

Projet d'initiation à la recherche

MODÉLISATION NEUROROBOTIQUE ET BIO-INSPIRÉE DE L'IMAGE CORPORELLE ET DE L'ESPACE PÉRIPERSONNEL PAR DES RÉSEAUX DE NEURONES GAIN-FIELD

Réalisé par Mélody ROBINET

Encadrant de projet : Alexandre PITTI

Modélisation neurorobotique et bio-inspirée de l'image corporelle et de l'espace péripersonnel par des réseaux de neurones Gain-field

Mélody Robinet, Alexandre Pitti

Résumé—Les représentations corporelles chez l'homme sont primordiales pour que celui-ci soit capable de contrôler son corps, d'effectuer des mouvements ou d'interagir avec son environnement. Ces représentations ne sont pas figées. Elles sont influencées par des modalités visuelles, auditives, tactiles, motrices ou proprioceptives, par exemple. En robotique bio-inspirée, la modélisation de ces représentations constitue un enjeu majeur si l'on souhaite doter un robot de la conscience de son corps pour que celui-ci soit en mesure d'effectuer des actions précises, d'interagir avec l'homme et de s'adapter aux changements de son environnement. Nous présentons un modèle permettant de modéliser l'image corporelle et l'espace péripersonnel pour un robot humanoïde : les réseaux de neurones gain-field sous la forme d'auto-encodeurs. Ces réseaux modélisent les neurones "gain-field" présents dans le cortex pariétal impliqué dans l'apprentissage des corrélations entre signaux multimodaux. Nous présentons les expériences que nous souhaitons réaliser pour modéliser au mieux l'image corporelle et l'espace péripersonnel à l'aide des réseaux de neurones gain-field.

Mots-clés—image corporelle, schéma corporel, espace péripersonnel, cortex pariétal, réseaux de neurones gain-field, transformations visuomotrices

I. INTRODUCTION

Au cours de son développement, l'homme acquiert des capacités lui permettant de contrôler son corps pour effectuer des mouvements précis et d'interagir avec son environnement. Sur le modèle de la distinction entre perception et action, les neuroscientifiques ont proposé de séparer les représentations corporelles en deux groupes : l'image corporelle et le schéma corporel. Ces deux représentations, parfois confondues, sont responsables des multiples capacités que possède l'homme ainsi que de nombreux mammifères. Parmi ces capacités, on compte notamment les facultés à traiter de multiples informations sensorielles (tactile, visuelle, proprioceptive, auditive), à contrôler son corps pour effectuer des actions, à s'adapter à l'utilisation d'outils.

Dans le domaine de la robotique, la modélisation des représentations corporelles chez les robots bio-inspirés constitue un enjeu majeur. En effet, ces représentations sont cruciales si l'on souhaite créer un robot autonome capable de générer des mouvements raisonnables et précis, d'apprendre et d'interagir avec l'homme en toute sécurité, de comprendre l'environnement en constante évolution qui l'entoure et de s'adapter aux changements de son corps. Il est donc nécessaire que celui-ci puisse maintenir des représentations internes précises de son corps : son schéma corporel et son image corporelle. Toutefois, le robot est dépourvu d'une plasticité cérébrale, et la difficulté principale sera dans la conception d'une représentation informatique fixe mais évolutive capable de décrire son corps.

Au cours de ce projet, on souhaite faire la modélisation de l'image corporelle et de l'espace péripersonnel d'un robot

humanoïde. En d'autres termes, on souhaite faire l'apprentissage des transformations visuomotrices afin que le robot soit capable de reconnaître la position et l'orientation de son corps dans l'espace, ainsi que la représentation de son espace péripersonnel dans lequel il sera capable d'effectuer des actions. La connaissance de ces informations lui permettra ensuite d'effectuer des mouvements à partir de ceux observés dans son champ visuel.

Dans la section II, nous présenterons un état de l'art sur les représentations corporelles et l'espace péripersonnel en neurosciences. Nous présenterons, dans la section III, un second état de l'art sur la modélisation de ces représentations dans le domaine de la robotique bio-inspirée. Nous décrirons ensuite, dans la section IV, l'architecture et les mécanismes des réseaux de neurones gain-field que nous avons choisi d'utiliser pour modéliser au mieux les mécanismes neuronaux dans le cortex pariétal. Puis nous présenterons en section V les expériences envisagées.

II. ETAT DE L'ART SUR LES REPRÉSENTATIONS CORPORELLES ET L'ESPACE PÉRIPERSONNEL EN NEUROSCIENCES

En 1911, Head et Holmes [1], neurologues britanniques, introduisent le terme de "schéma corporel". Lors de leurs études, ils se sont intéressés à la façon dont les lésions présentes sur différentes parties du système nerveux pouvaient induire des troubles psychologiques et comportementaux chez les patients. Ils ont alors mené des tests sur des patients souffrant de troubles neurologiques afin d'observer quelles capacités étaient diminuées par rapport au type de lésion. Par exemple, les patients devaient localiser leur bras dans certaines positions, ou encore déterminer le poids d'un objet dans leur main. Ils ont alors établi qu'il existait deux types de schéma pouvant décrire la représentation que l'on se fait de notre corps : le schéma corporel (ou schéma postural) qui correspond à la conscience que nous avons de la position de notre corps, et le schéma de surface qui correspond à la capacité que nous avons à localiser les stimuli à la surface de notre peau.

Les définitions énoncées par Head et Holmes ont subi de nombreuses modifications au cours des années et des recherches qui ont été menées dans le domaine des neurosciences. Aujourd'hui, les neurologues sont parvenus à établir la définition suivante du schéma corporel :

Le schéma corporel correspond à une représentation sensorimotrice du corps utilisée pour l'action. Elle permet d'intégrer les informations provenant à la fois de la proprioception, de la vision, de l'audition, du système vestibulaire, de la perception tactile ou encore du système moteur. Cette représentation se fait de façon inconsciente et englobe à la fois la dynamique à

court terme, telle que la représentation des positions des parties du corps dans l'espace, mais également à long terme, comme la représentation des dimensions des membres. Elle compte donc un groupe de représentations du corps essentielles au mouvement de celui-ci [2].

Le terme d'image corporelle, énoncé par Schilder [3] est apparu quelques années après. Schilder explique que l'image corporelle de l'humain correspond à l'image de notre propre corps que nous formons dans notre esprit, c'est-à-dire la manière dont le corps nous apparaît dans notre imagination.

Il existerait donc deux représentations bien distinctes du corps : le schéma corporel et l'image corporelle [4].

Pour mieux comprendre comment ces deux représentations interviennent chez l'adulte, il est nécessaire de s'intéresser à leur acquisition. En effet, l'apprentissage du schéma corporel et de l'image corporelle n'est pas inné, il est progressif. Des études ont montré que celui-ci est lié au développement sensorimoteur et s'acquiert dès le stade foetal [5]. En observant le comportement d'un fœtus in utero, il a été observé que le fœtus a tendance à se toucher les parties de son corps les plus innervées et donc les plus sensibles, c'est-à-dire sa bouche et ses pieds. De par ces actions, il prend conscience de la position de son corps et de ses membres, des sensations liées à son toucher et à ses mouvements, mais également de l'espace qui l'entoure. Après la naissance, le nourrisson se développe et devient peu à peu capable de percevoir les informations multisensorielles spatiales qui l'entourent, telles que les informations visuo-proprioceptives ou visuotactiles. Il passe la majeure partie de son temps à se toucher, s'observer et commence progressivement à explorer son environnement, ce qui lui permet de développer et "calibrer" son schéma corporel et son image corporelle. On considère que la formation d'un système multimodale des représentations du corps dans l'espace se produit autour de 2-3 mois [2] [6]. C'est dans cette même période que le nourrisson commence à entreprendre des mouvements pour l'atteinte d'objets.

