



HAL
open science

Coordination de tâches sensori-motrices par la coopération robuste d'entités distinctes chez les robots et chez l'homme

Jacques Ferber

► **To cite this version:**

Jacques Ferber. Coordination de tâches sensori-motrices par la coopération robuste d'entités distinctes chez les robots et chez l'homme. ROBEA 2005 - 5ièmes Journées Bilan du Programme Interdisciplinaire ROBotique et Entités Artificielles, Mar 2005, Montpellier, France. pp.169-177. lirmm-00268627

HAL Id: lirmm-00268627

<https://hal-lirmm.ccsd.cnrs.fr/lirmm-00268627>

Submitted on 21 Sep 2022

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.

Coordination de tâches sensori-motrices par la coopération robuste d'entités distinctes chez les robots et chez l'homme

Coordinateur scientifique : Jacques Ferber, UM2,
LIRMM, 161 rue Ada, 34392 Cedex 5,
Tel : 04 67 41 86 07 - Email : ferber@lirmm.fr

Equipes participantes :

- Département de Robotique et département d'Informatique du LIRMM (UMR 5506, Montpellier), Département STIC du CNRS (Automatique) 161 rue Ada, 34392 Montpellier cedex 05. Tel 04 67 41 85 85. <http://www.lirmm.fr>. Correspondant dans le cadre du projet: Robotique : René Zapata (zapata@lirmm.fr). Informatique : Jacques Ferber (ferber@lirmm.fr). Participants : J. Chapelle, J. Ferber, B.Jacquot, Ph. Lucidarme, R. Zapata.
- Equipes "Contrôles Sensori-moteurs" et "Information et Dynamique du Mouvement", Laboratoire Mouvement et Perception (UMR 6152, CNRS et, Université de la Méditerranée, Marseille), Département SDV du CNRS. (Neurosciences). CP 910, Av. de Luminy, 13288 Marseille cedex 09. Tél. 04 91 17 22 50. <http://www.laps.univ-mrs.fr>. Correspondant dans le cadre du projet: Jean-Louis Vercher (vercher@laps.univ-mrs.fr). Participants : R. Bootsma, J. Blouin, C. Bourdin, C. Craig, E. Daucé, E. Guillaud, J.-L. Vercher.

1. Objectifs du projet

Ce projet avait pour but de concevoir et de réaliser une plate-forme d'étude et de validation de nouveaux systèmes de commande de robots autonomes, robustes, apprenants et auto-adaptatifs, en s'inspirant de la biologie, notamment du comportement de l'Homme dans des actions sensori-motrices typiques : atteinte d'une cible, poursuite, etc. et ceci en présence de perturbations imprévues rapides affectant l'Univers.

Le projet a été mené en conjuguant de façon étroite une approche expérimentale chez l'homme et le robot, et en combinant différentes approches de modélisation (automatique pour les boucles de régulation, systèmes dynamiques non-linéaires pour le traitement de l'information, neuro-mimétique pour l'adaptation, et systèmes multi-agents pour la coordination intersegmentaire).

La partie expérimentale visait à déterminer 1) chez l'humain, le rôle spécifique des différentes modalités sensorielles et la nature de l'information véhiculée par celles-ci, ainsi que les paramètres adaptés et 2) sur le robot, la validité des modèles d'inspiration biologique afin de réaliser, par exemple, des systèmes de préhension à grande tolérance de panne et capables d'évoluer dans un environnement inconnu. L'approche expérimentale consiste à perturber (à modifier), en cours de mouvement, l'information véhiculée par les sens (ou les capteurs du robot) et concernant soit l'état de la cible à atteindre, soit celui du membre exécutant le geste.

La partie modélisation a permis 1) de proposer des modèles d'inspiration biologique et de formaliser les mécanismes mis en jeu, et 2) au travers de la simulation, de proposer des prédictions pour le comportement humain et robotique, et de les vérifier, par un retour à l'expérimentation chez l'Homme et le

robot. Ce projet, pluridisciplinaire (impliquant quatre équipes de deux laboratoires, eux-mêmes pluridisciplinaires), était fondamental par nature. Il ne s'agissait pas de réaliser un robot particulier, mais bien de développer un concept : celui d'une coordination (ou d'une coopération) entre éléments moteurs, réalisée à bas niveau, sur la base d'une communication entre des éléments autonomes.

Plus précisément, il s'agissait, pour les équipes SdV, de proposer des modèles d'intégrations multisensorielles pour le contrôle en ligne (le couplage perception-action) et l'adaptation (l'apprentissage) du contrôle des déplacements (ralliement de cible, poursuite, évitement d'obstacle), de déterminer les variables informationnelles pertinentes et d'identifier les paramètres. Pour les équipes STIC, il s'agissait de définir des architectures logicielles (multi-agents, réseaux de neurones) et matérielles (flottille de robots homogènes et hétérogènes) afin d'une part d'adapter sur les robots des modèles de comportement d'inspiration biologique, et d'autre part, en fournissant une incarnation matérielle pour les simulations, de tester la validité de ces modèles.

La première année a été consacrée à l'acquisition des données expérimentales (chez l'Homme) et au développement des modèles informatiques (multi-agents et neuromimétiques) et Robotique. La partie expérimentale, chez l'Homme, réalisée à Marseille, a eu la particularité de combiner (comme il était indiqué dans la demande de financement), les deux approches théoriques courantes en Sciences du Mouvement humain : l'approche écologique, ou des systèmes dynamiques non linéaires, donnant accès à la nature de l'information traitée, et l'approche systémique, ou biocybernétique, donnant accès au mécanisme de traitement de l'information. La deuxième année a permis le développement d'une

architecture de robot fondée sur la notion de colonne corticale et de multi-agents ainsi que le développement de techniques d'apprentissages de coordination sensori-motrices qui se relie aux comportements étudiés chez l'être humain.

