

PROGRAMME INTITUTS ET INITIATIVES
Appel à projet – campagne 2021
Proposition de projet de recherche doctoral (PRD)
SCAI - Sorbonne Center of Artificial Intelligence

Intitulé du projet de recherche doctoral (PRD): Induire un toucher socio-affectif par simulation visuo-auditive lors d'une interaction humain-agent

Directeur.rice de thèse porteur.euse du projet (titulaire d'une HDR) :

NOM : **Auvray** Prénom : **Malika**
Titre : Chargé de Recherche ou
e-mail : malika.auvray@sorbonne-universite.fr
Adresse professionnelle : ISIR, 4 place Jussieu, 75005 Paris
(site, adresse, bât., bureau)

Unité de Recherche :

Intitulé : Institut des Systèmes Intelligents et de Robotique
Code (ex. UMR xxxx) : UMR7222

École Doctorale de rattachement de l'équipe (future école doctorale du.de la doctorant.e) : ED158-Cerveau, cognition, comportement

Doctorant.e.s actuellement encadré.e.s par la.e directeur.rice de thèse (préciser le nombre de doctorant.e.s, leur année de 1^e inscription et la quotité d'encadrement) : 1 doctorant inscrit en 2020, encadrement à 50%

Co-encadrant.e :

NOM : **Pelachaud** Prénom : **Catherine**
Titre : Directeur de Recherche ou HDR
e-mail : catherine.pelachaud@upmc.fr

Unité de Recherche :

Intitulé : Institut des Systèmes Intelligents et de Robotique
Code (ex. UMR xxxx) : UMR 7222

École Doctorale de rattachement : **Choisissez un élément :**
Ou si ED non Alliance SU :

Doctorant.e.s actuellement encadré.e.s par la.e co-directeur.rice de thèse (préciser le nombre de doctorant.e.s, leur année de 1^e inscription et la quotité d'encadrement) : 5 doctorants dont : 2 à 100% (2017, 2018), 1 à 70% (2018), 2 à 50% (2020, 2020)

Co-encadrant.e :

NOM :

Prénom :

Titre : Choisissez un élément : ou

HDR

e-mail :

Unité de Recherche :

Intitulé :

Code (ex. UMR xxxx) :

Choisissez un élément :

École Doctorale de rattachement :

Ou si ED non Alliance SU :

Doctorant.e.s actuellement encadré.e.s par la.e co-directeur.rice de thèse (préciser le nombre de doctorant.e.s, leur année de 1^e inscription et la quotité d'encadrement) :

Cotutelle internationale : Non Oui, précisez Pays et Université :

Selon vous, ce projet est-il susceptible d'intéresser une autre Initiative ou un autre Institut ?

Non Oui, précisez Choisissez l'institut ou l'initiative :

Description du projet de recherche doctoral

Contexte. Le toucher est une composante centrale des interactions sociales et est primordial pour la communication non-verbale. Il permet de transmettre des signaux socio-affectifs, comme le réconfort et un sentiment de présence, qui ne peuvent pas être transmis par d'autres modalités sensorielles (Hertenstein et al., 2006). Cependant, les interactions tactiles sont de plus en plus limitées avec l'augmentation de la communication à distance. Ceci est décuplé dans le contexte particulier qu'est la pandémie COVID-19, où la distance physique est devenue la norme et par conséquent où le toucher est déconseillé. Il existe un besoin croissant de soutien social, de réduction du stress et du sentiment d'isolement ; aspects qui sont connus pour être diminués grâce aux interactions tactiles (Morrison, 2016). Il apparaît ainsi opportun de permettre des interactions socio-affectives à distance. L'objectif de ce projet est de susciter de tels sentiments de toucher socio-affectif, en l'absence de stimulations tactiles, lors d'interactions avec des agents virtuels. Cette recherche ouvrira la voie à l'amélioration des futurs systèmes de communication en permettant des interactions tactiles affectives à distance.