Du fait de leur apprentissage simultané, il est difficile de distinguer le schéma et l'image corporelle. Dans un article [4], de Vignemont établit les différences qui permettent de séparer ces deux représentations. En effet, des études réalisées sur des patients souffrant de problèmes neurologiques ou psychiatrique ont permis d'observer qu'un déficit du schéma corporel n'impliquait pas forcément un déficit de l'image corporelle, et inversement. Ces observations ont ainsi montré que le schéma corporel est une représentation sensorimotrice du corps, non consciente et nécessaire pour guider les actions, tandis que l'image corporelle est une représentation spatio-temporelle consciente et basée sur la perception (auditive, tactile, visuelle, somatique). Ainsi, le schéma corporel est associé à l'action et l'image corporelle est associée à la perception (voir table I).

De plus, ces deux représentations corporelles sont à la fois plastiques et capables de s'adapter à des changements importants du corps tels que la perte d'un membre, par exemple. L'illusion de la main en caoutchouc (The rubber hand illusion, Botvinick et Cohen 1998 [7]) en est un exemple. Lorsque l'on demande à des sujets d'observer une main en caoutchouc pendant que leur main est cachée, et qu'on utilise

la simulation tactile à la fois sur la main du sujet et sur la main en caoutchouc, on observe un conflit entre les informations visuelles et tactiles, ainsi que la modification de l'image corporelle (car liée à la perception et non à l'action) du sujet. En effet, le sujet incorpore la main en caoutchouc à son image corporelle et le sujet, trompé, perçoit la main en caoutchouc comme étant la sienne [2] [8] [9]. Ce phénomène a également été observé chez les singes [10].

D'autres expériences effectuées sur les singes ont permis la découverte de neurones visuo-tactiles bimodaux dans leur cerveau qui codent pour une zone autour du corps correspondant au prolongement du schéma corporel : l'espace péripersonnel. En effet, il a été observé que lorsque les singes utilisaient des râtaux pour attraper la nourriture qui n'était pas à leur portée, les radiofréquences visuelles de certains neurones visuo-tactiles bimodaux de la main étaient allongées jusqu'à inclure l'outil. L'outil était alors considéré comme une extension de la main et faisait parti à part entière du schéma corporel des singes. Deux types de neurones ont été mis en évidence : (i) les neurones de type distal qui répondent aux stimuli somatosensoriels de la main et aux stimuli visuels près de la main, et (ii) les neurones de type proximal dont le champ visuel n'est pas centré sur la main mais sur tout l'espace accessible avec l'outil. Une fois l'outil déposé ou tenu de manière passive, le schéma corporel retrouvait peu à peu un aspect "normal" (voir figure 1) [2] [8] [11].

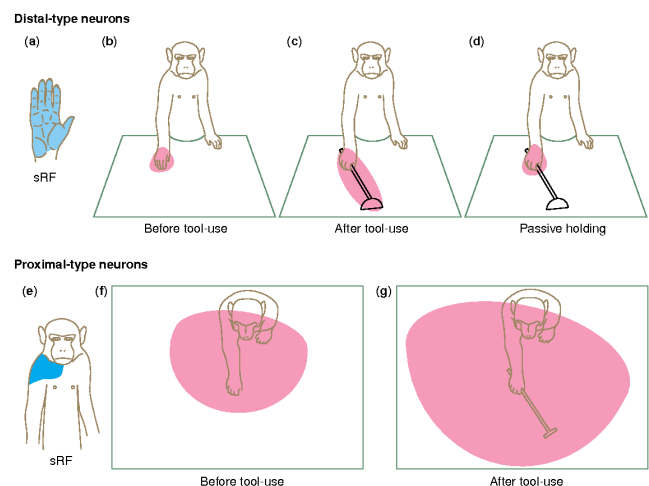


FIGURE 1: Modification de l'espace péripersonnel lors de l'utilisation d'outil (extrait de Maravita et Iriki [11])

L'un des premiers scientifiques à formuler le concept d'un espace particulier entourant le corps est Hediger [12], le directeur du Zoo de Zurich entre 1954 et 1973. En effet, en observant les animaux, il s'est rendu compte que lorsqu'un objet dangereux franchissait la marge de sécurité d'un animal, ce dernier s'enfuyait. Des recherches menées par Rizzolatti et al. (1981) [13] [14] sur les singes leur a permis de voir que lorsque des objets étaient très proches des singes, les neurones bimodaux répondaient de la même façon qu'une simulation visuelle ou tactile. Ils ont ainsi pu délimiter l'espace péripersonnel comme une zone se trouvant à environ un mètre du corps.

Finalement, l'espace péripersonnel peut être défini comme une région entourant immédiatement le corps, caractérisée par un degré élevé d'intégration multisensorielle (visuelle, tactile et auditif) vers laquelle se dirige l'information, et qui diffère des régions plus éloignées. Il constitue une interface privilégiée permettant à l'organisme d'interagir avec les objets [8]. Cet espace est donc assimilé à la fois à une marge de sécurité, à l'espace à notre portée mais également à l'espace dans lequel ont lieu les interactions sociales [15] [16]. Sa représentation peut s'ajouter à la taxonomie spatiale triadique (figure 2) qui peut être décrite selon un critère sensorimoteur : (i) l'espace personnel correspond à l'espace occupé par le corps lui-même, dont la représentation est principalement construite par le biais d'informations proprioceptives et tactiles, et de contribution visuelle sur les parties du corps dans l'espace ; (ii) l'espace extra-personnel, principalement basé sur les entrées visuelles et auditives qui transmettent les informations de l'espace lointain ; (iii) l'espace accessible, dans l'espace extra-personnel, mais à proximité du corps, défini de façon fonctionnelle en fonction de la distance à laquelle un objet peut-être atteint par la main d'un sujet sans bouger.

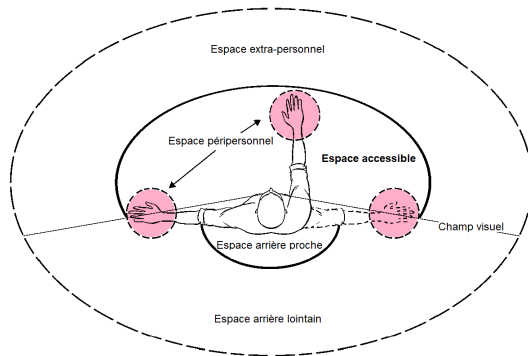


FIGURE 2: Vue d'ensemble des différents espaces de la taxonomie triadique spatiale

Dans leur revue, J. Cléry et S. Ben Hamed [17] résument les avancées récentes sur l'espace péripersonnel faites en neurosciences et suggèrent des pistes d'études pour le futur.