2. Contexte

L'auto-apprentissage, l'auto-adaptation et l'auto-organisation vis à vis des variations de l'environnement sont des facultés naturelles chez les êtres vivants. En particulier, les êtres humains manifestent une remarquable capacité à produire des comportements moteurs appropriés (déplacements, saisie et transport d'objets, manipulations diverses...), sous des conditions environnementales différentes et souvent changeantes ou incertaines. Ces facultés permettent à des individus isolés ou en groupe de réaliser de mieux en mieux et de manière sûre aussi bien des tâches directement vitales comme échapper aux prédateurs, se nourrir et procréer, que d'autres, indirectement vitales comme effectuer des travaux manuels, conduire un véhicule, etc. Les mécanismes qui entrent en jeu dans ces processus sont étudiés dans le domaine des Sciences de la Vie, afin :

- d'une part de localiser et d'identifier leurs diverses composantes dans le système nerveux (modèles fins)
- et d'autre part de modéliser les performances dans l'accomplissement d'une tâche, dans des conditions d'environnement variables, ce qui nécessite des régulations très complexes (modèles comportementaux globaux).

Ces mécanismes intéressent également depuis une cinquantaine d'années (début de la « Cybernétique ») les mathématiciens et les ingénieurs. L'approche des premiers est plutôt liée à l'établissement de modèles fins (de l'homéostat aux réseaux de neurones artificiels). Celle des automaticiens est plutôt fonctionnelle : ils cherchent à réaliser des machines et des robots ayant des propriétés d'apprentissage non guidé, et de robustesse automatique des performances devant des variations de l'environnement et même en cas de dérives et de pannes de leurs constituants. L'introduction de capacités auto-adaptatives, étape indispensable vers un robot (ou tout système artificiel) entièrement autonome, capable de s'adapter à tout type d'environnement, reste un vrai défi pour la Robotique, alors que les Neurosciences constatent que ces capacités existent naturellement chez les êtres vivants (l'Homme en particulier) sans être à même d'en proposer un modèle convaincant.

Les Neurosciences du mouvement tentent de comprendre comment, parmi toutes les solutions qui s'offrent à lui, l'être humain parvient à coordonner ses mouvements afin d'assurer le succès de ses actions. En se basant sur l'existence d'invariants dans le comportement moteur, et sur la connaissance et la compréhension naissante des réseaux nerveux

impliqués, ces disciplines ont proposé des modèles dans lesquels la plupart des paramètres et l'exécution elle-même du mouvement étaient pré-déterminés, laissant aux rétroactions sensorielles le seul rôle des ajustements fins terminaux, garants de la réussite de l'action (Figure 1, Schmidt, 1975). En réaction à cette vision essentiellement centraliste et séquentielle du contrôle du mouvement, est apparu au cours des années 80 un courant, issu de la psychologie écologique, qui voyait le comportement moteur comme émergeant de l'interaction entre l'acteur et l'environnement, la perception et l'action étant considérées comme indissociables: la perception n'a de sens que dans l'action est vice et versa. Cette approche met l'accent sur la nature de l'information plutôt que sur son traitement.

Sont particulièrement étudiés les processus adaptatifs permettant de maintenir la performance lorsque l'environnement ou l'acteur lui-même changent. Cette question de l'adaptation au contexte est bien entendu trop vaste pour être abordée ici dans son ensemble. Dans le cadre de ce projet, nous (roboticiens, informaticiens, neuroscientifiques) avons proposé de travailler ensemble et simultanément sur un problème volontairement limité à l'adaptation du bras (humain et robot) lors d'une tâche de pointage visuo-guidé. Cette adaptation est rendue nécessaire par une perturbation (ou la méconnaissance) de l'environnement, mais également par une dégradation sensorielle, suite à une pathologie ou le vieillissement chez l'homme, une panne ou un défaut chez le robot. En nous basant sur le schéma de la figure 2, nous sommes partis du postulat qu'un modèle (une formalisation) de la relation entre l'acteur et l'environnement est nécessaire (sans préjugé sur la nature et/ou la localisation de ce modèle) et qu'apprendre, c'est changer DE modèle (éventuellement la structure du modèle ou de la relation), alors que s'adapter, c'est modifier LE modèle (changer les paramètres).

Autour de la question de l'adaptation du geste de pointage (ou d'atteinte) aux contraintes (conflits, perturbations, situation « inattendues ») environnementales, notre proposition consiste à aborder ce problème, simultanément sous l'angle des deux approches théoriques (contrôlée et émergente), mais aussi par une synergie d'action entre roboticiens, informaticiens, et neuroscientifiques. Une approche similaire a été exploitée par l'équipe de Kawato (Wada et coll. 1999, 2001) dans un contexte d'apprentissage supervisé par imitation. Si ces dernières études ont permis, avec succès, d'obtenir des robots qui apprennent seuls, elles nécessitent une infrastructure et des procédures lourdes, et n'apportent pas réellement d'information sur la nature de l'adaptation chez l'Homme. Comme détaillé plus loin, nous travaillerons sur une infrastructure plus légère, plus réactive et non supervisée, que celle employée par les Japonais.

3. Coordination et apprentissage

La première étude s'est focalisée sur l'évitement d'obstacles en proposant une modélisation et une expérimentation humaine ainsi qu'une expérimentation robotique.

