

# Apprentissage de catégories sensori-motrices par un robot autonome

C. Joulain & P. Gaussier & A. Revel, ENSEA ETIS  
6 Av du Ponceau, 95014 Cergy Pontoise Cedex  
e-mail: gaussier or joulain@ensea.fr

## 1 Introduction

Nous présentons une architecture neuronale permettant à un robot autonome d'apprendre à créer "en ligne" des catégories sensori-motrices. Un signal de renforcement (douleur ou récompense) permet au robot d'associer une forme visuelle de son environnement avec l'action à effectuer. Notre architecture est basée sur l'utilisation de réseaux de neurones et suit le principe PerAc (Perception Action); c'est en agissant sur son environnement que le robot va pouvoir apprendre [5]. Nous décrivons comment le robot peut apprendre et reconnaître un objet dans une scène en introduisant un mécanisme de focalisation qui permet de ne s'intéresser qu'à une partie de l'image. Pour réaliser les associations, nous utilisons une règle probabiliste. Nous montrerons que la catégorisation des objets ne peut se faire qu'au niveau moteur.

## 2 L'Architecture

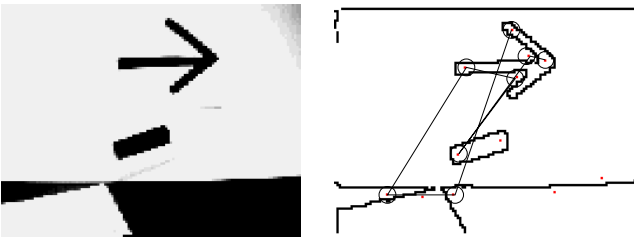


Figure 1: a) Exemple d'image avec un pictogramme et 2 distracteurs. b) Exploration de plusieurs points de focalisation (ici 8).

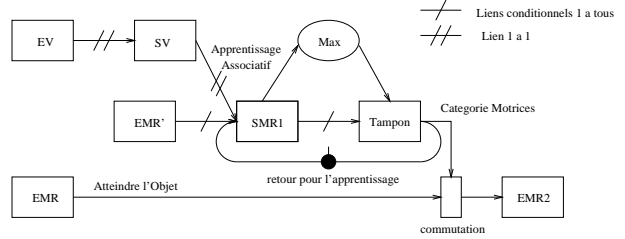


Figure 2: Implémentation neuronale simplifiée. Le robot apprend d'abord à créer des catégories sensori-motrices simples SMR1 (Sortie Mouvement du Robot 1) basées sur la reconnaissance d'une vue locale SV (Sortie Visuelle) et sur un mécanisme réflexe constant EMR' facilitant l'activité du neurone associé à "Aller tout droit". Ensuite le robot utilise SMR1 et le réflexe EMR pour obtenir SMR2.

### 2.1 Reconnaissance d'une vue

Le système sensoriel de notre robot repose sur des informations visuelles provenant d'une caméra CCD. Dans un premier temps, pour diminuer la quantité d'information et pour disposer de données plus facilement exploitables, une extraction de contours est réalisée. Nous espérons ainsi disposer des frontières des objets (figure 1b) ou tout au moins d'informations suffisamment pertinentes pour reconnaître les objets par leurs effets sur l'environnement (approche écologique de la vision [6]).

Les éléments pertinents permettant au robot de choisir son action peuvent se trouver n'importe où dans l'image, un traitement donnant une invariance par translation est donc nécessaire. Ceci est réalisé par un mécanisme de focalisation qui permet de centrer l'attention sur un point par-

ticulier de l’image. Ces points sont obtenus par l’utilisation d’un filtre qui détecte les angles entre les contours, c’est simplement la différence de deux filtres gaussiens appliquée à l’image de contours (les points sur la figure 1b) [2]. Le champ visuel du robot est maintenant réduit au voisinage du point de focalisation par une transformation log-polaire similaire à la projection rétinotopique de la rétine sur les aires corticales primaires [2]. Plus l’information est proche du centre, plus elle aura d’importance.

Cette “imagerie” constitue notre Entrée Visuelle (EV), elle est stockée par une carte topologique probabiliste (PTM)[4] qui constitue la Sortie Visuelle (SV). La carte permet d’avoir un apprentissage non supervisé immédiat et en ligne. Deux formes similaires seront codées sur des neurones voisins. Un paramètre nommé vigilance (équivalent à celui décrit dans Adaptive Resonance Theory (ART) [1]), contrôle le degré de nouveauté d’une forme pour qu’elle soit apprise.