Afin de mieux comprendre ces trois concepts que sont le schéma corporel, l'image corporelle et l'espace péripersonnel, il est intéressant de comprendre quelles parties du cerveau sont impliquées. Des études menées sur ces trois concepts par Sakata et al. [18] [19], ou encore Darmon [20] ont mis en évidence l'implication du cortex pariétal (voir figure 3) dans la représentation spatiale du corps et de l'espace péripersonnel, et sa participation à la planification et au contrôle des mouvements d'atteinte visuo-manuels.

Les différentes caractéristiques et fonctionnalités liées au schéma corporel, à l'image corporelle et à l'espace péripersonnel sont résumées dans le tableau I.

Alors que ces représentations sont en constante évolution et s'adaptent en fonction des actions de l'être humain, de son environnement et des différents stimuli qu'il peut percevoir, grâce à la plasticité cérébrale, la modélisation de ces représentations sur des systèmes robotiques humanoïdes est plus complexe puisque ceux-ci sont décrits par des modèles fixes et analytiques.

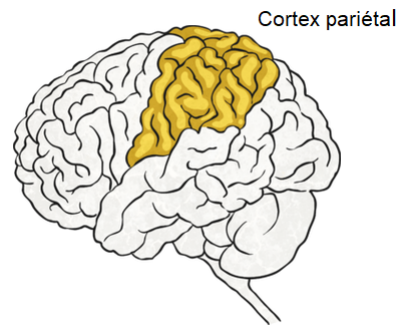


FIGURE 3: Vue latérale du cortex pariétal humain

III. MODÉLISATION EN ROBOTIQUE : ÉTAT DE L'ART

Selon Nguyen [6], le problème de l'apprentissage du schéma corporel chez un robot peut se décomposer en deux sous problèmes : (i) l'identification/la calibration des modèles cinématiques et (ii) l'apprentissage et la cartographie visuomotrices en fonction du type de signaux d'entrée.

De nombreux modèles ont été proposés pour solutionner ces deux problématiques. Roncone et al. [22] s'inspirent du comportement des nouveau-nés et de certains animaux en utilisant l'auto-simulation ou auto-toucher pour "calibrer" leur représentation corporelle. Ils présentent alors un modèle dans lequel le robot forme lui-même sa chaîne cinématique en touchant son propre corps pour apprendre son modèle cinématique mais également sa géométrie (sa surface dans l'espace) de façon autonome. Les données recueillies s'appuient sur les retours tactiles ou de force du robot sur sa peau. Pour mener leurs expériences, Roncone et al. utilisent le robot humanoïde iCub qui possède une peau artificielle dotée de points de détection tactile. La correspondance entre le point de contact prédit à partir de la cinématique directe existante et la position réelle sur la peau artificielle du robot fournit des exemples de données permettant d'affiner sa représentation cinématique, notamment les paramètres de Denavit-Hartenberg, qui décrivent pleinement la cinématique du robot à travers une série de rotations et translations de la base du robot jusqu'à l'effecteur terminal. Cependant des erreurs inobservables, liées aux articulations par exemple, ont limité le taux de réussite de leur approche.

Li et ses collègues [23] utilisent également l'auto-toucher pour l'apprentissage du schéma corporel. Et afin de fournir des informations de position plus précises, ils utilisent un servo-contrôleur en boucle fermée qui maintient la force de contact souhaitée lorsque le robot KUKA LWR fait glisser son doigt équipé de capteurs tactiles à la surface de sa peau. Cela leur permet d'obtenir plus rapidement une grande quantité de données d'entraînement représentant la forme du corps et présente moins de risque d'endommager le robot. Ils exploitent aussi le modèle CAO du robot comme initialisation pour la cinématique du bras. Le problème de calibrage devient un problème d'estimation des moindres carrés. Au cours de leurs expériences, ils ont pu observer que la combinaison de deux contrôleurs, de rétroaction et d'anticipation, permettait d'obtenir des résultats satisfaisants pour suivre l'emplacement

	Schéma corporel [2]	Image corporelle [21]	Espace péripersonnel [15]
Fonctionnalités	Posture Structure corporelle	Perception corporelle Conception corporelle	Actions volontaires Mouvements défensifs
Caractéristiques	Action Plastique Inconscient	Perception Plastique Conscient [4]	Action Plastique Conscient/inconscient
Simulations sensorielles	Proprioceptive Visuelle Tactile Auditive	Proprioceptive Visuelle Tactile Auditive	Proprioceptive Visuelle Tactile Auditive

TABLE I: Différences et similitudes entre schéma corporel, image corporelle et espace péripersonnel

du contact tout en conservant la force de contact souhaitée.

Vicente et al. [24] proposent une nouvelle approche pour obtenir une adaptation automatique du schéma corporel du robot humanoïde iCub, et pour améliorer sa perception et sa motricité sur la base de cette connaissance corporelle : il s'agit d'un problème de coordination main-oeil. Ils s'inspirent donc de l'apprentissage humain, et s'appuient sur un processus de simulation mentale et sur la stéréovision afin de générer des prédictions sur la façon dont la main apparaîtra dans les images de la caméra du robot. Ces prédictions sont ensuite comparées aux images réelles en utilisant des techniques séquentielles de Monte Carlo pour alimenter une méthode d'estimation bayésienne qui permet d'estimer les paramètres du schéma corporel. Celui-ci, mis à jour, améliorera les estimations de la pose de la mains 6D, et pourra ensuite être exploité pour des actions d'atteintes plus précises. Pour cette approche, la connaissance de la structure cinématique du robot, des matrices de transformations ainsi que des paramètres intrinsèques de la caméra doivent être connus.

Dans un autre article [25], Vicente, Zenha et leurs collègues utilisent une simulation réaliste du robot iCub et proposent une technique permettant à celui-ci d'estimer de manière incrémentielle et automatique ses inexactitudes de modèle en touchant des surfaces planes connues (par exemple, des murs) dans son voisinage grâce à l'exploration aléatoire. Il adapte ainsi efficacement son propre schéma corporel en fonction des informations de contact seules. Le problème est formulé sous la forme d'une estimation de paramètre adaptative, via l'utilisation du filtre de Kalman étendu, qui utilise des contraintes planes obtenues à chaque détection de contact. La comparaison de différentes méthodes de mise à jour incrémentielle à travers un vaste ensemble d'expériences, leur a permis de mettre en évidence que les inexactitudes du modèle proposé peuvent être réduites de plus de 80%.

Díaz Ledezma et Haddadin [26] utilisent le principe du premier ordre (First Order Principle, FOP), déduit à partir des équation de Newton-Euler, pour apprendre à la fois le schéma corporel, c'est-à-dire la topologie et la morphologie d'un robot, et sa dynamique inverse, c'est-à-dire ses propriétés inertielles. Pour leurs simulations, ils utilisent deux types de robot : le robot ATLAS et le robot Franka Emika Panda. L'utilisation d'un contrôleur PD permet la génération aléatoire de trajectoires qui vont permettre l'apprentissage des paramètres du FOP. Ils utilisent également les lois physiques ou encore la dynamique des articulations comme contraintes d'optimisation

afin de faciliter l'apprentissage du schéma corporel.