Les équipes SdV ont proposé un modèle de couplage informationnel entre une variable optique, disponible dans l'environnement, et l'action du sujet, dans une tâche de ralliement de cible ou d'évitement d'obstacle. Ce modèle a été ensuite implémenté et testé sur les robots, en utilisant en particulier le système de vision mobile. Cette étude, examinant l'organisation cinématique des gestes effectués face à des variations de la stabilité (visuelle) de l'environnement d'une part (à travers les tâches de capture) et des caractéristiques du mouvement du mobile à intercepter d'autre part (à travers les tâches de frappe et de relais), est complétée par l'étude psychophysique de la sensibilité d'observateurs aux sources d'informations visuelles portant sur le temps de pré-contact.

À partir de l'ensemble de ces données, une modélisation, basée sur le modèle de couplage information-mouvement proposé initialement par Peper et coll. (1994), a été poursuivie pour mettre en évidence le mécanisme de régulation de la vitesse courante par modélisation formelle de l'accélération à appliquer. Nous avons testé expérimentalement ce modèle à partir de la manipulation expérimentale de ses éléments constitutifs. Le modèle prédit, pour un lieu d'interception donné, une variation systématique de la cinématique du geste de capture, dès lors que les conditions d'exécution (angle d'approche du mobile, position de départ de la main) changent. Les résultats obtenus tant sur le plan qualitatif que quantitatif (Figure 1) nous ont permis de valider le modèle testé et d'envisager sa possible extension à des tâches de terrain.

Quatre expériences ont permis d'examiner la nature de l'information utilisée afin d'estimer si des événements vont conduire ou non à une collision entre un sujet mobile et un objet de son environnement. Des sujets ont été testés dans des situations simulant une approche entre deux objets (approche latérale), et une approche entre un objet et le point de vue (approche de la tête). Les objets se déplaçaient avec une décélération constante ou une décélération sinusoïdale. Les résultats montrent que l'estimation se base, dans une large mesure, sur la capacité de la décélération en cours à éviter la collision, spécifiée par la variable optique tau dot. Cependant, l'information spécifiée par la variable tau (liée à l'expansion du flux visuel, égale au temps restant avant l'occurrence de la collision), joue également un rôle important. Nous en avons déduit que le jugement de la collision était basé sur la détection de tau et de son évolution dans le temps (sa dérivée d'ordre 1). Les expérimentations sur la détection d'obstacles et les processus d'apprentissage

permettant d'améliorer cette détection ont été proposés.

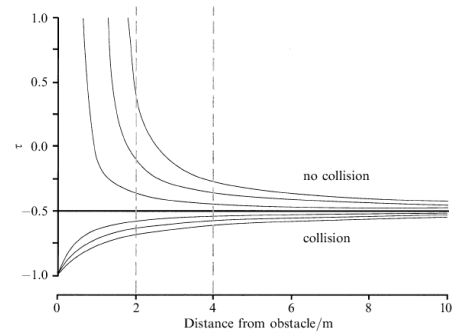


Figure 1 : variable tau dot en fonction de la distance restant avant collision selon une approche avec une décélération constante (Bootsma et Craig, 2003).

Apprentissage de l'évitement d'obstacles

La conception d'équipes de robots coopérant, en particulier de robots mobiles, est envisagée pour prolonger ou remplacer la présence humaine dans des activités pénibles ou dangereuses en environnement hostile. Les applications les plus souvent visées concernent la défense et la sécurité, l'exploration des planètes et autres corps spatiaux ainsi que les activités en milieu sous-marin, nucléaire, etc. Mais les applications aux tâches manufacturières, agricoles et domestiques sont aussi intéressantes, bien que les contraintes de coût soient alors très dures.

Le travail réalisé par les roboticiens (essentiellement dans le cadre de la thèse de Doctorat de Philippe Lucidarme) se place dans le contexte des systèmes multi-agents distribués (voir la partie informatique ci-dessous). L'objectif est l'étude de méthodes d'auto-apprentissage appliquées à des ensembles de robots réactifs. Ces travaux se focalisent sur l'apprentissage de comportements sensorimoteurs de bas niveaux.

Nous avons appliqué les principes trouvés à un type particulier de processus robotique considéré comme un système multi-agent : un robot mobile dans lequel les divers agents, hétérogènes et complémentaires, sont les diverses ressources en capteurs proprioceptifs et extéroceptifs, les motorisations pour la locomotion, un bras manipulateur embarqué, et un système de vision, le robot M3 (Figure 2).

Les résultats ont été validés par des expérimentations sur robots réels disponibles au LIRMM. Il nous a en effet semblé important que les méthodes d'apprentissage proposées puissent être appliquées sur des systèmes réels, dont les contraintes sont parfois loin de la simulation. C'est pour cette raison que nous avons imaginé et conçu une plateforme expérimentale composée de quatre robots mobiles, un manipulateur (bras) mobile miniature et un système de vision stéréoscopique.



Figure 2. Le robot M3 développé au LIRMM dans le cadre du projet Robea

Cette partie du projet s'est découpée en deux étapes. La première, appliquée aux systèmes homogènes, présente l'étude de méthodes évolutionnistes appliquées aux systèmes multi-robots. La deuxième étape des travaux a consisté à rechercher des algorithmes d'auto-apprentissage par renforcement pour les objectifs plus globaux. Dans ce cas il s'agit de récompense retardée et des méthodes classiques seront d'abord testées et améliorées : différence temporelle, programmation dynamique stochastique, "Q-learning" par exemple. Les performances ont été ici aussi testées par des simulations, par l'utilisation de modèles théoriques et par des expérimentations.

Les travaux sur la coordination main-caméra se prolongent actuellement dans une collaboration entre le Département de Robotique du LIRMM et l'Université d'Angers où a été recruté Philippe Lucidarme en octobre 2004. Un stagiaire de DEA travaille sur la saisie d'objet par le robot M3. Pour saisir l'objet, le robot pourra utiliser deux sources de perception : la caméra et les capteurs de proximité. A la fin de l'apprentissage, le comportement appris sera analysé afin d'étudier comment le robot a appris à fusionner les données entre la caméra et les capteurs.

