Hands Up: Who Knows Something About Performance and Ergonomics of Mid-Air Hand Gestures

Anna Feit, Mathieu Nancel

Hands Up: Who Knows Something About Performance and Ergonomics of Mid-Air Hand Gestures

Anna Maria Feit and Mathieu Nancel
Aalto University, Helsinki, Finland

Abstract
Advances in markerless and un-instrumented hand tracking allow us to make full use of the hands' dexterity for interaction with computers. However, the biomechanics of hand movements remain to be thoroughly studied in HCI. The large number of degrees of freedom of the hand (25) presents us with a huge design space of possible gestures, which is hard to fully explore with traditional methods like elicitation studies or design heuristics. We propose an approach to develop a model of fatigue and stress of manual mid-air input, inspired by prior work on the ergonomics of arm movements and on the performance of multi-finger gestures. Along with our vision of the incoming challenges in mid-air interaction, we describe a design framework for mid-air input that, given such models, can be used to automatically evaluate any given gesture set, or propose an optimal gesture vocabulary for a given set of tasks.

Motivation: What We Do and Don't Know
Technological advances in tracking and sensing of hand, finger and arm movements have equipped designers with a vast choice of interaction methods. For example, the un-instrumented tracking of all 25 degrees of freedom (DoFs) of the hand by only a single camera (e.g. [7, 9]) finally allows us to make full use of the hands' dexterity for computer input. However, these 25 DoFs construct a design space far larger than that of any other input "device". This poses a big challenge for interaction designers.

Currently, Human-Computer interactions through mid-air gestures[1] is mostly limited to specific scenarios and small gesture sets. Their design often relies on elicitation studies, the emulation of 2D methods in 3D, or heuristics derived from other modalities or common sense. While useful, these methods are limited when designing for usability criteria such as high performance, low muscle fatigue, good memorability, or ease of use. They also make it unrealistic to fully explore the combinatorially immense design space of mid-air gestures[2]. Yet it is necessary to understand which gestures are fast to perform, easy to memorize or even executable in terms of anatomical constraints.

Previous research introduced systematic methods to quantify and model the performance and ergonomics of arm [2, 5] and finger [8] movements. The advantage of such models over design heuristics or guidelines is that they can be used to quantitatively evaluate any given set of gestures and even optimize it for specific tasks. For instance, Sridhar *et al.* [8] developed a model of execution time for hand gestures (expressed as combinations of flexing or extending fingers), which they used to compute an optimal gesture set

[1] Here, any arm, finger or hand movement that is not constrained by an input surface, e.g. simple pointing or drawing with a finger, static hand postures [6], or dynamic combinations of multiple end-effectors [8].

[2] For instance, with only 3 discretization levels for each of the 25 DoFs of the hand we obtain 10^{11} possible combinations, and that is ignoring associated arm movements or bimanual interactions.

for high performance text entry in mid-air. Other work made use of existing knowledge on the musculoskeletal structure of the upper limbs to infer muscle fatigue, either by mechanical modeling [5] or by deducing muscle activation from motion tracking data using biomechanical analysis [1, 2].

Approach: What We Propose

While useful, the models presented above are far from capturing the full design space of mid-air gestures, characterizing only arm movements [2, 5], or simple finger combinations [8]. Our goal is to advance the state of the art on performance and motor capabilities of the hand [8] and expand existing models of arm fatigue [5] to hand gesture input. In addition to muscle fatigue, we also want to explore questions of tendon stress and strain that are likely to occur with prolonged use of hand gestures.

Biomechanical analysis [1, 2] yields insightful data about the means and consequences of user movements, but has only been applied to the arms so far. We propose to use and adapt it to the specifics of hand and wrist movements, seeking three contributions: (1) a new model of muscle fatigue applicable to wrist and finger movements; (2) an analysis of tendon stress, that we can map to user feedback to accurately model comfort and ease of use; (3) a longitudinal study involving repeated hand gestures, to assess both learnability and the evolution of fatigue and joint stress.

Accurate finger detection offers incomparable expressiveness [6, 8] but comes with a very large design space, making it challenging to model the full gesture space in a manner both meaningful and useful. To address this we envision an approach similar to [3, 8]. There, the gesture space was divided into independent DoFs that could be studied separately. The resulting models were then combined to predict the performance of more complex gestures.