Gama et Hoffmann [27] présentent un modèle d'apprentissage de la représentation spatiale du corps se basant sur les informations somatosensorielles (tactiles et proprioceptives) acquises lors de l'auto-toucher. Pour ce faire, ils utilisent un modèle amélioré d'une carte auto-organisée (Self-Organizing maps, SOM), le MRF-SOM (Maximum Receptive Field). Cette carte prend en entrée les angles des articulations du robot collectés lorsque celui-ci se trouve dans différentes positions, et sa sortie correspond à l'entrée d'un réseau 2D. Les paramètres MRF sont choisis afin d'imiter au mieux les mécanismes neuronaux se produisant dans le cortex sensoriel primaire observés chez les primates. Toutefois les résultats présentés restent très préliminaires.

Ulbrich et ses collègues [28] s'inspirent des PSOMs (Parametrized Self-Organizing Maps) pour apprendre la cinématique inverse dans le cadre du schéma corporel à partir de la configuration des articulations du robot et de la position du membre effecteur. Ils proposent des cartes cinématiques de Bézier (Kinematic Bezier Maps, KB-Maps) afin d'obtenir un problème linéaire plus facilement solvable et dont la cinématique exacte peut être obtenue en l'absence de bruit. Ils montrent après simulation virtuelle et matériel du robot ARMA-IIIa que le modèle est capable d'apprendre plus rapidement une bonne approximation de la cinématique à partir de mouvements articulaires aléatoires avec une optimisation linéaire des moindres carrés à partir de mesures inexacts bruitées.

Fuke et al. [21] s'inspirent du comportement d'imitation des enfants et proposent une méthode d'apprentissage qui permet à un robot d'acquérir une image corporelle pour des parties de son corps qui lui sont invisibles, telles que son visage. Ce modèle associe à la fois la perception spatiale et la perception basée sur les activations de capteurs tactiles et d'images tactiles. Le robot estime les positions invisibles de sa main en se basant sur le jacobien entre le déplacement des angles articulaires et le flux optique de sa main. Ainsi lorsque sa main touche l'un de ses capteurs, le robot associe le capteur à la position estimée de sa main en créant une carte auto-organisée (SOM). Les expériences réalisées montrent que le robot est capable d'apprendre la disposition spatiale des capteurs tactiles de son visage et donc son image corporelle.

Schillaci et al. [29] proposent un modèle qui s'inspire des propriétés d'auto-organisation des zones du cerveau humain, dont les topologies sont structurées par les informations pro-

duites par l'interaction de l'individu avec le monde extérieur, pour coder les représentations internes de l'expérience sensorimotrice pouvant être alimentée par des données visuelles, tactiles et auditives. Ils utilisent donc deux cartes dynamiques auto-organisées (Dynamic Self-Organizing Maps, DSOM), codant les postures des bras et de la tête du robot, liés par des liens Hebbien pour permettre un apprentissage continu et modéliser au mieux la plasticité cérébrale. Ils montrent alors que le robot humanoïde Nao améliore la qualité de sa coordination visuo-motrice lors du processus d'exploration aléatoire et que la cartographie possède une plasticité.

Wijesinghe et ses collègues [30] s'inspirent de la biologie pour permettre à un robot humanoïde de suivre son effecteur terminal en intégrant des indices visuels et proprioceptifs lorsqu'il interagit avec l'environnement. Pour cela, ils utilisent des cartes auto-organisées de sous-espace adaptatif génératif (Generative Adaptive Subspace Self Organizing Maps, GAS-SOM) pour coder les stimuli visuels avant de les associer avec la proprioception pour générer une prédiction des stimuli codés, et permettre de contrôler les mouvements de l'oeil du robot pour le suivi des mouvements de son bras. L'utilisation d'entrées visuelles texturées ont permis de mettre en évidence la robustesse du modèle face aux changements d'apparence de l'effecteur terminal.

Lanillos et Cheng [31] mettent en place un modèle de perception informatique basé sur un traitement prédictif qui permet à tout robot multisensoriel d'apprendre, de déduire et de mettre à jour la configuration de son corps. Basées sur le modèle de bruit additif gaussien et la minimisation de l'énergie libre variationnelle, leurs expériences sur un véritable bras robotique multisensoriel ont montré que le modèle est capable de fournir dynamiquement la configuration corporelle la plus plausible et la localisation de l'effecteur terminal, même lors d'injections de plusieurs sources bruyantes d'informations sensorielles de manière évolutive.

Nguyen et ses collègues [6] proposent une architecture pour permettre à un agent artificiel de développer un soi, c'est-à-dire la connaissance de sa représentation et son espace péri-personnel, en explorant son environnement. Ils considèrent que cela ne peut se faire que si l'agent possède une représentation sensorimotrice multimodale et une association actions/effets bidirectionnelle. La dernière condition est modélisée par deux modules : le prédicteur, un modèle prévisionnel multimodal qui prédit les effets d'une sensation à l'instant $t + 1$ à partir d'une action en cours et de la sensation associée à l'instant t , et le générateur d'action qui génère des actions motrices en tenant compte de l'erreur commise par le prédicteur. Les deux modules apprennent de façon simultanée et leur apprentissage est motivé par la minimisation des erreurs de prédictions et de leurs actions. Les deux modules sont ensuite couplés par un processus d'intégration multisensoriel. Cependant, un problème de surapprentissage a été observé lorsque l'agent n'apprend qu'à partir d'une seule tâche ou que d'un seul environnement.

Abrossimoff et al. [32] s'inspirent des mécanismes neurobiologiques observables dans le cortex pariétal au niveau des neurones pariétaux, qui intègrent les informations visuelles, motrices et proprioceptives grâce à des processus que l'on

pourrait assimiler à des produits. Ils utilisent les réseaux gain-field networks afin de pouvoir réaliser des mouvements pour l'atteinte d'objets dans l'espace péri-personnel et l'apprentissage du schéma corporel. En fournissant au réseau une image visuelle de la position du bras du robot et une image visuelle de la position de la cible à atteindre en entrée, celui-ci est capable de retrouver les ordres moteurs qui permettront de générer ce mouvement. Ceci correspond à la construction d'une direction visuelle privilégiée en coordonnées centrées sur la main, c'est-à-dire la construction du schéma corporel. Les tests réalisés avec un bras manipulateur ont montré de faibles erreurs d'estimation, le réseau est robuste pour la majorité des exemples appris dans l'espace de travail.

Pugach et al. [33] utilisent aussi les gain-field networks dans leurs expériences avec un bras robotique Jaco de Kinova recouvert de peau artificielle qu'ils ont développé. Les entrées visuelles sont acquises par une caméra à fil fixe en hauteur afin de visualiser le mouvement complet du bras. Ils montrent à travers leurs expériences comment les gain-field networks reproduisent les comportements des neurones pariétaux pour encoder dynamiquement le schéma corporel du bras robotique et pour estimer l'orientation relative et la distance des cibles par rapport à lui. Nous verrons plus en détails le fonctionnement des gain-field networks dans la section IV.

