Exploring Social Mobile Music with Tiny Touch-Screen Performances

Charles P. Martin
Department of Informatics,
University of Oslo, Norway
charlepm@ifi.uio.no

Jim Torresen
Department of Informatics,
University of Oslo, Norway
jimtoer@ifi.uio.no

ABSTRACT

Touch-screen musical performance has become commonplace since the widespread adoption of mobile devices such as smartphones and tablets. However, mobile digital musical instruments are rarely designed to emphasise collaborative musical creation, particularly when it occurs between performers who are separated in space and time. In this article, we introduce an app that enables users to perform together asynchronously. The app takes inspiration from popular social media applications, such as a timeline of contributions from other users, deliberately constrained creative contributions, and the concept of a reply, to emphasise frequent and casual musical performance. Users' touch-screen performances are automatically uploaded for others to play back and add reply performances which are layered as musical parts. We describe the motivations, design, and early experiences with this app and discuss how musical performance and collaboration could form a part of social media interactions.

1. INTRODUCTION

Popular social media apps for mobile devices have allowed millions of users to engage with creative production of images and text. These devices' cameras, touch-screens, powerful processors, and portability suggest on-the-go creativity, and it would appear that straightforward sharing with friends, or a wider network of followers, is a key factor in encouraging users to create content of all forms. Given the many affordances of mobile devices, it has been well noted that they are suitable platforms for mobile music making [1]. Despite many creative mobile digital musical instruments (DMIs) appearing in recent years, we have yet to see the widespread adoption of musical creation as an integrated element of social media. Furthermore, few musical apps have attempted to emphasise ensemble, rather than individual, performance, even though group music-making is often seen as a valuable social activity.

In this article, we present the design for *MicroJam* [1] [2], a collaborative and social mobile music-making app. This design emphasises casual, frequent, and social performance. As shown in Figure 1, the app features a very simple touch-screen interface for making electronic music where skill is not a necessary prerequisite for interaction. Performances are limited to five seconds and uploaded automatically to encourage improvisation and creation rather than editing. Users can reply to others' performances by recording a new layer, this combines social interaction with musical ensemble interaction. In sections 1.1 and 1.2, we will motivate MicroJam's design with a discussion of music-making in social media, and the possibilities for asynchronous and distributed collaborations with mobile musical interfaces. In Section 2 we will describe the app in detail and, by way of a preliminary evaluation, we will explore the performance contributions of early users and testers of the app.

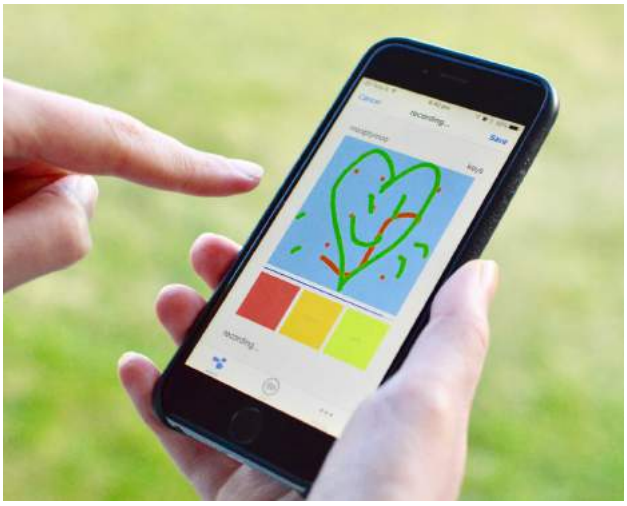

Figure 1. MicroJam allows users to create tiny touch-screen performances that are shared to other users. Replies to performances form distributed and asynchronous duets. A video demo of MicroJam is available online: https://youtu.be/SkUjjQd13KU

1.1 Social Media and Music-Making

Many social media platforms emphasise the value of concise and frequent contributions by users. Twitter famously limits written notes to 140 characters. Instagram used a square image format and film-like processing to emphasise the creative possibilities of mobile phone images. Both of these services show users a timeline of contributions from others that they follow, who might be friends, celebrities, or just interesting strangers. Posts in these services are in-

[1] Source code and further information about MicroJam is available online at DOI:10.5281/zenodo.322364

