Towards a taxonomy for video analysis on collaborative musical tabletops

Anna Xambó
The Open University
a.xambo@open.ac.uk

Robin Laney
The Open University
r.c.laney@open.ac.uk

Chris Dobbyn
The Open University
c.h.dobbyn@open.ac.uk

Sergi Jordà
Universitat Pompeu Fabra
sergi.jorda@upf.edu

This position paper summarises some themes encountered when analysing video data in the context of music performance with interactive tabletops. It presents methodological approaches and coding schemes used for a set of experiments on musical tabletops and collaboration. Finally, it outlines an initial taxonomy based on the outcomes of the projects introduced, which can be used for video annotation of collaborative music interaction.

coding schemes, collaboration, evaluation, musical tabletops, taxonomies, video analysis

1. INTRODUCTION

In recent years there has been a proliferation of studies on interactive tabletops and collaboration in several contexts such as museums (Horn et al 2012) or educational institutions (Piper and Hollan 2009; Rick et al 2011). The general perspective is that interactive tabletops are suited for collaborative activities because these can engage discussion and problem solving, mainly through verbal communication. There are a number of studies of interactive tabletops for music performance (e.g. Reactable, Audiopad), yet there are few studies interested in the assessment of collaboration from the musical angle (Fiebrink et al 2009; Mealla et al 2011). This approach may inform the music technology and HCI domains about how to deal with potentially nonverbal communication during collaborative activities on interactive tabletops.

Video analysis is a tool traditionally used in social sciences for understanding human interactions with technology from an interdisciplinary perspective (Heath et al 2010). It specially supports the understanding of nonverbal communication due to the detailed recording of gestures and motion apart from dialogues. All these qualities make video analysis a useful tool for understanding collaboration on musical tabletops. As reported in Xambó et al (2012b), it is a flexible tool which can be employed on different musical experiments ranging from task-oriented to open tasks. Yet, one major criticism is that it can become highly time consuming. Not only is a matter of refining the research question to be as focused as possible on what to investigate during the analysis, but it is also about not reinventing the wheel when replicating a similar approach to different data.

We here review how video analysis was applied to different case studies of music performance and improvisation with interactive tabletops. Then, we present an initial taxonomy for video annotation. The purpose of this theoretical framework is to facilitate the understanding of real-time collaboration. Lessons learned can inform what are the essential aspects to be considered when designing and evaluating collaborative tabletops in creative and nonverbal situations such as music, but it may also be of utility to other areas such as gaming, design or brainstorming.

2. CASE STUDIES

2.1. Exploratory case study

We conducted a task-oriented experiment on a minimal multi-touch interface (Laney et al 2010). The interface was designed for four players, with four buttons each that would represent a musical instrument (e.g. bass, drums, keyboard and percussion). Each button triggered one different sample, which was consistent with the rest. We gathered three groups of four participants each with mixed levels of expertise. Participants were asked to perform three musical tasks constrained in time and their performances were video-recorded with one general view camera. Of the three tasks, one was a structured task with a score and a coordinator, whereas the other two were unstructured tasks (i.e. sound exploration and musical improvisation). We were interested in understanding collaboration in terms of engagement. For video analysis, we used two coding schemes and techniques: the first approach was bottom-up, adapted from Grounded Theory (Glaser and Strauss 1967; Lazar et al 2009), and based on themes that emerged
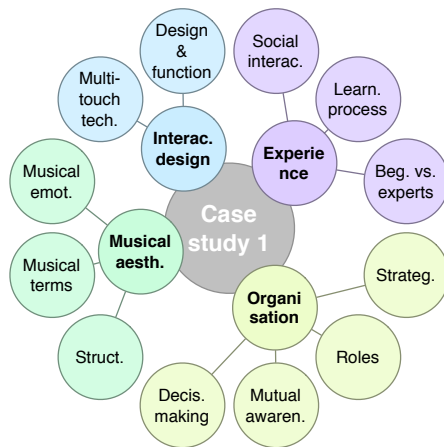
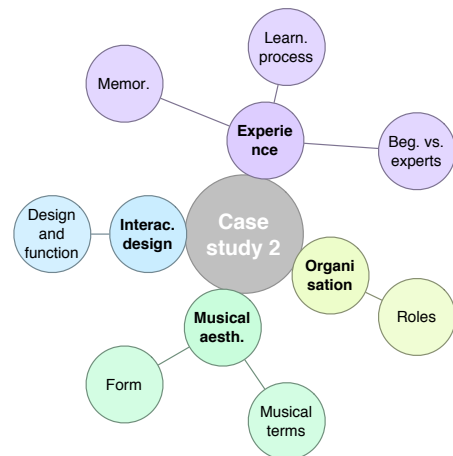**Figure 1:** *Themes emerged from case study 1*

