pp. 135-140, 1982.
- [KOH-84] T. Kohonen. "Self organization and associative memory". *Springer-Verlag*, Berlin, 1984.
- [KOH-88] T. Kohonen. "The neural phonetic typewriter". *Computer*, Vol. 21, n° 3, pp. 11-22, 1988.
- [LEC-85] Y. Le Cun. "A learning scheme for asymmetric threshold network". *In cognitiva 85, CESTA-AFCET Ed.*, Vol. 2, pp. 599-604, 1985.
- [LEM-88] Lemberg. "Les réseaux neuronaux artificiels". *Langages & Systèmes, Info. P.C*, Vol. 71, 1988.
- [LIN-86a] R. Linsker. "From basic network principles to neural architectures, emergence of orientation column cells". *Proc. Natl. Acad. of Science, USA*, Vol. 83, pp. 8779-8783, 1986.
- [LIN-86b] R. Linsker. "From basic network principles to neural architectures, emergence of orientation-selective cells". *Proc. Natl. Acad. of Science, USA*, Vol. 83, pp. 8390-8394, 1986.
- [LIN-86c] R. Linsker. "From basic network principles to neural architectures, emergence of spatial-opponent cells". *Proc. Natl. Acad. of Science, USA*, Vol. 83, pp. 7508-7512, 1986.
- [LON-90] J. Loncelle. "Contribution des réseaux connexionnistes au traitement d'images bas-niveau". *Thèse de Doctorat, Université Paris XI*, 1990.
- [McC-43] W.S. McCulloch and W. Pitts. "A logical calculus of the ideas immanent in nervous activity". *Bulletin of Mathematical Biophysics*, Vol. 5, pp. 115-133, 1943.
- [MET-53] N. Metropolis, A. W. Rosenbluth, A. Teller. "Equations of state calculations by fast computing machines". *J. Chem. Phys*, 21, pp. 1087-1092, 1953.
- [MOU-92] M. Mougeot. "Modèles connexionnistes appliqués à la compression d'images et à l'auto-organisation du système visuel des mammifères". *Thèse de Doctorat en Sciences, Université Paris-Sud, Centre d'Orsay*, 1992.

- [NAS-88] N. M. Nasrabadi, Y. Feng. "Vector quantization of images based upon the Kohonen self-organizing features maps". *Int. Joint Conf. on Neural Networks*, Vol. 1, pp.101-108, 1988.
- [PAR-85] D. B. Parker. "Learning logic". *TR47, Sloan School of Management, MIT, Cambridge, Mass.*, 1985.
- [ROS-57] F. Rosenblatt. "The perceptron: a perceiving and recognising automaton". *Project PARA, Cornell Aeronautical Lab. Report*, n° 85-460-1, 1957.
- [ROS-62] F. Rosenblatt. "Principles of neurodynamics: perceptrons and theory of brain mechanisms". *Spartan Books, New-York*, 1962..
- [RUM-86] D. E. Rumelhart, G. Hinton, and R. J. Williams. "Learning internal representations by error propagation". In *Parallel Distributed Processing - Exploration in the Microstructure of Cognition. D. E. Rumelhart & J. C. McClelland (Eds), MIT Press Cambridge, USA*, Vol. 1, pp. 318-362, 1986.
- [SEJ-87] T. Sejnowsky, and C. Rosenberg. "Parallel networks that learn to pronounce english text". *Complex Systems* 1 (1): 145-68, 1987.
- [SPI-88] D. N. Spinelli. "A trace of Memory: An evolutionary perspective on the visual system." - In *"Vision, Brain, and Cooperative Computation."* Ed. by M.A. Arbib and A.R. Hanson, a Bradford book, the MIT Press, Cambridge, 2nd pr., pp. 165-182, 1988.
- [SZU-87] H. Szu. "Non-convex optimization by fast simulated annealing". *Proc. of the IEEE*, Vol. 75, n° 11, pp. 1538-1540, 1987.
- [THI-90] S. Thiria. "L'apprentissage supervisé dans les modèles connexionnistes". *Thèse de Doctorat de l'Université Paris V*, 1990.
- [TRY-89] V. Tryba, S. Metzen and K. Goser. "Designing basic integrated circuits by self-organizing feature map". In *Neuro-Nîmes, International Workshop Neural Networks and their applications*, Nanterre, France, pp. 225-235, 1989.
- [WEI-89] G. Weisbuch. "Dynamique des systèmes complexes. Une introduction aux réseaux d'automates". *Savoirs Actuels, InterEditions/Éditions du CNRS*, 1989.
- [ULT-90] A. Ultsch. "Kohonen's self-organizing feature maps for exploratory data analysis". *Int. Neural Network Conference of Paris*, Vol.1, pp. 305-308, 1990.
- [WID-60] B. Widrow, and M. E. Hoff. "Adaptative switching circuits". *IRE WESCON Conv. Record*, partie 4, pp. 96-104, Août 1960.