## 2.2 Choix d’une action

Chaque SV est connectée par un lien inhibiteur et un lien activateur à la sortie motrice (SMR1). Ces connections sont des liens probabilistes, c’est une version analogique des liens PCR (Probabilistic Conditioning Rule [3]). Dans notre cas, la Sortie Motrice du Robot numéro 1 (SMR1) est réduite à trois neurones, “tourner de 90° à gauche”, “aller tout droit” et “tourner de 90° à droite”. Comme le montre la figure 1b, le robot fait une exploration séquentielle de plusieurs zones, à chacune de ces zones correspond une SV qui va proposer un mouvement (groupe SMR1). Le mécanisme du maximum (max) et du groupe de neurones servant de mémoire à court terme (tampon) de la figure 2 permet de garder en mémoire la copie de SMR1 ayant eu la plus grande activité. Nous avons donc une sélection en fonction du mouvement le plus actif et non en fonction de l’EV la mieux reconnue (c.a.d la SV la plus active). Par exemple, lorsqu’un distracteur est mieux reconnu

qu’un pictogramme, le mouvement proposé par le distracteur sera moins actif que celui proposé par le pictogramme si le lien entre le distracteur et la mauvaise action est plus faible qu’entre le pictogramme et la bonne action.

## 2.3 Apprentissage de catégories

Le processus expérimental est le suivant: le robot est placé dans une certaine position et voit une image contenant un motif à apprendre. Le robot analyse l’image et effectue un mouvement. Après chaque mouvement, le robot est remis dans sa position initiale. Si le mouvement effectué est celui désiré, le robot reçoit un signal de récompense, sinon il reçoit un signal de douleur. Ce signal lui permet d’associer ou de dissocier l’ensemble des SV activées lors de la phase d’exploration séquentielle avec le mouvement à effectuer (figure 3). C’est l’apprentissage associatif mentionné dans la figure 2 entre SV et SMR1. La règle utilisée est le PCR [7], les poids sont modifiés selon une règle probabiliste quand un signal de renforcement est présent.

### Algorithme PCR Analogique

Si  $\left| \frac{\partial P(t)}{\partial t} \right| > \xi$  :

Mise à jour des probabilités

$$\begin{aligned} \Delta p_{ij}(t) &= \alpha \cdot \frac{\partial P}{\partial t} \cdot C_{ij} \cdot \text{Sign}(W_{ij}) - \lambda p_{ij}(t) \\ p_{ij}(t+1) &= p_{ij}(t) + \Delta p_{ij}(t) \end{aligned} \quad (1)$$

Si  $W_{ij}=0$  alors  $\text{Sign}(W_{ij})=-1$   
Si  $W_{ij}=1$  alors  $\text{Sign}(W_{ij})=1$

Tirage aléatoire

$$\begin{aligned} \text{Si } Alea > p_{ij} \text{ et } C_{ij} \neq 0 \\ \text{alors } \begin{cases} W_{ij}=1-W_{ij} \\ p_{ij}=1-p_{ij} \end{cases} \end{aligned} \quad (2)$$

$P(t)$  signal de renforcement, satisfaction du robot.

$\alpha$  coefficient d’apprentissage, récompense différée.

$\lambda$  coefficient d’oubli.

$\xi$  constante fixée par l’expérimentateur.

$Alea$  valeur aléatoire  $\in [0, 1]$ .

PCR consiste à utiliser des poids binaires associés à une probabilité qui mesure la confiance dans la valeur du poids ( $p \in [0, 1]$ ). Pour être en mesure de modifier le terme de confiance, une mesure de la corrélation est con-

servée et mise à jour à chaque itération. Trois paramètres intégrés dans le temps sont respectivement associés à l'entrée  $I_i$ , la sortie  $O_j$  et le produit entrée-sortie. Les notations utilisées sont:  $\overline{I}$ ,  $\overline{O}$  et  $\overline{IO}$ . La corrélation peut être approximée par  $\mathcal{C} = \frac{\overline{IO}}{\sqrt{\overline{I} \cdot \overline{O}}}$ . Dans notre cas, nous utilisons PCR avec des poids analogiques. La valeur du poids n'est plus  $W_{ij}$  mais  $W_{ij} \cdot p_{ij}$ . Ce produit nous donne un poids analogique en accord avec le terme de confiance. La règle pour la sortie d'un neurone est :

$$W'_{ij} = W_{ij} \cdot p_{ij}$$

$$O_j = \max_{i=1}^n (W'_{ij} \cdot I_i) + \min_{i=1}^n (W'_{ij} \cdot I_i) + \text{bruit}$$

