Mining Interaction Log Data in a Creative Arts Practice using Macro Gestures and Transition Matrices

1st Author Name
Affiliation
City, Country
e-mail address

2nd Author Name
Affiliation
City, Country
e-mail address

3rd Author Name
Affiliation
City, Country
e-mail address

ABSTRACT

In this note we describe the analysis of a data archive for an ensemble musical performance practice on touch screens where transition matrix measures are able to distinguish between sessions of different contexts, and the internal structure of these sessions. Traditional methodologies for studying interfaces that support collaborative creativity can be time-consuming and intrusive to implement but our methods may provide a practical alternative for large scale analysis of creative interactions.

Our archive includes 95 touch data logs of ensemble touch-screen interactions in concerts, rehearsals and formal HCI studies of groups of musicians over more than two years. The data analysis has been undertaken by first classifying these data into higher level gesture types and summarising the performers' transitions between these gestures. We find that two transition-matrix measure, flux, and entropy, can be used to distinguish different performance types and contexts, as well as the internal structure (beginning, middle and end) of the performances in our corpus.

Author Keywords

Creativity Support Tools; Agent; Design; Methodology

ACM Classification Keywords

H.5.5. Information Interfaces and Presentation (e.g. HCI): Sound and Music Computing Systems

INTRODUCTION

Modern computing devices and interfaces continue to present opportunities to support new kinds of creative expression and collaborative interaction [10]. However, it has been acknowledged that understanding and evaluating interfaces in creative contexts can be challenging [12]. In the context of live performance, interface evaluation using questionnaires, or ethnographic data gathering can miss the dynamic needs of artists during collaborative interactions and may be impractical in live performative contexts. When performers use a combuter-based artistic interface, logged interaction data which has been collected for curation and preservation [3], might also allow quantitative investigation into these performances. In this note we will show how an archive of collaborative performance data logs can be subjected to a quantitative analysis based on transition probabilities between gestural states. We will demonstrate that two measures on transition matrices, flux and entropy, show significant differences between different kinds of performances, and give insight into the internal structure of these performances.

Our archive of interaction data logs has been accumulated to support a long-term practice-led research program into simple iPad interfaces for group music-making similar to those described in Martin et al [6]. Over a period of more than two years (April 2013–June 2015), we collected data from group music-making sessions where the participants used our custom-built, touch-screen apps on Apple iPad devices. Although several different apps have been included in our corpus, they shared a common mode of free-form touch interaction where tapping produced short sounds and swirling or swiping produced long sounds. Apart from this commonality, the apps featured a variety of sound palettes, different kinds of visual feedback, and different kinds of networked interactions (between the iPads in an ensemble and with a central server).

In all of our 95 interaction sessions, every touch interaction—every touch down, dragging movement and release—was time-stamped and logged to disk for later analysis. These sessions included formally-structured HCI experiments, where participants were interviewed or surveyed. They also included rehearsals without an audience and performances at arts festivals where creative goals were paramount; in these situations surveys or interviews would have been inappropriate and possibly disruptive. Some of the performances were iPad-only and others used iPads together with other acoustic percussion instruments. Although the majority of the sessions were free-form improvisations, a number of rehearsals and performances involved composed pieces where the musicians read and performed percussion notation using the apps.

Representative photographs of some of these session contexts are are shown Figure 1) and other descriptive statistics about the sessions are shown in Table 1. In our corpus, the median length of sessions is 7m26s and the median number of participants per ensemble is four. Each session record is labelled with metadata about the session *context* (performance, rehearsal, study), session *type* (improvisation or composition), and *instrumentation* (iPads-only, or iPads with acoustic per-

Every submission will be assigned their own unique DOI string to be included here.

Figure 1: Collaborative musical interaction with our touch-screen interfaces in three different contexts: (left to right) rehearsal in a studio, performance with percussion instruments, and performance with just iPads.

