Re-designing the Input Pipeline in Interactive Systems

Duration: 36 months
Teams: Loki (Inria centre at the University of Lille & CRIStAL, University of Lille, France) & Exii (Cheriton School of Computer Science, UWaterloo, Canada)
Supervisors: Géry Casiez, Mathieu Nancel & Daniel Vogel (gery.casiez@univ-lille.fr, mathieu.nancel@inria.fr, daniel.vogel@uwaterloo.ca
Application deadline: 30/04/2024

The Loki and Exii research groups are looking for a PhD student in co-tutelle starting fall 2024.

Context

Routine tasks such as controlling a system cursor or moving a virtual camera involve continuous visuo-motor control from the user, to which the system has to respond accurately and with minimal latency. Modern Human-Computer Interfaces use multi-step input pipelines between the user's movements and the system's feedback that are frequently opaque to interaction designers. Each of these steps (sensing, filtering, transforming, predicting, etc.) strongly affects the next, as well as the pipeline's outcome. And yet, many of them are designed in isolation and calibrated by trial and error using legacy or ad-hoc approaches, or limited knowledge of the underlying psychomotor phenomena. This limits user performance and experience in everyday computer use, and hinders the design and adoption of new devices and sensing methods.

Objectives

This Ph.D. will explore and address the knowledge gaps in the design of continuous input pipelines: in the way we design their steps, but also in our understanding of how users perceive, decide, and act in these real-time tasks. It will involve methods and tools from HCI, experimental psychology, machine learning, or control theory. The goal is to discover and establish generalizable guidelines, methods and algorithms to build input pipelines whose steps can fit and adapt to any user, system, or device.

This is a vast topic that can take any of the following paths, depending on the interests of the candidate:

Input filtering / signal processing:

Raw input data needs to be processed to become usable in a UI, especially when captured by IMUs or optical tracking as it is often the case in Mixed Reality. This data is often noisy, and prone to sensing or timing inaccuracies that can strongly affect the quality of the system's response. On the other hand, filtering methods often introduce latency in the pipeline, which is also detrimental to performance even when users do not notice it (see related work [3, 8]). Our team developed the 1-euro filter [5] which has become a standard in research and industry [https://gery.casiez.net/1euro/], based on the observation that the effects of input noise become less noticeable at high speeds. Its parameters are currently tuned by hand, however, so the next stage in this direction could consist in developing a better understanding of the perception and characterization of motion noise to assist or automatize the tuning of this filter. Other venues for research can include developing faster methods to infer the user's instant velocity in real time from noisy displacement and timing data, without adding delay, based on previous results by our research group [7].

Inferring / transforming:

End-to-end latency in the input pipeline derives mostly from the display side [2, 3] which is generally irreducible in a given system. A promising approach consists in predicting the near future of the user's ongoing movements in order to display the system's response to its *expected* current location, rather than to the last *sensed* location. Our research group has developed one such predictor [7] and there are many ways to develop this topic further. One is to improve existing predictors using better models of movement, another is to characterize the expected errors of a predictor in order to adapt the visual feedback to the system's confidence [8].

Once processed, inputs are generally transformed (e.g. with pointer acceleration) and converted into the output's frame of reference (e.g. the screen's) by so-called "transfer functions" [4, 6] at the system level before being sent to graphics libraries as exploitable events. These functions help adapt the unitless input events to the size and resolution of the display, and can improve performance and usability on certain tasks. However, the design and tuning of these functions remains poorly understood and ad-hoc. A possible venue for this PhD is to explore new methods to help interaction designers choose and tune these functions based on functional models [1] that can take into account the specific capabilities of the user and of the system, and the task at hand.

Sensing / hardware / fine motor actions:

New and existing sensing devices can be investigated to improve the user's degree of control, with different form factors or sensing characteristics better adapted to the task and context. This could involve exploring the effects of friction on classic touch devices like touch-pads or touch-screens, or designing new input methods based on Optical Finger Navigation (OFN) modules that can be used in a variety of mobile context, including Mixed Reality for which prolonged use of mid-air gestures can be tiring and imprecise.