Enfin, nous avons étendu les travaux sur l'apprentissage pour l'amélioration des méthodes de traitement d'image et plus spécifiquement en vision stéréoscopique (travaux de thèse de Bastien Jacquot). S'il existe un grand nombre de méthode d'extraction d'information de profondeur dans une image, ces méthodes souffrent cependant d'un choix souvent arbitraire de paramètres résultant de la non existence d'un critère d'optimisation. Cependant, on peut, dans certains cas, définir un critère permettant d'évaluer un indice de satisfaction de la solution obtenue. Ainsi, il est possible, en utilisant un algorithme d'apprentissage adapté, de trouver une solution (l'ensemble des paramètres) optimisant le traitement. Les résultats obtenus sur la mise en correspondance d'amers entre 2 images sont particulièrement

intéressant puisque nous associons ainsi jusqu'à 20% d'amers supplémentaires suivant les images.

Concrètement, la première phase consiste à optimiser un masque de convolution ayant pour but de détecter des amers « intéressants » et ensuite de rechercher le meilleur correspondant dans l'autre image en tenant compte du voisinage du point. La fonction à optimiser a été définie à partir du nombre de points correctement associé sur les images.

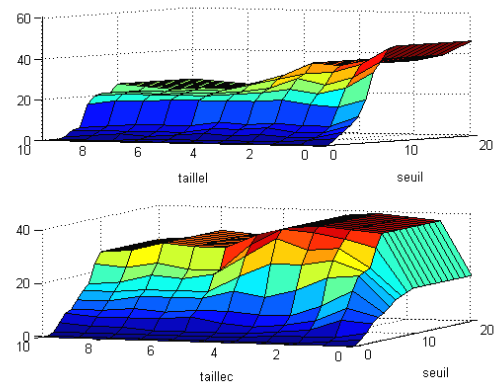


Figure 3 : Performances des algorithmes d'apprentissage d'amers.

On peut voir figure 3, les performances obtenues pour différentes tailles de voisinage en nombre de colonnes et de lignes et la valeur du seuil de comparaison. Il ressort de ce graphique une plus forte influence du nombre de colonnes utilisées dans la comparaison et une valeur optimale autour de 3.

La figure 4 montre le résultat obtenu après optimisation de la taille du voisinage et du masque de convolution utilisé pour la détection des amers, avec :
-> en vert les points amers.
-> en bleu points non retenus
-> en rouge points associés.
Plus que de régler certains paramètres, cette méthode est utilisable en aval de tout traitement d'image en modifiant les couleurs traitées. Il s'agit de créer à partir de l'image d'origine une nouvelle image dont chaque pixel est représenté par un n-uplets.

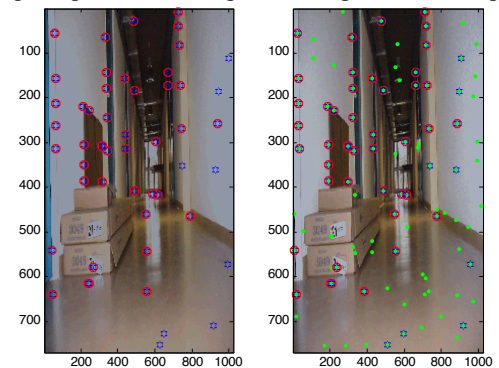


Figure 4. Détection d'amers.

Cette transformation étant apprise dans le but d'optimiser les traitements futurs appliqués à la

nouvelle image. Dans le cadre de projet d'évitement, nous avons étudié la possibilité de transfert d'information en cours d'apprentissage entre différents agents. Il nous a paru intéressant d'utiliser la technique du Q-Learning pour permettre de réel transfert de connaissance. L'idée est de stocker 2 modèles de la Q-valeur, le premier étant propre à l'agent et le second étant donné par des agents ayant de fortes récompenses. A chaque état donné, l'agent choisira parmi les 2 actions optimales résultantes. Ainsi, on peut autoriser une phase d'exploration plus rapide sachant que l'échange d'information permettra le transfert des meilleures solutions au sein de la population.

4. Interaction mutisensorielle dans le contrôle des mouvements

L'objectif de cette seconde étude était d'observer l'interaction des signaux vestibulaires et oculaires pour le contrôle des mouvements du bras, lors d'une tâche de compensation à une rotation passive du tronc autour d'un axe vertical dans un premier temps, puis dans une tâche de pointage en conditions gravito-inertielles modifiées, dans un second temps. Le but final était de proposer un modèle de coordination qui serait testé pour la régulation multisensorielle des déplacements de robots hétérogènes. Pour cela nous avons comparé (chez l'homme) la précision de mouvements du bras dans une tâche de poursuite manuelle, dans laquelle les sources d'informations relatives à la rotation étaient variées. Nous avons créé trois conditions dans lesquelles la détection des rotations de la cible relativement aux participants était soit strictement vestibulaire, soit strictement visuelle, soit la combinaison de ces deux sources (Figure 5). La fréquence de rotation au sein de chaque condition était variée de manière à mettre en avant les plages de précision de chacun des modes de contrôle. Les résultats montrent une intégration non linéaire des sources informationnelles, fonction de la phase du mouvement. Les informations vestibulaires jouent un grand rôle dans la première phase du mouvement, une grande importance des informations visuelles apparaît ensuite. Les signaux oculaires permettent de reproduire correctement les déplacements du tronc, mais avec un grand délai. Les signaux vestibulaires permettent des compensations moins précises à basses fréquences mais plus rapides. La combinaison des deux sources d'information permet des performances optimales sur tout le mouvement et à toutes les fréquences de rotation, les signaux vestibulaires, liés à l'accélération, apportent la précision spatiale, alors que les signaux oculaires, liés à la position, apportent la précision spatiale.