	Total	Min	Median	Max
N	95			
Length	13H9M19S	1M19S	7M26S	22M20S
Participants	353	2	4	9
Flux		0.02239	0.1445	0.3246
Entropy		0.7683	3.342	4.442

Table 1: Representative data about the corpus of musical touch-screen interaction sessions used in this study. These sessions were rehearsals, performances, or studies and participants improvised or performed compositions. Each record consisted of touch-data recorded in CSV format and was later analysed with a gestural classification system. Some participants were present in multiple sessions, however this data was not tracked in the corpus.

cussion). The distribution of recorded sessions by context and type is shown in Figure 2.

MACRO GESTURES, TRANSITION MATRICES AND MATRIX MEASURES

A useful way of modelling user interfaces is to consider their theoretical configurations as state vectors and to deploy the methods of matrix algebra to describe transitions between these states during an interaction sequence [15, 14]. An inverted approach is to record users' interactions and manually code interaction states to produce a transition model. This method has been applied to usability analysis of resource planning software for security applications [4] where the transition probability matrix of these state sequences was used to draw conclusions about the nature of this interactions. In the creative arts, editing commands at a computer interface to "live code" computer music has also been coded as states and analysed using transition matrices [13], in both the textual and musical dimensions, in order to arrive at conclusions regarding artistic style.

In our corpus of iPad musical performances, logs of low-level touch data record each performer's micro-gestures. However, it has been established that musical performances on touch-screens can be seen in terms of vocabularies of continuous macro gestures that unfold over several seconds of touch interaction [6]. Improvised performances can be then thought of as a sequence of such states [9] for each performer.

If the performers' touch interactions can be classified as belonging to one of m distinct states, then over the course of a session each participant X generates a length N sequence of these interaction states X_n ($n = 1, \ldots, N$) which can be thought of as Markov process.

If there is an (automated) classification process for identifying each participant's state sequence from their interaction data, then we can estimate the $m \times m$ transition matrix P which characterises this process by setting each entry p_{ij} to be the number of times state j follows state i in the sequence. The transition activity of the whole ensemble can be calculated by averaging the TM for each performer. The matrix can be normalised so that the sum of all elements is equal to one.

To generate state sequences for our corpus we have used an automated method to classify logs of touch interaction, at a regular interval, into nine continuous touch-screen gestures, as described by Martin et al [7]. A Random Forest [2] classifier[1] was applied to 5-second windows of touch-data in our corpus to provide chains of gesture-states for each participant in each session. These chains can be used to create 9×9 transition matrices for individual performers and the whole ensemble in each session.

QUANTITATIVE TRANSITION MATRIX STATISTICS

While the interaction sessions recorded in our corpus were performed by many different participants with varying experience levels, in different contexts and on different instrumental setups, the common interface and method of recording the data allows all the sessions to be transformed into gesture sequences, and transition matrices. While these matrices can be used to visualise the structure of individual performances [13], we wish to extract quantitative measures from these TMs to compare the large number of sessions in our corpus.

To do this we calculate two high-level quantities from each session: "flux" and "entropy". The flux of a sequence is defined as the ratio of state transitions (e.g. $A \rightarrow B$ where

[1]The Random Forest classifier used was from Python's Scikit-learn package[8].

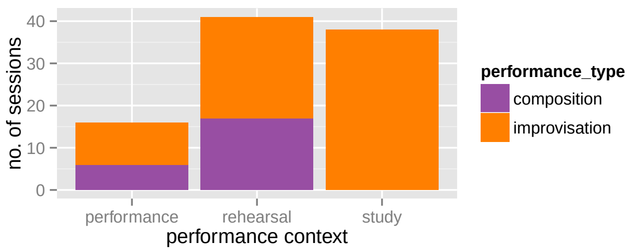


Figure 2: Distribution of the sessions in our interaction corpus by performance context. Improvisations or compositions were repeated several times in rehearsals and study sessions but only performed once in concerts.

$A \neq B$), to self transitions (e.g. $A \rightarrow A$). The flux of a TM P is given by: $\text{flux}(P) = 1 - \text{tr}(P)$ where the trace $\text{tr}(P)$ is the sum of the diagonal entries of P. Intuitively, flux is a measure of how frequently the participant/group changes state. Flux returns a value in the interval $[0, 1]$ and will return 0 when participants never change state, and 1 when participants never stay in the same state for two measurements in a row.

