

Thèse de doctorat

Apprentissage et évaluation de la prise de parole en public en réalité virtuelle

Date limite pour l'envoi de la candidature : 8 mai 2020

Laboratoire d'Informatique et des Systèmes (LIS) et l'Institut des Sciences du Mouvement (ISM)
Aix-Marseille Université & CNRS

Mots-clés : Interaction humain-machine, Recueil de données comportementales, Réalité virtuelle, Agent Conversationnel Animé, Intelligence Artificielle, Apprentissage automatique

Dans le domaine de l'e-éducation et de la formation, un intérêt croissant est apparu autour de la « formation par la simulation » s'appuyant sur des environnements virtuels. Plusieurs travaux ont été ainsi conduits autour de dispositifs permettant de simuler une interaction sociale avec un Agent Conversationnel Animé (ACA) pour entraîner ses propres compétences sociales (« Virtual Agents for Social Skills Training », Bruijnes et al., 2019; Ochs et al., 2019). Certains travaux de recherche, encore très peu nombreux, ont exploré l'utilisation d'une audience virtuelle composée d'un ensemble d'ACAs pour entraîner les individus à la prise de parole en public (Chollet et al., 2015; Pertaub et al., 2002; North et al., 1998).

Un des enjeux dans le développement des *environnements virtuels pour la formation des compétences sociales* est de pouvoir automatiquement évaluer la performance de l'utilisateur et son engagement dans la tâche. Dans le contexte de la prise de parole en public, un certain nombre d'indices verbaux et non-verbaux objectifs reflète la qualité de la prise de parole en public ; par exemple le nombre de disfluences dans le discours, les gestes et la direction du regard. L'engagement de l'utilisateur, et plus particulièrement dans le domaine de la réalité virtuelle son sentiment de présence et de co-présence, se traduit aussi par des indices verbaux et non-verbaux objectifs (Ochs et al, 2019). L'identification des indices comportementaux de la qualité d'une interaction humain-machine est aujourd'hui une problématique de recherche majeure. L'objectif de ce projet de thèse est d'explorer les indices objectifs comportementaux et physiologiques de la qualité de l'interaction avec comme cas d'usage plus spécifiquement la prise de parole en public face à une audience virtuelle dans un environnement de Réalité Virtuelle (RV). Il s'agit d'un défi scientifique important, reposant à la fois sur les travaux en sciences du mouvement humain et en informatique. La méthodologie proposée reposera en effet sur une analyse interdisciplinaire des indices objectifs de l'Interaction Humain-Machine dans cette situation notamment à partir de l'analyse du langage (e.g. prosodie, lexique, syntaxe), une analyse des gestes et des postures (e.g. mesure biomécanique, expression faciale) et des indices physiologiques (e.g. mesure de la fréquence cardiaque, de la conductance electro-dermale). Ces données d'interactions seront analysées à partir de méthodes d'apprentissage automatique pour extraire les indices comportementaux et physiologiques caractéristiques de la qualité de la prise de parole.

La thèse s'organisera autour de 4 grandes étapes : 1/ Collecte et annotation semi-automatique d'un corpus d'interaction humain-audience virtuelle en RV du comportement de prise de parole; 2/ identification des compétences sociales à partir de l'étude du cas d'usage ; 3/ identification des indices comportementaux et physiologiques des compétences sociales à l'aide de techniques d'apprentissage automatique ; et 4/ développement d'un système d'évaluation automatique des compétences sociales adapté au cas d'usage.

Cette thèse interdisciplinaire sera réalisée en collaboration avec **deux laboratoires de recherche de Aix Marseille Université : l'ISM et le LIS**. La plateforme technologique de simulation développée au Centre de Réalité Virtuelle Méditerranée (CRVM) de l'ISM pour l'entraînement à la prise de parole face à une audience virtuelle sera utilisée. L'évaluation sera réalisée en collaboration avec les experts du domaine (coach en prise de parole) pour comparer la sortie du modèle avec l'expertise des coaches.

Compétences requises

La-e candidat-e devra avoir une formation initiale en Informatique, ou en Sciences du Mouvement ou en Sciences Cognitives.

Des compétences en apprentissage automatique, en méthodes expérimentales et en analyses statistiques sont attendues.

Des connaissances en Traitement Automatique du Langage, en traitement du signal, et en analyse du mouvement (motion capture) en Python et R seraient appréciées.

Des connaissances techniques en programmation (Java, C++, C Sharp) et dans le logiciel Unity seraient un plus.

Une ouverture pluridisciplinaire est incontournable.

La thèse est entièrement financée par un contrat doctoral de 3 ans (salaire brut entre 1700 et 2000€/mois) dans le contexte de l'appel Inter-Ed interdisciplinaire d'Aix Marseille Université. Les laboratoires assurent le financement des déplacements professionnels, de la formation du candidat et de la participation aux conférences internationales.