Chapitre III

- [ACK-85] D. Ackley, G. Hinton and T. Sejnowski. "A learning algorithm for Boltzmann machines". *Cognitive Science*, Vol. 9, pp. 147-169, 1985.
- [AMA-74] S. Amari. "A method of statistical neurodynamics". *Kybernetik*, Vol. 14, pp. 201-215, 1974.
- [AMA-92] S. C. Amari, D. Piraino, and Y. Takefuji. "Optimization neural networks for the segmentation of magnetic resonance images". *IEEE Transactions on Medical Imaging*, Vol. 11, n° 2, pp. 215-220, June 1992.
- [AMI-85] D. J. Amit, H. Gutfreund and H. Sompolinsky. "Spin-glass models of neural networks" *Physical Review A*, Vol. 32, n° 2, pp. 1007-1018, 1985.
- [AMI-87] D. J. Amit, H. Gutfreund and H. Sompolinsky. "Statistical mechanics of neural networks near saturation". *Annals of Physics*, Vol. 173, pp. 30-67, 1987.
- [BRU-87] J. Bruck, and J. W. Goodmann. "A generalized convergence theorem for neural networks and its application in combinatorial optimization". *Proc. of IEEE 1st Conf. on Neural Networks*, San Diego, pp. III649-III656, 1987.
- [CAS-92] D. Casaent, B. Telfer. "High capacity pattern recognition associative processors". *Neural Networks*, Vol. 5, pp. 687-698, 1992.
- [CIA-88] P. G. Ciarlet. "Introduction à l'analyse numérique matricielle et à l'optimisation". *Ed. Masson*, 1988.
- [COO-95] B. S. Cooper. "Higher order neural networks for combinatorial optimization - improving the scaling properties of the Hopfield network". *In IEEE International Conference on Neural Networks (ICNN'95)*, Perth, Western Australia, Vol. 5, pp. 1855-1860, 27 Nov - 01 Dec. 1995.
- [DAV-90] E. Davalo et P. Naim., "Des réseaux de neurones". *Ed. Eyrolles*, 1990.
- [FAR-89] N. H. Farhat. "Opto-electronic neural networks and learning machines". *IEEE Circuits and Device Magazine*, pp. 32-41, 1989.

- [FOG-87] F. Fogelman-Soulie, and G. Weisduch. "Random iterations of threshold networks and associative memory". *SIAM. J. Computing*, Vol. 16, pp. 203-220, 1987.
- [GOL-85] E. Golès. "Comportement dynamique des réseaux d'automates" *Thèse de Doctorat de l'USMG, IMAG, Grenoble, France, 1985.*
- [GOL-86] E. Golès, and G. Y. Vichniac. "Lyapunov function for parallel neural network". *Proc. of AIP Conf. on Neural Networks for Computing*, Vol. 151, pp. 165-180, 1986.
- [GOL-90] G. H. Golub, C. F. Van Loan. "Matrix computation". *The John Hopkins University Press, Baltimore, London, second edition, 1990.*
- [GRA-89] H. Graf, and L. D. Jackel. "Analog electronic neural network circuits". *IEEE Circuits and Device Magazine*, Vol. 5, n° 4, pp. 44-49, 1989.
- [HIN-84] G. E. Hinton, T. Sejnowsky, D.H. Ackley. "Boltzmann machines: constraint satisfaction networks that learn". *Technical report Carnegie Mellon Univ.*, 1984.
- [HIN-89] G. E. Hinton. "Deterministic Boltzmann learning performs steepest descent in weight space". *Neural Computation*, Vol. 1, pp. 143-150, 1989.
- [HOP-82] J.J. Hopfield. "Neural networks and physical systems with emergent collective computational abilities". *Proc. Natl. Acad. Sci. USA*, Vol. 79, pp. 2554-2558, 1982.
- [HOP-84] J.J. Hopfield. " Neurons with graded response have collective computational properties lie those of two-states neurons." *Proc. Natl. Acad. Sci. USA*, Vol. 81, pp. 3088-3092, 1984.
- [HOP-85] J. J. Hopfield, D.W. Tank.. "Neural computation of decisions in optimisation problems". *Biological Cybernetic*, Vol. 52, pp. 141-152, 1985.
- [HOP-86] J. J. Hopfield, D. W. Tank. "Computing with neural circuits: a model". *Science*, Vol. 233, pp. 625, 1986.
- [HSU-90] K. Y. Hsu, H. Y. Li, and D. Psaltis. "Holographic implementation of a fully connected neural network". *Proceedings of the IEEE*, Vol. 78, n° 10, pp. 1637-1645, 1990.
- [KIR-82] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. "Optimization by simulated annealing". *IBM Thomas J. Watson Res. Center, Yorktown Heights, NY, 1982.*
- [KOH-88] T. Kohonen. "Self Organization and Associative Memory", *Springer Verlag, 1988.*