Copyright: © 2017 Charles P. Martin et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 Unported License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Same Location	Mobile Phone Orchestra [10]; Ensemble Metatone [11] Viscotheque Jams [12].	Locative Performance
Different Location	Networked Musical Performance Magic Piano [14]	Glee Karaoke [13]; MicroJam Performances.
	Same Time	**Different Time**

Table 1. Ensemble performances typically occur with all participants in the same time and place; however, other situations are possible. MicroJam focuses on performances that are distributed (different place) and asynchronous (different time).

tended to be frequent, casual, and ephemeral. Snapchat introduced an interface that optimised replying to many image messages in one session. This app further emphasised ephemerality by introducing images that expire and can only be viewed for a few seconds after being opened. These and other social applications have attracted many millions of users and encouraged their creativity in the written word and photography. While social media is often used to promote music [3], music *making* has yet to become an important creative part of the social media landscape.

While music is often seen as an activity where accomplishment takes practice and concerted effort, casual musical experiences are well-known to be valuable and rewarding creative activities. Accessible music making, such as percussion drum circles, can be used for music therapy [4]. Digital Musical Instruments (DMIs) such as augmented reality instruments [5] and touch-screen instruments [6] have also been used for therapeutic and casual music-making. In the case of DMIs, the interface can be designed to support creativity of those without musical experience or with limitations on motor control. Apps such as *Ocarina* [7] and *Pyxis Minor* [8] have shown that simple touch-screen interfaces can be successful for exploration by novice users as well as supporting sophisticated expressions by practised performers.

Some mobile music apps have included aspects of social music-making. Smule's *Leaf Trombone* app introduced the idea of a "world stage" [9]. In this app, users would perform renditions of well-known tunes on a unique trombone-like DMI. Users from around the world were then invited to critique renditions with emoticons and short text comments. World stage emphasised the idea that while the accuracy of a rendition could be rated by the computer, only a human critic could tell if it was ironic or funny. Indeed, *Leaf Trombone*, and other Smule apps have made much progress in integrating musical creation with social media.

1.2 Jamming through Space and Time

While performance and criticism is an important social part of music-making, true musical collaboration involves performing music together. These experiences of group creativity can lead to the emergence of qualities, ideas, and experiences that cannot be easily explained by the actions of the individual participants [15]. Mobile devices have often been used in ensemble situations such as *MoPho* (Stanford Mobile Phone Orchestra) [10], *Pocket Gamelan* [16], and *Ensemble Metatone* [11]; however, in these examples, the musicians played together in a standard concert situation.

Given that mobile devices are often carried by users at all times, it would be natural to ask whether mobile device ensemble experiences can be achieved even when performers are not in a rehearsal space or concert venue. Could users contribute to ensemble experiences at a time and place that is convenient to them? The use of computer interfaces to work collaboratively even when not in the same space and time has been extensively discussed. In HCI, groupware systems have been framed using a time-space matrix to address how they allow users to collaborate in the same and different times and places [17]. For many work tasks, it is now common to collaborate remotely and at different times using tools such as Google Docs or Git; however, distributed and asynchronous musical collaboration is not as widely accepted.

In Table 1, we have applied the time-space matrix to mobile musical performance. Conventional collaborative performances happen at the same time and location. Even with mobile devices, most collaboration has occurred in this configuration. Collaborations with performers distributed in different locations but performing at the same time are often called networked musical performances [18]. Early versions of Smule's *Magic Piano* [14] iPad app included the possibility of randomly assigned, real-time duets with other users. Networked performances are also possible with conventional mobile DMIs and systems for real-time audio and video streaming.

Performance with participants in different times, the right side of Table 1, are less well-explored than those on the left. Performances where participants are in the same place, but at different times, could come under the banner of locative performances, such as *Net_dérive* [19] or *Sonic City* [20], where geographical location is an important input to a musical process.