Il existe des projets visant à induire une sensation de toucher chez les humains lorsqu'ils interagissent avec des objets (Lecuyer, 2017). D'autres projets ont pour objectif de simuler un toucher social d'un agent virtuel vers un humain grâce à des interfaces haptiques (e.g., Huisman et al., 2014). En France, le projet ANR "Social Touch" développe des systèmes pour simuler le toucher d'un humain sur un agent virtuel dans des environnements interactifs (Boucaud et al., 2021). Cependant, aucun travail n'a encore été mené sur la manière de transmettre artificiellement un toucher socio-affectif à un humain (que ce soit par un agent virtuel ou par un humain distant) sans interfaces haptiques complexes. Dans ce projet de thèse, notre objectif est d'induire un sentiment d'être touché sans dispositif, par d'autres modalités sensorielles. Bien qu'ambitieux, cet objectif trouve des racines prometteuses dans trois ensembles de résultats scientifiques : 1) l'augmentation sensorielle, notamment liée aux correspondances intermodales, 2) les modifications perceptives tactiles induites par les représentations visuelles et 3) l'interface humain-machine, en particulier les agents virtuels autonomes et les modèles de décision qui les sous-tendent. Nous détaillerons ces trois aspects tour à tour ci-dessous.

Premièrement, l'augmentation sensorielle vise à améliorer la perception dans une modalité sensorielle par le biais d'autres modalités sensorielles. De nombreux travaux ont été menés dans notre équipe sur ces techniques (e.g., Auvray, 2019 ; Deroy et al., 2016 ; Kirsch, Job, & Auvray, 2020). Nous nous appuyons sur une correspondance intermodale audio-tactile écologique, qui commence tout juste à être étudiée dans notre équipe et qui apparaît comme la plus pertinente pour induire auditivement une expérience tactile socio-affective. Plus précisément, le contact peau à peau crée des vibrations qui peuvent être enregistrées au moyen d'accéléromètres (Kirsch et al., 2020). Les vibrations obtenues peuvent ensuite être amplifiées au moyen de traitements de signaux classiques, puis être rejouées auditivement. Nous avons commencé à établir une base de données des traductions auditives d'interactions tactiles et étudions actuellement les seuils de discriminations de ces "touchers auditifs". Le travail de thèse s'appuiera sur cette base de données. Deuxièmement, des travaux montrent que la perception du toucher peut être modulée par des représentations visuelles de l'interaction tactile. Par exemple, les seuils de détection tactiles sont augmentés par la présentation simultanée de vidéos montrant la main touchée (Schaefer et al., 2005). Les stimuli visuels non seulement augmentent la perception tactile, mais ils donnent aussi l'impression que le toucher provient d'une autre localisation (e.g., Pavani et al., 2000). Le travail de thèse s'appuiera sur ces travaux en variant la représentation visuelle des agents virtuels et leurs comportements multimodaux. Troisièmement, dans le cadre d'interactions avec des agents virtuels, des travaux de notre équipe ont dégagé des modèles computationnels (Biancardi et al., 2019 ; Dermouche &

Pelachaud, 2020) : modèle de décision qui prend en compte la relation entre les interlocuteurs (humain et agent), la fonction communicative qui doit être transmise (e.g., réconforter, attirer l'attention), l'état émotionnel de l'agent, celui de l'humain tel que perçu par l'agent, ainsi que la sensibilité au toucher de l'humain, afin de décider si l'agent doit toucher l'humain et avec quel type de toucher (Boucaud et al., 2021).

Objectif scientifique. L'objectif de ce projet de thèse est d'étudier jusqu'où il est possible d'induire un toucher socio-affectif lors d'interactions avec des agents 3D animés en réalité virtuelle et ce, uniquement à partir de stimulations auditives et visuelles.