Une seconde partie du travail concerne l'organisation du mouvement spatialement dirigé et son adaptation à des conditions environnementales changeantes. Le but est de proposer un modèle d'adaptation sensorimotrice basé sur l'intégration des

informations visuelles, vestibulaires (inertielles) et proprioceptives.

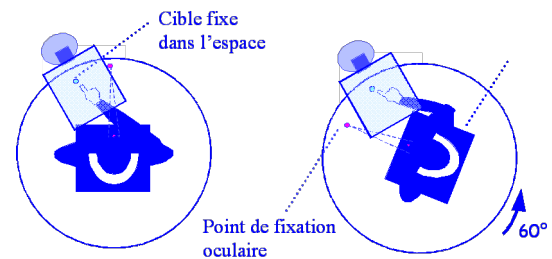


Figure 5 : Dispositif expérimental pour l'expérience de stabilisation

Nos études montrent que des conditions gravito-inertielles modifiées entraînent des erreurs systématiques dans des pointages effectués par des sujets en boucle ouverte, c'est-à-dire sans vision du bras pointeur. Une adaptation du geste de pointage, c'est-à-dire une réduction significative de la quantité d'erreur, semble par ailleurs être directement dépendante de la disponibilité du feedback visuel. Au contraire de l'adaptation à la seule force de Coriolis, l'adaptation aux forces de Coriolis et centrifuge ne semble pas possible sur la seule base de l'information proprioceptive, laissant une place prépondérante à l'information visuelle. Les erreurs observées peuvent avoir deux origines citées dans la littérature : une origine perceptive (mauvais codage du bras, et mauvaise localisation de cible) et une origine motrice (inadéquation de la production de force au mouvement réalisé en conditions inertielles inhabituelles. Cette adaptation différenciée pourrait faire appel à la dissociation entre adaptation dynamique et adaptation statique, déjà suggérée dans la littérature. Elle pourrait aussi être basée sur des processus sensoriels distincts, à la base de la mise à jour des modèles internes. Des données obtenues sur une tâche de localisation visuelle de cible sans implication motrice ont permis de confirmer l'influence de l'hypergravité sur l'activité perceptive des sujets, suggérant une piste supplémentaire pour expliquer l'imprécision des pointages de cibles visuelles mémorisées lors de phases hypergravitaires. Dans ce cas, une partie au moins des erreurs de pointage serait due par une localisation altérée de la cible.

Ces expériences nous permettent entre autre d'estimer le degré d'interdépendance des représentations centrales des caractéristiques intrinsèques (e.g. forme, dimension: système ventral) et extrinsèques (e.g. direction, distance: système dorsal) des objets au cours de leur saisie alors que l'acteur est en déplacement et d'évaluer la possibilité de mettre à jours ces représentations par différentes sources sensorielles (vestibulaire, proprioception, copie d'efférence). Nous avons étudié la contribution des signaux sensoriels générés au cours du déplacement du corps (vestibulaire, proprioceptif,

visuel) dans le contrôle du mouvement du bras et des doigts. Une attention particulière est portée sur la manière dont les déplacements du corps, du bras et des doigts sont influencés par la situation dans laquelle l'acteur se trouve (par exemple, objets à saisir qui se déplacent, qui sont fragiles ou qui ont des formes complexes, situation où le maintien postural est précaire, etc) ainsi que par les intentions de l'acteur.

Ph. Lucidarme, dans le cadre de son postdoctorat au Japon a développé une technique qui permettrait vraisemblablement à terme de donner un modèle plus satisfaisant de cette interaction sensori-motrice. Lorsqu'il existe un grand nombre de degrés de liberté, il existe une infinité de position du corps permettant d'atteindre une position donnée de la main ou d'un autre membre.

L'approche utilisée est basée sur les systèmes multi-agents. On représente chacun des degrés de liberté du robot sous la forme d'un agent. Chaque agent commande l'actionneur associé et reçoit des informations de l'agent qui le précède. Le premier agent (le plus proche de la main) cherche constamment à minimiser la distance entre la main et l'objet à atteindre. Les résultats de simulation ont permis de montrer les propriétés inédites d'une telle méthode. Outre la simplicité et la rapidité qui incombent aux multi agent, d'autres propriétés ont été mis en évidence : la comparaison qualitative avec la gestuelle humaine (réalisé par P. Lucidarme et J.L. Vercher) est assez intéressante et pourrait confirmer certaines hypothèses biologiques. Ainsi, la calibration (l'apprentissage) de la métrique du système se fait durant le mouvement lui-même, sans nécessité d'un pré-mapping, le recrutement des degrés de liberté se fait de façon progressive (lisse) en fonction d'éléments exogènes (position de la cible) et endogènes (disponibilité des actionneurs). Le système est notamment capable de s'adapter suite à un changement de position de la cible ou à la perte d'un agent ou d'un actionneur ; les agents voisins de l'actionneur défaillant compensant l'absence de mouvement. Enfin, l'ajout d'une contrainte d'optimisation (dérivée de la théorie du *minimum jerk*) produit une trajectoire biologiquement plausible (profil de vitesse en cloche). Il est important de noter que toutes ces propriétés sont des propriétés naturelles, émergentes, de la méthode et non une couche adaptative de niveau supérieur.

Afin d'affiner le comportement d'un robot de type humanoïde, une méthode évolutionniste a été utilisée afin d'optimiser les paramètres du systèmes. Plusieurs critères ont été expérimentés : la distance entre la main et l'objet, le temps de saisie et l'énergie consommée. Les deux premiers critères donnent de bons résultats d'un point de vue dynamique, mais le troisième (l'énergie) offre à la machine un comportement très proche de la gestuelle humaine. Ces travaux sont actuellement en cours de publication.

