Piano Genie as an Interactive System

Shauna Morrisey*
Kaushal Sali*
smorrisey@gatech.edu
ksali3@gatech.edu
Georgia Institute of Technology
Atlanta, Georgia

ABSTRACT

This paper examines Piano Genie as it fits into the larger category of interactive music systems. Piano Genie is a research project out of Google Magenta. The main feature of Piano Genie is it's ability to map a full piano down to just eight keys using a recurrent LSTM. Using a dimension space analysis model, Piano Genie is compared to two similar systems on seven axes. Advantages and disadvantages of Piano Genie are also discussed.

ACM Reference Format:
Shauna Morrisey and Kaushal Sali. 2019. Piano Genie as an Interactive System. In *Proceedings of ACM Conference (Conference'17).* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Piano Genie [6] is an interactive music generation system [4] that makes it possible for people to play the piano without prior training. Piano Genie was developed by the Google Magenta group and is available as a web demo [5]. Piano genie provides users with a very basic input interface: 8 keys, which get mapped to the entire range 88 of keys on a piano. Having such a simplistic input allows users to map a high level idea of the flow of music to actual musically relevant contours. Of course, here we consider 'musically relevant' in the context of traditional western classical music. Piano Genie strikes a balance between the automatic generation of music and user's interaction with the instrument. This is achieved through the use of deep learning, specifically recurrent auto-encoders.

2 DESCRIPTION

Piano Genie presents users with an input interface having 8 buttons. In spite of being such a minimal interface it is able to capture the essence of the player's musical intention. A button press triggers a note that lasts untill the button is released just like a regular piano. Each button press is mapped to one of the 88 keys of a piano. So, to play chords or produce polyphony the user can press multiple buttons simultaneously. The user can drive the melody by pressing higher buttons to play higher notes and lower buttons to play lower notes. Thus a high level idea of the flow of music can be realized in the form of a meaningful piano composition.

Figure 1: A picture of the hardware version of Piano Genie connected directly to interface with an acoustic piano. [6]

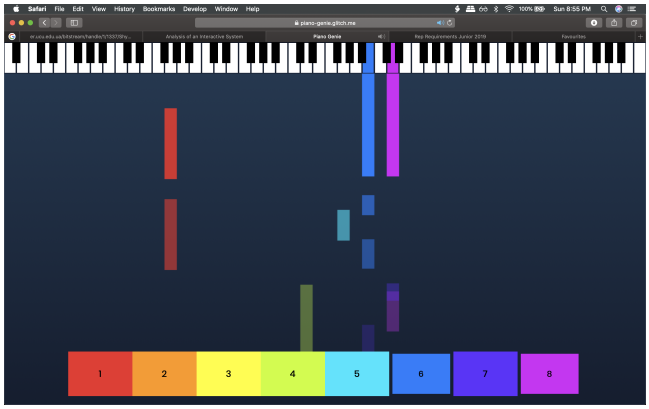

Figure 2: A screenshot of the online version of Piano Genie being used. [5]

The system uses a dynamic one to one mapping between the input and the output notes. It is dynamic in the sense that the mapping is decided based on the previous button sequences the user has pressed. This process of updating mappings is determined by the decoder of a recurrent auto encoder which has been trained on musical pieces played by skilled musicians. The auto encoder is trained by providing it midi note sequences of the pieces from Piano-e-competition dataset[1]. This input (belonging to a set of 88 values) is mapped to 8 output values by the encoder. The decoder decodes these 8 values back to the 88 keys. The network is trained such that the decoder generates the input sequence provided to the

*Both authors contributed equally to this research.

Conference'17, July 2017, Washington, DC, USA

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

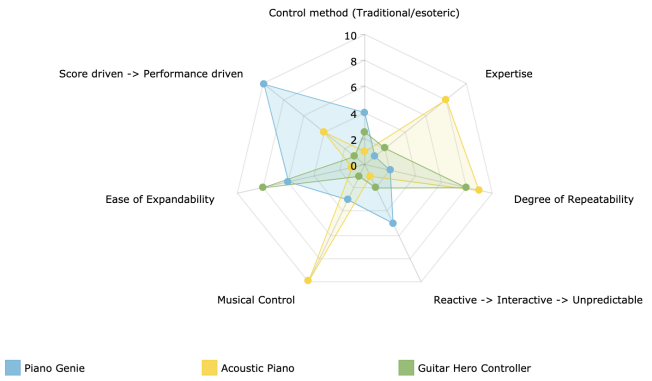

Figure 3: Classification Chart based on the Birnbaum Dimension Space Analysis method[3] comparing Piano Genie, an acoustic piano, and a Guitar Hero Controller.

encoder with higher probability. The recurrent LSTM (Long Short Term Memory) network is able to learn and predict short musical phrases which are mapped to the encoder output of 8 values and are reconstructed by the decoder.