The final area of the matrix involves music-making with performers that are in different places and different times. *Glee Karaoke* [13] allows users to upload their sung renditions of popular songs, and add layers to other performers' contributions. The focus in this app, however, is on singing along with a backing track, and the mobile device does not really function as a DMI but an audio recorder. These limitations rule out many musical possibilities; for instance, touch-screen DMIs can create a variety of sounds from similar interfaces, so orchestras of varying instruments could be assembled from many remote participants to improvise original music. It remains to be seen whether large scale musical collaboration between distributed users is possible. It seems likely, however, that such collaborations would uncover hidden affordances of the medium as has been seen in other distributed online media such

as "Twitch Plays Pokémon" [21]. Our app, MicroJam, also fits into this lower-right quadrant. In the next section, we will describe how this new app enables distributed and asynchronous collaboration on original musical material.

2. MICROJAM

MicroJam is an app for creating, sharing, and collaborating with tiny touch-screen musical performances. This app has been specifically created to interrogate the possibilities for collaborative mobile musical performances that span space and time. While these goals are lofty, the design of MicroJam has been kept deliberately simple. The main screen recalls social-media apps for sharing images. Musical performances in MicroJam are limited to very short interactions, encouraging frequent and ephemeral creative contribution. MicroJam is an iOS app written in Swift and uses Apple's CloudKit service for backend cloud storage. The source code is freely available for use and modification by other researchers and performers [2]. In the following sections we will discuss the design of the app, the format of the tiny musical performances that can be created, and some early experiences and musical performances recorded with users.

2.1 Design

The primary screen for MicroJam is the performance interface, shown in the centre of Figure 2 and called *jam!*. This screen features a square touch-area which is initially blank. Tapping, swirling, or swiping anywhere in this area will create sounds and also start recording touch activity. All touch interaction in this area is visualised with a simple paint-style drawing that follows the user's touches, and is simultaneously sent to a Pure Data patch running under `libpd` [22] to be sonified. After five seconds of touch interaction, the recording is automatically stopped (although the performer can continue to interact with the touch area). The recording can be subsequently replayed with the *play* button, or looped with the *jam* button.

Users of MicroJam can choose the sound-scheme used to sonify their interactions in the jam interface. At present, four options are available through the settings screen: *chirp*, a simple theremin-like sound with pitch mapped to the x-dimension of interactions; *keys*, a Rhodes-like keyboard sound with pitch and timbre mapped to the x- and y-dimensions; *strings*, a simple modelled string sound that performs mandolin rolls and changes timbre as the performer moves around the screen; and *drums*, a simple drum machine with different sounds mapped to quadrants of the screen. As each of these sound-schemes is implemented as a Pure Data patch, it is straightforward to add more to MicroJam in future.

Previously recorded performances, and those recorded by other users and downloaded from the server, are listed in the *world* screen as shown on the left side of Figure 2. Each performance is represented by a visual trace of the touch-drawing captured during recording as well as the name of the contributor and date of the performance. Any one of these performances can tapped which opens the recording back up in the jam screen for playback. When playing back, both the sound and visualised touch-interactions are replayed in the touch-area.

When viewing a previously saved performance, the user can tap *reply*, to open a new layer on top of the recording. As shown in the right side of Figure 2, the previous as well as current touch-visualisations are shown as well as separate sonifications for each layer. At present, only one reply is possible in MicroJam; however, the ability to reply to replies, creating multi-layered performances is planned for future versions.

2.2 Tiny touch-screen performances

MicroJam is intended to provide a musical experience that is similar to other mobile social creativity applications. One innovation in this field has been to constrain contributions, leading to more frequent interactions and possibly higher creativity due to the lower stakes and effort. Musical interactions in MicroJam are similarly constrained to be tiny touch-screen performances: those that are limited in the area and duration of interaction. We define a *tiny* performance as follows:

1. All touches take place in a square subset of the touch-screen.
2. Duration of the performance is five seconds.
3. Only one simultaneous touch is recorded at a time.

Such performances require very little effort on the part of users. While some users may find it difficult to conceive and perform several minutes of music on a touch-screen device, five seconds is long enough to express a short idea, but short enough to leave users wanting to create another recording. It has been argued that five seconds is enough time to express a sonic object and other salient musical phenomena. [23]. While the limitation to a single touch may seem unnecessary on today's multi-touch devices, this stipulation limits tiny performances to monophony. In order to create more complex texture or harmony, performers must collaborate, or record multiple layers themselves.