Le max est pris sur les liens activateurs et le min sur les liens inhibiteurs. Le bruit peut être aussi faible que désiré, il est uniquement utilisé pour permettre un choix aléatoire lorsque plusieurs neurones ont la même sortie.

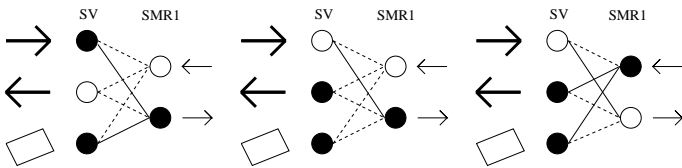


Figure 3: Exemple d'évolution des liens sensori-moteurs durant l'apprentissage, a)  $t=0$ : association correcte avec la flèche "tourner à droite", b)  $t=1$ : mauvaise association avec "tourner à gauche", c)  $t=2$ : le robot effectue les bonnes associations

## 2.4 Le mouvement réflexe

Un mécanisme réflexe permet d'éviter au robot des comportements erratiques voire dangereux lorsqu'il n'a encore rien appris de son environnement. Dans notre cas, le réflexe est d'aller dans la direction de l'objet sur lequel l'attention est focalisée. Pour ce faire les coordonnées du centre de la zone de vision sont projetées sur l'axe des x. Un ensemble de neurones code cette projection. Le neurone central correspond à un objet au centre de l'image. Le premier et le dernier neurone correspondent respectivement à un objet complètement sur la gauche et un objet complètement sur la droite. Le champ de vision du robot étant de  $70^\circ$ , si la projection est codée sur le premier neurone (l'objet est

à gauche) le robot effectuera une rotation de  $35^\circ$  sur la gauche dans le but d'avoir l'objet juste en face de lui.

Ce mécanisme permet de limiter le nombre de mouvements à trois dans le groupe moteur SMR1. C'est la combinaison du réflexe (EMR) et de SMR1 qui donne des possibilités de mouvements beaucoup plus grandes (figure 4). Tout ceci est totalement "gratuit", le robot n'a rien à apprendre en plus des associations entre SV et SMR1. Nous pouvons voir les mouvements effectués par le robot en fonction du pictogramme sur les figures 4 et 5.

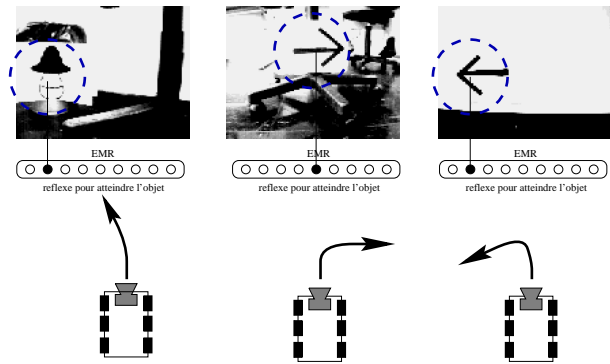


Figure 4: Différents mouvements du robot en fonction des associations sensori-motrices apprises et de la position de l'objet dans le champ de vision. Le cercle représente la zone de focalisation utilisée par EV.

## 3 Résultats

Durant l'expérience nous présentons 7 images différentes au robot (trois flèches "tourner à droite", deux flèches "tourner à gauche", une image représentant un champignon et un autre motif). Sur chaque image le robot explore 6 points de focalisation. Le réseau reçoit donc au total 42 vues locales SV. Grâce aux propriétés topologiques de la carte PTM seulement 13 SV sont apprises. Les autres vues sont suffisamment ressemblantes pour activer un neurone dans le voisinage d'une SV précédemment apprise. Nous voyons donc l'efficacité de la généralisation a priori de PTM. De plus, la vitesse d'apprentissage des associations sensori-motrices est nettement augmentée, c'est toujours le même neurone de

SV ou de son voisinage qui réagit, donc toujours le même qui est associé au mouvement.