Tâche 1 : Traduction auditive du toucher. La première partie de la thèse sera consacrée à l'étude des stimuli auditifs correspondant à la traduction de mouvements tactiles. Il s'agira ici, à partir de notre base de données de traduction auditive d'interactions tactiles d'étudier l'interprétation et la valence attribuée à ces stimuli. Dans un second temps, en collaboration avec Frédéric Bevilacqua (équipe STMS, IRCAM), il s'agira de dégager les possibilités de traitements du signal auditif appropriés pour avoir un signal optimal avec un contenu émotionnel. Divers types de sonification seront implémentés et comparés, comme des sonifications avec des liens directs entre signaux détectés et paramètres sonores (Boyer et al., 2020) ou des sonifications avec des liens métaphoriques entre toucher et retour sonore (Bevilacqua et al., 2017), récemment utilisées dans des cas d'exercices physiques ou de rééducation. Au niveau méthodologique, dans cette tâche, ainsi que dans les deux suivantes, les types de mesures enregistrées sur l'interprétation des signaux et l'interaction incluront des mesures comportementales, physiologiques (électrodermales et électromyographique faciales) et subjectives.

Tâche 2 : Traduction audio-visuelle du toucher. Dans cette seconde partie, les stimuli auditifs seront couplés aux agents virtuels de la plateforme Greta. L'objectif est de mesurer l'intensité, la valence et l'interprétation des signaux audio-visuels en fonction des représentations visuelles. En particulier, nous manipulerons les paramètres de l'agent virtuel : expression du visage, regard, dynamique du mouvement du toucher (amplitude, vitesse, accélération) afin de varier la cohérence entre le retour visuel et les signaux auditifs (Teyssier et al., 2020). Ceci permettra de déterminer comment ces paramètres impactent à la fois l'interprétation et la valence des signaux. Les participants visualiseront l'agent en réalité virtuelle avec un casque HMD. Nous manipulerons aussi la représentation du participant (perspectives en 1^{ère} et 3^{ème} personne). Dans toutes ces conditions, le participant est passif et le contexte d'interaction demeure neutre.

Tâche 3 : Interaction en temps réel avec l'agent. Dans cette troisième partie, les participants seront placés en contexte d'interaction en temps-réel avec l'agent animé en réalité virtuelle. Les agents de la plateforme Greta seront contrôlés par un modèle de décision (Boucaud et al., 2021) qui permet de décider si l'agent virtuel doit toucher le participant et comment, comme décrit plus haut. L'objectif est de mesurer l'intensité, la valence et l'interprétation des signaux audiovisuels en faisant varier des paramètres de l'agent tels que son état émotionnel (exprimé par son expression faciale, son regard ou encore sa posture), ses fonctions communicatives (e.g., attirer l'attention, réconforter). Ces évaluations permettront aussi de mesurer l'impact du comportement de l'agent sur la perception qu'en a le participant et sur la qualité de l'interaction. Les mesures nous permettront de définir des directives pour mieux adapter le modèle : sur le choix du type de toucher de l'agent (modèle de décision), sur le comportement multimodal de l'agent et sur les paramètres acoustiques du toucher.

Adéquation à l'initiative et à l'institut : Comme décrit ci-dessus, l'approche scientifique nécessite des expertises en neurosciences cognitives (M. Auvray), en IA, en particulier en interaction humain-agent (C. Pelachaud), ainsi qu'en systèmes interactifs sonores et sonification (collaboration avec F. Bevilacqua). L'expertise des chercheurs dans leurs domaines respectifs ainsi que leurs complémentarités assurent une collaboration fructueuse vers des résultats fondamentaux et des

applications importantes. Le caractère intrinsèquement interdisciplinaire du projet le place au cœur de la forte exigence d'interdisciplinarité de SCAI. De plus, ce projet s'inscrit dans un des défis de SCAI : les interactions humain-machine. Les applications possibles de ce projet seront nombreuses et diverses, de la recherche fondamentale sur la perception multisensorielle et les interactions tactiles, au développement de la communication tactile avec des agents virtuels animés, avec un intérêt brûlant pour les interactions tactiles à distance en période d'éloignement physique.