- [KOR-89] J. H. Korst, and E. H. Aarts. "Combinatorial optimization on a Boltzmann machine". *J. Parallel Distributed Computing*, n° 2, pp. 331-357, April 1989.
- [LIT-74] W.A. Little. "The existence of persistent states in the brain". *Math. Biosci.*, 19, pp. 101-120, 1974.
- [MAT-95] N. Matsui, and K. Nakabayashi. "Minimum searching by extended Hopfield model". In *IEEE International Conference on Neural Networks (ICNN'95)*, Perth, Western Australia, Vol. 6, pp. 2648-2651, 27 Nov - 01 Dec. 1995.
- [PAR-89] S. Park. "Signal space interpretations of Hopfield neural network for optimization". In *IEEE ISCAS Int. Symp. Circuits Syst.*, pp. 2181-2184, 1989.
- [SOM-86] H. Sompolinsky. "Neural networks with nonlinear synapses and a static noise". *Physical Review A*, Vol. 34, pp. 2571-2574, 1986.
- [STY-90] M. A. Styblinski, T. S. Tang. "Experiments in nonconvex optimization: stochastic approximation with function smoothing and simulated annealing". *Neural Networks*, Vol. 3, pp. 467-483, 1990.
- [SZU-86] H. Szu. "Nonconvex optimization". *Proceedings of the SPIE, Real Time Signal Processing*, Vol. 698, n° 9, pp. 59-65, 1986.
- [SZU-87] H. Szu. "Nonconvex optimization by fast simulated annealing". *Proceedings of the IEEE*, Vol. 75, n° 11, pp. 1538-1540, 1987.
- [TAG-87] G. A. Tagliarini, and E. W. Page. "Solving constraint satisfaction problems with neural networks". *Proceedings of the first IEEE Conf. on Neural Networks*, San Diego, CA, Vol. 3, pp. 741-747, 1987.
- [TIK-76] A. Tikhonov et V. Arsenine. "Méthodes de résolution de problèmes mal posés". *MIR*, 1976.
- [WEI-89] G. Weisbuch. "Dynamique des systèmes complexes. Une introduction aux réseaux d'automates". *Savoirs Actuels, InterEditions/Éditions du CNRS*, 1989.
- [ZHO-92] Y. T. Zhou, and R. Chellappa. "Artificial neural networks for computer vision". *Springer-Verlag*, New York, 1992.

