

## Stage M2

# ACtive Multimodal mErging: from psychophysics to computational modeling to robotics

## Modélisation de données psychophysiques de fusion de données

### 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. En particulier, nous avons mesuré sur 45 participants les effets de la fiabilité des stimuli sensoriels et des saccades oculaires (dans une perspective de perception active) sur le biais crossmodal.

### Sujet

L'objectif de ce stage consiste en la modélisation de la condition passive (i.e. sans saccade) avec des champs neuronaux dynamiques (DNF) [1] et en particulier l'étude de l'influence des hyperparamètres sur l'adéquation avec les données. Les DNF permettent de modéliser l'activité neuronale par champ moyen et offrent ainsi des possibilités de modélisation plus riche et fine des mécanismes décisionnels, avec la capacité de simuler la dynamique de décisions individuelles, à l'inverse de modèles descriptifs s'intéressant davantage aux distributions sur des centaines d'essais.

Il s'agit de l'extension directe du travail fait dans [3] mais appliqué aux nouvelles données psychophysiques que nous avons à notre disposition. Cette modélisation computationnelle se fera conjointement avec une modélisation statistique fréquentiste et bayésienne des données, classiquement utilisée dans le cadre multi-sensoriel [2], réalisée au CRNL.

## Profil

Les compétences suivantes sont indispensables :

- master en intelligence artificielle, machine learning, science cognitives ou équivalent
- bonne capacité de programmation (Python, ...)
- autonomie
- curiosité scientifique et interdisciplinaire
- esprit d'équipe (2 autres stages auront lieu en parallèle, un sur la modélisation de la condition active, un sur l'utilisation de ce type de modèle sur robot)

## 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 CRNL (Bron) et au LJK (Grenoble) pourront avoir lieu.

## Encadrants

- Mathieu Lefort : MCF au LIRIS
- Jean-Charles Quinton : MCF au LJK
- Marie Avillac : MCF au CRNL

## 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](mailto:mathieu.lefort@liris.cnrs.fr)) et Jean-Charles Quinton ([quintonj@univ-grenoble-alpes.fr](mailto: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] Marc O Ernst and Martin S Banks. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415(6870) :429–433, 2002.
- [3] Simon Forest, Jean-Charles Quinton, and Mathieu Lefort. A dynamic neural field model of multimodal merging : application to the ventriloquist effect. *Neural Computation*, 2022.
- [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.