Comme énoncé précédemment les liens entre SV et SMR1 sont doubles, avec un lien inhibiteur qui sera toujours inhibiteur ou nul et un lien activateur qui lui aussi sera toujours activateur ou nul. Ces liens inhibiteurs sont indispensables, ils évitent que le premier mouvement appris soit toujours le mouvement effectué (pour une explication plus détaillée voir [3]). Mais ces liens apportent aussi une plus grande vitesse d'apprentissage car l'activité liée à un mauvais mouvement sera diminuée, ainsi les chances qu'il soit de nouveau activé par la forme sont plus faibles.

L'expérimentation montre aussi que le robot est en mesure de déterminer quelles sont les formes pertinentes dans son champ de vision. Si nous prenons comme exemple la figure 1a) nous voyons deux distracteurs et une flèche "tourner à droite". Si en plus de cette flèche nous présentons une flèche "tourner à gauche" avec les mêmes distracteurs le robot est en mesure d'associer les mouvements "tourner à droite" et "tourner à gauche" à une vue locale centrée sur la flèche et non sur les éléments distracteurs. Ceux-ci ont une activité motrice moindre que les éléments pertinents.

Nous avons aussi émergence d'un comportement d'évitement comme le montre la figure 5. Quels que soient l'orientation du robot et la position de la flèche dans le champ de vision le robot va contourner l'obstacle soit par la droite soit par la gauche en fonction de la direction indiquée par la flèche.

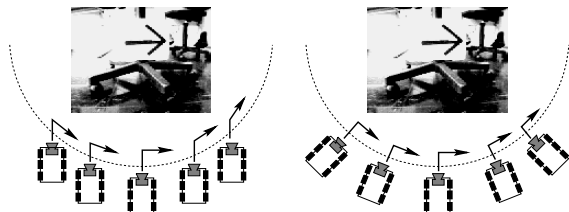


Figure 5: L'ordre "tourner à droite" associé à la reconnaissance de la flèche "tourner à droite" permet au robot d'éviter toujours la flèche en passant par la gauche. Le cercle en pointillé représente le début de la zone où la flèche est reconnue.

## 4 Conclusion

Notre architecture permet "l'émergence" d'un comportement d'évitement par la simple association d'un pictogramme et d'un mouvement. De plus, nous avons montré que l'identification des éléments pertinents d'une image n'est pas vraiment complexe si le robot dispose d'un mécanisme de focalisation sur les différents objets possibles. Un autre point important est la contribution de la topologie aussi bien pour la réduction des informations apprises que pour la vitesse d'apprentissage. Elle permet d'avoir "gratuitement" des informations a priori sur la topologie du problème.

Les résultats obtenus valident notre architecture. Elle sert en ce moment de base pour l'implémentation d'un modèle de navigation en temps réel [5].

## References

- [1] G.A. Carpenter, S. Grossberg, and D.B. Rosen. Art2-a : an adaptive resonance algorithm for rapid category learning and recognition. *Neural Networks*, 4:493–504, 1991.
- [2] P. Gaussier. *Simulation d'un système visuel comprenant plusieurs aires corticales : Application à l'analyse de scènes*. PhD thesis, Université d'Orsay, 1992.
- [3] P. Gaussier, A. Revel, C. Joulain, and B. Gas. Living in a partially structured environment: How to bypass the limitation of classical reinforcement techniques. *submitted to Robotics and Autonomous Systems*, 1996.
- [4] P. Gaussier and S. Zrehen. The probabilistic topological map (ptm) : A self-organizing and fast learning neural map that preserves topology. *Submitted to Neural Networks*, 1994.
- [5] P. Gaussier and S. Zrehen. Perac: A neural architecture to control artificial animals. *Robotics and Autonomous Systems*, 16:291–320, 1995.
- [6] J.J. Gibson. *The ecological approach to visual perception*. Laurence erlbaum associates, Hillsdale, New Jersey / London, 1986.
- [7] A. Revel, P. Gaussier, and C. Joulain. Une règle de conditionnement probabiliste pour le contrôle de robots autonomes. In *submitted to NSI*, 1996.