Typical input devices such as mice and touch-pads express displacements in unitless *counts* or *mickeys*, or finger pressure in ad-hoc units, which do not convey the actual scale of user movements and therefore make it harder to exploit them optimally in the next steps of the input pipeline. One approach is to design better event protocols and standards to include this information, another is to develop methods to determine it automatically by using machine learning approaches to infer, e.g., the resolution of a mouse using only the unitless events that it emits.

At a higher perspective, the knowledge gained on these topics will be used to help to redesign the input pipeline in current interactive systems.

The three supervisors have recognized expertise at the international level on these topics with papers published in leading conferences of the domain. The avenues of research described above expand beyond one thesis so the specific topics covered by the candidate will be tailored based on preference and profile.

Location

The PhD candidate will join the LOKI research group based in the Inria centre at the University of Lille, and the Exii group of the Cheriton School of Computer Science at the University of Waterloo. Lille is at the northern tip of France and its metropolitan area, situated at the crossroads of northern continental Europe, is the 5th biggest in France. Kitchener-Waterloo is an international technology hub located 1 hour from Toronto. It's where Google chose to establish their Canadia headquarters (with 1500 employees) and it boasts numerous start-up incubation centres.

Both Loki and Exii are dynamic and multicultural teams with members coming from different countries (Germany, Colombia, Canada, China, France, etc.) and communicating daily in English.

The candidate

A successful candidate must hold a MSc in Human-Computer Interaction, Computer Science, or Control Theory, and show a great interest in performing high quality research. The candidate must demonstrate experience or strong interest in software development. Creativity, independence, team working and communication skills are valuable advantages. It is not required to speak French. A good level of technical and scientific English is also a plus. If interested in this project, simply e-mail Géry Casiez (gery.casiez@univ-lille.fr), Mathieu Nancel (mathieu.nancel@inria.fr) and Daniel Vogel (daniel.vogel@uwaterloo.ca).

References

[1] Bye, R.T. and Neilson, P.D. 2008. The BUMP model of response planning: Variable horizon predictive control accounts for the speed–accuracy tradeoffs and velocity profiles of aimed movement. *Human Movement Science*. 27, 5 (Oct. 2008), 771–798. DOI:https://doi.org/10.1016/j.humov.2008.04.003.

[2] Casiez, G., Conversy, S., Falce, M., Huot, S. and Roussel, N. 2015. Looking Through the Eye of the Mouse: A Simple Method for Measuring End-to-end Latency Using an Optical Mouse. *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology* (New York, NY, USA, 2015), 629–636.

[3] Casiez, G., Pietrzak, T., Marchal, D., Poulmane, S., Falce, M. and Roussel, N. 2017. Characterizing Latency in Touch and Button-Equipped Interactive Systems. *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology* (New York, NY, USA, Oct. 2017), 29–39.

[4] Casiez, G. and Roussel, N. 2011. No More Bricolage!: Methods and Tools to Characterize, Replicate and Compare Pointing Transfer Functions. *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology* (New York, NY, USA, 2011), 603–614.

[5] Casiez, G., Roussel, N. and Vogel, D. 2012. 1 € filter: a simple speed-based low-pass filter for noisy input in interactive systems. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, mai 2012), 2527–2530.

[6] Casiez, G., Vogel, D., Balakrishnan, R. and Cockburn, A. 2008. The Impact of Control-Display Gain on User Performance in Pointing Tasks. *Human–Computer Interaction*. 23, 3 (Aug. 2008), 215–250. DOI:https://doi.org/10.1080/07370020802278163.

[7] Nancel, M., Aranovskiy, S., Ushirobira, R., Efimov, D., Poulmane, S., Roussel, N. and Casiez, G. 2018. Next-Point Prediction for Direct Touch Using Finite-Time Derivative Estimation. *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology* (New York, NY, USA, Oct. 2018), 793–807.

[8] Nancel, M., Vogel, D., De Araujo, B., Jota, R. and Casiez, G. 2016. Next-Point Prediction Metrics for Perceived Spatial Errors. *Proceedings of the 29th Annual Symposium on User Interface Software and Technology* (New York, NY, USA, Oct. 2016), 271–285.