Chapitre IV

- [ADI-85] G. Adiv. "Determining 3D motion and structure from optical flow generated by several moving objects". *IEEE PAMI*, Vol. 7, n° 4, pp. 384-401, Juin 1985.
- [AIS-89] J. Aisbett. "Optical flow with an intensity-weighted smoothing". *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 11, n° 5, pp. 512-522, 1989.
- [AYA-89] N. Ayache. "Vision stéréoscopique et perception multisensorielle (Applications à la robotique mobile)". *Interéditions, Collection Sciences Informatique*, 1989.
- [BEN-92] M. Benjelloun, A. Sdigui. "A method for line segment matching in an image sequence". *IFAC, IMAC, IEEE, IUTAM, Workshop on Motion Control for Intelligent Automation*, Perugia, Italie, pp. 395-400, October 1992.
- [BOU-88] P. Bouthemy. "Modèles et méthodes pour l'analyse du mouvement dans une séquence d'images". *Technique et Science Informatiques*, Vol. 7, n° 6, pp. 527-546, 1988.
- [BOU-89] P. Bouthemy. "A maximum-likelihood framework for determining moving edge". *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 11, n° 5, pp. 499-511, May 1989.
- [CAB-92] F. Cabestaing. "Détection des contours en mouvement dans une séquence d'images. Conception et réalisation d'un processeur câblé temps-réel". *Thèse de Doctorat de Lille I*, 1992.
- [CAF-76] C. Cafforio, F. Rocca. "Methods for measuring small displacements of television images". *IEEE Trans. Inf. Theory*, Vol. 22, n° 5, pp. 573-579, 1976.
- [DAU-88] I. Daubechies. "Orthonormal bases of compactly supported wavelets". *In Pure and Applied Mathematics*, Vol. 41, 1988.
- [DEH-94] I. Dehon. "Etude de l'utilisation de la transformée de Wigner-Ville à l'estimation de mouvement dans une séquence d'images réelles". *Thèse de Doctorat de Lille I*, Oct. 1994.
- [DID-82] E. Diday et Coll. "Eléments d'analyse de données". *Dunod, Bordas*, 1982.

- [EAG-89] R. Eagleson. "Group-theoretic analysis of local flow characteristics while visually tracking a textured surface". *Int. Conf. on Image Analysis and Processing, Positano, Italy*, Sept. 1989.
- [ENK-86] W. Enkelmann. "Investigations of multigrid algorithms for the estimation of optical flow fields in image sequences". *Workshop on motion: Rep. and Analysis, Charleston*, pp.81-87, May 1986.
- [FEN-79] C. L. Fennema, W. B. Thompson. "Velocity determination in scenes containing several moving objects". *Computer Graphics and Image Processing*, Vol. 9, pp. 301-315, (1979).
- [FLE-85] D. J. Fleet, A. D. Jepson. "Velocity extraction without form interpretation". *Proc. 3rd Workshop on Computer Vision: Representation and Control*, pp. 179-185, Oct. 1985.
- [GIV-89] D. D. Giusto, and G. Vernazza. "Optical flow calculation from feature space analysis through an automatic segmentation process". *Signal Processing*, Vol. 16, n° 1, 1989.
- [HAY-66] J. C. Hay. "Optical motion and space perception: an extension of Gibson analysis". *Psychological Review*, Vol. 73, pp. 550-564, 1966.
- [HEE-88] D. J. Heegler. "Optical flow using spatio-temporal filters". *Int. Journal of Computer Vision*, Vol. 1, n° 4, pp. 279-302, Jan. 1988.
- [HEI-89] F. Heitz, P. Bouthemy. "Estimation et segmentation du mouvement. Approche Bayésienne et modélisation Markovienne des occlusions". *AFCET/INRIA, 7^{ème} Congrès sur Reconnaissance des Formes et Intelligence Artificielle*, Tome 3, pp. 1359-1368, Nov. 1989.
- [HEI-91] F. Heitz, E. Menin, P. Bouthemy. "Markov random fields and parallel algorithms for 2D motion analysis". *Proc. 13 th. Imacs Word Congress on Computation and Applied Mathematics*, Dublin, July 1991.
- [HEI-93] F. Heitz & P. Bouthemy. "Multimodal estimation of discontinuous optical flow using Markov random fields". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 15, n° 12, pp. 1217-1232, 1993.
- [HOP-82] J.J. Hopfield. "Neural networks and physical systems with emergent collective computational abilities". *Proc. Natl. Acad. Sci. USA*, Vol. 79, pp. 2554-2558, 1982.
- [HOR-81] B. K. P. Horn, B. G. Schunck. "Determining optical flow". *Artificial Intelligence*, Vol. 17, pp. 185-203, 1981.
- [HUA-81] T. S. Huang. "Image sequence analysis". *Springer-Verlag, Berlin*, 1981.

