Can Machine Learning Apply to Musical Ensembles?

Charles Martin
Research School of Computer Science, CECS
Australian National University,
Canberra, ACT, 0200, Australia
charles.martin@anu.edu.au

Henry Gardner
Research School of Computer Science, CECS
Australian National University,
Canberra, ACT, 0200, Australia
henry.gardner@anu.edu.au

Abstract
In this paper we ask whether machine learning can apply to musical ensembles as well as it does to the individual musical interfaces that are frequently demonstrated at NIME and CHI. While using machine learning to map individual gestures and sensor data to musical output is becoming a major theme of computer music research, these techniques are only rarely applied to ensembles as a whole. We have developed a server-based system that tracks the touch-data of an iPad ensemble and have used such techniques to identify touch-gestures and to characterise ensemble interactions in real-time. We ask whether further analysis of this data can reveal unknown dimensions of collaborative musical interaction and enhance the experience of performers.

Author Keywords
machine learning; music; ensemble performance; collaborative creativity; ensemble director agent

ACM Classification Keywords
H.5.m [Information interfaces and presentation (e.g., HCI)]: Miscellaneous

Introduction
Musical ensembles are complicated but fascinating examples of creative, collaborative interaction. When taking part in free-improvisation [1], where the music is created with-

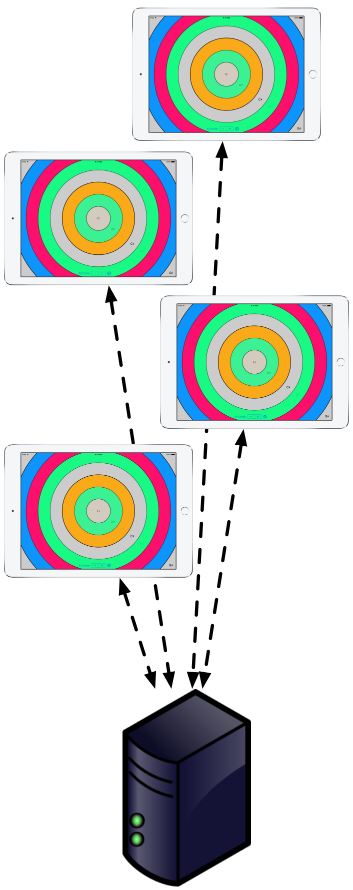

Figure 1: ML techniques have been used inside digital musical instruments to translate gestures into sound, we suggest that these techniques could be used to track ensembles of performers and to uncover new musical interactions.

out boundary during the performance, ensembles display continual negotiation and game-like interaction that has been well-documented [2]. Machine Learning (ML) has recently become a major theme in computer music practice, where ML algorithms have been used to map gestural data captured from sensors to sound [4]. Frameworks such as Wekinator [6], that assist with this process, have led to it becoming widespread. We propose that ML can be applied to a whole ensemble of performers as well. This could lead to real-time understandings of ensemble negotiations and interactions, and the large-scale musical structure of improvised performances.

Improvising algorithms that participate in musical ensembles as performers are often explored in computer music [5]. An alternative role for a computer system in improvising groups is as an ensemble director agent [11], where the system tracks all performers simultaneously and directs high level aspects of the performance such as musical structure. We suggest that information sourced from ensemble-application of ML could be used by ensemble director agents for high-level interactions with performers.

Challenges and Rewards of Ensemble Performance

Ensemble performance is a complex phenomena to track. Not only does each performer produce a potential stream of interaction and audio data from their instrument, but performers make and break connections with each other through (mostly non-verbal) communication, and simply by playing in certain ways. Sensing and collecting this data is an ongoing challenge. Experiencing this process, as well as the resultant music, makes ensemble performance rewarding for participants and audiences.

Finding ways to track some of these interactions through ML could be used to create systems that enhance ensemble performances by assisting with the formation of particular structures. These systems could direct ensembles to unusual artistic outcomes, and could, themselves, be composed to support a variety of interactions. Enhancing the connections between ensemble performers could be especially useful in musical education, recreational music-making, or for assisting collaborative performances over the internet.

Metatone Classifier

Metatone Classifier is a server application designed to connect to an ensemble of touch-screen performers and to automatically classify their touch gestures throughout a performance. Previous research showed that performers with our apps use continuous percussive touch-gestures to express musical ideas [8]. These gestures have no specific start or end point, so rather than attempt to *recognize* command gestures, our system makes classifications every second based on the previous 5 seconds of touch-data to build a sequences of touch gestures for each performer.

Metatone Classifier uses a Random Forest classifier [3] from the Python Scikit-Learn package [10] to identify each window of touch-data as one of nine basic performance gestures in five simple groups: nothing, taps, swipes, swirls, and combinations. Training data was sourced from a formally collected process that is detailed in previous work [9].