The system however does not control the output temporally. So the user needs to press the buttons in time so as to give the output a sense of rhythm.

3 CONTEXT

Piano Genie is an intelligent controller that is intended to be used by non-musicians. It promotes users to improvise on a piano. According to Donahue et al.[6], "It keeps the players in an improvisational loop as they respond to the generative procedure in real-time".

Piano Genie was a research project at Google Magenta as a part of a summer internship[4]. More than being an end product, the system is a demonstration of a technology. It is a demonstration of the innovative way of mapping. It shows the capability of autoencoders to learn and formulate mappings between complex and simple interfaces.

4 CLASSIFICATION

For the purpose of analyzing this system, the Birnbaum et al. Dimension Space Analysis method [3] provided flexibility and precision to discuss Piano Genie and to compare Piano Genie to other systems. Dimension Space Analysis allows for multiple systems to be compared easily among multiple axes. The axes were selected because they would "meaningfully display design differences among devices" [3].

Two instruments were selected from the space in which Piano Genie occupies for comparison on these axes. Acoustic piano was selected for it's obvious connection to Piano Genie and to compare and contrast the method of performing on both systems. On the other end of the spectrum, a Guitar Hero Controller [2] was selected. It contains a similar simplistic, accessible version of a standard guitar interface. The Guitar Hero Controller is similar to a guitar in the same way that Piano Genie is similar to a piano. The Guitar Hero Controller is different than Piano Genie in that each song is hard coded into a game with accompanying tracks. Different parameters were chosen to highlight strengths and weaknesses of Piano Genie, an acoustic piano, and a Guitar Hero Controller. All axes values were determined subjectively.

4.1 Classification Axes

- The *Control Method* axis specifies in what way input is received. Examples of traditional Control Methods would be buttons, keys, or knobs, where esoteric Control Methods would be sensors.
- The *Expertise* axis is used to visualize how much background knowledge is expected. High Expertise values would represent a large time commitment on the user to learn how to use the system.
- The *Degree of Repeatability* axis measures the consistency of the system. Some systems may be more directly mapped, with the same gesture causing the exact same result.
- The *Reactive -> Interactive -> Unpredictable* axis is based on the definition of an interactive system provided by Drummond [7]. He describes the definition of an interactive system as a balancing act. On one side of the balancing act, a system could be too reactive, with one to one mappings that feel more like an instrument than an interactive system. On the other side, the resulting output may be so unpredictable that the user feels no control or connection to the system[7].
- The *Musical Control* axis is a subjective measurement of how much expressive control the user has.
- The *Ease of Expandability* axis has been defined as how easy it would be for someone to expand the current system. Low values convey that it would be difficult to expand the given system and high values convey that it would be easy to expand the system.
- The *Score driven -> Performance driven* axis describes a specific type of predictability, repeatability, and the performance practice surrounding the system. This axis also describes how users most often interact with the system in a meaningful way.

4.2 Discussion

In this section, each axis will be discussed as it relates to the three systems being compared. All three systems, Piano Genie, acoustic piano, and the Guitar Hero Controller, all use traditional control methods. However, the lack of a direct one-to-one mapping caused Piano Genie's eight keys to lay closer to the center of this continuum.

One goal of Piano Genie was to create an authentic experience for the inexperienced musician[4]. Piano Genie met this goal, so it scored very low on the Expertise axis. For the Guitar Hero Controller, there is a sense of skill required. In some cases, it might be reasonable to practice a song to improve these skills. Acoustic piano obviously requires the most expertise and practice from performers.

Because of the strict one-to-one mapping of the piano, it also provides a very repeatable or predictable system for players. The Guitar Hero Controller will provide the same output when playing the same song, but if the same sequence is played in a different song, the output will be different. Piano Genie takes this to an extreme, as there is no expectation that the same sequence will provide the

same output. Piano Genie does not consistently map the exact same gestures to the same resultant audio.

No systems were considered to lack a connection between gesture and output as to classify it as unpredictable. The piano, however, is a classic example of a reactive system. There is a clear expectation of what and how will play when interacting with the system. The Guitar Hero Controller was closer to an interactive system, but still quite reactive. Each song is mapped differently; mappings may even change within one song, but the gestures are still maintained. Playing the same song again will result in the same gestures leading to the same output. Piano Genie was in the sweet spot of an interactive system. Dependent on user input, Piano Genie does real time calculations and generates output while maintaining the input gestures.

The Guitar Hero Controller provides the least musical control to the user. They are interacting with a game and focus only on hitting the correct buttons at the correct time. Piano Genie provides authentic expressive experiences to the user, but in a way that does not give the user all control. The constraints provided by the interactive system limit the musical control. The piano provides the user with all the musical and expressive control.

