

Mécanisme connexionniste pour l'anticipation visuelle

Jérémy Fix, Nicolas P. Rougier, Frédéric Alexandre

► **To cite this version:**

Jérémy Fix, Nicolas P. Rougier, Frédéric Alexandre. Mécanisme connexionniste pour l'anticipation visuelle. [Interne] 2005. inria-00000245

HAL Id: inria-00000245

<https://hal.inria.fr/inria-00000245>

Submitted on 16 Sep 2005

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Mécanisme connexionniste pour l'anticipation visuelle

Connectionist Mechanism of visual prediction for saccadic eye movements

Jeremy Fix

Nicolas Rougier

Frédéric Alexandre

Laboratoire LORIA
Campus Scientifique, B.P. 239
54506 Vandoeuvre-lès-Nancy Cedex, France

Contact : Jeremy.Fix@Loria.fr

Résumé

On présente un modèle d'anticipation visuelle utilisant un modèle de neurone Sigma-Pi. Ce modèle permet de prédire les modifications de la perception visuelle à la suite de saccades oculaires, avant qu'elles ne soient réalisées ; ce mécanisme a été mis en évidence par des données biologiques ([1], [2]). On souhaite appliquer ce mécanisme pour construire des représentations motrices centrées sur l'oeil et sur la tête des objets visuellement saillants, mises à jour par la perception visuelle et le mouvement oculaire. On peut ainsi déterminer la consigne motrice permettant de recentrer un stimulus visuel sorti du champ visuel à la suite d'un mouvement des yeux.

Mots Clef

Anticipation visuelle, saccade oculaire

Abstract

We present a model of visual anticipation using a Sigma-Pi model of neuron. This model allows to predict the modifications of the visual perception after saccadic eye movements, even before the movement is done ; this mechanism has been highlighted by biological evidences ([1], [2]). We wish to apply this mechanism to build eye-centered and head-centered motor representations of visually salient objects, updated by the visual perception and the eye movement. We can then determine the motor command to focus on a visual stimulus out of the visual field because of a saccade.

Keywords

Visual anticipation, saccadic eye movement

1 Introduction

L'homme perçoit visuellement une scène par l'intermédiaire de la lumière que lui renvoie le monde. Cette information lumineuse est captée par des photorécepteurs qui

tapissent le fond de la rétine qui la convertissent en signal nerveux. Le signal nerveux transite, via le nerf optique, vers le cortex visuel primaire. Il traverse alors hiérarchiquement plusieurs aires qui lui appliquent différents filtres afin d'en extraire des informations relatives aux couleurs, contours, contrastes et mouvements. Ces informations visuelles sont alors combinées au sein d'une carte de saillance ([4]). Cette carte de saillance représente les stimuli saillants de la scène observée. Le système visuel humain ne permet pas de voir avec précision au delà de quelques degrés d'angle : l'accuité visuelle est nettement moins bonne en périphérie de la rétine qu'à la fovéa. Pour pallier ce handicap, l'oeil effectue des mouvements rapides (20 à 150 ms) et de grande amplitude (jusqu'à 180 degrés par seconde) que l'on appelle les *saccades oculaires*. Des données biologiques ([3]) sur le cerveau du singe suggèrent que le cerveau dispose d'un mécanisme qui permet de prédire la position de la projection de stimuli saillants après une saccade oculaire, avant même que cette saccade ne soit réalisée.

Au delà de la motivation d'émuler des mécanismes cognitifs, il est intéressant de constater que dans l'exemple d'une application robotique, il peut être nécessaire d'avoir une représentation stable du monde, malgré les mouvements de caméra nécessaires pour l'explorer. C'est pour construire cette représentation de l'environnement qu'il est nécessaire de disposer d'un mécanisme permettant de prédire les modifications de la perception visuelle avant un mouvement de caméra, ce que l'on appelle *anticipation visuelle*.

On commence par présenter un mécanisme générique dont l'objectif est de déplacer un profil d'activité d'une consigne donnée. Ce mécanisme est étudié sous deux approches. La première utilise une carte associative dont les dimensions grandissent avec le carré des dimensions des cartes d'entrée. Cette première approche n'est pas efficace en temps de calcul mais permet de mieux comprendre le principe qui est à la base de la seconde approche, utilisant des neurones Sigma-Pi. Ce mécanisme générique peut alors être

appliqué au problème de l'anticipation visuelle tel que nous l'avons introduit. Il peut aussi s'étendre aux problèmes de changement de coordonnées puisqu'il calcule un profil d'activité déplacé d'une consigne fixée, par exemple pour passer d'un référentiel d'une modalité sensorielle à un référentiel d'une modalité motrice.

Enfin, nous appliquerons ce modèle sur une plateforme robotique, pour construire, à partir d'une perception visuelle, un référentiel moteur qui permet de déterminer le mouvement de caméra nécessaire pour recentrer un stimulus déjà observé dans le champ visuel.