Dans un autre article, Abrossimoff et al. [34] utilisent les gated networks, similaires aux gain-field networks, sous forme d'auto-encodeur pour apprendre les transformations visuomotrices dans le but d'atteindre des cibles et représenter l'emplacement du bras du robot. Les gated networks, couramment utilisés par Memisevic [35] pour l'apprentissage du flux optique ; ou encore Droniou et Sigaud pour la modélisation d'une hiérarchie de représentations multimodales à partir des capteurs des robots [36] et l'apprentissage d'un répertoire d'actions tel que l'écriture manuscrite [37], utilisent des réseaux multiplicatifs pour lier des images corrélées les unes aux autres et pour apprendre leurs changements relatifs. La partie encodeur sert à apprendre les différentes transformations visuomotrices en estimant l'ordre moteur engendré par le déplacement visuel, et la partie décodeur permet la reconstruction du signal visuel en anticipant le déplacement visuel et la localisation du bras provoqué par l'ordre moteur. La combinaison des deux parties permet le contrôle du bras robotique en utilisant la génération de son image corporelle. Avec cette méthode, ils parviennent également à construire l'espace péri-personnel du robot. En effet, en approchant la cible suffisamment près du bras du robot, le réseau va l'assimiler à un déplacement visuel du bras et va retourner les ordres moteurs associés.

Roncone et al. [38] proposent un modèle efficace et en temps réel pour l'apprentissage de la représentation distribuée de l'espace autour de son corps en exploitant une peau artificielle et un contact physique entre celle-ci et l'environnement. Les stimuli, sous forme d'événements moteurs ou visuels, sont détectés et enregistrés. Lorsqu'ils correspondent à un contact physique avec la peau artificielle, les taxels mettent à jour leur représentation en remontant dans le temps le stimulus qui approche et en augmentant l'estimation de la prédiction, c'est-à-dire la vraisemblance que le stimulus touche une partie du corps. Ils ont réalisé différentes expériences avec le robot

Article	Type de représentation	Modèle	Robot utilisé	Information sensorielle
Roncone et al. [22]	schéma corporel	chaîne cinématique	robot humanoïde iCub	proprioceptive, tactile
Li et al. [23]	schéma corporel	chaîne cinématique	bras robot KUKA LWR	proprioceptive, tactile
Vicente et al. [24]	schéma corporel		robot humanoïde iCub	proprioceptive, visuelle
Zenha et al. [25]	schéma corporel	filtre de Kalman étendu	simulateur robot iCub	proprioceptive, tactile
Díaz Ledezma et Haddadin [26]	schéma corporel	FOP networks	simulateur ATLAS, Franka Emika	proprioceptive
Gama et Hoffmann [27]	schéma corporel	MRF-SOM	robot humanoïde Nao	proprioception, tactile
Ulbrich et al. [28]	schéma corporel	cartes cinématiques de Bézier KB-Maps	ARMA-IIIa	visuelle, proprioceptive
Fuke et al. [21]	image corporelle	SOM, Jacobien	simulation robot humanoïde	visuelle, tactile, proprioceptive
Schillaci et al. [29]	schéma corporel	DSOM	robot humanoïde Nao	visuelle, proprioceptive
Wijesinghe et al. [30]	schéma corporel	GASSOM	simulateur robot iCub	visuelle, proprioceptive
Lanillos et Cheng [31]	schéma corporel	codage prédictif basé sur la régression de bruit gaussien	bras robotique	visuelle, tactile, proprioceptive
Nguyen et al. [6]	schéma corporel, espace péripersonnel	réseau de neurones prédictif	agent artificiel	visuelle, tactile, proprioceptive
Abrossimoff et al. [32]	schéma corporel, espace péripersonnel	gain-field networks	simulateur bras robotique	visuelle, proprioceptive
Pugach et al. [33]	schéma corporel, espace péripersonnel	gain-field networks	bras robot Jaco de Kinova	visuelle, tactile, proprioceptive
Abrossimoff et al. [34]	schéma corporel, espace péripersonnel	gated networks	bras robot Jaco de Kinova	visuelle, tactile
Rocone et al. [38]	espace péripersonnel	champs visuels réceptifs distribués	robot humanoïde iCub	visuelle, tactile, proprioceptive
Serino et al. [39]	espace péripersonnel	réseau de neurones	/	visuelle, tactile, auditive
Chinellato et al. [40]	espace péripersonnel	réseau de fonctions à base radiale	robot humanoïde	visuelle, proprioceptive
Juett et Kuipers [41]	espace péripersonnel	Graphe	robot Baxter	visuelle, proprioceptive
Nguyen et al. [42]	espace péripersonnel	DNN	simulateur robot iCub	visuelle, tactile, proprioceptive

TABLE II: Récapitulatif des modélisations du schéma corporel et de l'espace péripersonnel en robotique

iCub, tels que de l'évitement d'obstacles pour l'apprentissage d'une marge de sécurité autour du robot, ou l'atteinte d'objets dans l'espace proche.

Serino et al. [39] un réseau de neurones computationnel décrivant la représentation visuelle-tactile de l'espace péripersonnel autour de la main. Le modèle a pu reproduire le comportement caractéristique des neurones, à savoir qu'ils répondent aux stimuli tactiles et visuels présentés à proximité du corps. Il a également pu reproduire l'extension de l'espace péripersonnel grâce à l'utilisation d'outils. Le modèle génère également une prédiction testable pour expliquer comment une telle plasticité dans la représentation de l'espace péripersonnel peut être générée. En effet, ils ont pu constater que la simple présentation de stimuli tactiles et visuels (ou auditifs) proches et synchrones suffisait à étendre la représentation de l'espace péripersonnel, ce qui n'est pas le cas lorsque les stimuli sont asynchrones.

Chinellato et al. [40] s'appuient sur les comportements des primates et des nourrissons pour construire une représentation implicite de l'espace péripersonnel. Ils construisent une carte sensori-motrice 3D partagée de l'environnement, construite

sur un cadre fonctionnel de base radiale, configurée et entraînée par le contrôle coordonné des mouvements des yeux et des bras d'un robot humanoïde. La carte est accessible et modifiable par deux types d'informations : les informations visuelles et proprioceptives (oeil et mouvements des bras). Elle est mise à jour au fur et à mesure en ajustant les poids en fonction des erreurs observées. L'agent doit être en mesure de construire une telle carte visuomotrice de son environnement en 3D, apprenant simultanément à regarder et à atteindre différentes cibles visuelles.

Juett et Kuipers [41] s'inspirent de l'apprentissage précoce chez les nourrissons, et présentent un modèle qui permet à un robot d'apprendre la structure de son espace péripersonnel dans lequel l'atteinte et la saisie ont lieu, avec une connaissance spatiale préalable minimale. Ils représentent l'espace péripersonnel sous forme graphique : chaque noeud représentent un vecteur de sens visuel et un vecteur de sens proprioceptif correspondant à l'état du monde, et les arêtes reliant les noeuds correspondent à la faisabilité du mouvement entre ces deux états. L'apprentissage et la création du graphe se fait par exploration aléatoire de l'espace des vecteurs d'angles

d'articulation. Avec le graphe appris, il est ensuite possible d'utiliser des algorithmes de recherche du plus court chemin reliant l'état actuel à l'état final désiré en utilisant Dijkstra par exemple. Le fait d'exiger que chaque nœud du graphe stocke des images constitue une solution gourmande en mémoire.