5. Modélisation multi-agent de l'interaction sensori-motrice

L'objectif de cette partie est de développer une architecture multi-agent adaptative pour des agents situés hétérogènes dans le cadre de la thèse de J. Chapelle. Nous avons à terme comme objectif d'obtenir des robots capables de se construire une représentation du monde qui leur est propre à partir de modèles essentiellement sensori-moteurs. Il faudra donc montrer à travers les différentes expériences que le robot développe et améliore ses connaissances tout en conservant un temps de délibération court. Ceci afin de conserver une des propriétés initiales du modèle : la réactivité.

Nous avons proposé dans le cadre de ce projet de réaliser une architecture de robot fondée sur la notion de système multi-agents et de développer des techniques et un début de méthodologie appropriée. En effet, le choix d'un modèle d'apprentissage plus souple et l'utilisation des propriétés et des méthodes de la programmation multi-agent, permet de définir une méthodologie d'apprentissage plus modulaire, notamment en permettant l'ajout de perceptions ou d'actionneurs à un « cerveau » déjà existant sans avoir à reconstruire toutes les connaissances. De la même manière, la panne ou la destruction d'un capteur se déroule sans bloquer le système. De plus il est imaginable de reporter l'utilité du capteur tombé en panne sur une autre perception de type éventuellement différent. Pour cela nous sommes nous avons fondé notre technique d'apprentissage non supervisé sur la notion de motivation liée à une satisfaction.

Nous sommes partis du modèle de Satisfaction-Altruisme [Simonin, 2001], développé en partie dans le cadre de l'avant projet Robea, qui est une extension des progress estimators de M.J. Mataric [Mataric, 1997]. Notre architecture s'inspire des colonnes corticales : les colonnes corticales [Alexandre, 2001]. Cela consiste à regrouper le fonctionnement de plusieurs neurones au sein d'une seule entité apprenante, appelée *colonne corticale*, d'un niveau fonctionnel plus élevé que le simple neurone formel.

Les colonnes corticales se regroupent alors en *aires corticales*. Chaque aire corticale est typée par la nature des informations qu'elle traite. On distingue alors quatre catégories d'aires: Les *aires sensorielles* qui sont connectées aux perceptions ; les *aires motrices* qui réalisent les sorties vers le monde extérieur ; les *aires associatives* qui connectent les aires sensorielles et motrices entre elles (par exemple pour la coordination œil - main) ; les *aires frontales* qui contrôlent et coordonnent le fonctionnement des aires associatives et déterminent l'ordonnement des comportements.

Nous avons modélisé les colonnes corticales comme des agents autonomes et les aires corticales comme des groupes dans le modèles AGR [Ferber et

al. 03].

- Chaque colonne est autonome : elle peut choisir de filtrer les signaux qu'elle reçoit sur ses entrées, et la modulation qu'elle va appliquer à ces informations.
- Les colonnes se rassemblent dans des groupes appelés aires.
- Chaque colonne a une perception locale de l'information. En effet même si une colonne peut avoir beaucoup de connections, elle ne peut avoir une vue complète du système.
- Chaque colonne peut envoyer ou recevoir des informations vers ou depuis les autres colonnes qui sont dans son voisinage, comme des messages échangés entre agents du même

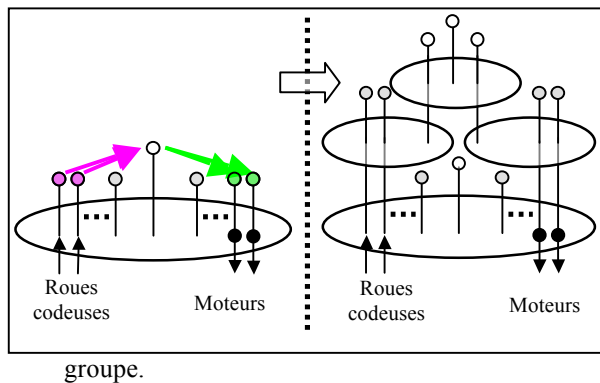


Figure 6. La phase de babbling learning dans une architecture SMA.

Pour chaque type de perception, il existe une aire corticale sensorielle et pour chaque type d'effecteur, une aire motrice. Une *aire associative* fait alors la relation entre chaque paire d'aire motrice et sensorielle. Enfin, l'aire frontale contrôle toutes les aires associatives. C'est cette dernière qui reçoit les émotions / motivations et qui coordonne l'activité des aires associatives (l'apprentissage et la sélection d'actions) en diffusant ou pas les émotions perçues.

1^{ère} phase : le babillage

La phase de *babbling learning* consiste à générer des mouvements aléatoires et à observer leurs impacts sur les perceptions. Pour réaliser ceci, un gestionnaire envoie des valeurs aléatoires aux agents moteurs, et observe leurs impacts en écoutant les messages envoyés par les agents capteurs. On associe, par ce processus, le vecteur cible (le vecteur des mouvements perçus en entrée) au vecteur des commandes motrices générées.

A la fin de cette phase de babillage on obtient une structure de contrôle qui est proche du schéma sensori-moteur du robot et qui prends en compte les contraintes physiques et la dynamique du système, comme sur la Figure 6. Sur la gauche de la figure, tous les capteurs et les moteurs (i.e. les roues droite et gauche) appartiennent au même groupe qui

représente la première couche du cerveau. Cette couche est directement connectée aux éléments physiques du robot. Le gestionnaire de groupe (l'agent central sur la figure) stimule les effecteurs (i.e. les moteurs des roues droite et gauche), et observe les nouvelles valeurs envoyées par les capteurs. Quand le gestionnaire remarque une similarité de fonctionnement, il crée un nouveau groupe correspondant. Sur la partie gauche de la Figure 3, les effecteurs similaires ont été regroupés dans ce que l'on pourrait considérer comme le groupe responsable du déplacement. Un gestionnaire de groupe est alors placé dans celui ci dans le but d'interagir avec les couches plus hautes en utilisant un nombre limité d'ordres synthétisés (par exemple avancer, reculer, tourner à droite, à gauche).