Encadrement. **Malika Auvray**, chercheuse CNRS à l'ISIR (UMR 7222) est spécialiste de la perception multisensorielle et de la substitution sensorielle qu'elle aborde sous l'angle des neurosciences cognitives. Elle a développé des systèmes de substitution sensorielle visuo-tactiles et visuo-auditifs. Elle sera l'encadrante principale de la thèse et supervisera les aspects scientifiques liés à la perception auditive et multisensorielle ainsi que les aspects méthodologiques qui s'inscrivent en grande partie dans le champ des neurosciences cognitives. La composante prédominante de neurosciences cognitives de cette thèse implique une inscription doctorale dans l'ED3C. **Catherine Pelachaud**, directrice de recherche CNRS à l'ISIR est spécialiste des agents virtuels, notamment conversationnels, et de la communication multimodale. Elle a développé la plateforme Greta utilisée dans ce projet. Elle sera la co-encadrante de la thèse et supervisera les aspects liés à la plateforme Greta et de manière plus globale les aspects liés à l'interaction humain-machine. **Frédéric Bevilacqua**, directeur de recherche IRCAM à STMS, est spécialiste des systèmes interactifs sonores et de sonification des gestes. Il collaborera au projet sur les aspects de sonification et interviendra principalement sur les aspects de traduction auditive et audio-visuelle du toucher. **Les publications** des encadrants et du collaborateur figurent dans la section référence, indiqués en gras.

Profil d'étudiant(e) recherché : Le/La candidat.e doit avoir une solide expérience en psychologie expérimentale, en particulier dans le domaine du toucher socio-affectif et de la perception multisensorielle, tout en montrant une ouverture importante au domaine de l'IHM. De solides compétences statistiques sont requises et un bon niveau de programmation sera un plus. Le/La candidat.e devra avoir le profil pour s'intégrer au sein d'une équipe pluridisciplinaire. A noter que nous avons identifié un candidat potentiel qui possède ces compétences, Jordan Cohen, actuellement en M2 sur le sujet, qui participe notamment à la création de la base de données de traduction auditive d'interactions tactiles.

Références

- Auvray, M.** (2019). *Restorative Neurology and Neuroscience*, 37, 609-619.
- Bevilacqua, F.**, Schnell, N. et al. (2017). *Routledge Companion to Embodied Music Interaction* (pp. 391-401). Taylor & Francis.
- Biancardi, B., Lerner, P., Mancini, M. & **Pelachaud, C.** (2019). *Frontiers in Robotics and AI*, 6, 93.
- Boucaud, F., **Pelachaud, C.**, & Thouvenin, I. (2021). *Autonomous Agents and Multi-Agent Systems*.
- Boyer, E. O., **Bevilacqua, F.** et al. (2020). *Experimental Brain Research*, 238, 1-14.
- Dermouche, S., **Pelachaud, C.** (2020). *IEEE Transactions on Affective Computing*, 14.
- Deroy, O., Fasiello, I., Hayward, V., & **Auvray, M.** (2016). *JEP: HPP*, 42, 1204-1214.
- Hertenstein, M. J., Keltner et al. (2006). *Emotion*, 6, 528-533.
- Huisman, G., Kolkmeier, J., & Heylen, D. (2014). *Intelligent Virtual Agents*, 204-213.
- Kirsch, L., Job, X., **Auvray, M.**, & Hayward, V. (2020). *Behavior Research Methods*.
- Lecuyer, A. (2017). *IEEE computer graphics and applications*, 37, 20-26.
- Morrison, I. (2016). *Adaptive Human Behavior and Physiology*, 2, 344-362.
- Pavani, F., Spence, C., & Driver, J. (2000). *Psychological Science*, 11, 353-359.
- Schaefer, M., Heinze, H. J., & Rotte, M. (2005). *Neuroreport*, 16, 367-370.
- Teyssier, M., Bailly, G., **Pelachaud, C.**, Lecolinet, E. (2020). *IEEE Transactions on Affective Computing*.