It would be the easiest to expand the Guitar Hero System. Certain aspects, like encoding new songs, are extremely easy. The controller recently added a sixth button, in a three by two array, demonstrating another way the system could be expanded from the older five button system. Piano Genie could be expanded in several ways. Adding new training data to the dataset, implementing sound control methods, and adding more buttons are all ways to expand the current system. The acoustic piano was deemed the most difficult of the three to expand. Adding a string would require changing the size of the body, creating the correct tension on the string, and adding the mechanics behind each piano key. It is possible that these changes could affect the timbre of the piano. Some specific ways to expand the piano would be by making a prepared piano. A prepared piano is the technique of adding items to the interior of the piano. Cage is one notable user of the prepared piano.

Guitar Hero was the most score driven approach of the three systems. The controller is entirely dependent on the score of the song being played for output. The piano was scored near the center, as there are use cases for both approaches. From classical to jazz improvisation, piano covers the entire spectrum. Piano Genie is a performance based system, as all the online performances are improvisations [8] [9] [10]. It would probably be very difficult to practice and perform a score on Piano Genie given the inconsistent output.

5 AFFORDANCES

As we can see from Section 2, Piano genie provides a user with a very simple interface that can be used by anybody without prior musical knowledge or training. It therefore has a very low floor. Yet, the system is designed in such a way that it respects the role human interaction. Anybody can operate a system that automatically generates the complete melody at a push of a button, but this disregards the satisfaction that one gets out of creating something of their own. Creators of Piano Genie understands this and therefore provide a note-level control to the users. There is an immediate output for each input which makes the system highly responsive. There is no temporal modification, which gives users a sense of high control over the system. Since users can drive the musical contour up or down in pitch by pressing higher buttons or lower buttons, high level ideas can be formally expressed on a piano. All these factors argue for a high ceiling design. There are some things that limit the user's expression: The output phrase might not always turn out to be as expected which can be seen as limiting one's control over the precise notes. Another limitation of the system is the user control over melodic resolution. It completely depends on what the system has learned from the training data and the user's input. As a result, the user may not always be able to end the piece when they want and might feel like the phrase is being dragged out.

6 ASSESSMENT

Piano Genie provides for a wide range of interactions with specific strengths, limitations, and opportunities for growth. Piano Genie has a low floor, or requires little expertise to use. It is easy for an untrained musician to approach and use Piano Genie. This was an important goal set out by Donahue et al [6]. While the interface is minimalistic, there is a high level of expressivity. One of the largest limitations was a lack of genres. Based on the sourcing for their data set, a classical piano competition, the genres Piano Genie can express is limited. Piano Genie is almost certainly a solo instrument. To another user, the output might sound subjectively good, but be impossible to play along with. Several opportunities have arisen throughout interacting with Piano Genie. More keys, for example, could be added to further gestural control. Playing chords and melodies together feels congested on the 8 keys. An interesting modification that can be done to the system is to have two controllers which span the entire piano range. One can be used by to play chords and another to play the lead. However the system would have to careful here to not map the two buttons on the controllers to same key simultaneously. A larger data set, potentially with more genres, could allow for different experiences. An entirely new direction Piano Genie could implement would be to leverage sound design mechanics. This would act mostly as a setting that could be preset or changed on the fly from a computer, but it probably could not be contained or edited from the eight key interface.

7 CONCLUSION

In this paper, we looked at the interactive music system Piano Genie. We described it's control mechanism and how the system works. We analyzed the system in context of the Birnbaum's classification method. The Birnbaum classification method provided us a way to compare Piano Genie with similar systems like an acoustic piano and a Guitar Hero Controller.

REFERENCES

[1] 2018. Piano-e-Competition. http://www.piano-e-competition.com
[2] 2019. Guitar Hero Controller. https://www.guitarhero.com/game/controller
[3] David Birnbaum, Rebecca Fiebrink, Joseph Malloch, and Marcelo M Wanderley. 2005. Towards a dimension space for musical devices. In *Proceedings of the 2005 conference on New interfaces for musical expression*. National University of Singapore, 192–195.
[4] Chris Donahue. 2019. Piano Genie lecture at ACM SIGCHI. https://www.youtube.com/watch?v=OCc4V2aWpHQ

[5] Chris Donahue, Ian Simon, and Sander Dieleman. [n.d.]. Piano Genie Web Demo. https://piano-genie.glitch.me/
[6] Chris Donahue, Ian Simon, and Sander Dieleman. 2018. Piano genie. *arXiv preprint arXiv:1810.05246* (2018).
[7] Jon Drummond. 2009. Understanding interactive systems. *Organised Sound* 14, 2 (2009), 124–133.
[8] Magenta. 2018. Piano Genie Improvisation 1. https://www.youtube.com/watch?v=YRb0XAnUpIk
[9] Magenta. 2018. Piano Genie Improvisation 2. https://www.youtube.com/watch?v=0hk7Imdvh9s
[10] Magenta. 2018. Piano Genie Improvisation 3. https://www.youtube.com/watch?v=_QlzbodchAc