Figure 2: *Themes emerged from case study 2*

from the video annotations of the dialogues and interactions. The second approach was top-down, closer to ethnographic content analysis (Altheide 1987), and based on codes coming from two theoretical frameworks: one of tangible social interaction (Hornecker and Buur 2006) and the second of collaborative musical engagement (Bryan-Kinns and Hamilton 2012). The idea was to use two complementary analytical techniques in order to obtain a more detailed overall picture. Video annotation was done by one of the authors.

The themes that emerged from the transcription of conversations and interactions can be divided into four groups: experience, interaction design, musical aesthetics and organisation (see Figure 1). We found consistency between the content analysed using the two coding schemes. In experience, there were themes related to differences between beginners vs. experts' goals, learning process (e.g. thinking aloud about individual progress) and social interaction (e.g. potential contexts for this device). In interaction design, there were themes about system design and function (e.g. improvements, needs) and multi-touch technology (e.g. responsiveness). In musical aesthetics the themes were mainly related to musical emotions (e.g. playfulness, emotiveness), musical terms (e.g. dynamics, harmony, timbre) and compositional structure (e.g. finale, coda). Finally, in organisation, the themes were about decision-making, mutual awareness, roles and collaborative strategies. In this case study (CS1), video analysis was useful to get an initial picture of collaborative music making on musical tabletops.

2.2. Participatory case study

The purpose of this study was to conduct a participatory design experiment with a basic multi-touch prototype (Xambó et al 2011). The interface consisted of two modules, one for recording tracks

and the second for applying effects on the overall musical output. The prototype was designed for two players. We conducted an informal evaluation with two groups of two people: one expert and one beginner group. Participants were asked to spontaneously play and report their thinking aloud, which was annotated. With the beginner group, the session was video-recorded with a hand-held camera, and video data (e.g. conversations, interactions) was transcribed by one of the authors. This group was provided with a Stylophone to record sounds from this instrument. Participants were asked to try two layouts: a fixed and a flexible layout (adaptable interface). We were interested in what were the interface aspects that could support better the collaboration between participants. The analysis approach was inspired by participatory design practices (Schuler and Namioka 1993), which invites participants to form part of the design process.

The themes that emerged were mainly related to interaction design, but there were also themes on experience, musical aesthetics and organisation (see Figure 2). In experience, the themes were related to differences between beginners vs. experts's expectations, sharing the discoveries as part of the learning process and memories associated to other occasions of making music. In interaction design, the themes were related to system design and function (e.g. features to improve precision and control, flexible vs. fixed layout). In musical aesthetics, there were comments about the form (e.g. musical output) and musical terms (e.g. timbre). Finally, in organisation, there were themes about roles, which were more flexible than expected: there were situations where the two participants divided one of the two tasks into two (e.g. recording was divided into who played the Stylophone and who recorded it, or applying effects was also performed by the two participants). In this case study (CS2),

video analysis was useful for getting informed about how to improve the interaction design with the aim of enhancing the collaboration experience.

2.3. Improvisational case study

The aim of this study (Xambó et al 2012a) was to conduct a longitudinal study of collaborative learning with expert musicians on the Reactable, a well-known tangible and multi-touch tabletop. Reactable is a real-time virtual modular synthesizer with a round shape. It has local and global objects, which allows users to create music by building audio threads (i.e. audio channels) with these objects (Jordà 2008). The study focused on understanding how collaborative learning and the development of expertise can emerge over time with a novel interface in an unconstrained environment such as musical improvisation. We conducted a longitudinal study of four sessions that lasted from 35-45 minutes with 12 expert musicians divided into four groups. We video-recorded all sessions with two cameras (general view and close-up view). We used two complementary approaches for video analysis with the aim at annotating dialogues and interactions. First, a territorial coding scheme was used for quantitative analysis based on tabletop literature (Scott et al 2004). Second, a coding scheme based on social units and behaviours was used for qualitative analysis in accordance with interaction analysis, an analytical tool for understanding nonverbal phenomena based on physical tasks (Jordan and Henderson 1995). Video analysis was done by three of the authors.