During performances Metatone Classifier generates sequences of touch gestures for each performer. Transition Matrices (TMs) can be used to characterise these sequences for all performances simultaneously [12]. We have applied simple measures on TMs of these sequences to understand the musical behaviour of the ensemble and, in particular, to react to sudden changes of musical direction that segment free-improvised performance [9]. Gesture classifications

Figure 2: We have collected data from more than 150 collaborative interaction sessions using our ensemble director agent that applies ML techniques to the ensemble's touch-movements. These sessions include rehearsals, performances, installations, and demonstrations. This corpus constitutes a significant dataset for further investigations of creative interaction.

and structural data are returned to the touch-screen apps which respond by changing their user interface according to several designs.

Having implemented this system and performed with it for around two years, we have collected data from more than 150 collaborative interaction sessions in ensembles of between two and seven members. This constitutes around 275000 gestural classifications at one second intervals. Although this is far from a big-data corpus, it represents a significant dataset of collaborative interaction that could be interrogated with statistical techniques. What such an investigation would reveal about creative collaboration and how DMIs and ensembles should respond to this data could be the topic of much future work.

Although our recent work with Metatone Classifier has demonstrated that this system can enhance some aspects of performance, and although our artistic explorations of agent-directed performance has been successful and rewarding, we feel that our design could be further improved. Gestural sequences could yield further information about interactions between ensemble members that could be used in computer musical interfaces. Our scheme for classifying touch gestures has only been used with a small number of iPad apps developed within our research group and could be useful for other touch-screen apps or even other styles of instrument. Although Metatone Classifier communicates with connected devices through a web-server, it has only been used in local performances. Experiments with distributed, agent-mediated performance, could produce interesting artistic and research results. Metatone Classifier is available online as a git repository [7] and we welcome suggestions for further development!

Conclusion

Our experience with Metatone Classifier suggests that applying ML techniques to musical ensemble performances can lead to the discovery of new musical interactions, and enhance free-improvised performances. Creative use of these techniques could help create new Ensemble Director Agents that interact with established performers, and recreational musicians. The emergence of large sets of ensemble performance data suggests many questions for future investigations: What group interactions can be found in this data? Can group behaviours in these sessions be predicted? How should DMIs react to new information about ensemble interactions?

References

[1] Derek Bailey. 1993. *Improvisation: Its Nature and Practice in Music*. Da Capo, Cambridge, MA.

[2] David Borgo. 2006. Sync or Swarm: Musical Improvisation and the Complex Dynamics of Group Creativity. In *Algebra, Meaning, and Computation*, Kokichi Futatsugi, Jean-Pierre Jouannaud, and José Meseguer (Eds.). Lecture Notes in Computer Science, Vol. 4060. Springer, Berlin, Heidelberg, 1–24. DOI: http://dx.doi.org/10.1007/11780274_1

[3] Leo Breiman. 2001. Random Forests. *Machine Learning* 45, 1 (2001), 5–32. DOI: http://dx.doi.org/10.1023/A:1010933404324

[4] Baptiste Caramiaux and Atau Tanaka. 2013. Machine Learning of Musical Gestures. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME '13)*, Woon Seung Yeo, Kyogu Lee, Alexander Sigman, Haru Ji, and Graham Wakefield (Eds.). KAIST, Daejeon, Republic of Korea, 513–518. http://nime.org/proceedings/2013/nime2013_84.pdf

[5] Arne Eigenfeldt, Oliver Bown, Philippe Pasquier, and Aengus Martin. 2013. Towards a Taxonomy of Musical Metacreation: Reflections on the First Musical Metacreation Weekend. In *Proceedings of the Musical Metacreation Workshop (MUME 2013)*. AAAI, Palo Alto, CA, USA.

[6] Rebecca Fiebrink, Dan Trueman, and Perry R Cook. 2009. A Meta-Instrument for Interactive, On-the-fly Machine Learning. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME '09)*. Pittsburgh, PA, USA, 280–285. http://www.nime.org/proceedings/2009/nime2009_280.pdf

[7] Charles Martin. 2014. Metatone Classifier: Ensemble director agent application. Git Repository. (2014). https://github.com/cpmpercussion/MetatoneClassifier

[8] 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

[9] Charles Martin, Henry Gardner, and Ben Swift. 2015. Tracking Ensemble Performance on Touch-Screens with Gesture Classification and Transition Matrices. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME '15)*, Edgar Berdahl and Jesse Allison (Eds.). Louisiana State University, Baton Rouge, LA, USA, 359–364. http://www.nime.org/proceedings/2015/nime2015_242.pdf

[10] 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.

[11] Jeff Pressing. 1990. Cybernetic Issues in Interactive Performance Systems. *Computer Music Journal* 14, 1 (1990), 12–25. DOI: http://dx.doi.org/10.2307/3680113

[12] Ben Swift, Andrew Sorensen, Michael Martin, and Henry J 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