For transmission and long-term storage, tiny touch-screen performances are stored as simple comma-separated values files. The data format records each touch interactions time (as an offset in seconds from the start of the performance), whether the touch was moving or not, x and y locations, as well as touch pressure. In MicroJam, the visual trace of performances is also stored for later use in the app, although this could be reconstructed from the touch data.

2.3 User Data and Early Experiences

The prototype version of MicroJam has been distributed and demonstrated to small numbers of researchers and students. These early experiences have allowed us to streamline the creative workflow of the app and to observe the kind of tiny performances that can be created in the jam interface. Since the release of the first prototype, around 200 tiny performances have been collected. Most of the

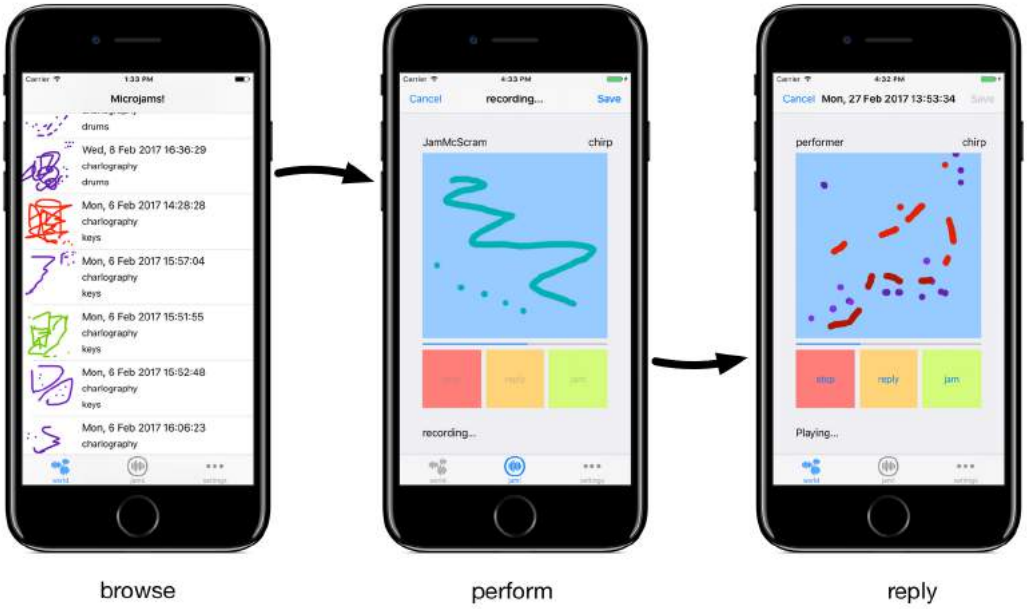

Figure 2. The MicroJam interface allows users to browse a timeline of performances (left), create new performances (centre), and reply, or play in duet, with previously recorded performances (right).

participants were computer science and engineering students with little computer music experience, a few music technology students were also included.

The visualisations of a subset of these performances are reproduced in Figure 3. This figure shows the variety of touch-interaction styles that have already been observed in performers. Many of the interactions are abstract, resembling scribbles that show the user experimenting with the synthesis mapping of the jamming interface. In some performances, repeated patterns can be seen, where performers have repeated rhythmic motions in different parts of the touch-area. A number of the performances are recognisable images: figures, faces, and words. These users were particularly interested in how the interface maps between drawing, a visual and temporal activity, and sound. Observing the users, it seems that some focused on what a particular drawing might *sound* like, while others were interested in what particular sounds *look* like. At present, these performance recordings have not been analysed with respect to which sound-scheme was in use, and how reply layers fit together. Further experiments could seek to understand these relationships.

It has been gratifying to hear that several early users of the app greatly enjoyed the experience of creating tiny performances, and wished that similar interactions could be integrated into existing social apps. These users immediately set about recording multiple performances, exploring the sound-schemes and the creative possibilities of the touch-screen interface. Other users, however, had a lukewarm reaction to the concept of free-form touch-screen musical performance. It could be that casual users do not expect to be able to create original music. After all, musical performance is often (erroneously, we feel) seen as a task only for specialists with a high level of training or talent. It may be a hard sell to ask users to create their own music, and to collaborate with others. Rather than a discouragement, we see this as an opportunity to continue developing mobile music experiences that push the boundaries of everyday music making. Understanding how comfortable users would be composing original tiny performances could be addressed in future studies. The app design could also include more guidance, such as a training system, or more extensive visual feedback, to help users who are unsure about making touch-screen music.

