

Stage M2

ACtive Multimodal mErging: from psychophysics to computational modeling to robotics

Fusion bio inspirée de données sur robot

Contexte

La fusion de données est une question cruciale en intelligence artificielle et en sciences cognitives. En effet, elle doit être résolue par tout agent, biologique ou artificiel, qui perçoit son environnement à travers plusieurs capteurs différenciés et complémentaires. Le projet ACME vise à mieux comprendre la perception chez l'humain afin d'améliorer les systèmes d'IA pour des agents plus autonomes. Il se positionne dans le cadre des théories sensori-motrices [7] qui s'appuient sur un certain nombre d'indices provenant des neurosciences, de la psychologie, de la philosophie, ... Plus particulièrement, nous considérons la théorie des contingences sensori-motrices [8] qui propose un cadre unifié fondé sur 2 principes :

1. L'apprentissage de corrélations sensori-motrices qui pose en particulier la question de l'intégration multi-sensorielle, dès lors que l'on considère plusieurs capteurs, et même multimodale en considérant l'action. Une des questions centrales est de savoir quand et comment fusionner les informations disponibles de manière autonome en s'appuyant par exemple sur l'apprentissage passé des relations entre les différentes données [6].
2. La perception active qui formalise que l'être humain interagit avec son environnement non seulement pour atteindre ses objectifs ou maintenir son intégrité physique, mais également afin d'obtenir plus d'informations sur son environnement, par exemple via les saccades oculaires [5, 4]. Cela nécessite à la fois de disposer d'un modèle prédictif de la relation à l'environnement (en conjonction avec l'apprentissage de représentation sensori-motrices) et de décider dans quelle(s) modalité(s) et dans quelle(s) zone(s) de l'espace aller chercher la prochaine information pertinente.

Lors d'un projet précédent, nous avons réalisé des expériences de psychophysique basées sur un paradigme de ventriloquie. On présente aux sujets un stimulus visuel (patch de Gabor de différents niveaux de contraste ou nuages de points avec densité variable) et un stimulus auditif (bruit rose de différents niveaux d'intensité dans un bruit blanc constant). Les sujets produisent des erreurs de localisation du son en direction du stimulus visuel (nommées biais crossmodal) si celui-ci est présenté de manière synchrone mais à une position spatiale différente du stimulus auditif. Ce type de résultats peut être modélisé [2] utilisant le paradigme des champs neuronaux dynamiques (DNF) [1], celui-ci étant compatible avec l'intégration d'anticipations sur la dynamique de l'environnement [9].

Sujet

L'objectif de ce stage sera de porter ces modèles sur robot Pepper disponible au LIRIS (les DNF ayant déjà été utilisés dans un cadre robotique, voir [10] par exemple), avec des stimuli auditifs et visuels simples, inspirés de ceux des expériences psychophysiques. L'objectif sera non seulement d'évaluer les architectures en condition d'interaction réelle mais également d'étudier plus en profondeur la dimension active de la perception dans l'estimation de la pertinence de chaque modalité. Cela permettra ainsi de proposer une solution de fusion multimodale originale qui pourra être éventuellement couplé avec des approches basées sur l'apprentissage profond pour traiter des stimuli plus complexes [3].

Profil

Les compétences suivantes sont indispensables :

- master en intelligence artificielle, machine learning ou équivalent
- bonne capacité de programmation (Python, ...)
- autonomie
- curiosité scientifique et interdisciplinaire
- esprit d'équipe (2 autres stages auront lieu en parallèle sur la modélisation des données psychophysiques, conditions active et passive)

Durée

Le stage commencera en Février-Mars pour une durée de 5-6 mois.

Gratification

4.05€/h, 35h/semaine (i.e. environ 580€/mois)

Localisation

Laboratoire LIRIS, Lyon, France. Des déplacements ponctuels au LJK (Grenoble) pourront avoir lieu.

Encadrants

- Mathieu Lefort : MCF au LIRIS
- Jean-Charles Quinton : MCF au LJK

Candidature

Merci d'envoyer un CV, une lettre de motivation et les relevés de notes de l'année en cours et précédente à Mathieu Lefort (mathieu.lefort@liris.cnrs.fr) et Jean-Charles Quinton (quintonj@univ-grenoble-alpes.fr).

Références

- [1] Shun-ichi Amari. Dynamics of pattern formation in lateral-inhibition type neural fields. *Biological cybernetics*, 27(2) :77–87, 1977.
- [2] Simon Forest, Jean-Charles Quinton, and Mathieu Lefort. A dynamic neural field model of multimodal merging : application to the ventriloquist effect. *Neural Computation*, 2022.
- [3] Simon Forest, Jean-Charles Quinton, and Mathieu Lefort. Champ neuronal et apprentissage profond de topologies pour la fusion multimodale. In *CNIA 2023-Conférence Nationale en Intelligence Artificielle, PFIA*, pages 40–49, 2023.
- [4] Karl Friston, Jérémie Mattout, and James Kilner. Action understanding and active inference. *Biological cybernetics*, 104(1) :137–160, 2011.
- [5] David Kirsh and Paul Maglio. On distinguishing epistemic from pragmatic action. *Cognitive science*, 18(4) :513–549, 1994.
- [6] Mathieu Lefort, Yann Boniface, and Bernard Girau. Somma : Cortically inspired paradigms for multimodal processing. In *The 2013 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE, 2013.
- [7] Matteo Mossio and Dario Taraborelli. Action-dependent perceptual invariants : From ecological to sensorimotor approaches. *Consciousness and cognition*, 17(4) :1324–1340, 2008.
- [8] J Kevin O’regan and Alva Noë. A sensorimotor account of vision and visual consciousness. *Behavioral and brain sciences*, 24(5) :939–973, 2001.
- [9] Jean-Charles Quinton and Bernard Girau. Predictive neural fields for improved tracking and attentional properties. In *The 2011 International Joint Conference on Neural Networks*, pages 1629–1636. IEEE, 2011.
- [10] Gregor Schöner and John P Spencer. *Dynamic thinking : A primer on dynamic field theory*. Oxford University Press, 2016.