Another useful measure that can be used on a transition matrix is its entropy, defined in the information-theoretic[11] sense: $H(P) = -\sum_{i,j} p_{ij} \log_2(p_{ij})$. This measure is small when the matrix is sparse, and largest when each matrix element is equal. It offers a different perspective on collaborative interactions than the flux measure by capturing the breadth of the gestural space explored by the participants throughout the course of an interaction. Consider the degenerate case of a participant who alternates between two states for a whole performance: the flux in this case will be maximal ($\text{flux}(P) = 1$) since there are no self-transitions, (only $A \rightarrow B$ or $B \rightarrow A$) even though the participant has only used a small subset of the state space. The entropy of this interaction, on the other hand, will be low. Entropy, therefore, is a measure of how broad the participant's exploration of the state space is in a given interaction.

In the rest of this note we will use these matrix measures to answer two questions. First, do the transition matrix measures allow us to distinguish between different performance contexts, types and instrumental setups? Secondly, do these measures allow us further insight into the internal structure of performances?

Differentiating interaction sessions

The distribution of flux and entropy with respect to session contexts, types and instrumentation are shown in Figure 3. A three-way ANOVA procedure was performed on the data set to find whether the metadata could significantly predict the entropy and flux values.

All main effects on flux were found to be significant: performance context ($F(2, 86) = 6.28, p < 0.01$), performance type ($F(1, 86) = 7.21, p < 0.01$), and instrumentation ($F(1, 86) = 20.06, p < 0.001$). Significant interaction effects were also found for performance context and type ($F(1, 86) = 6.18, p < 0.05$), and performance context and instrumentation ($F(1, 86) = 10.48, p < 0.01$).

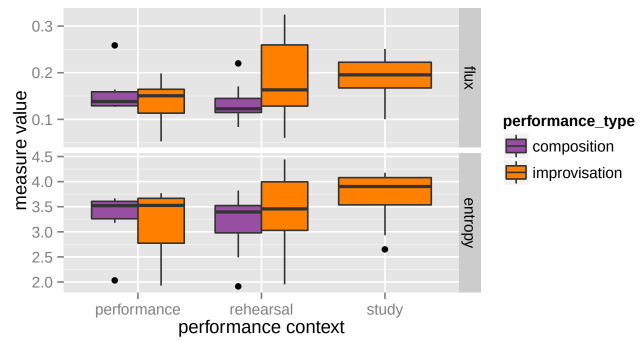


Figure 3: Distributions of flux and entropy values by performance context. Study and rehearsal sessions had the highest flux for improvisations. Compositions had lower flux in rehearsals than performances.

For entropy, significant main effects were found due to performance context ($F(2, 86) = 12.10, p < 0.001$), and instrumentation ($F(1, 86) = 25.86$). The interaction of performance type and instrumentation was also found to be significant ($F(1, 86) = 9.93, p < 0.01$).

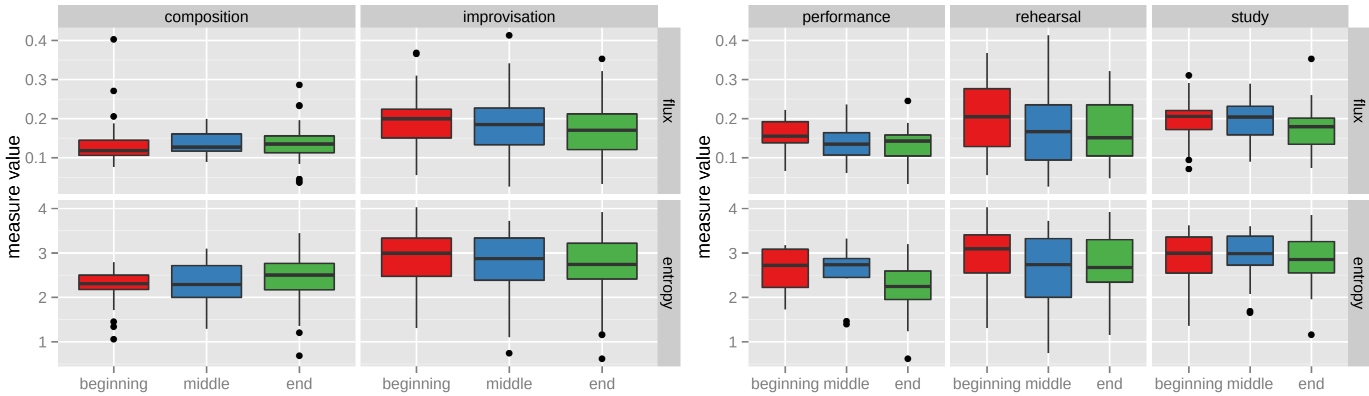
These tests show that the different interaction styles that might be present in performance contexts, types and instrumentations are reflected in the values of flux and entropy of their transition matrices.