Dans la suite de l'article, nous appellerons *rétilne* la carte neuronale représentant le profil d'activité que l'on souhaite déplacer et *consigne* la carte neuronale représentant l'amplitude et la direction de ce déplacement.

2 Problématique

Représentons la rétine par un plan sur lequel se projettent des stimuli visuels. Notons F le point de focalisation des yeux. Un stimulus visuel se projette en un point M avant une saccade oculaire et en un point M' après cette saccade. Le déplacement du stimulus visuel entre les moments pré et post saccadiques est caractérisé par le vecteur \vec{u} (cf. fig. 1) et correspond à l'opposé du mouvement des yeux (cf. fig. 2). On fait l'hypothèse que le déplacement relatif d'un stimulus sur la rétine est proportionnel au mouvement angulaire des yeux. On peut alors prédire la position du point M' à partir du point M et du vecteur \vec{u} :

$$F\vec{M}' = F\vec{M} + \vec{u} \quad (1)$$

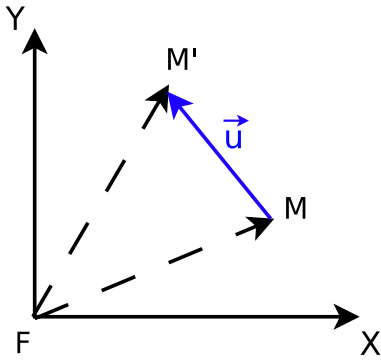


FIG. 1 – La rétine est modélisée par un plan où se projette en F le point de focalisation des yeux, en M la position d'un objet avant une saccade et en M' la position du même objet après une saccade. L'amplitude et la direction du mouvement oculaire sont représentées par le vecteur $-\vec{u}$

Le modèle que nous présentons dans cet article est composé de cartes neuronales : une carte représente l'entrée visuelle, dans laquelle sont codés un nombre arbitraire de stimuli visuels, une autre code la consigne motrice (dans notre étude, l'amplitude et la direction du déplacement des

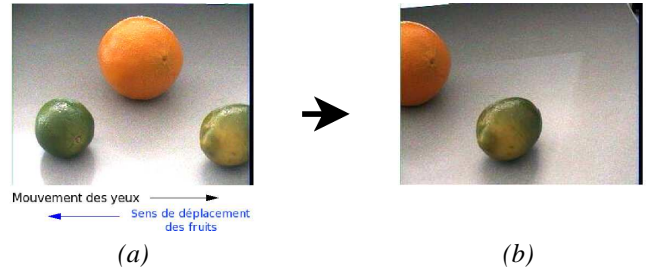


FIG. 2 – (a) Une scène visuelle avant une saccade oculaire. (b) La même scène après une saccade vers la droite. La scène visuelle se déplace dans le sens opposé au mouvement des yeux.

stimuli) et une dernière code la position prédite de la projection des stimuli, résultant d'un mouvement oculaire. On ne traitera pas dans cet article de la plausibilité biologique de ces différentes cartes ; l'objectif, en construisant des cartes d'entrées et une carte de sortie est de dissocier les entrées de la sortie du mécanisme.

3 Conventions de représentation

3.1 Entrée visuelle

La carte *rétilne*, constituée de $n \times n$ neurones, représente les stimuli visuels. Chaque neurone est caractérisé par deux composantes $(x, y) \in \llbracket 1, n \rrbracket^2$. De même, un stimulus est caractérisé par deux coordonnées $(x_v, y_v) \in \llbracket 1, n \rrbracket^2$. Un neurone de la *rétilne* calcule son entrée en évaluant une distance qui le sépare du stimulus à coder. La distance choisie dans notre étude (cf. eq. 2) permet de représenter les stimuli d'entrée par des gaussiennes d'activité.

$$d((x, y), (x_v, y_v)) = e^{-\left(\frac{x-x_v}{r}\right)^2 - \left(\frac{y-y_v}{r}\right)^2} \quad (2)$$

où r est un paramètre à fixer.

3.2 Consigne motrice

La consigne motrice représente le déplacement relatif des stimuli visuels. La carte *consigne*, constituée de $m \times m$ neurones, représente un stimulus qui code l'amplitude et la direction du déplacement des stimuli visuels entre les moments pré et post saccadiques, c'est à dire l'opposé du mouvement moteur. Ce déplacement relatif est caractérisé par le vecteur \vec{u} (cf. fig. 1) de composantes $(x_c, y_c) \in \llbracket 1, m \rrbracket^2$. Chaque neurone de la carte *consigne* calcule son activité en évaluant une distance qui le sépare du stimulus à coder sachant que le neurone $(m/2, m/2)$ code une consigne nulle (cf. eq. 3).

$$d((x, y), (x_c, y_c)) = e^{-\left(\frac{x-x_c+\frac{m}{2}}{r}\right)^2 - \left(\frac{y-y_c+\frac{m}{2}}{r}\right)^2} \quad (3)$$

où r est un paramètre à fixer.

Un stimulus centré sur la carte *consigne* représente une consigne nulle (cf. fig. 3). Dans la suite de l'article, les dimensions de la carte *rétilne* et de la carte *consigne* sont égales ($n = m$).

