PhD Proposal

ANR Project ELEMENT: *Enabling Learnability in Embodied Movement Interaction*
Supervision: Yacine Bellik, Jules Françoise
Lab: AMI Team, LIMSI-CNRS, Orsay (https://www.limsi.fr/)
Dates: Start Nov 2018 - Jan 2019, 3 years.

Adaptive Computational Models of Human Movement to Support Learnability in Embodied Interaction

Titre en français : Modèles computationnels adaptatifs du mouvement humain pour faciliter l'apprentissage sensori-moteur en interaction par le mouvement.

CONTEXT

Because memorizing and executing gestures is challenging for users, most current approaches to movement-based interaction consider intuitive interfaces and trivial gesture vocabularies. While these facilitate adoption, they also limit users' potential for more complex, expressive and truly embodied interaction. Considering movement-based interaction beyond the mouse-keyboard paradigm, the ANR project ELEMENT (*Enabling Learnability in Embodied Movement Interaction*) proposes to shift the focus from *intuitiveness* towards *learnability*: new interaction paradigms require users to develop specific sensorimotor skills compatible with – and transferable between, – digital interfaces (including video interface, mobile devices, internet of things, game interfaces). With learnable interactions, novice users should be able to approach a new system with a difficulty adapted to their expertise, then the system should be able to carefully adapt to the improving motor skills, and eventually enable complex, expressive and engaging interactions. The long-term aim is to foster innovation in multimodal interaction, from assistive technologies to media interaction in creative applications.

The project ELEMENT is coordinated by Ircam (Paris), and also involves LRI (Orsay) and LIMSI (Orsay). The PhD candidate is expected to strongly interact with other PhD students and postdocs in the project, and will work in collaboration with all partners.

RESEARCH QUESTION & GOALS

Recent progress in computational modelling of human movement has capitalized on advances in machine learning. Typical approaches to movement recognition, regression, spotting, or tracking, involve statistical models trained on datasets from a wide range of users [Rautaray & Agrawal, 2015; Alemi et al., 2015, Escalera et al 2016]. To be robust to the task at end, like recognition, these models often require a large amount of data to produce models that, eventually, average movement variability across individuals. As a result, standard machine learning methodologies currently fail at characterizing the differences between individual users in terms of mastery, expression and learning abilities; and user-specific training of these methods is practically unfeasible. One direct consequence for the design of gesture-based interaction involving machine learning is the tendency to simplify the gesture vocabulary to fit recognizers' abilities [Wobbrock et al., 2009].

This has led the community to research ways of designing movement models that can be used with more complex gesture vocabularies. Considering our use cases, the training procedures for movement models must fit radical specifications such as scarce data resources, high variability and fast inference. In this context, Rajko proposed a model called SNM (Semantic network models) that uses Hidden Markov Models with significantly reduced training phase [Rajko et al., 2006]. Bevilacqua developed a similar series of models, such as a one-shot learning HMM called gesture follower (GF) [Bevilacqua et al., 2010]. It demonstrated its ability to perform real-time recognition in the context of dance and musical gesture. Developing further this approach to characterize movement qualities, [Caramiaux et al., 2015] developed a system called Gesture Variation Follower (GVF) that uses particle filtering for the adaptive estimation of gesture variation in scale, speed, direction size and orientation based on single gesture templates. Françoise also expanded HMM models toward multimodal and hierarchical HMMs trained with very limited number of examples [Francoise et al., 2014, 2018]. These approaches are aligned with the emerging research field of Interactive Machine Learning (IML) [Fails et al., 2003], or Human-centered Machine Learning (see the CHI'16 workshop[1]), which propose to better leverage end-user knowledge during the machine learning process [Amershi et al., 2014].

A core contribution of this doctorate is to propose adaptive computational movement models. Specifically, we propose to consider human-centered approaches to machine learning that can be truly adaptive: adaptive to the user expertise and their expressive signature, adaptive to varying movement vocabularies, and adaptive to the interfaces used to capture the movement. The doctorate will address the design of computational models of movements and gestures as continuous multiscale time processes that vary according to a number of factors (or contextual variables) representing the individual characteristics of a user, their skill level, or the context of use. Such models should also integrate theoretical knowledge about human movement, such as biomechanical constraints. These models will rely on approaches such as Dynamic Bayesian Networks or Deep Neural Networks which can integrate context variables and multilevel representations of temporal processes. The thesis will also consider longer-term adaptation of particular interactive systems involving a user interacting within a feedback loop. It consists in the development of adaptive feedback strategies, that evolve over time to adapt to the user's skill level. In this case, the models should adapt feedback and guidance strategies to the learning pathways of an individual user. Methods such as reinforcement learning and active learning will be considered.

Applications

The ANR project ELEMENT will consider three complementary use-cases where learnability is essential for the development of expressive or efficient interaction: gesture-based interaction techniques for communication ; music and dance, from novice to experts ; and assistive technologies. The present doctorate will focus on the latter use-case, aiming to develop assistive technologies for people with sensorimotor disabilities. The AMI team has expertise in multimodal

[1] See for example the CHI 2016 Workshop (http://hcml2016.goldsmithsdigital.com/) and associated ACM TiiS Special issue on Human-Centred Machine Learning.

interaction and assistive technologies. In particular, the AMI team recently conducted research on new spatial interfaces for the blind [Bellik & Clavel, 2017] and is investigating novel touch-based interface for wheelchair control designed for people with motor disabilities [Guedira et al., 2016]. In this use case, individual adaptation is critical because of the large variety of sensorimotor abilities. We will consider how adaptive technologies and multimodal feedback can facilitate the learning and/or appropriation of gesture-based interfaces for disabled persons. While assistive technologies will be the main application of the project, the doctoral candidate will work in close collaboration with all partners and contribute to other use-cases.