The emergent themes from the two coding schemes can be divided into two axes: behaviours (e.g. verbal communication vs. nonverbal communication) and social units (e.g. group vs. individual). In nonverbal communication, we found codes divided into interactional and musical aspects, both collective and individual. Nonverbal interactional codes were about events concerning the reciprocal action between humans and the tangible musical instrument (e.g. interaction techniques). Here we can also situate the territorial codes, which were related to those interactions on personal vs. shared spaces. Nonverbal musical codes concerned with actions related to musical language and musical improvisation practices (e.g. solos, dialogues). In this case study (CS3), video analysis was useful for understanding collaboration over time.

3. TAXONOMY

Informed by the coding schemes for video analysis presented above, we here propose a set of common principles derived from these that may be useful as a starting point for understanding collaborative musical tabletops when analysing video data.

3.1. Organizational

As stated by Weinberg (2005), the principles behind social organizations may help us to understand musical networks, which are based on two main aspects: the level of central control (from centralized to decentralized) and the level of equality between participants (from equality to inequality). Here we talk about levels of control (from local to global) since musicians are co-located and not distributed in a network. In CS1, there is lack of global controls that govern the overall musical output. Depending on the musical task, participants had higher or lower level of equality. Musical roles were fixed: they were determined by the participant's position. In CS2, there is presence of global and local controls. An open task together with the user interface promoted a high level of equality. Musical roles were dynamic: participants spontaneously explored musical roles, and there was mutual modifiability. In CS3, there are global and local controls, and with multiple configuration possibilities. A long-term open task on such a modular interface promoted equal participation and dynamic musical roles between all groups. There was as much mutual modifiability as individual driven actions.

3.2. Interactional

Jordan and Henderson (1995) propose a set of principles to understand non-talk driven interaction with technology, named instrumental interaction. Instrumental interaction involves activities that require the manipulation of physical objects, in this case musical tabletops by means of tangible and/or multi-touch interaction. In CS1, multi-touch interaction was a repeated topic of discussion, mainly because the system responsiveness was less accurate than expected, but also because more features were suggested. In CS2, participants informed about how to improve the prototype in terms of having more control by incorporating multi-touch features. In CS3, the territorial coding scheme informed about participants' territorial behaviour over time with space and objects, whereas the second coding scheme informed about individual and group progress. This progress was reported in terms of development of more complex interactional techniques and configurations of objects by tangible and multi-touch interaction.

3.3. Aesthetical

New aesthetics have been linked to collaborative practices with technologies such as network music (Kim-Boyle 2009). Collaborative music on tabletops can be seen similar to local networks, and thus potential novel aesthetics of play are worth noting. In CS1, musical aesthetics was a major concern that was both verbally discussed, but it was also expressed by gestures (e.g. playfulness,

emotiveness). In CS2, beginners verbally associated the musical output to certain timbres and explored the interface using their own collaboration styles. In CS3, groups and individuals could develop over time different techniques and collaboration styles related to the technology used and each group dynamics.

3.4. Experiential

Since the origins of The International Conference on New Interfaces for Musical Expression (NIME), a major concern has been how to assess user experience (Poupyrev et al 2001). Either explicit or implicit, user experience can inform about interaction design, cognitive processes, or musical aesthetics, among others. In CS1, different themes emerged related to the user experience. These themes tended to be verbally discussed and ranged from beginners vs. experts' goals to sharing the learning process. In CS2, apart from differences between beginners and experts' needs, and also a shared learning process, there were associations with personal memories. In CS3, verbal communication events informed about explicit collaborative learning, which was determined and particularly developed over time by each group.

4. CONCLUSION

In this paper we presented an initial taxonomy for understanding real-time collaboration on tabletops for music performance and improvisation. This theoretical framework has emerged from video analysis of three specific case studies. This taxonomy attempts to provide a useful frame for future studies using video analysis, with the aim at reducing coding time. Yet this generic approach lacks specific codes which should still be developed for each particular case. We attempted to shed light towards a generalizable perspective for design and evaluation of tabletops aiming at collaborative, creative activities such as music. As reported in the three case studies, when supporting collaboration on musical tabletops we should consider aspects such as multi-player, multi-interface control, non-verbal communication, or different levels of expertise (e.g. beginners vs. experts), a set of characteristics which can also be found in the wider area of real-time collaboration on tabletops based on creative activities such as gaming, collaborative sketching or brainstorming.