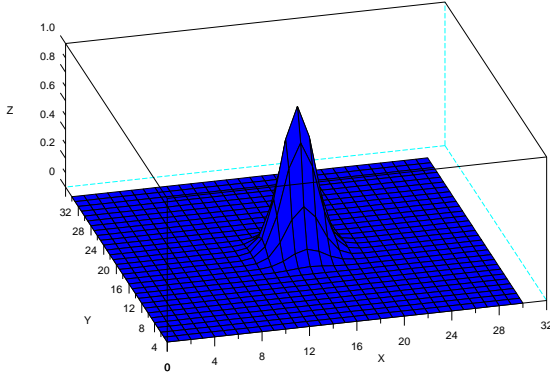


FIG. 3 – Un stimulus centré sur la carte représente une consigne nulle. Pour l'illustration, on a considéré une carte de 30x30 neurones et la gaussienne est centrée en $x_c = y_c = 15$ avec un rayon $r = 2$.

4 Dynamique des neurones

Les réseaux de neurones permettent d'effectuer des calculs numériques et distribués et il est intéressant de les considérer comme des systèmes dynamiques régis par des équations faisant intervenir le temps. La théorie de la CNFT (*Continuum Neural Field Theory*) étudie les mécanismes de compétition au sein d'un réseau de neurones (grâce à des connexions entre les neurones d'un même réseau) permettant de faire émerger des propriétés "intéressantes". La théorie de la CNFT a été introduite par S. Amari ([5]), qu'il a étudié dans le cas de réseau de neurones à une dimension. J.G. Taylor en a fait l'étude dans le cas à deux dimensions dans [6]. La théorie de la CNFT propose une équation d'évolution de l'activité des neurones, continue dans le temps et l'espace (cf. eq. 4).

$$\tau \cdot \frac{\partial u(x,t)}{\partial t} = -\alpha \cdot u(x,t) + \oint w(x,x') \cdot f(u(x',t)) dx' + I(x,t) + h \quad (4)$$

où f est la fonction d'activation des neurones, h une constante représentant leur seuil d'excitation, α et τ des paramètres à fixer, $I(x,t)$ l'entrée perceptive à la coordonnée x et à l'instant t , et w un jeu de poids représentant des connexions latérales au sein d'une même carte. En pratique, et dans notre étude, l'entrée perceptive est issue d'un calcul (cf. eq. 5), il n'y a pas de connexions latérales, et l'équation 4 et discrétisée dans le temps et l'espace (cf. eq. 6).

$$I(x,t) = \oint s(x,y) \cdot i(y,t) dy \quad (5)$$

$$\tau \cdot \Delta u_j(t + \delta t) = -\alpha \cdot u_j(t) + \sum_{i \in E} s(i,j) \cdot I_i(t) + h \quad (6)$$

5 Modèles

Le principe des modèles que l'on présente consiste à caractériser de manière unique tous les couples formés d'un neurone de la carte *ré tine* et d'un neurone de la carte *consigne*. On va considérer deux modèles : le premier illustre le mécanisme et a une vertu pédagogique et le second modèle est une solution plus efficace que la première, dont une application robotique possible est présentée à la fin de l'article.

5.1 Avec une carte associative

Architecture

Considérons une carte *associative* constituée de $(n \times n) \times (n \times n)$ neurones. Chaque neurone a pour entrée un neurone de la *ré tine* et un neurone de la *consigne* tels que le couple formé par ces deux neurones ne soit représenté que par un unique neurone de la carte *associative* (cf. fig. 4).

Les deux entrées de chaque neurone de la carte *associative* sont additives et excitatrices. Ainsi, chaque neurone de la carte associative est d'autant plus excité que ses deux neurones d'entrée sont excités. Si $u_j(t)$ l'activité du neurone j de la carte associative, à l'instant t , $e(t)$ et $c(t)$ les activités à l'instant t des neurones de la *ré tine* et de la *consigne* qui sont les deux entrées du neurone j , l'équation d'évolution de $u_j(t)$ s'écrit :

$$\tau \cdot \Delta u_j(t + \delta t) = -\alpha \cdot u_j(t) + e(t) + c(t) + h \quad (7)$$

où $\Delta u_j(t + \delta t) = u_j(t + \delta t) - u_j(t)$, τ et α sont des paramètres à fixer, et h est le seuil d'excitation du neurone.

Considérons une carte de sortie qui reçoit ses entrées de la carte associative. La carte de *sortie* à des dimensions identiques à celles de la *ré tine*, et représente les stimuli visuels déplacés de la quantité codée par la *consigne*. Un neurone de coordonnées $(x,y) \in [1;n]^2$ de la carte *sortie*, a pour entrée le neurone de la carte *associative*, lui-même recevant ses entrées du neurone (x_v, y_v) de la *ré tine* et du neurone (x_c, y_c) de la *consigne*, si et seulement si :

$$\begin{cases} x = x_v + x_c - \frac{n}{2} \\ y = y_v + y_c - \frac{n}{2} \end{cases} \quad (8)$$

A l'aide des équations précédentes, on détermine les entrées de chaque neurone de la carte de sortie.