3. CONCLUSIONS AND FUTURE WORK

In this paper we have advocated for social apps for creating music, as opposed to more popular written and visual media. We have argued that such apps could take advantage of the ubiquity of mobile devices by allowing users to collaborate asynchronously and in different locations, and shown that such modes of interaction are relatively unexplored compared to more conventional ensemble performances. Our app, MicroJam, represents a new approach to asynchronous musical collaboration. Taking inspiration from the ephemeral contributions that typify social media apps, MicroJam limits performers to tiny five-second touch-screen performances, but affords them extensive opportunities to browse, playback, and collaborate through responses. MicroJam's tiny performance format includes a complete listing of the touch interactions as well as a simple visualisation of touch interactions. This format allows performances to be easily distributed, viewed, and studied. Early experiences with MicroJam have shown that users can engage with the interface to create a range of interesting performances. While some see potential to include music-making in their social media activities, others may lack confidence about producing music, even in the tiny performance format.

There is much scope for refinement and development of

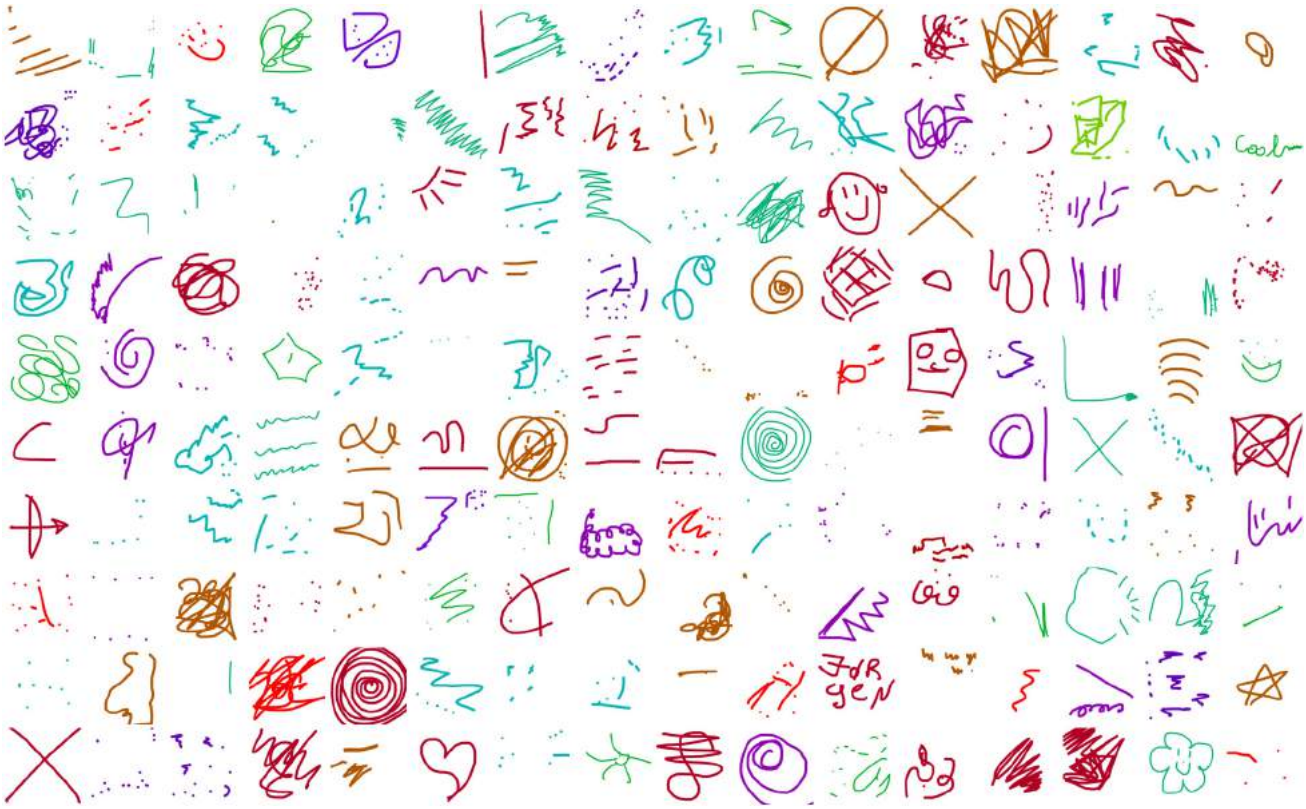

Figure 3. The visual trace of performances in the prototype for MicroJam. Some performances focus on similar touch gestures in different locations of the jamming area, others appear to be exploratory scribbles, and several focus on visual interaction.

MicroJam in future work. Enhancements such as more varied and refined sound-schemes and visualisations could be appealing to users. Future efforts could also focus on enhancing distributed collaborations. Allowing more than one reply to MicroJam performances could generate very large collaborations between users. If multiple reply threads were available, users might be able to generate complex performance structures. Automatic traversal of such structures could constitute a kind of generative composition with users' original musical material. As MicroJam affords a high quantity of short interactions, data collected from the app could be used to train generative models for tiny performances. It may be possible to predict potential replies to a given performance or to generate performances that extend beyond five seconds while keeping within a user's style. Exploring these interactions with experienced, as well as novice, musicians could point the way to more expressive and musically powerful interactions in MicroJam.