- [KAO-93] K. Kaoula, M. Benjelloun, B. Dubuisson. "3D region matching and tracking in a long binocular image sequence". *Symposium on Visual Communications and Image Processing'93*, 8-11, Cambridge, Massachusetts, USA, Vol. 2094, pp. 1148-1158, November 1993.
- [KON-90] J. Konrad, E. Dubois. "A comparison of stochastic and deterministic solution methods in bayesian estimation of 2d motion". *Proc. First. European Conf. on Computer Vision, Antibes, France*, pp.149-160, April 1990.
- [LAL-89] P. Lalande & P. Bouthemy "Détection de zones en mouvement dans une séquence d'images selon une approche Markovienne". *12ème Colloque GRETSI, Juan-Les-Pins*, pp. 707-710, Juin 1989.
- [LI-86] H. Li, M. A. Lavin & R. J. Master. "Fast Hough transform: A hierarchical approach". *Computer Vision, Graphics and Image Processing*, Vol. 36, pp. 139-161, 1986.
- [LIM-74] J. O. Limb, J. A. Murphy. "Estimating the velocity of moving images in television signals". *Computer Vision, Graphics and Image Processing*, Vol. 4, pp. 311-327, 1974.
- [MAL-89] S. Mallat. "A theory for multiresolution signal decomposition: The wavelet representation". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 2, n°7, 1989.
- [MIL-89] C. Millour. "Contribution à la vision dynamique: une approche multi-résolutions et multi-traitements". *Thèse de Doctorat ès Sciences, Université de Paris-Sud, Centre d'Orsay, Mars 1989*.
- [MIT-92] A. Mitiche & P. Bouthemy. "Computation and analysis of visual motion: A review". *Rapport technique de l'INRS-Télécommunications*, n°92.26, 1992.
- [MIT-87] A. Mitiche, Y.F. Wang, J. K. Aggarwal. "Experiments in computing optical flow with the gradient-based, multiconstraint method". *Pattern Recognition*, Vol. 20, n° 2, pp. 173-179, 1987.
- [MOU-90] D. Moulet. "Extraction et suivi de contours dans les séquences d'images animées". *Thèse de Doctorat, Université de Nantes, 1990*.
- [NAG-87] H. H. Nagel. "On the estimation of optical flow: relations between different approaches and some new results". *Artificial Intelligence*, Vol. 33, pp. 299-324, 1987.
- [NET-79] A. N. Netravali, J. D. Robbins. "Motion-compensated television coding". *Part 1, Bell System. Tech. Journal*, Vol. 58, pp. 631-670, 1979.

- [ORO-81] J. O'rourke. "Motion detection using Hough techniques". *Proc. of the IEEE Conf. on Pattern Recognition and Image Processing*, pp. 82-87, 1981.
- [POG-85] T. Poggio, V. Torre, et C. Koch. "Computational vision and regularization theory". *Nature*, pp. 314-319, 1985.
- [ROU-87] A. Rougee, B. Levy, A. S. Willsky. "Reconstruction of two-dimensional velocity fields as a linear estimation problem". *Proc. 1st Int. Conf. on Computer Vision, London*, pp. 646-650, June 1987.
- [SAM-84] R. Samy et C. Bozzo. "Extension de la transformée de Hough au traitement d'une séquence d'images". *Proc. Premier Colloque Image CESTA-GRETSI, Biarritz*, pp. 559-564, Mai 1984.
- [SAG-89] D. Sage. "Poursuite d'objets par analyse d'une séquence d'images couleur". *Thèse de Doctorat de l'Institut National Polytechnique de Grenoble*, Avril 1989.
- [SCH-86] B. J. Schunck. "The image flow constraint equation". *Computer Vision, Graphics and Image Processing*, Vol. 35, pp. 20-46, 1986.
- [SDI-94] A. Sdigui, M. Benjelloun. "Tracking of a line segment in a monocular image sequence". *Progress in Image Analysis and Processing*, Edited by Sebastiano Impedovo, World Scientific, pp. 661-668, 1994.
- [TER-86] D. Terzopoulos. "Image analysis using multigrid relaxation methods". *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. PAMI-8, n° 2, pp. 129-139, Mars 1986.
- [THO-87] W. B. Thompson, J. K. Kearney. "Optical flow estimation: an error analysis of gradient-based methods with local optimization". *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 9, n° 2, pp. 229-244, 1987.
- [TIK-76] A. Tikhonov et V. Arsenine. "Méthodes de résolution de problèmes mal posés". *MIR*, 1976.
- [VER-86] A. Verri, T. Poggio. "Against quantitative optical Flow". *M.I.T., A.I. Memo 917 et C.B.I.P. Memo22*, Dec. 1986.
- [ZHO-88] Y.T. Zhou., R. Chellappa. "Computation of optical flow using a neural network". *IEEE, Proc.Intl. Conf. on Neural Network*, San Diego, Vol. 2, pp. 71-78, 1988.