Nous n'irons pas plus loin dans la description de ce modèle parce qu'il a un inconvénient majeur : les dimensions de la carte *associative*. Néanmoins, il permet de comprendre le principe du second modèle que l'on va exposer. Pour résumer, on a cherché à caractériser de manière unique tous les couples possibles d'un neurone de la *ré tine* et d'un neurone de la *consigne*. Pour des cartes d'entrées de dimension $n \times n$, il est nécessaire de construire une carte *associative* de dimension $(n \times n) \times (n \times n)$. Une fois cette carte construite, chaque neurone caractérisant de manière unique un couple d'un neurone de la *ré tine* et d'un neurone de la *consigne*, en

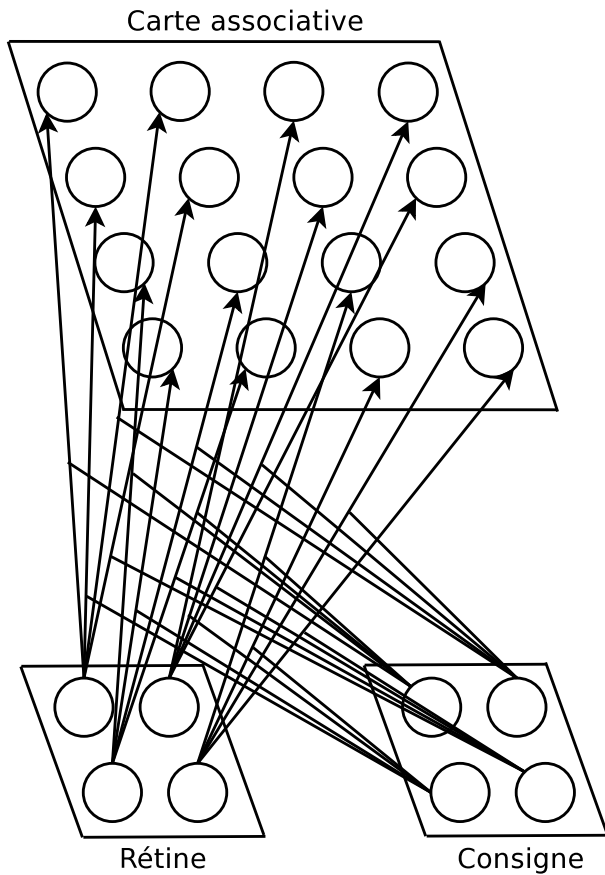


FIG. 4 – Chaque neurone de la carte *associative* caractérise de manière unique un couple constitué d’un neurone de la *rétine* et d’un neurone de la *consigne*. La contribution de ces deux entrées est additive. Pour l’illustration, la rétine et la consigne sont des cartes de 4 neurones et la carte associative est constituée de 16 neurones. Le principe des connexions est identique dans le cas de réseaux plus grands. Un neurone de la carte *associative* reçoit une entrée de la *rétine* avec un poids synaptique de 1.0 et une entrée de la *consigne* avec un poids synaptique de 1.0.

utilisant l’équation 8, on détermine quelles sont les entrées des neurones de la carte de sortie. On décompose donc le principe en deux étapes :

1. Caractériser de manière unique tous les couples constitués d’un neurone de la *rétine* et d’un neurone de la *consigne*
2. Déterminer pour chaque neurone de la carte de sortie les couples précédemment caractérisés qui l’excitent

Simulation

Considérons une entrée visuelle constituée de deux stimuli et une consigne fixée. La figure 5 présente la simulation du modèle que l’on vient de présenter. On reconnaît la connectivité présentée sur la figure 4 en s’intéressant à l’activité dans la carte *associative*. La carte *associative*

peut être vue comme un ensemble de copies de la *consigne* dont l’amplitude de la copie dépend de l’activité d’un neurone de la *rétine*. Chaque neurone de la rétine excite un pavé de neurones de la carte associative (ce qui explique les pavés de neurones excités), pavé qui reçoit également ses entrées de la *consigne*. Le seuil d’excitation des neurones de la carte *associative* est ajusté de telle sorte qu’une activité émerge si et seulement si ses deux neurones d’entrée sont excités. On peut ainsi caractériser la double information visuelle et motrice. Le profil d’activité obtenu sur la carte *sortie* est le profil d’activité de la *rétine* déplacé de la quantité codée par la *consigne* (un stimulus centré sur la carte *consigne* représente une consigne nulle, cf. section 3.2).