Integrating music-making, as opposed to the more conventional music appreciation, or music promotion, into social media calls into question the musical confidence and creative aspirations of users. Future studies could examine how users could potentially include mobile music-making in everyday social media interactions. The precedent set by other successful mobile music apps suggest that users do seek out musical outlets for their creativity. Future work with MicroJam may focus on guiding beginner users towards more musical confidence and rewarding their exploratory improvisations.

Acknowledgments

This work is supported by The Research Council of Norway as a part of the Engineering Predictability with Embodied Cognition (EPEC) project, under grant agreement 240862.

4. REFERENCES

[1] A. Tanaka, "Mapping out instruments, affordances, and mobiles," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, ser. NIME '10, K. Beilharz, A. Johnston, S. Ferguson, and A. Y.-C. Chen, Eds. Sydney, Australia: University of Technology Sydney, 2010, pp. 88–93. [Online]. Available: http://www.nime.org/proceedings/2010/nime2010_088.pdf

[2] C. P. Martin, "Microjam v0.1-alpha," Git Repository, Feb. 2017. [Online]. Available: https://doi.org/10.5281/zenodo.322364

[3] S. Dewan and J. Ramaprasad, "Social media, traditional media, and music sales." *MIS Quarterly*, vol. 38, no. 1, pp. 101–121, 2014.

[4] S. Scheffel and B. Matney, "Percussion use and training: A survey of music therapy clinicians," *Journal of Music Therapy*, vol. 51, no. 1, p. 39, 2014. DOI: 10.1093/jmt/thu006

[5] A. G. D. Correa, I. K. Ficheman, M. d. Nascimento, and R. d. D. Lopes, "Computer assisted music therapy: A case study of an augmented reality musical system for children with cerebral palsy rehabilitation," in *International Conference on Advanced Learning Technologies*. IEEE, 2009. DOI:10.1109/ICALT.2009.111 pp. 218–220.

[6] S. Favilla and S. Pedell, "Touch screen ensemble music: Collaborative interaction for older people with dementia," in *Proceedings of the 25th Australian Computer-Human Interaction Conference*, ser. OzCHI '13. New York, NY, USA: ACM, 2013. DOI:10.1145/2541016.2541088 pp. 481–484.