Nguyen et al. [42] s'inspirent des découvertes sur le développement des nourrissons et proposent un modèle pour l'atteinte d'objets dans l'espace péripersonnel. Pour cela, ils réalisent leurs expériences sur le simulateur du robot iCub, qui présente des proportions anthropomorphes assimilables à celles d'un enfant, équipé de 2 caméras/yeux dans un arrangement biomimétique et une peau artificielle recouvrant tout son corps. Ils génèrent différents objets, de paramètres et textures modifiables, et par babillage moteur, le robot doit les toucher. Le contact avec l'objet est réussi lorsque celui-ci est détecté par les capteurs tactiles de sa main ou son avant-bras. Ils utilisent un réseau de neurones profond DNN dont les entrées codent les images brutes sans pré-traitement, prises par les yeux, et les articulations de la tête, et les sorties sont la configuration du bras et le toucher. Après apprentissage, le réseau de neurones est capable de déduire avec succès les configurations du bras qui aboutiraient à une atteinte réussie et la prédiction de l'activation tactile, c'est-à-dire quelle partie du corps entrera en contact avec l'objet.

L'ensemble de modèles présentés ainsi que leurs caractéristiques sont résumés dans le tableau II.

IV. MÉTHODE GAIN-FIELD NETWORKS : ARCHITECTURE ET MÉCANISMES

Pour effectuer certaines actions telles que se déplacer ou encore atteindre un point dans l'environnement avec le bras, il est essentiel d'estimer où se situe visuellement le corps dans l'espace. Pour cela, il est nécessaire de corrélérer les signaux sensoriels et moteurs afin d'apprendre les transformations sensorimotrices qui permettront par la suite d'anticiper un contact avec des objets ou des obstacles ou d'estimer la position du corps sans information visuelle par exemple [32] [33] [34].

Dans le cerveau, au niveau du cortex pariétal, des circuits neuronaux sont impliqués dans l'apprentissage des corrélations entre signaux multimodaux (visuel, tactile, auditif, proprioceptif ou encore vestibulaire) : ce sont les neurones "gain-field" ou neurones à modulation de gain. Ceux-ci vont permettre la modulation des neurones moteurs par l'information sensorielle, et ainsi permettre de fusionner les informations multisensorielles afin de construire des représentations centrées sur la main, le bras, la tête ou une autre partie du corps.

En robotique, ces neurones peuvent être modélisés par les gain-field networks présentés par Abrossimoff et al. [32] et Pugach et al [33]. Ces réseaux de neurones sont des instances des réseaux sigma-pi constitués de fonctions à base radiale, prédéfinis paramétriquement ou apprises, qui produisent une somme pondérée de distributions de probabilités conjointes en sortie.

Un neurone gain-field peut être représenté par la figure 4.

La sortie H est une combinaison linéaire du produit des variables d'entrées X et Y , de taille respective n_H , n_X et n_Y .

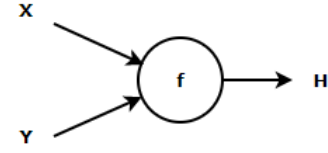


FIGURE 4: Représentation d'un neurone gain-field

La prédiction de H , \hat{H} correspond donc au calcul de toutes les valeurs H_k de H , $k \in n_H$:

$$\forall k, H_k = \sum_i^{n_X} \sum_j^{n_Y} W_{ijk} X_i \cdot Y_j, \quad (1)$$

avec W les poids synaptiques en $n_X \times n_Y \times n_H$. Pour réduire les dimensions de la matrice du réseau gain-field, il est possible de factoriser les variables d'entrées telles que $f_X = \sum W_X X$ et $f_Y = \sum W_Y Y$. La sortie Y , auparavant définie par (1), devient alors :

$$H = \sum W_H f_X \cdot f_Y \quad (2)$$

$$= \sum W_H \left(\sum W_X X \right) \cdot \left(\sum W_Y Y \right) \quad (3)$$

L'erreur globale E est définie comme la distance euclidienne calculée entre H et \hat{H} pour tous les exemples d'entrée.

Pour l'apprentissage des poids synaptiques W , la fonction d'optimisation utilisée est la descente de gradient stochastique classique. Le calcul de chaque sous-réseau X et Y se fait de façon indépendante, avant de calculer H , en utilisant la règle d'apprentissage de Widrow-Hoff :

$$\Delta W_H = \epsilon (H^* - H) (f_X \cdot f_Y) \quad (4)$$

Il est ensuite possible de reconstruire l'une des variables d'entrée X (respectivement Y), à partir de la sortie \hat{H} calculée par le premier réseau, en utilisant la même logique développée dans l'équation (3) et en remplaçant la sortie H par l'entrée \hat{X} (respectivement \hat{Y}) que l'on souhaite obtenir :

$$\hat{X} = \sum W_{\hat{X}} \hat{H} \cdot Y \quad (5)$$

$$= \sum W_{\hat{X}} \left(\sum W_H X \cdot Y \right) \cdot Y \quad (6)$$

L'erreur globale correspond alors à la distance euclidienne entre les variables réelles X (respectivement Y) et celles reconstruites \hat{X} (respectivement \hat{Y}).

L'association du premier réseau gain-field et du second constitue un auto-encodeur (voir figure 5). Celui-ci permet alors de reconstituer une modalité manquante à partir de deux autres modalités connues.

Ici, on considère que X et Y sont des entrées visuelles et H correspond à un ordre moteur. Le premier réseau encodeur réalise donc un mapping des changements visuels entre afin d'obtenir l'ordre moteur correspondant alors que le second réseau decodeur prédit le déplacement visuel induit par l'ordre moteur (c'est-à-dire l'image corporelle).

Dans la section V, on cherchera par exemple à retrouver l'ordre moteur H pour permettre au poignet du robot Reachy de se déplacer du point visuel X au point visuel Y .

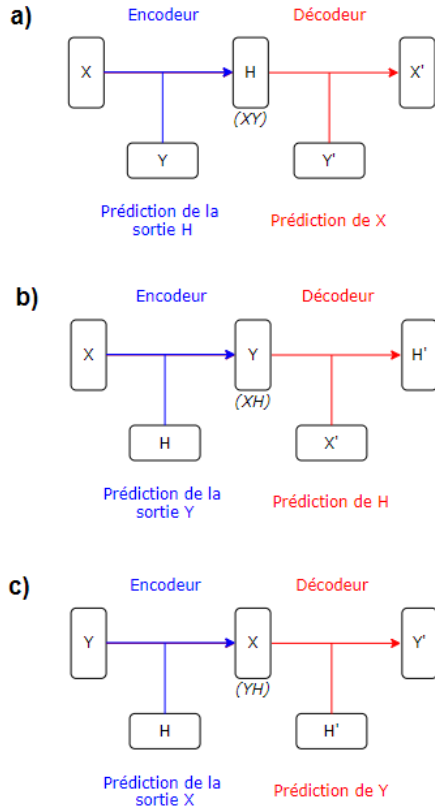


FIGURE 5: Architecture du gain-field network auto-encodeur

V. EXPÉRIENCES ENVISAGÉES

Nous réaliserons nos expériences sur le simulateur ainsi que sur le robot humanoïde Reachy fabriqué par Pollen Robotics. Celui-ci est composé d'un buste, d'un micro et d'un haut-parleur, d'une tête rotative équipée de deux antennes et de deux caméras : l'une pour observer l'environnement et l'autre pour se concentrer sur la tâche de manipulation. Le robot Reachy possède également deux bras bio-inspirés avec à leur bout une pince.