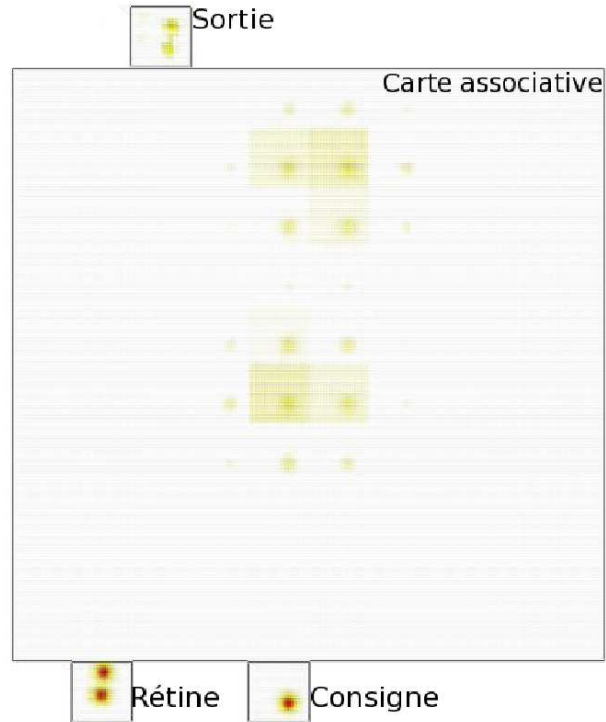


FIG. 5 – La carte *sortie* représente le profil d’activité de la carte *rétine*, déplacé de la quantité codée dans la carte *consigne* (un stimulus centré sur la carte *consigne* représente une consigne nulle, cf. section 3.2). Pour l’illustration, la *rétine*, la *consigne* et la *sortie* sont constituées de 20x20 neurones. La carte *associative* est constituée de 400x400 neurones.

5.2 Avec des neurones Sigma-Pi

Architecture

On constate que plusieurs couples constitués d’un neurone de la *rétine* et d’un neurone de la *consigne* excitent un même neurone de la carte de sortie : ils sont équivalents. Dans le modèle précédant, on utilisait une carte *associative* qui faisait figurer de manière explicite la caractérisation de ces couples en introduisant un neurone par couple pos-

sible. Pour chacun de ces neurones, la *ré tine* et la *consigne* avaient des contributions additives. On propose désormais de caractériser un couple de neurones d'entrée en faisant le produit de leur activité et en intégrant ce calcul dans la mise à jour de l'activité des neurones de sortie. On utilise pour cela des neurones de type Sigma-Pi (cf. fig. 6). Si $u_j(t)$ désigne l'activité du neurone j de la couche de sortie, à l'instant t , $e_i(t)$ et $c_i(t)$ les activités des neurones de la *ré tine* et de la *consigne*, du couple indiqué i pour le neurone de sortie j (formés en respectant l'équation 8), à l'instant t , l'équation d'évolution de $u_j(t)$ s'écrit :

$$\tau \cdot \Delta u_j(t + \delta t) = -\alpha \cdot u_j(t) + \sum_{i=1}^r e_i(t) \cdot c_i(t) \quad (9)$$

où $\Delta u_j(t + \delta t) = u_j(t + \delta t) - u_j(t)$ et τ et α sont des paramètres à fixer.

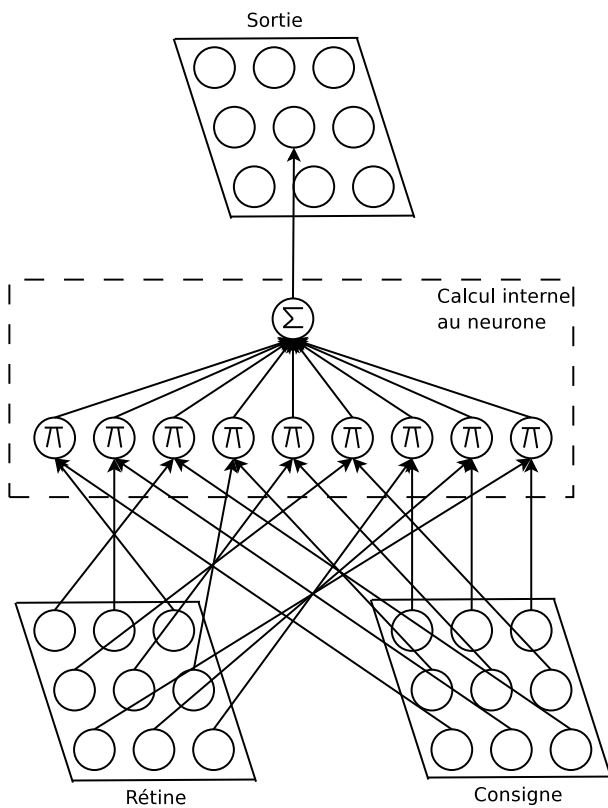


FIG. 6 – Modèle utilisant des neurones de sortie de type Sigma-Pi. On a représenté les connexions du neurone central de la carte de *sortie*. Chaque neurone de sortie calcule en **interne** une somme de produit d'activités. Les couples constitués d'un neurone de la *ré tine* et d'un neurone de la *consigne* sont caractérisés par le produit de leur activité.

