Understanding virtuosi as expert users and celebrating variation as expressive difference in embodied interaction

Atau Tanaka
Goldsmiths, University of London
London SE14 6NW UK
a.tanaka@gold.ac.uk

Introduction

Interactive music in its different forms has close ties to HCI. The field of NIME got its start as a CHI workshop [11], and benefits greatly from advances in human-computer interaction technologies and theory. The NIME-HCI dynamics is, however, of mutual benefit, and we explore ways in which NIME can give back to CHI. First, notions of the user differ, and we discuss ways in which the music virtuoso can be considered an "expert user" in HCI. Second, we present approaches to measure variation in embodied interaction, and report that information as an "expressive" vector. In this workshop we would like to share with workshop participants work in progress and late breaking results in two studies that apply machine learning to musical gesture: One looking at the measurement of expressive gesture variation by physiological signals from the arm muscles; and the second using motion capture systems to trace the differences of two virtuoso pianists as they perform the same repertoire piece by Schubert.

Interactive Music

Interactive music practice is not confined to one form of human-machine interaction, but situates interaction between a composer, performer, or listener, and a computer-based musical system, at different points. In the field of algorithmic composition or generative music, the user (be they a composer or a listener) might interact with a system that interactively produces melodic and rhythmic musical structures [4]. This has been extended in exciting new ways in scenes like Live

Coding, where the performance interaction onstage is the act of programming [8, 9].

While new interfaces and input devices for music performance remain at the heart of NIME, the study of motion and movement qualities for interactive music, dance, and other artforms is the subject of new conferences such as Moco[1]. The of musical content in music consumption and related research in music information retrieval may seem to fall outside the scope of interactive music for performance, but Pachet uses analysis and modelling of melodic and acoustic performance to create conversational partners for music performance [10]. This in effect can be thought of as a kind of musical Turing test, and connects to the rapidly growing field of Computational Creativity. Meanwhile modelling of both gesture and content in parallel, and machine learning of the relationship between sound and evoked gesture as some of the new directions in NIME [5].

While interaction takes place in different ways in all these examples, research in interactive music, despite their diversity, share goals that distinguish them from the objectives of HCI research. The systems created in music research are often unique and idiosyncratic. Indeed, in line with traditions of Western art and cultural production, the individual identify and subjective judgement of the person creating the work drive the design of the system. The result to be assessed is not the system itself, but the resulting music. Scientific generalization can be difficult.

Technology for all, including experts

Recent work in pedagogy as well as the study of musical practices as gone beyond the focus on individual performance-based systems. Interactive systems and haptic interaction have been used to make music making more accessible [1]. Ethnographic approaches from HCI have been used to study technology enhanced sketching as part of the compositional process [6].

There is a tension between these approaches that warrants further investigation. Enabling technologies are understood to facilitate the achievement of complex tasks. In this ethos, it is understood that technologies have a democratising power to broaden access and lower barriers of entry to otherwise difficult activities. Music and music pedagogy stands to benefit greatly from this [14].

Meanwhile, technology cannot be seen as a panacea that obviates the human effort and investment (often of years or even a lifetime), to develop musical craft, technique, and virtuosity. If interactive systems can facilitate difficult tasks to support a beginner user, could they also be configured to support the expert user? What can be learned from virtuosos in order to help us design interactive systems for all levels of user that provide profound forms of interaction and satisfaction, and allow us to go beyond the “instant gratification” assumed in many technological systems? If acoustic instruments are able to accommodate, without modification, beginner and virtuoso play, could we design digital interactive systems that, unlike the user level settings of a typical computer game, naturally respond in interaction sophistication automatically to level of use?

Machine learning brittleness

Machine learning technologies have the potential to be applied in HCI contexts to allow the kind of fluidity described above. However, problems as they are formulated in machine learning traditionally operate in more brittle ways, with a certain intended levelling effect. Classification algorithms such as Support Vector Machines attempt to distinguish between varying input providing a binary result. Traditional approaches to invariance, for example in handwriting recognition, suppress variation to best match idiosyncratic input to canonical examples. In music, on the other hand, we

[1] http://moco.iat.sfu.ca/

seek to identify and exploit the minute and subtle differences in varying input, and describe this as the personal style and expressiveness of an individual. Instead of the "one size fits all" paradigm, can we find approaches to celebrate individual difference?

We have used probabilistic approaches to study gestural variation, and individual idiosyncratic difference in different people achieving the same musical task.

