

Fiche de stage

Sujet du stage : Étude de terrain sur le pointage grâce à la vision par ordinateur

Encadrant : Julien Gori

Date de début du stage : dès que possible

Durée du stage : 6 mois

Niveau d'études souhaité : en Master 2 ou dernière année d'École d'Ingénieur·e

Laboratoire d'accueil : ISIR (*Institut des Systèmes Intelligents et de Robotique*), Campus Pierre et Marie Curie, 4 place Jussieu, 75005 Paris.

Personne à contacter

Prénom Nom : Julien Gori

Tel : +33 6 51 40 64 04

Email : gori(at)isir.upmc.fr

Envoyer votre candidature par mail, avec [*sujet du stage*] en objet, un CV et une lettre de motivation.

Date limite de dépôt de la candidature : fin janvier 2023

Description du stage (en français)

Résumé :

La sélection de cibles à l'aide d'un curseur (pointage), par exemple avec une souris, est l'une des interactions les plus fréquentes pour un utilisateur. Pourtant, étonnamment, l'utilisateur typique pointe toujours comme il y a 30 ans, malgré de nombreuses avancées publiées dans les recherches sur le pointage dans le domaine de l'interaction homme-machine (IHM) – certaines prédisant des gains de performance de plus de 30%. Qu'est-ce qui peut expliquer cela ? À mon avis, il y a un décalage entre les travaux de recherche actuels sur le pointage, qui sont uniquement basés sur des études contrôlées en laboratoire, et le pointage « réel ». On ne sait qu'évaluer la performance de nouvelles méthodes de pointage à travers des expériences contrôlées parce qu'on ne sait pas observer les grandeurs dont on a besoin pour les caractériser (ni en général implémenter ces méthodes) sur des vrais systèmes.

Le but de ce stage est d'instrumenter un ordinateur pour pouvoir acquérir les données nécessaires, particulièrement les tailles et positions des cibles sélectionnées par l'utilisateur, pour caractériser le pointage directement dans l'environnement de ce dernier. On utilise souvent une loi, dite loi de Fitts, pour décrire le pointage dans un cadre d'expérience contrôlée. Une des questions à laquelle on répondra est de savoir si cette description est toujours pertinente en dehors du cadre contrôlé. Ce travail servira aussi de première étape pour l'implémentation future de techniques d'interaction qui nécessitent de connaître les caractéristiques de cibles pour fonctionner.

Sous la co-tutelle de :

Objectifs du stage :

Concrètement, l'étudiant exploitera les API d'accessibilités qui sont disponibles pour certaines applications. Par exemple, l'environnement Gnome (disponible sur Linux, et par exemple standard sur Ubuntu) expose des informations de taille et position de certaines icônes à travers l'interface AT-SPI. Ceci implique que pour un nombre donné d'applications, on est capable de récupérer les informations sur les cibles, nécessaires à l'évaluation de la loi de Fitts. En même temps, on peut faire une capture d'écran ; on peut donc constituer une base de données, avec des captures d'écran de différentes interfaces et en correspondance des informations sur les cibles de cette interface. A partir de cette base de donnée, on peut alors utiliser de l'apprentissage supervisé pour extraire automatiquement d'une capture d'écran les cibles potentielles. L'étudiant-e constituera le jeu de donnée et réalisera l'apprentissage supervisé et son évaluation.

Profil recherché :

L'étudiant-e recherché aura des connaissances en programmation (C ou C++ ou Python), et Linux. Des connaissances de bases en apprentissage supervisé sont aussi demandées. Des connaissances en détection d'objets, en programmation d'interfaces utilisateurs et en Xlib/Wayland seront particulièrement appréciées.

Description du stage (en anglais)

Subject: Field study on pointing with computer vision

Abstract:

Selecting targets with a cursor, for example with a mouse, is one of the most common interactions (pointing) for a user. Yet, surprisingly, the typical user still points as they did 30 years ago, despite the many published advances in human-computer interaction (HCI) pointing research - some predicting performance gains of more than 30%. What can explain this? In my opinion, there is a disconnect between current pointing research, which is based solely on controlled laboratory studies, and "real" pointing. We only know how to evaluate the performance of new pointing methods through controlled experiments because we don't know how to observe the quantities we need to characterize them (nor generally implement these methods on real systems).

The goal of this internship is to instrument a computer to acquire the necessary data, particularly the sizes and positions of the targets selected by the user, to characterize the pointing directly in the user's environment. A law, known as Fitts' law, is often used to describe pointing in a controlled experiment setting. One of the questions that will be answered is whether this description is still relevant outside the controlled setting. This work will also serve as a first step for the future implementation of interaction techniques that require knowledge of target characteristics to work.

Internship Objectives:

Concretely, the student will exploit the accessibility APIs that are available for some applications. For example, the Gnome environment (available on Linux, and for example standard on Ubuntu) exposes size and position information of some icons through the AT-SPI interface. This implies that for a given number of applications, we are able to retrieve the information about the targets, necessary for the evaluation of Fitts' law. At the same time, we can make a screen capture; we can thus constitute a database, with screen captures of various interfaces and in correspondence of the information on the targets of this interface. From this database, we can

Sous la co-tutelle de :

then use supervised learning to automatically extract potential targets from a screenshot. The student will build the dataset and perform the supervised learning and its evaluation.

Required Profile:

The desired student will have knowledge of C or C+ or Python programming, and Linux. Basic knowledge of supervised learning is also required.

Knowledge of object detection, user interface programming, and Xlib/Wayland will be particularly appreciated.