Dans le cadre de ce projet, nous nous focaliserons uniquement sur le mouvement du poignet du bras du robot, et sur les informations visuelles recueillies par les caméras. En reprenant les équations données dans la section IV et la figure 5, on associe X et Y aux images prises par les caméras à deux instants différents, et H à l'ordre moteur nécessaire pour induire le mouvement du poignet.

Nous réaliserons trois expériences décrites ci-dessous :

A. Expérience 1 : apprentissage des transformations visuomotrices

La première expérience consiste à apprendre au réseau gain-field toutes les transformations visuomotrices possibles du poignet du robot Reachy. Pour cela, on laisse le robot effectuer des mouvements du poignet de façon aléatoire. On capture ensuite des images à partir des caméras à intervalle de temps régulier dt déterminé au préalable (voir figure 6a). Les deux images, séparées par un intervalle de temps dt , seront ensuite envoyées au premier réseau gain-field décrit par l'équation 1.

Celui-ci doit être capable de prédire en sortie l'ordre moteur H qui a permis au poignet de passer de la configuration t à la configuration $t + dt$ (voir figure 6b).

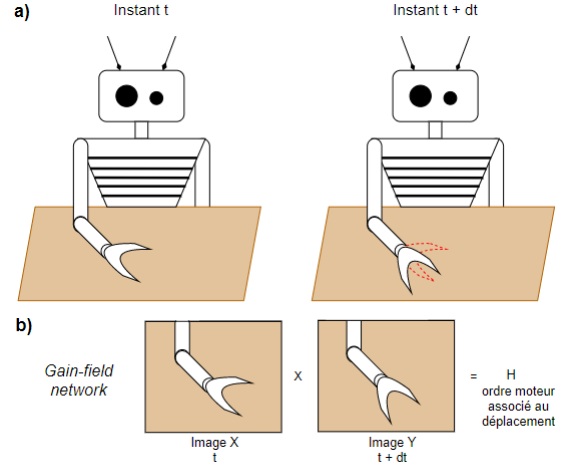


FIGURE 6: Apprentissage des transformations visuomotrices par le réseau de neurones gain-field

On pourra alors observer le comportement du premier réseau de neurones après la phase d'apprentissage, ainsi que la justesse des prédictions des ordres moteurs H en calculant les erreurs de prédiction.

B. Expérience 2 : reconstruction de l'image corporelle dans l'espace visuel

Dans une seconde expérience, on utilisera la seconde partie du réseau gain-field, décrite par l'équation 6. Celle-ci correspond à l'inverse de l'expérience décrite précédemment. A partir de la position de la main à un instant t donné et d'un ordre moteur, le réseau gain-field doit être capable d'estimer l'orientation du poignet à l'instant $t + 1$ (voir figure 7). Cette expérience permettra également d'obtenir une représentation de l'espace péripersonnel associé au poignet du robot Reachy.

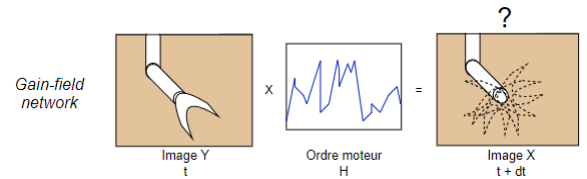


FIGURE 7: Reconstruction de l'image corporelle par le réseau de neurones gain-field

A nouveau, on pourra observer le comportement du réseau en calculant les erreurs de prédiction. Et on proposera des optimisations possibles afin de minimiser les erreurs de prédictions.

C. Expérience 3 : imitation

Une fois que le robot aura appris toutes les transformations visuomotrices, et aura une connaissance de son schéma corporel et son espace accessible, on lui fera effectuer des mouvements du poignet par mimétisme. Le principe est le

suivant : on se place devant le robot et on lui présente une main dont le poignet a une certaine orientation. Celui-ci doit être capable de reconnaître le mouvement observé parmi les exemples de sa base d'apprentissage et le reproduire (voir figure 8b). L'orientation du poignet du robot Reachy peut être identifiée à celle de la main comme le montre la figure 8a.

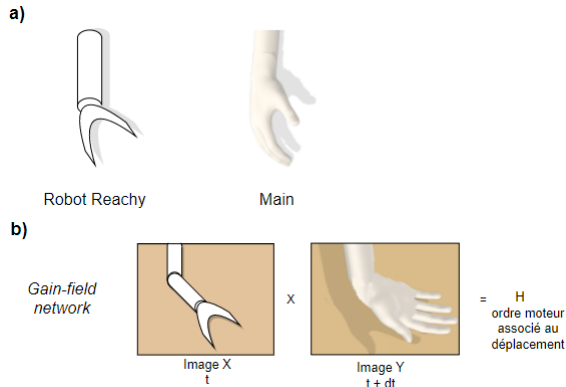


FIGURE 8: Imitation de mouvements visuels par le réseau de neurones gain-field

Cette dernière expérience validera la capacité du réseau gain-field à prédire correctement les transformations visuomotrices nécessaires pour effectuer un mouvement visuel.