Chapitre V

- [DAU-88] J. G. Daugman. "Networks for image analysis: Motion and Texture". *Proc. First Int'l. Joint Conf. Neural Networks, Washington, D.C.*, pp. 189-193, 1988.
- [GRZ-87] N.M. Grzywacz and A.L. Yuille. "Massively parallel implementations of theories for apparent motion". *Technical Report AI Memo N° 888, CBIP Memo N° 016, MIT Artificial Intelligence Lab. and Center for Biological Information Processing*, June 1987.
- [GRZ-90] N.M. Grzywacz and T. Poggio. "Computation of motion by real neurons". *An Introduction of Neural and Electronic Networks*, Ed. Academic Press, Inc., 1990.
- [HOP-85] J. J. Hopfield, D. W. Tank. "Neural computation of decisions in optimization problems". *Biological Cybernetics*, Vol. 52, pp. 141-152, 1985.
- [HUB-65] D. H. Hubel and T. N. Wiesel. "Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat". *J. Neurophysiol.*, Vol. 128, pp. 229-289, 1965.
- [HUB-77] D. H. Hubel and T. N. Wiesel. "Functional architecture of macaque monkey visual cortex". *Proc. R. Soc. Lond. Ser. B*, Vol. 198, pp. 1-59, 1977.
- [HUT-88] J. Hutchinson, C. Koch, J. Luo, and C. Mead. "Computing motion using analog and binary resistive networks". *IEEE Computer Magazine*, pp. 52-63, 1988.
- [KOC-87] C. Koch. "Analog neuronal networks for real-time vision systems". In *Proc. Workshop on Neural Network Devices and Applications, Los Angeles, CA*, February 1987.
- [KOC-88] C. Koch. "Computing Motion in the Presence of Discontinuities: Algorithm and Analog Networks". in: *Ch. v.d. Malsburg, Neural Computers, NATO ASI Series*, Vol. F41, Springer-Verlag, Berlin, Heidelberg, 1988.
- [LAP-93] M. Lappe and J.P. Rauschecker. "A neural network for the processing of optical flow from egomotion in man and higher mammals". *Neural Computation*, Vol. 5, n° 3, pp. 374-391, 1993.

- [LEE-93] J-C. Lee, B. J. Sheu, W-C. Fang and R. Chellappa. "VLSI neuroprocessors for video motion detection". *IEEE Transactions on Neural Networks*, Vol. 4, n° 2, pp. 178-190, March 1993.
- [OGM-90] H. Ogmen and S. Gagne. "Neural network architectures for motion perception and elementary motion detection in the fly visual system". *Neural Networks*, Vol. 3, n° 5, pp. 487-506, 1990.
- [POG-88] T. Poggio, W. Yang, & V. Torre. "Optical flow: computational properties and networks, biological analog". In R. Durbin, C. Miall, & G. Mitchison (Eds.), *The computing neuron*, PA: Addison-Wesley, 1988.
- [SAM-90] R. Samy. "Infared target motion estimation using a neural network". *SPIE, Applications of Artificial Neural Networks*, Vol. 1294, pp. 238-245, Orlando, 1990.
- [SHI-94] T. Shibata and T. Fukuda. "Hierarchical intelligent control for robotic motion". *IEEE Transactions on Neural Networks*, Vol. 5, n° 5, pp. 823-832, 1994.
- [TSA-91] Tsao et al T.-R. Tsao, H.-J. Shyu, J.M. Libert, and V.C. Chen. "A Lie group approach to a neural system for 3-D interpretation of visual motion". *IEEE Transaction on Neural Networks*, Vol. 2, n° 1, pp. 149-155, 1991.
- [YAC-95] A. Yacoubi, L. Lelaurin, V. Devlaminck, J. P. Dubus. "Contribution des réseaux neuromimétiques à l'estimation du mouvement des cavités cardiaques". *Innovation et Technologie en Biologie et Médecine (ITBM)*, France, Vol. 16, n° 4, pp. 441-453, 1995.
- [ZHO-88] Y.T. Zhou., R. Chellappa. "Computation of optical flow using a neural network". *IEEE, Proc.Intl. Conf. on Neural Network*, San Diego, CA, Vol. 2, pp. 71-78, 1988.
- [ZHO-90] Y.T. Zhou., R. Chellappa. "A network for motion perception ". In *IEEE, Proc. Intl. Joint Conf. on Neural Networks*, Vol. 2, pp. 875-884, San Diego, CA., June 1990.