Simulation

Considérons une entrée visuelle constituée de 4 stimuli et une consigne fixée. La figure 7 présente la simulation

du modèle que l'on vient de présenter. Le profil d'activité obtenu sur la carte *sortie* est le profil d'activité de la *ré tine* déplacée de la quantité codée dans la carte *consigne* (un stimulus centré sur la carte *consigne* représente une consigne nulle, cf. section 3.2).

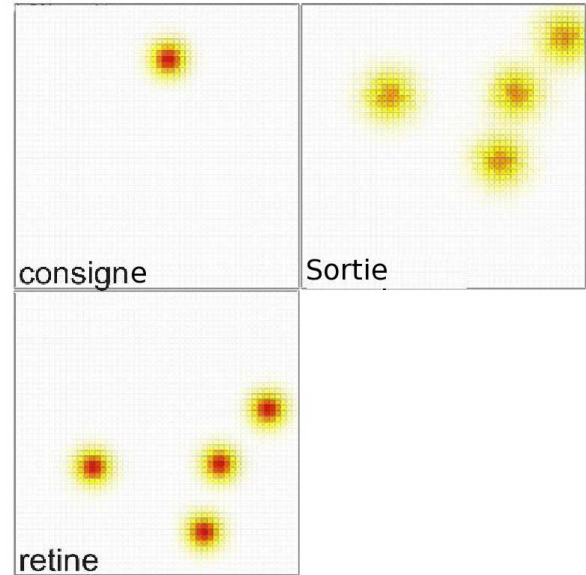


FIG. 7 – Simulation du modèle utilisant des neurones de type Sigma-Pi. L'activité de la carte de *sortie* est l'activité de la *ré tine* déplacée de la quantité codée par la *consigne* (un stimulus centré sur la carte *consigne* représente une consigne nulle, cf. section 3.2). Les cartes sont constituées de 50x50 neurones.

Discussion

Ce modèle présente l'avantage d'être efficace en temps de calcul puisque le nombre de neurones utilisés pour le calcul est égal à la taille des cartes d'entrée. On peut reboucler la carte *anticipation* sur la *ré tine* pour faire apparaître sur la même carte, à savoir la *ré tine*, les stimuli visuels et la prédiction de leur position. On peut même envisager de supprimer complètement la carte *anticipation*, en considérant que la *ré tine* sert à la fois d'entrée et de sortie du mécanisme. Seulement, il est nécessaire dans ces deux cas d'ajuster les seuils d'excitation des neurones pour s'assurer qu'il n'y a pas de phénomène en chaîne de type : prédire la position d'un stimulus qui lui-même est la prédiction d'un stimulus visuel.

6 Validation sur une plateforme robotique

On se propose d'utiliser le modèle Sigma-Pi dans une application robotique. L'objectif est de créer une représentation motrice du monde, par l'intermédiaire d'une perception visuelle. Le champ visuel de la caméra est limité. Dans notre expérience, une caméra fixe initialement un ob-

jet saillant qui est une pomme verte sur un fond blanc. La caméra effectue alors successivement 2 mouvements, prédéterminés, qui sortent l'objet saillant du champ visuel (cf. fig. 8).

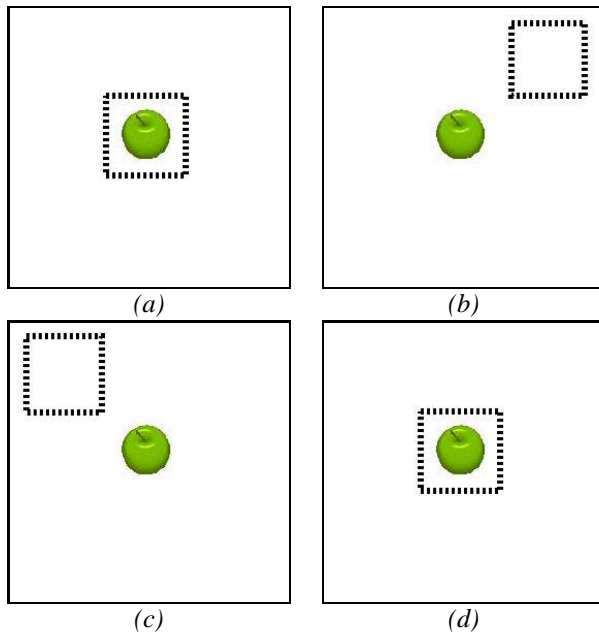


FIG. 8 – Le champ visuel, limité, est représenté par la région en pointillés. (a) Un objet saillant est initialement dans le champ visuel de la caméra. (b) puis (c) La caméra fait 2 mouvements successifs, prédéterminés qui sortent l'objet saillant du champ visuel. (d) Le mouvement moteur décodé de la représentation centrée sur l'oeil permet de recentrer la pomme sur le champ visuel.