[7] G. Wang, "Ocarina: Designing the iPhone's magic flute," *Computer Music Journal*, vol. 38, no. 2, pp. 8–21, 2014. DOI:10.1162/COMJ_a_00236

[8] T. Barraclough, D. Carnegie, and A. Kapur, "Musical instrument design process for mobile technology," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, E. Berdahl and J. Allison, Eds. Baton Rouge, Louisiana, USA: Louisiana State University, May 2015, pp. 289–292. [Online]. Available: http://www.nime.org/proceedings/2015/nime2015_313.pdf

[9] G. Wang, S. Salazar, J. Oh, and R. Hamilton, "World stage: Crowdsourcing paradigm for expressive social mobile music," *Journal of New Music Research*, vol. 44, no. 2, pp. 112–128, 2015. DOI:10.1080/09298215.2014.991739

[10] J. Oh, J. Herrera, N. J. Bryan, L. Dahl, and G. Wang, "Evolving the mobile phone orchestra," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, ser. NIME '10, K. Beilharz, A. Johnston, S. Ferguson, and A. Y.-C. Chen, Eds. Sydney, Australia: University of Technology Sydney, 2010, pp. 82–87. [Online]. Available: http://www.nime.org/proceedings/2010/nime2010_082.pdf

[11] C. Martin, H. Gardner, and B. Swift, "Tracking ensemble performance on touch-screens with gesture classification and transition matrices," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, ser. NIME '15, E. Berdahl and J. Allison, Eds. Baton Rouge, LA, USA: Louisiana State University, 2015, pp. 359–364. [Online]. Available: http://www.nime.org/proceedings/2015/nime2015_242.pdf

[12] B. Swift, "Chasing a feeling: Experience in computer supported jamming," in *Music and Human-Computer Interaction*, ser. Springer Series on Cultural Computing, S. Holland, K. Wilkie, P. Mulholland, and A. Seago, Eds. London, UK: Springer, 2013, pp. 85–99.

[13] R. Hamilton, J. Smith, and G. Wang, "Social composition: Musical data systems for expressive mobile music," *Leonardo Music Journal*, vol. 21, pp. 57–64, December 2011. DOI:10.1162/LMJ_a_00062

[14] G. Wang, "Game design for expressive mobile music," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, vol. 16. Brisbane, Australia: Queensland Conservatorium Griffith University, 2016, pp. 182–187. [Online]. Available: http://www.nime.org/proceedings/2016/nime2016_paper0038.pdf

[15] R. K. Sawyer, "Group creativity: Musical performance and collaboration," *Psychology of Music*, vol. 34, no. 2, pp. 148–165, 2006. DOI:10.1177/0305735606061850

[16] G. Schiemer and M. Havryliv, "Pocket Gamelan: Swinging phones and ad-hoc standards," in *Proceedings of the 4th International Mobile Music Workshop*, Amsterdam, May 2007.

[17] S. Greenberg and M. Roseman, "Using a room metaphor to ease transitions in groupware," Department of Computer Science, University of Calgary, Tech. Rep. 98/611/02, 1998.

[18] A. Carôt, P. Rebelo, and A. Renaud, "Networked music performance: State of the art," in *Audio Engineering Society 30th International Conference*, Mar 2007. [Online]. Available: http://www.aes.org/e-lib/browse.cfm?elib=13914

[19] A. Tanaka and P. Gemeinboeck, "Net_dérive: Conceiving and producing a locative media artwork," in *Mobile Technologies: From Telecommunications to Media*, G. Goggin and L. Hjorth, Eds. London, UK: Routledge, 2008.

[20] L. Gaye, R. Mazé, and L. E. Holmquist, "Sonic City: the urban environment as a musical interface," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, ser. NIME '03. Montreal, Canada: McGill University, 2003, pp. 109–115. [Online]. Available: http://www.nime.org/proceedings/2003/nime2003_109.pdf

[21] S. D. Chen, "A crude analysis of twitch plays pokemon," *arXiv preprint arXiv:1408.4925*, 2014. [Online]. Available: https://arxiv.org/abs/1408.4925

[22] P. Brinkmann, P. Kirn, R. Lawler, C. McCormick, M. Roth, and H.-C. Steiner, "Embedding Pure Data with libpd," in *Proceedings of the Pure Data Convention*. Weimar, Germany: Bauhaus-Universität Weimar, 2011. [Online]. Available: http://www.uni-weimar.de/medien/wiki/PDCON:Conference/Embedding_Pure_Data_with_libpd:_Design_and_Workflow

[23] R. I. Godøy, A. R. Jensenius, and K. Nymoen, "Chunking in music by coarticulation," *Acta Acustica united with Acustica*, vol. 96, no. 4, pp. 690–700, 2010. DOI:10.3813/AAA.918323