RÉFÉRENCES

- [1] H. Head and G. Holmes. Sensory disturbances from cerebral lesions. *Brain*, 34 :102–254, 1911.
- [2] M. Hoffmann, H. Gravato Marques, A. Hernandez Arieta, H. Sumioka, M. Lungarella, and R. Pfeifer. Body Schema : A review. *IEEE Transactions On Autonomous Mental Development*, 2(4) :304–324, 2010.
- [3] P. Schilder. *The Image and Appearance of the Human Body : Studies in the Constructive Energies of the Psyche*. K. Paul, Trench, Trubner Company Limited, 1935.
- [4] F. de Vignemont. Body schema and body image - pros and cons. *Neuropsychologia*, 48(3) :669–680, 2010.
- [5] G. Orioli, A. Santoni, D. Dragovic, and T. Farroni. Identifying peripersonal space boundaries in newborns. *Scientific Reports*, 9(1) :1–11, 2019.
- [6] D.H.P. Nguyen, Y.K. Georgie, E. Kayhan, M. Eppe, V.V. Hafner, and S. Wermter. Sensorimotor representation learning for active self emergence in robots : A model survey. 2020.
- [7] M. Botvinick and J. Cohen. Rubber hands "feel" touch that eyes see. *Nature*, 391(6669) :756, 1998.
- [8] L. Cardinali, C. Brozzoli, and A. Farnè. Peripersonal Space and Body Schema : Two Labels for the Same Concept? *Brain Topogr*, 21(3-4) :252–260, 2009.
- [9] N.P. Holmes and C. Spence. The body schema and multisensory representation(s) of peripersonal space. *Cognitive Processing*, 5(2) :94–105, 2004.
- [10] M.S.A. Graziano. Where is my arm? The relative role of vision and proprioception in the neuronal representation of limb position. *Proceedings of the National Academy of Sciences*, 18(2) :10418–21, 1999.
- [11] A. Maravita and A. Iriki. Tools for the body (schema). *TRENDS in Cognitive Sciences*, 8(2) :79–86, 2004.
- [12] H. Hediger. *Studies of the psychology and behaviour of captive animals in zoos and circuses*. New York : Criterion Books, 1955.
- [13] G. Rizzolatti, C. Scandolara, and R. Camarda M. Gentilucci. Response properties and behavioral modulation of 'mouth' neurons of the postarcuate cortex (area 6) in macaque monkeys. *Brain Research*, 225 :421–424, 1981a.
- [14] G. Rizzolatti, C. Scandolara, and M. Gentilucci M. Matelli. Afferent properties of periarculate neurons in macaque monkeys : I. somatosensory responses. *Behavioral Brain Research*, 2 :125–146, 1981b.
- [15] E. Vagnoni and M.R. Longo. Peripersonal space : its function, plasticity and neural basis. 2019.
- [16] M. Hoffmann, P. Lanillos, L. Jamone, A. Pitti, and E. Somogyi. Editorial : Body Representations, Peripersonal Space, and the Self : Humans, Animals, Robots. *Frontiers in Neurorobotics*, 14 :35, 2020.
- [17] J. Cléry and S. Ben Hamed. Frontier of Self and Impact Prediction. *Frontiers in Psychology*, 9, 2018.
- [18] H. Sakata, M. Taira, M. Kusunoki, A. Murata, and Y. Tanaka. The TINS Lecture. The parietal association cortex in depth perception and visual control of hand action. *Trends in Neurosciences*, 20(8) :350–7, 1997.
- [19] H. Sakata, M. Taira, M. Kusunoki, A. Murata, K. Tsutsui, Y. Tanaka, W.N. Shein, and Y. Miyashita. Neural coding of 3D features of objects for hand action in the parietal cortex of the monkey. *Philosophical Transactions of The Royal Society B Biological Sciences*, 353(1373) :1363–73, 1998.
- [20] C. Darmon. Réactivité visuo-manuelle, Contrôle du geste et Expertise sensorimotrice : étude en IRM fonctionnelle événementielle. pages 1–331, 2006.
- [21] S. Fuke, M. Asada, and M. Ogino. Body Image Constructed from Motor and Tactile Images with Visual Information. *International Journal of Humanoid Robotics*, 4(2) :347–364, 2007.
- [22] A. Roncone, M. Hoffmann, U. Pattacini, and G. Metta. Automatic kinematic chain calibration using artificial skin : self-touch in the iCub humanoid robot. *Robotics and Automation (ICRA), 2014 IEEE International Conference on*, pages 2305–2312, 2014.
- [23] Q. Li, R. Haschke, and H. Ritter. Towards Body Schema Learning using Training Data Acquired by Continuous Self-touch. *IEEE-RAS International Conference on Humanoid Robots*, 2015-Decem :1109–1114, 2015.
- [24] P. Vicente, L. Jamone, and A. Bernardino. Online Body Schema Adaptation Based on Internal Mental Simulation and Multisensory Feedback. *Frontiers Robotics AI*, 3(MAR), 2016.
- [25] R. Zenha, P. Vicente, L. Jamone, and A. Bernardino. Incremental adaptation of a robot body schema based on touch events. *2018 Joint IEEE 8th International Conference on Development and Learning and Epigenetic Robotics, ICDL-EpiRob 2018*, pages 119–124, 2018.
- [26] F. Díaz Ledezma and S. Haddadin. FOP Networks for Learning Humanoid Body Schema and Dynamics. *IEEE-RAS International Conference on Humanoid Robots*, 2018-Novem :1121–1127, 2018.
- [27] F. Gama and M. Hoffman. The homunculus for proprioception : Toward learning the representation of a humanoid robot's joint space using self-organizing maps. *arXiv preprint arXiv :1909.02295*, 2019.
- [28] S. Ulbrich, V. Ruiz de Angulot, and T. Asfour et al. Rapid Learning of Humanoid Body Schemas with Kinematic Bezier Maps. *9th IEEE-RAS International Conference on Humanoid Robots, HUMANOIDS09*, pages 431–438, 2009.
- [29] G. Schillaci, V.V. Hafner, and B. Lara. Online Learning of Visuo-Motor Coordination in a Humanoid Robot. A Biologically Inspired Model. *IEEE ICDL-EPIROB 2014 - 4th Joint IEEE International Conference on Development and Learning and on Epigenetic Robotics*, pages 130–136, 2014.
- [30] L.P. Wijesinghe, J. Triesch, and B.E. Shi. Robot end effector tracking using predictive multisensory integration. *Frontiers in Neurorobotics*, 12(October) :1–16, 2018.
- [31] P. Lanillos and G. Cheng. Adaptive robot body learning and estimation through predictive coding. *IEEE International Conference on Intelligent Robots and Systems*, pages 4083–4090, 2018.
- [32] J. Abrossimoff, A. Pitti, and P. Gaussier. Visual Learning for Reaching and Body-Schema with Gain-Field Networks. *IEEE International Joint Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob)*, 2018.
- [33] G. Pugach, A. Pitti, O. Tolochko, and P. Gaussier. Brain-inspired coding of robot body schema through visuo-motor integration of touched events. *Frontiers in Neurorobotics*, 13(March), 2019.
- [34] J. Abrossimoff, A. Pitti, and P. Gaussier. Visuo-Motor Control Using Body Representation of a Robotic Arm with Gated Auto-Encoders.
- [35] R. Memisevic. Learning to relate images. *IEEE Transactions on Software Engineering*, 35(8) :1829–1846, 2013.
- [36] O. Sigaud and A. Droniou. Towards Deep Developmental Learning. *IEEE Transactions on Cognitive and Developmental Systems*, 8(2) :99–114, 2014.

- [37] A. Droniou, S. Ivaldi, and O. Sigaud. Learning a Repertoire of Actions with Deep Neural Networks. *4th International Conference on Development and Learning and on Epigenetic Robotics*, 2014.
- [38] A. Roncone, M. Hoffmann, U. Pattacini, L. Fadiga, and G. Metta. Peripersonal Space and Margin of Safety around the Body : Learning Visuo-Tactile Associations in a Humanoid Robot with Artificial Skin. *PLOS ONE*, 11(10) :e0163713, 2016.
- [39] A. Serino, E. Canzoneri, M. Marzolla, G. di Pellegrino, and E. Magosso. Extending peripersonal space representation without tool-use : evidence from a combined behavioral/computational approach. *Frontiers in Behavioral Neuroscience*, 9, 2015.
- [40] E. Chinellato, M. Antonelli, B.J. Grzyb, and A.P. del Pobil. Implicit mapping of the peripersonal space by gazing and reaching. *IEEE Transactions on Autonomous Mental Development*, 3(1) :43–53, 2011.
- [41] J. Juett and B. Kuipers. Learning to Reach by Building a Representation of Peri-Personal Space. *IEEE-RAS 16th International Conference on Humanoid Robots (Humanoids)*, pages 1141–1148, 2016.
- [42] P.D.H. Nguyen, M. Hoffmann, U. Pattacini, and G. Metta. Reaching development through visuo-proprioceptive-tactile integration on a humanoid robot - a deep learning approach. *IEE 9th International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob)*, pages 163–170, 2019.