Chapitre VI

- [DEV-95] V. Devlaminck, L. Lelaurin, A. Yacoubi, J. P. Dubus. "Estimation of discontinuous optical flow using a neural network for echocardiographic images". *Medical Imaging, Proceedings of the SPIE, Image Processing*, San Diego, California, USA, pp. 210-211, 26 Feb - 02 Mar., 1995.
- [GEM-88] S. Geman, and D. Geman. "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images". *IEEE, Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-6, n° 6, Nov.1984.
- [HOP-85] J. J. Hopfield, D. W. Tank. "Neural computation of decisions in optimization problems". *Biological Cybernetics*, Vol. 52, pp. 141-152, 1985.
- [HUT-88] J. Hutchinson, C. Koch, J. Luo, and C. Mead. "Computing motion using analog and binary resistive networks". *IEEE Computer Magazine*, pp. 52-63, 1988.
- [KAM-90] Y. Kamp, M. Hasler., "Réseaux de neurones récursifs pour mémoires associatives". *Ed. Presses polytechniques et Universitaires Romandes, Collection électricité*, 1990.
- [LEE-93] J-C. Lee, B. J. Sheu, W-C. Fang and R. Chellappa. "VLSI neuroprocessors for video motion detection". *IEEE Transactions on Neural Networks*, Vol. 4, n° 2, pp. 178-190, March 1993.
- [LEL-95] L. Lelaurin, A. Yacoubi, V. Devlaminck, J. P. Dubus. "Motion estimation by using a neuromimetic approach". *Proceedings of the IEEE International Conference on Neural Networks (ICNN'95)*, Perth, Western Australia, Vol. 6, pp. 2917-2921, 27 Nov - 01 Dec., 1995.
- [YAC-95] A. Yacoubi, L. Lelaurin, V. Devlaminck, J. P. Dubus. "Optimization by neural networks: Application for optical flow estimation in echocardiographic images". *Proceedings of the 17th Annual International Conference IEEE Engineering in Medecine and Biology Society*, Montréal, Québec, Canada, September 20-23, 1995.



Résumé

Le travail présenté dans cette thèse porte sur l'utilisation des réseaux de neurones pour la minimisation de fonctionnelles avec comme application particulière l'estimation du mouvement dans des séquences d'images.

Dans une première étape, nous décrivons les notions fondamentales indispensables à la définition d'un neurone formel et ses variantes, les principaux types de réseaux constitués de tels neurones parmi les plus connus, et nous présentons plus particulièrement le modèle de Hopfield. Dans une seconde étape, nous rappelons l'état des diverses techniques classiques existantes sur l'analyse du mouvement 2D.

Après avoir étudié dans un premier temps, un réseau de neurones destiné à l'estimation du mouvement basé sur des considérations physiologiques et anatomiques du système visuel, qui reprend des travaux antérieurs sur le sujet, nous mettons en évidence les limitations de ce réseau, ce qui nous a amené à proposer et à développer une autre approche du problème. Celle-ci ne prend pas en compte la complexité biologique. Le réseau proposé est vu comme un système destiné à la minimisation de fonctionnelles. La méthode consiste à transcrire le problème en termes d'états qui sont les variables discrètes représentant les deux composantes du paramètre à estimer. Chaque composante est alors déduite par une somme des états des neurones d'un réseau associé. Nous caractérisons la fonction d'énergie associée au réseau et nous appliquons celui-ci au cas particulier des fonctionnelles basées sur les méthodes différentielles en estimation de mouvement.

Afin de satisfaire le critère de rapidité du processus de convergence, nous étudions différentes stratégies de parcours du réseau.

Enfin, nous comparons les résultats avec ceux fournis par les méthodes classiques d'estimation du mouvement de façon à valider les solutions obtenues par le réseau proposé.

MOTS CLES: RESEAUX DE NEURONES - MOUVEMENT - MODELE DE HOPFIELD - FLOT OPTIQUE - MINIMISATION DE FONCTIONNELLES.