Deuxième phase: *satisfying learning*

Dans cette phase, on utilise les abstractions des couches physiques faites pendant la phase de *babbling*. Cette phase de construction des connaissances est réalisée en essayant de satisfaire les émotions ressenties à l'aide de « délégués d'émotions » qui recherchent les situations leur permettant d'être satisfaits. Pour mémoriser une configuration de capteurs et d'effecteurs intéressante, le délégué crée un groupe dans lequel les agents concernés sont invités à se joindre, et enregistre les valeurs de chacun d'entre eux. Au fur et à mesure que le robot évolue et se déplace, il va apprendre les objectifs intermédiaires qui lui amène satisfaction.

Quand le robot est branché à la base, il ressent un plaisir de satiété (satisfaction de la faim) et le délégué correspondant à cette émotion mémorise donc la configuration qui correspond à cette situation. Ceci peut être interprété par "la situation dans laquelle je vois deux sources lumineuses, et où je touche devant moi, est un but intermédiaire qui satisfait ma faim".

Nous utilisons des techniques d'apprentissage par renforcements. Le principal problème de ces approches est l'évaluation des récompenses. Dans des systèmes réels, les agents doivent percevoir par eux mêmes la réussite ou l'échec de leurs actions. Dans notre approche émotionnelle la réussite ou l'échec est donnée par la valeur de l'émotion qui est traitée. Pour permettre un tel apprentissage, les comportements de l'agent sont définis en utilisant des règles pondérées:

$$(p_i, a_i) \rightarrow (w_i, t_i)$$

Pour une perception donnée p , la probabilité que l'action a soit choisie est proportionnelle à la récompense w qu'il a reçu dans le passé, et inversement proportionnelle au nombre d'essais t . Cette pondération assure l'exploration de tous les comportements possibles.

L'architecture des robots a été simulée sur un simulateur multirobots [Simonin *et al.*, 2002] qui s'exécute sur la plateforme multi-agent MadKit [Madkit; Gutknecht *et al.*, 2000]. La structure et les caractéristiques mécaniques des robots sont les

mêmes que les robots développés dans le projet CoRoM [CoRoM] et dans le pré-programme Robea, déjà utilisés dans d'autres travaux [Simonin and Ferber, 2000; Simonin, 2001; Lucidarme *et al.* 2002]. Le simulateur peut être consulté à l'adresse <http://www.lirmm.fr/~chapelle>.

Nous avons réalisé des expérimentations pour montrer que le robot développe et améliore ses connaissances tout en préservant un court temps de délibération. Dans ces expérimentations nous avons remarqué l'émergence de certains comportements connus comme l'évitement d'obstacle ou la recherche de la base de recharge quand le besoin se fait ressentir.

Conclusion et perspectives

En conclusion, la confrontation des méthodes et théories issues d'une part des Sciences du Mouvement Humain et d'autre part de l'informatique et de la robotique distribuée (multi-agent) se révèle être une piste extrêmement intéressante à creuser afin de doter à terme les robots de capacités plastiques (auto-adaptatives) se rapprochant le plus possible de celles observées chez les humains.

Nous envisageons de poursuivre les collaborations entre les deux laboratoires (plus le laboratoire d'Angers avec Ph. Lucidarme) pour développer un modèle plus complet des interactions sensori-motrices applicables à la robotique et qui permettent de mieux comprendre les mécanismes humains. Dans cette approche, l'utilisation de techniques multi-agents à la fois pour contrôler les mouvements et pour superviser le comportement global d'un robot semblent très prometteurs.

6. Références et principales publications des participants, liées au projet

Les publications réalisées dans le cadre du projet sont notées en gras

[Alexandre, 2001] Frédéric Alexandre. *Un survol de la modélisation corticale*, CORTEX Team, 2001.

Bootsma RJ, Craig CM. Global and local contributions to the optical specification of time to contact: Observer sensitivity to composite tau. PERCEPTION 31 (8): 901-924 2002

Bootsma RJ, Craig CM. Information used in detecting upcoming collision. PERCEPTION 32 (5): 525-544 2003

Bourdin, C., Gauthier, G.M., Blouin, J., Vercher, J.L. (2001) Visual feedback of the moving arm allows complete adaptation of pointing movements to centrifugal and Coriolis forces in human subjects. Neuroscience Letters, 301: 25-28

Bresciani, J.-P.; Blouin, J., Popov, K., Bourdin, C., Sarlegna, F., Vercher, J.-L., Gauthier, G.M. (2002) Galvanic vestibular stimulation in humans produces online arm movement deviations when

reaching towards memorized visual targets. Neuroscience letters, 318: 34-38

Chapelle J., Apprentissage de la coopération dans les systèmes multi-agents situés hétérogènes, DEA Informatique, Université Montpellier II, Juillet 2001.

Chapelle J., Simonin O. et Ferber J. (2002), How Situated Agents can Learn to Cooperate by Monitoring their Neighbors' Satisfaction, ECAI'2002, Lyon, 68-72 .

Daucé, E., Dynamic Retention in Recurrent Networks of Spiking Neurons, 1st Workshop on the Future of Neural Networks (FUNN 2003), July 5, 2003, Eindhoven, The Netherlands.