Aix Marseille Université (<http://www.univ-amu.fr/en>), la plus grande université française, est idéalement localisée sur la côte méditerranée à seulement 1h30 des Alpes.

Le dossier du candidat doit comporter les documents suivants :

- 1 CV
- Une lettre de motivation
- Les relevés de notes et copies des diplômes (Licence et master)
- Une lettre de recommandation du responsable de stage ou du directeur de mémoire de master.
- Le mémoire de master 1 et/ou 2.

Le dossier doit être envoyé avant le 8 mai 2020 à

Magalie Ochs: magalie.ochs@lis-lab.fr

Rémy Casanova: remy.casanova@univ-amu.fr

Pour toute question, vous pouvez nous contacter par email.

PhD Position

Learning and Assessment of Public Speaking in Virtual Reality

Deadline for application: 8 May 2020

Laboratoire d'Informatique et des Systèmes (LIS) and Institut des Sciences du Mouvement (ISM)

Aix-Marseille Université & CNRS

Keywords: Human-machine interaction, Behavioural data collection, Virtual Reality, Embodied Conversational Agent, Artificial intelligence, Machine Learning

In the field of e-education and training, a growing interest has emerged around "training by simulation" based on virtual environments. Several studies have been conducted on devices to simulate social interaction with Embodied Conversational Agents (ECAs) to train one's own social skills ("Virtual Agents for Social Skills Training", Bruijnes et al., 2019; Ochs et al., 2019). Some research works have explored the use of a virtual audience composed of a set of ECAs to train individuals to speak in public (Chollet et al., 2015; Pertaub et al., 2002; North et al., 1998).

One of the challenges in the development of *virtual environments for the social skills training* is to be able to automatically evaluate the user's performance and engagement in the task. In the context of public speaking, certain objective verbal and non-verbal cues may reflect the quality of public speaking; for example, the number of disfluencies in speech, the gestures and the gaze direction. User engagement, and more specifically in the field of virtual reality her/his sense of presence and co-presence, is also reflected through objective verbal and non-verbal cues (Ochs et al., 2019). The identification of behavioural cues of the quality of human-machine interaction is a major research issue. The objective of this thesis project is to explore objective behavioural and physiological cues of the quality of interaction in the context of the specific use case of public speaking in front of a virtual audience in a virtual reality (VR) environment. This is a major scientific challenge, based on work in both human movement science and computer science. The proposed methodology will be based on an interdisciplinary analysis of the objective cues of human-machine interaction in this situation, including language analysis (e.g. prosody, lexicon, syntax), analysis of gestures and postures (e.g. biomechanical measurement, facial expression) and physiological clues (e.g. measurement of heart rate, electro-dermal conductance). These interaction data will be analyzed using machine learning methods to extract behavioural and physiological cues characteristic of the public speaking quality.

The thesis will be organized around 4 main steps: 1/ Collection and semi-automatic annotation of a corpus of human-virtual auditory interaction in VR environment; 2/ Identification of social skills from the use case studied; 3/ Identification of the behavioural and physiological cues of social skills using machine learning techniques; and 4/ Development of an automatic evaluation system of social skills adapted to the use case.

This interdisciplinary thesis will be carried out in collaboration with **two research laboratories of Aix Marseille University**: the [ISM](#) and the [LIS](#). The simulation technology platform developed at the [Centre de Réalité Virtuelle Méditerranée](#) (CRVM) of the ISM for public speaking training in front of a virtual audience will be used. The evaluation will be carried out in collaboration with experts in the field (public speaking coach) to compare the output of the model with the expertise of the coaches.

Required skills

The candidate must have an initial training in Computer Science, Movement Sciences or Cognitive Sciences.

Skills in machine learning, experimental methods and statistical analysis are expected.

Knowledge on Natural Language Processing (NLP), signal processing, and motion capture would be appreciated.

Technical knowledge in programming (Java, C++, C Sharp) and Unity software would be a plus.

A multidisciplinary background is a must.

The thesis is fully funded by a 3-year doctoral contract (salary between 1400 and 1600 €/month) in the context of the Inter-Ed Interdisciplinary call of Aix Marseille University.

The laboratories provide funding for professional travel, the candidate's training and participation in international conferences.

French language is not required.

Aix Marseille University (<http://www.univ-amu.fr/en>), the largest French University, is ideally located on the Mediterranean coast, and only 1h30 away from the Alps.

The application files consists of the following documents:

- A detailed curriculum,
- A motivation letter,
- A description of the academic background and copy of academic records and most recent diploma,
- A letter of recommendation from the internship supervisor or the director of the master's thesis.
- The master thesis 1 and/or 2.

The application files should be sent before the **8 may** to

Magalie Ochs: magalie.ochs@lis-lab.fr

Rémy Casanova: remy.casanova@univ-amu.fr

For any question, you can contact us by mail.