Muscular subtleties
Sensing muscle tension is one way to capture the subtlety of gesture. We have shown that users are able to reproduce and vary the power of gesture as measured by the EMG [2]. In a study we are currently running, we are analysing expressive differences of EMG using a probabilistic technique that adapts to real time variation [3]. We conducted a study of hand gesture measured by the electromyogram (EMG) in a scenario based on the familiar Rock, Papers, Scissors game. Users train the system, then different expressive variations of the three gestures. GVF provides realtime, continuous classification, allowing early recognition, creating a low latency system distinct from traditional classifiers that recognize static posture after completion. GVF also reports scale, adapting to weak/strong variants of the gestures. This takes game into a new dimension where new rules harnessing gesture variants could be imagined. The system provides audiovisual feedback, creating an action-perception loop, aiding the user to enhance reproducibility, and even modifying their gesture midstream. This raises important interaction design questions of whether the machine "learns" user gesture or whether the user "learns" the system. As a follow up to this study, we are currently mapping EMG gesture variation as reported by GVF to sound synthesis output for musical performance.

Motion capture and Schubert
We have used Gaussian Mixture Modeling (GMM) on consumer motion capture systems (the Kinect) to study the different ways in which different novice, amateur, and trained musicians gesticulate staccato and legato articulation while emulating an orchestra conductor [Anon]. We are currently extending this work to use the GVF with a high-spec motion capture system to look at virtuoso pianist gesture. This builds upon the notion of ancillary gesture in instrumental performance practice [13], and connects to musicological studies of differences in performance tradition [7]. A single phrase in Schubert's *Träumerei* was played by two virtuoso pianists. Each was wearing a full body motion capture system. We captured joint trajectories, as well as audio, video, and MIDI data from the performance. We focused on a phrase in the left hand as it plays two grace notes and jumps an interval of a musical 10^{th}. The two pianists played this with different expressive instructions, and we are currently looking at encoding these gestural differences and seeing if GVF can report on variations of gestural scale and speed as a way to characterise expressive difference.

Conclusions
The work in progress reported here builds on our prior published work [Anon] connecting embodied interaction in HCI to interactive music. In the first example, a non-musical task is used to assess gesture variation, as detected using EMG and machine learning. This will be then rolled out into a music performance scenario for both expert and novice musical use. In the second example, multimodal interaction techniques from HCI were used to capture musical gestures from two pianists. By gleaning expressive difference and applying machine learning to characterise the gestural variation we hope to learn not just about pianistic gesture, but explore the potential for expressive gesticulation to be detected and exploited in non-musical interaction.

References

1. Bouwer, A., Holland, S., and Dalgleish, M. (2013). “Song walker harmony space: embodied interaction design for complex musical skills.” In Holland, S., Wilkie, K., Mulholland, P., and Seago, A. (Eds.) Music and Human-Computer Interaction. Springer.
2. Caramiaux, B., Donnarumma, M., Tanaka, A. “Understanding Gesture Expressivity through Muscle Sensing” *ACM Transactions on Computer-Human Interaction (ACM ToCHI)*. 2015.
3. Tanaka, A., Ortiz, M., Bernardo, F. ” Machine Learning with Early Recognition of Muscle Gesture. DIS2016 submitted.
4. Cope, D. (1996).Experiments in Musical Intelligence. A-R Editions, Madison, WI.
5. Françoise, J., Caramiaux, B., Bevilacque, F. “A Hierarchical Approach For The Design Of Gesture-To-Sound Mappings”. In Proceedings of the 9th Sound and Music Computing Conference, Copenhagen, Denmark, 11-14 July, 2012.
6. Garcia, J., Tsandilas, T., Agon, C., and Mackay, W. 2012. “Interactive paper substrates to support musical creation.” In Proc. CHI’12.
7. Lourenço, S., Martins, L. G., Baltazar, A., Megre, R. “European Piano Schools of Music Performance: Building up a multimodal analysis.” In Proc. ISSEI 2014.
8. Magnusson, T. (2009). Of epistemic tools: Musical instruments as cognitive extensions. Organised Sound, 14(02):168–176.
9. McLean, A., Griffths, D. ,Collins, N., and Wiggins, G. (2010). Visualisation of live code. In Proceedings of the 2010 International Conference on Electronic Visualisation and the Arts, EVA’10, Swinton, UK.
10. Pachet, F. “Enhancing Individual Creativity with Interactive Musical Reflective Systems.” In Musical Creativity: Current Research in Theory and Practice, Deliège, I. And Wiggins, G. Editors, Psychology Press, 2004.
11. Poupyrev, I., Lyons, M. J., Fels, S., and Blaine, T. (2001b). New inter- faces for musical expression. In Extended Abstracts of CHI ’01, Seattle, USA.
12. Anon
13. Wanderley, M., Vines, B. W., Middleton, N., McKay, C., and Hatch, W. (2005). The musical significance of clarinetists’ ancillary gestures: An explo- ration of the field. Journal of New Music Research, 34(1):97—113.
14. Yee-King, M., D’Inverno, M. (2014) “A Social Timeline for Exchanging Feedback about Musical Performances”. In Proc. NIME’14.