On maintient, pendant les mouvements de la caméra, une représentation motrice du monde, par rapport la position actuelle de la caméra. Pour cela, on utilise en entrée deux informations : d'une part l'information visuelle et d'autre part la position de la caméra. Initialement, la caméra est centrée et l'objet est dans le champ visuel. Cette représentation motrice du monde permet à chaque instant de déterminer le mouvement moteur nécessaire pour recentrer l'objet saillant dans le champ visuel.

Le modèle (cf. fig. 9) construit deux représentations. Ces deux représentations sont motrices et sont construites successivement. La première représentation est appelée "représentation centrée sur la tête" et correspond à une représentation motrice de la scène visuelle, par rapport à la position initiale de la caméra. La caméra est initialement centrée par rapport à la tête (au corps) du robot qui, lui, est fixe dans la scène durant toute l'expérience; c'est pourquoi on appelle cette représentation "centrée sur la tête". La deuxième représentation, centrée sur l'oeil (ou la caméra), est construite à partir de la précédente et de l'opposé de la consigne motrice. Dans cette expérience, la

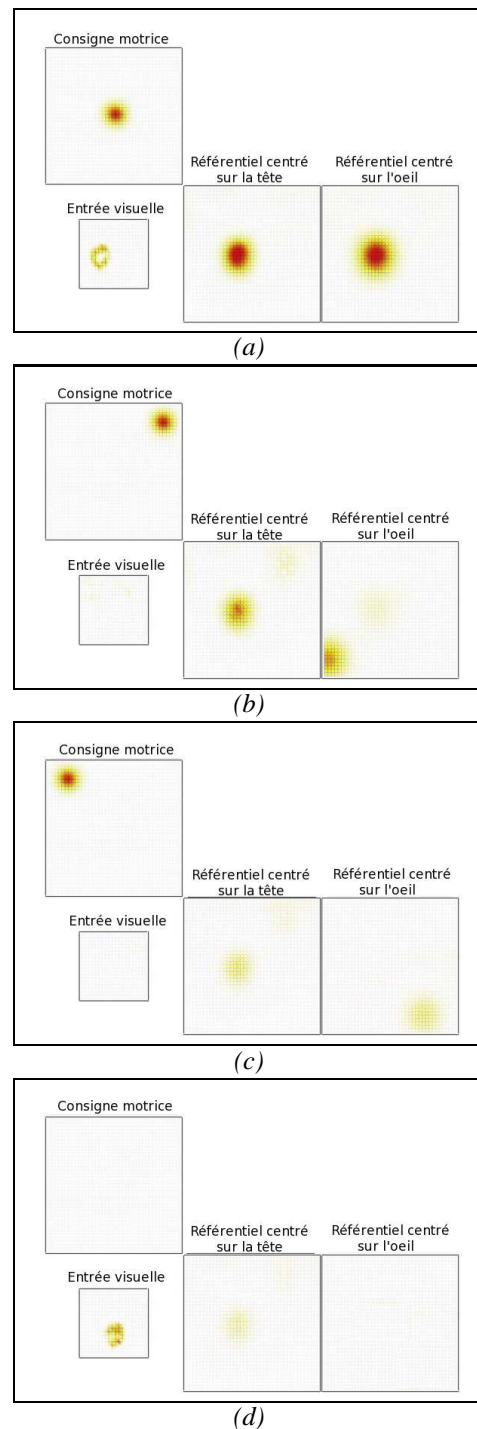


FIG. 9 – (a) Un objet saillant est présent dans le champ visuel à l'instant initial. La consigne motrice est nulle et les deux représentations motrices centrées sur la tête et sur l'oeil sont identiques. (b) puis (c) Deux mouvements moteurs, caractérisés par un stimulus dans la carte consigne sont effectués et sortent l'objet du champ visuel. Le corps du robot étant fixe dans la scène, la représentation centrée sur la tête n'a pas changée d'allure. La représentation centrée sur l'oeil est mise à jour. (d) Le mouvement moteur, décodé du référentiel centré sur l'oeil de la figure (c), permet de recentrer le stimulus dans le champ visuel (cf. Entrée visuelle).

consigne motrice représente la position de la caméra, dans un référentiel moteur, par rapport à sa position initiale (centrée par rapport au corps du robot). Or, les stimuli visuels se déplacent sur la rétine, dans le sens opposé au mouvement moteur, ce qui justifie qu'on utilise l'opposé de la consigne motrice afin de construire cette seconde représentation du monde.