Breaking performances into sections

It is clear from these results that flux and entropy can be used to differentiate between performance styles and types, but can they be used to understand some of the internal structure of performances? A classical model for understanding temporal art-forms is the "beginning, middle, end" structure [1]. To investigate how this structure fits our corpus, we can divide the gestural-sequences of each session into thirds and calculate the transition matrices given by each section.

Figure 4a shows that the flux and entropy of compositions are much more stable with respect to section than for improvisations. The stability of compositions is to be expected since performers use the same gestures in each rehearsal and performance of these written scores, however it may be that different kinds of compositions could result in different structural differences. Both measures have a downward trend through sections in improvisations with Kruskal-Wallis tests showing a significant effect of section on flux ($\chi^2(2) = 5.8, p = 0.05$). This could be explained by performers entering an exploration stage at the start of an improvisation, where they change gesture frequently to experiment with different ideas. At the end of an improvisation, performers may change gesture less frequently while "winding up" the performance.

The gestural change in improvisation can be further explored by dividing the sessions by performance context. Figure 4b shows this division for improvisations only. In rehearsals and studies the distribution of flux and entropy over sections has a similar shape, both measures drop in the ending section of



(a) Performances divided by performance type.

(b) Sessions divided by performance context.

Figure 4: The flux and entropy of different interaction sessions divided into beginning, middle, and end.

studies and after the beginning section in rehearsals. In performances however, flux drops after the beginning while entropy drops substantially in the ending section. This indicates a more constrained performance style in this closing section, where performers transition between fewer gestures in the final section of performances.

Performance styles over time

While the interaction session in our corpus have many different contexts, they are all part of a program of artistic and HCI research that has evolved over time. Figure 5 shows the values of flux and entropy for each session in the corpus. Over the course of this research program, different ensembles and activities have been emphasised and some of these differences can be seen as clusters of performances with similar flux and entropy in the two plots. For instance, the earliest series of rehearsals and performances had flux between 0.05 and 0.15, while the period of activity after July 2014 had flux between 0.1 and 0.25. This change corresponded with an intensive series of study sessions and rehearsals where the participants developed a more "fluctuating" style of performance. The entropy between these two periods, however, remains at a similar level, showing the the space of available gestures that the performers explored was relatively stable.

CONCLUSIONS

In this note we have extended previous work using transition matrices to understand HCI processes with the application of two simple matrix measures, flux and entropy, both of which have been demonstrated to have significant discriminatory power with regards to a corpus of 95 collaborative touch-screen interactions. Not only do these measures differentiate between different types of sessions and the internal structure of these sessions, but they have explanatory power that has thrown new light on the musical interactions under examination. It is notable that, for improvised performances, live performances had the lowest measure in both flux and entropy, while lab-studies had the highest. This may be due to the increased pressure on performers when improvising in front of an audience. The internal investigation of performances session revealed more structural variety in improvisations than

Figure 5: Distribution of flux and entropy values of sessions through time. Different performance styles are visible as clusters of similar flux regardless of performance context.

compositions. In response, new works could be composed that specifically explore different gestural structures in touch-screen performance

HCI researchers are increasingly concerned with interactions that happen in-the-wild, where the creative collaborations of many users may be recorded, but where conducting surveys or interviews would be innapropriate. We suggest that analysing interaction-state transitions using matrices and simple measures can lead to new insights into HCI interactions outside the lab.