Daucé, E., Quoy, M. & Doyon, B. Resonant spatio-temporal learning in large random recurrent networks. Accepted dans Biological Cybernetics.

Daucé, E., Quoy, M., Cessac, B., Doyon, B. & Samuelides, M. (1998) Self-organization and dynamics reduction in recurrent networks: stimulus presentation and learning. Neural Networks, 11 : 521-533

Ferber J. (1995), Les systèmes multi-agents. Vers une intelligence collective. Interéditions.

Guédon, O., Gauthier, G.M., Cole, J., Vercher, J.L., & Blouin, (1998) J. The role of arm afferent information in the adaptation to altered visuo-manual relationships in a two dimensional tracking task. J. Mot. Behav, 30: 234-248

Guillaud E., Blouin J., Gauthier GM & Vercher J.L., Ocular and vestibular signals integration in hand trajectories stabilization, en preparation

Ingram, H., van Donkelaar, P., Cole, J., Vercher, J.L., Gauthier G.M. The role of proprioception and attention in a visuo-motor adaptation task. Experimental Brain Research, 132(1): 114-126

Lucidarme P. & Liégeois A., Learning reactive neurcontrollers using simulated annealing for mobile robots. Proc. IROS 2003, Las Vegas, p.674-679.

Lucidarme P., Apprentissage et adaptation pour des ensembles de robots réactifs coopérants. Thèse de Doctorat. Université de Montpellier. Nov. 2003.

Lucidarme P., Apprentissage de comportements réactifs pour des ensembles de robots mobiles, JJCR17, 2003, Paris, p. 33-39 (prix IEEE du meilleur article).

Lucidarme P., Apprentissage et adaptation pour des ensembles de robots réactifs coopérants. Thèse de Doctorat. Université de Montpellier. Nov. 2003.

Lucidarme P., Liégeois A., Vercher J.L. & Bootsma R. Un algorithme évolutionniste pour l'auto-apprentissage de groupes de robots mobiles autonomes. Proc. NSI 2002, La Londe des Maures, art. 56, session Modèles 2.

- Lucidarme P., Simonin O. and Liégeois, A.** (2002), "Implementation and Evaluation of a Satisfaction/Altruism Based Architecture for Multi-Robot Systems", Proc. ICRA'02, Washington, 1007-1012.
- Lucidarme, P., Rongier, P., Liégeois, A. Implementation and Evaluation of a Reactive Multi-Robot System, 2001 IEEE/ASME Int. Conf on Advanced Intelligent Mechatronics, Como, p. 165-170, 8-11 July 2001
- Mataric M. J. (1997), Behavior-Based Control : Examples from Navigation, Learning, and Group Behavior. *J. of Experimental and Theoretical Artificial Intelligence*, 9(2-3), 323-336.
- Michel F., Ferber J., Gutknecht O., Generic simulation Tools based on MAS organization, MAAMAW'2001 Annecy, May 2001
- Montagne, G., Laurent, M., Durey, A., & Bootsma, R.J. (1999). Movement reversals in catching. *Exp. Brain Res.*, 129 : 87-92.
- Mottet, D & Bootsma, R.J. (1999). The dynamics of goal-directed rhythmic aiming. *Biol Cybern.* 80:235-45.
- Peper, C.E., Bootsma, R.J., Mestre, D.R., & Bakker, F.C. (1994). Catching balls: How to get the hand to the right place at the right time. *Journal of Experimental Psychology: Human Perception and Performance*, 20, 591-612.
- Prieur, J.M., Bourdin, C., Sarès, F., Gauthier, G.M., Blouin, J. Vercher, J.L.** Accuracy of spatial localization depending on head posture in a perturbed gravito-inertial force field. *Experimental Brain Research* (en revision)
- Rongier, P. Développement et évaluation de méthodes multi-agents pour les applications aux véhicules robotisés, Thèse de Doctorat, Université Montpellier 2, 20 décembre 2000.
- Rongier, P., Liégeois, A., Simonin, O. Markovian Analysis of a Heterogeneous System: Application to a Cooperation Task for Multiple Consumer Robots, Proc of the IEEE Int. Conf. On Systems, Man and Cybernetics, p. 3033-3038, Nashville, Oct. 8-11, 2000.
- Sarlégnas, F., Blouin, J., Bresciani, J.P., Bourdin, C., Jean-Louis Vercher, J.L. & M. Gauthier G. Role of Target and hand position information in the online control of goal-directed arm movements. (soumis à *Experimental Brain Research*).
- Sarchilli K., Vercher J.L. (1999) Oculo-manual coordination: taking into account the dynamical properties of the arm. *Experimental Brain Research*, 124, 42-52
- Sarchilli, K., Vercher, J.L., Gauthier, G.M. & Cole, J. (1999). Does the oculo-manual coordination control system use an internal model of the arm dynamics? *Neurosciences Letters*. 265, 2: 139-142
- Simonin O. (2001), Le modèle satisfaction-altruisme : coopération et résolution de conflits entre agents situés réactifs, application à la robotique Thèse de Doctorat, Univ. Montpellier 2, décembre 2001.
- Simonin, O., Ferber, J., Modélisation des satisfactions personnelle et interactive d'agents situés coopératifs, 9es Journées francophones d'Intelligence Artificielle distribuée et systèmes multi-agents, Montréal, 12-14 novembre 2001, p. 215-226.
- Vercher J.L., Sarès F, Blouin J, Bourdin C, & Gauthier G. Role of sensory information in updating internal models of the effector during arm tracking. *Brain Research*.
- Villebrun O. (2001). Etude et expérimentation d'un ensemble de mini-robots mobiles coopérant par réactivité et altruisme. Rapport de stage de DEA SyAM, Université Montpellier 2, 6 juillet 2001.