A la suite des deux mouvements oculaires qui ont fait sortir l'objet saillant du champ visuel, le mouvement moteur nécessaire pour recentrer le stimulus est décodé de la représentation centrée sur l'oeil, à l'aide des estimateurs décrit dans [7] :

$$\begin{aligned}\widehat{x}_0 &= \frac{\oint_R x \cdot u_{x,y} dx dy}{\oint_R u_{x,y} dx dy} \\ \widehat{y}_0 &= \frac{\oint_R y \cdot u_{x,y} dx dy}{\oint_R u_{x,y} dx dy}\end{aligned}\quad (10)$$

7 Discussion

Nous venons de présenter un modèle qui permet de déplacer un profil d'activité, étant donnée une consigne. Ce modèle a été utilisé pour une application robotique, afin de construire une représentation motrice d'un environnement à partir d'une perception visuelle, ce que nous avons introduit comme un mécanisme d'anticipation visuelle. Il permet ainsi de conserver une représentation cohérente du monde, malgré les mouvements de caméra qui modifient la perception. Ce mécanisme de prédiction des modifications de la perception visuelle peut être utilisé pour augmenter les possibilités d'autres mécanismes déjà existants. Par exemple, le mécanisme d'attention sélective présenté dans [8] permet de parcourir successivement plusieurs objets saillants d'une scène visuelle. Afin de tous les parcourir successivement, les stimuli déjà observés sont stockés dans une mémoire de travail. Pour maintenir cette mémoire de travail cohérente avec l'environnement, l'entrée visuelle ne doit pas changer brusquement. L'un des intérêts d'étendre ce mécanisme avec le mécanisme de prédiction présenté dans l'article est que l'on peut prévoir les modifications de la perception à la suite d'un mouvement oculaire volontaire, avant même que ce mouvement ne soit réalisé.

Plus généralement, ce mécanisme peut être utilisé pour effectuer des calculs de changement de coordonnées, par exemple dans le cadre de ce que A. Pouget et T.J. Sejnowski appellent le *Dynamic Remapping* dans [9]. Différents modèles proposent déjà des solutions pour ce problème. D. Zipser et R.A. Andersen propose dans [10] (voir aussi [11]) un modèle à trois couches (une couche d'entrée, une couche cachée et une couche de sortie), pour lequel les connexions afférentes sont apprises par rétropropagation du gradient de l'erreur. On a vu en 5.1 qu'une telle solution nécessite une couche cachée dont les dimensions grandissent avec le carré de la dimension des cartes d'entrée. Un autre modèle proposé par J.

Droulez dans [12], mais dont le principe rejoint ce que propose K. Zhang dans [13], repose sur l'asymétrie de connexions latérales, dont l'influence doit être contrôlée. Par ailleurs, dans ce modèle, le profil d'activité est déplacé continuellement, contrairement au modèle Sigma-Pi présenté dans l'article.

Différentes questions ne sont pas encore résolues. Par exemple, il pourrait être intéressant d'apprendre les connexions qui sont pour le moment définies explicitement. L'apprentissage ressemblerait à un apprentissage supervisé puisque les stimuli visuels pré et post saccadiques correspondent aux mêmes stimuli, projetés sur la rétine à des instants différents. Mais ce ne serait pas vraiment un apprentissage supervisé en ce sens que le modèle devrait confronter lui même le résultat de la prédiction à la réalité.

Références

- [1] A. BERTHOZ. *Le sens du mouvement*. Editions Odile Jacob
- [2] K. NAKAMURA, C.L. COLBY. Visual, Saccade-Related, and Cognitive Activation of Single Neurons in Monkey Extrastriate Area V3A. *Journal of Neurophysiology*. Aug 2000 ; 84 : 677 - 692.
- [3] K. NAKAMURA and C.L. COLBY. *Updating of the visual representation in monkey striate and extrastriate cortex during saccades*. PNAS, vol. 99, no. 6. March 19, 2002
- [4] L. ITTI and C. KOCH. *Computational modeling of visual attention*. Nature Reviews Neuroscience. 2, 194-203, 2001.
- [5] S. AMARI. *Dynamic of pattern formation in lateral-inhibition type neural fields*. Biological Cybernetics, 27, 77-78, 1977.
- [6] J.G. TAYLOR. *Neural bubble dynamics in two dimensions : foundations*. Biological Cybernetics, 80, 5167-5174, 1999.
- [7] E. SALINAS, L.F. ABBOTT. *Vector Reconstruction from Firing Rates*. Journal of Computational Neuroscience 1, 89-107, 1994.
- [8] J. VITAY, N.P. ROUGIER, F. ALEXANDRE. *A Distributed Model of Spatial Visual Attention*. In Biomimetic Neural Learning for Intelligent Robotics, S. Wermter and G. Palm Eds. Springer-Verlag, 2005.
- [9] A. POUGET, T.J. SEJNOWSKI ; *Dynamic Remapping in "The Handbook of Brain Theory"*, 335-338.
- [10] D. ZIPSER, R.A. ANDERSEN. *A back-propagation programmed network that simulates response properties of a subset of posterior parietal neurons*. Nature, vol. 331, 25 February 1988.
- [11] P. MAZZONI, R.A. ANDERSEN. *Gaze Coding in the Posterior Parietal Cortex* in "The Handbook of Brain Theory". p. 423-426.

- [12] J. DROULEZ, A. BERTHOZ. *A neural network model of sensoritopic maps with predictive short-term memory properties*. *Neurobiology*, Vol 88, pp. 9653-9657, November 1991.
- [13] K. ZHANG. *Representation of Spatial Orientation by the Intrinsic Dynamics of the Head-Direction cell Ensemble : A Theory*. 1996.