Data-mining logs of interactions with creative interfaces sugggest further approaches for future work. The macro-gestures that these processes reveal could be used as the basis for new works of new media art [5]. Additionally, a wider dissemination of such log data could better support the replication of findings in HCI research [16]

REFERENCES

1. Aristotle. 350 B.C.E. *Poetics*. The Internet Classics Archive. `http://classics.mit.edu/Aristotle/poetics.1.1.html` Translated by S. H. Butcher.

2. Leo Breiman. 2001. Random Forests. *Machine Learning* 45, 1 (2001), 5–32.

3. David England, Jocelyn Spence, Celine Latulipe, Ernest Edmonds, Linda Candy, Thecla Schiphorst, Nick Bryan-Kinns, and Kirk Woolford. 2014. Curating the Digital: Spaces for Art and Interaction. In *CHI '14 Extended Abstracts on Human Factors in Computing Systems (CHI EA '14)*. ACM, New York, NY, USA, 21–24. DOI: http://dx.doi.org/10.1145/2559206.2559222

4. Thomas Kannampallil and Steven Haynes. 2007. Using Event Based Markov Model Simulation to Analyze Interactive Human Behavior. In *Americas Conference on Information Systems*. Article 142. http://aisel.aisnet.org/amcis2007/142

5. Lev Manovich. 2002. *The Language of New Media*. MIT Press, Cambridge, MA, USA.

6. Charles Martin, Henry Gardner, and Ben Swift. 2014. Exploring Percussive Gesture on iPads with Ensemble Metatone. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 1025–1028. DOI: http://dx.doi.org/10.1145/2556288.2557226

7. Charles Martin, Henry Gardner, and Ben Swift. 2015. Tracking Ensemble Performance on Touch-Screens with Gesture Classification and Transition Matrices. In *Proc. NIME '15*.

8. F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. 2011. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* 12 (2011), 2825–2830.

9. Jeff Pressing. 1988. Improvisation: Methods and Models. In *Generative Processes in Music*, J. Sloboda (Ed.). Oxford University Press, Oxford, UK.

10. Mitchel Resnick, Brad Myers, Kumiyo Nakakoji, Ben Shneiderman, Randy Pausch, Ted Selker, and Mike Eisenberg. 2005. *Design principles for tools to support creative thinking*. Technical Report. Institute for Software Research, School of Computer Science, Carnegie Mellon University. http://repository.cmu.edu/isr/816

11. C. E. Shannon. 1948. A Mathematical Theory of Communcation. *The Bells Systems Technical Journal* 27 (October 1948), 379–423.

12. Ben Shneiderman. 2007. Creativity Support Tools: Accelerating Discovery and Innovation. *Commun. ACM* 50, 12 (Dec. 2007), 20–32. DOI: http://dx.doi.org/10.1145/1323688.1323689

13. Ben Swift, Andrew Sorensen, Michael Martin, and Henry Gardner. 2014. Coding Livecoding. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 1021–1024. DOI: http://dx.doi.org/10.1145/2556288.2557049

14. Harold Thimbleby. 2004. User Interface Design with Matrix Algebra. *ACM Trans. Comput.-Hum. Interact.* 11, 2 (June 2004), 181–236.

15. Harold Thimbleby, Paul Cairns, and Matt Jones. 2001. Usability Analysis with Markov Models. *ACM Trans. Comput.-Hum. Interact.* 8, 2 (June 2001), 99–132.

16. Max L. Wilson, Wendy Mackay, Ed Chi, Michael Bernstein, Dan Russell, and Harold Thimbleby. 2011. RepliCHI - CHI Should Be Replicating and Validating Results More: Discuss. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems (CHI EA '11)*. ACM, New York, NY, USA, 463–466. DOI: http://dx.doi.org/10.1145/1979742.1979491