Work Program

1. Perform a state of the art on the field of adaptive methods in movement modeling and machine learning.
2. Conduct participatory workshops in collaboration with the project partners to identify relevant use-cases and formalize the relevant types of adaptation, and participate in the creation of new movement datasets.
3. Design and evaluate several adaptive computational models addressing different types of adaptation (to the user's expertise and sensori-motor abilities, the their learning process, short/long-term, etc) identified in the previous step.
4. Design and conduct user evaluation studies to evaluate the benefits of user adaptation in real-world use-cases.
5. Propose guidelines for the design and implementation of adaptive computational models for embodied interaction.

Equipment

The PhD student will be provided with a laptop computer and accessories. LIMSI has acquired several wearable movement sensors (IMUs, EMG sensors), and owns a full-body motion capture system that will be used for data collection, prototyping, and for the evaluation of the developed computational models.

CANDIDATE PROFILE

We are looking for passionate candidates with creativity and scientific curiosity qualities, strong problem solving skills, good knowledge of signal processing, machine learning and HCI, as well as experimental methodologies and statistical analysis. Candidates should be proficiency in at least one programming language and good scientific writing abilities in English and/or French.

HOW TO APPLY

Send a CV, a cover letter and your recent marks to:

- Jules Françoise <jules.francoise@limsi.fr>
- Yacine Bellik <Yacine.Bellik@limsi.fr>

REFERENCES

[**Alemi et al., 2015**] Alemi, O., Li, W., & Pasquier, P. (2015, September). Affect-expressive movement generation with factored conditional Restricted Boltzmann Machines. In Affective Computing and Intelligent Interaction (ACII), 2015 International Conference on (pp. 442-448). IEEE.

[**Amershi et al., 2014**] Amershi, S., Cakmak, M., Knox, W. B., and Kulesza, T. (2014). Power to the people: The role of humans in interactive machine learning. AI Magazine, 35(4):105–120.

[**Bellik & Clavel, 2017**] Y. Bellik, C. Clavel, "Geometrical Shapes Rendering on a Dot-Matrix Display", in Proc. of the 9th International conference on Intelligent Human Computer Interaction, IHCI 2017, 11-13 December 2017, Evry, Paris, France, 10 pages

[**Bevilacqua et al., 2010**] Bevilacqua, F., Zamborlin, B., Sypniewski, A., Schnell, N., Gu ́edy, F., and Rasamimanana, N. (2010). Continuous realtime gesture following and recognition. In Lecture Notes on Computer Science, Gesture Workshop, pages 73–84. Springer.

[**Caramiaux et al., 2015**] Caramiaux, B., Montecchio, N., Tanaka, A., and Bevilacqua, F. (2015). Adaptive gesture recognition with variation estimation for interactive systems. ACM Transactions on Interactive Intelligent Systems (TiiS), 4(4):18.

[**Escalera et al., 2016**] Escalera, S., Athitsos, V., and Guyon, I. (2016). Challenges in multi-modal gesture recognition. In Gesture Recognition, pages 1–60. Springer.

[**Fails and Olsen Jr, 2003**] Fails, J. A. and Olsen Jr, D. R. (2003). Interactive machine learning. In Proceedings of the 8th international conference on Intelligent user interfaces, pages 39–45. ACM.

[**Françoise et al., 2014**] Françoise, J., Schnell, N., Borghesi, R., and Bevilacqua, F. (2014). Probabilistic models for designing motion and sound relationships. In Proceedings of the 2014 international conference on new interfaces for musical expression, pages 287–292.

[**Françoise et al., 2018**] Françoise, J. and Bevilacqua, F.. 2018. Motion-Sound Mapping through Interaction: An Approach to User-Centered Design of Auditory Feedback Using Machine Learning. ACM Transactions on Interactive Intelligent Systems. 8, 2, Article 16 (June 2018).

[**Guedira et al., 2016**] Y. Guedira, L. Jordan, C. Favey, R. Farcy, & Y. Bellik, "Tactile Interface for Electric Wheelchair", in Proc. of the 18th International ACM SIGACCESS Conference on Computers and Accessibility, ASSETS 2016, ACM, October 2016, Reno, NV, USA

[**Klemmer et al., 2006**] Klemmer, S. R., Hartmann, B., and Takayama, L. (2006). How bodies matter: five themes for interaction design. In Proceedings of the 6th conference on Designing Interactive systems, pages 140–149. ACM.

[**Rajko et al., 2007**] Rajko, S., Qian, G., Ingalls, T., and James, J. (2007). Real-time gesture recognition with minimal training requirements and online learning. In Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on, pages 1–8. IEEE.

[**Rautaray & Agrawal, 2015**] Rautaray, S. S., & Agrawal, A. (2015). Vision based hand gesture recognition for human computer interaction: a survey. Artificial Intelligence Review, 43(1), 1-54.

[**Wobbrock et al., 2009**] Wobbrock, J. O., Morris, M. R., and Wilson, A. D. (2009). User-defined gestures for surface computing. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pages 1083–1092. ACM.