5. REFERENCES

Altheide, D. L. (1987). *Ethnographic content analysis*. Qual. Sociol., 10(1). 65–77.

Bryan-Kinns, N. and Hamilton, F. (2012). *Identifying Mutual Engagement*. Behav. Inform. Technol., 31(2). 101–125.

Fiebrink, R., Morris, D. and Morris, M. R. (2009). *Dynamic Mapping of Physical Controls for Tabletop Groupware*. Proc. of CHI '09. 471–480.

Glaser, B. G. and Strauss, A. L. (1967). *Discovery of Grounded Theory: Strategies for Qualitative Research*. New York: Aldine de Gruyter.

Heath, C., Hindmarsh, J. and Luff, P. (2010). *Video in Qualitative Research*. London: SAGE.

Horn, M., Atrash Leong, Z., Block, F., Diamond, J., Evans, E. M., Phillips, B. and Shen, C. (2012). *Of BATs and APEs: an interactive tabletop game for natural history museums*. Proc. of CHI '12. 2059–2068.

Hornecker, E. and Buur, J. (2010). *Getting a grip on tangible interaction: A framework on physical space and social interaction*. Proc. of CHI '06. 437–446.

Jordà, S. (2008). *On stage: the reactable and other musical tangibles go real*. Int. J. Art. Tech. 1(3/4). 268–287.

Jordan, B. and Henderson, A. (1995). *Interaction analysis: Foundations and Practice*. J. Learn. Sci. 4(1). 39–103.

Kim-Boyle, D. (2009). *Network Musics: Play, Engagement and the Democratization of Performance*. Contemp. Music Rev. 28 (4-5). 363–375.

Laney, R., Dobbyn, C., Xambó, A., Schirosa, M., Miell, D., Littleton, K. and Dalton, S. (2010). *Issues and Techniques for Collaborative Music Making on Multi-touch Surfaces*. Proc. of SMC 2010. 146–153.

Lazar, J., Feng, J. H. and Hochheiser, H. (2009). *Research Methods In Human-Computer Interaction*. Chichester: Wiley.

Mealla, S., Väljamäe, A., Bosi, M. and Jordà, S. (2011). *Listening to Your Brain: Implicit Interaction in Collaborative Music Performances*. Proc. of NIME 2011. 149–154.

Piper, A. M. and Hollan, J. D. (2009). *Tabletop displays for small group study: affordances of paper and digital materials*. Proc. of CHI '09. 1227–1236.

Poupyrev, I., Lyons, M. J., Fels, S. and Blaine, T. (2001). *New interfaces for musical expression*. Proc. of CHI EA '01. 491–492.

Rick, J., Marshall, P. and Yuill, N. (2011). *Beyond one-size-fits-all: how interactive tabletops support collaborative learning*. Proc. of IDC '11. 109–117.

Schuler, D. and Namioka, A. (1993). *Participatory Design: Principles and Practices*. Hillsdale NJ: CRC / Lawrence Erlbaum Associates.

Scott, S. D., Sheelagh, M., Carpendale, T. and Inkpen, K. M. (2004). *Territoriality in collaborative tabletop workspaces*. Proc. of CSCW '04. 294–303.

Weinberg, G. (2005). *Interconnected Musical Networks: Toward a Theoretical Framework*. Comput. Music J. 29(2). 23–39.

Xambó, A., Laney, R. and Dobbyn, C. (2011). *TOUCHtr4ck: democratic collaborative music*. Proc. of TEI '11. 309–312.

Xambó, A., Hornecker, E., Marshall, P., Jordà, S., Dobbyn, C. and Laney, R. (2012a) (In Preparation). *Let's jam the Reactable: peer learning during musical improvisation on a tabletop*.

Xambó, A., Laney, R., Dobbyn, C. and Jordà, S. (2012b) (In Press). *Video analysis for evaluating music interaction: musical tabletops*. In: Holland, S., Wilkie, K., Mulholland, P. and Seago, A. (eds.). Music and Human-Computer Interaction. London: Springer Verlag.