Outlook: What We Aim For

A deep understanding of arm and hand movement capabilities will allow us to assert the feasibility of any set of mid-air gestures, balancing performance and fatigue for realistic scenarios of short- and long-term use. The approach we propose could further be extended to study the impact of haptic feedback on ergonomic factors and learnability. It will also inform the definition of a formal design space, that needs to be descriptive enough to encompass any possible gesture and combination thereof, and at the same time simple enough to be usable and understandable despite the combinatorial explosion of possible gesture sets.

In the long term, we envision tools to quickly evaluate and optimize mid-air input vocabularies. Interaction designers would enter different sets of gestures and define their envisioned use-cases, the sensing capabilities of a platform, and the actions available in the application. Then they would receive model- and simulation-based feedback on the trade-offs of these vocabularies, as well as suggestions for improvements. This tool would consider multiple features such as performance, learnability, memorability, technological and human limitations, subjective experience and others, and balance them depending on the type and criticality of the tasks, or the users' needs and requests[3].

While an ambitious goal, the field of multi-objective interface optimization is advancing [3, 4, 8], and some of these objectives have been extensively studied in psychology and HCI. However, the expressiveness and constraints of hand gesture input are not understood well enough to match its otherwise great assets for interaction. Thus, we believe that focusing on hand gestures is the appropriate next step.

[3] For instance: fatigue, performance and memorability would be favored for scientific or professional applications intended for extended periods of use and long-term work, whereas learnability and fast skill acquisition would be the main goals for entertainment-oriented applications.

References

[1] Myroslav Bachynskyi, Antti Oulasvirta, Gregorio Palmas, and Tino Weinkauf. 2014. Is Motion Capture-based Biomechanical Simulation Valid for HCI Studies?: Study and Implications. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 3215–3224. DOI: http://dx.doi.org/10.1145/2556288.2557027

[2] Myroslav Bachynskyi, Gregorio Palmas, Antti Oulasvirta, Jürgen Steimle, and Tino Weinkauf. 2015. Performance and Ergonomics of Touch Surfaces: A Comparative Study Using Biomechanical Simulation. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 1817–1826. DOI: http://dx.doi.org/10.1145/2702123.2702607

[3] Xiaojun Bi, Tom Ouyang, and Shumin Zhai. 2014. Both Complete and Correct?: Multi-objective Optimization of Touchscreen Keyboard. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 2297–2306. DOI: http://dx.doi.org/10.1145/2556288.2557414

[4] Anna Maria Feit, Myroslav Bachynskyi, and Srinath Sridhar. 2015. Towards Multi-Objective Optimization for UI Design. (April 2015). http://annafeit.de/resources/papers/Multiobjective_Optimization2015.pdf Presented at CHI 2015 Workshop on Principles, Techniques and Perspectives on Optimization and HCI.

[5] Juan David Hincapié-Ramos, Xiang Guo, Paymahn Moghadasian, and Pourang Irani. 2014. Consumed Endurance: A Metric to Quantify Arm Fatigue of Mid-air Interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 1063–1072. DOI: http://dx.doi.org/10.1145/2556288.2557130

[6] Mingyu Liu, Mathieu Nancel, and Daniel Vogel. 2015. Gunslinger: Subtle Arms-down Mid-air Interaction. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology (UIST '15)*. ACM, New York, NY, USA, 63–71. DOI: http://dx.doi.org/10.1145/2807442.2807489

[7] Toby Sharp, Cem Keskin, Duncan Robertson, Jonathan Taylor, Jamie Shotton, David Kim, Christoph Rhemann, Ido Leichter, Alon Vinnikov, Yichen Wei, Daniel Freedman, Pushmeet Kohli, Eyal Krupka, Andrew Fitzgibbon, and Shahram Izadi. 2015. Accurate, Robust, and Flexible Real-time Hand Tracking. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 3633–3642. DOI: http://dx.doi.org/10.1145/2702123.2702179

[8] Srinath Sridhar, Anna Maria Feit, Christian Theobalt, and Antti Oulasvirta. 2015. Investigating the Dexterity of Multi-Finger Input for Mid-Air Text Entry. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 3643–3652. DOI: http://dx.doi.org/10.1145/2702123.2702136

[9] Srinath Sridhar, Franziska Mueller, Antti Oulasvirta, and Christian Theobalt. Fast and Robust Hand Tracking Using Detection-Guided Optimization. In *Proceedings of Computer Vision and Pattern Recognition (CVPR)*. DOI: http://dx.doi.org/10.1109/CVPR.2015.7298941