Dueto: Accessible, Gaze-Operated Musical Expression

Stephanie Valencia
Carnegie Mellon University
Human-Computer Interaction Institute
svalenci@cs.cmu.edu

Dwayne Lamb, Shane Williams, Harish S. Kulkarni, Ann Paradiso, Meredith Ringel Morris
Microsoft Research
dwayne@visualbytestudios.com
{shanewil, annpar, harish.kulkarni, merrie}@microsoft.com

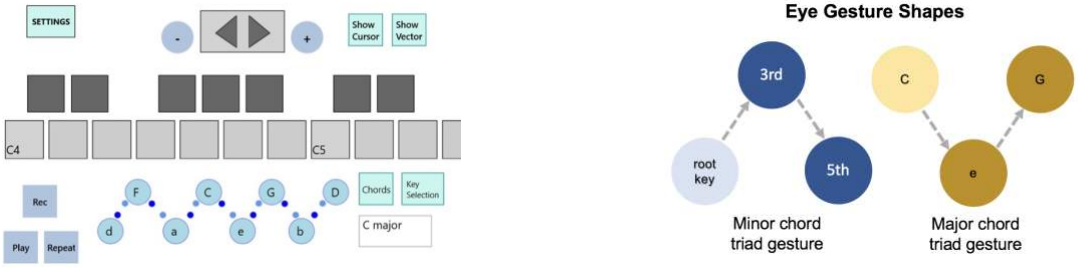

Figure 1. Dueto explores multimodality combined with gaze-based interactions for the creation of accessible musical interfaces. Dueto can be played with eye-gaze alone, gaze and switch, or gaze and multi-touch, where multi-touch input is carried out by a partner. The harmony ladder is located on the lower part of Dueto's UI (left) showing notes corresponding to the C major key. The small circles that connect the main notes appear and disappear in a cascade animation, hinting to the user how to perform the shape of the gestures that enable playing multiple notes at the same time (chords) with just the eyes (right).

ABSTRACT
Gaze-tracking technologies can enable computer access for users who are unable to use standard input devices. However, using gaze as input poses challenges for interactions that require visual planning, like playing a digital instrument. We explore how multimodality can support eye-controlled musical expression by designing different multi-modal gaze interactions around a digital instrument we call Dueto. We tackle three design goals: creating an instrument that is explorable, easy to learn, and allows feature controllability. We showcase three different multimodal interactions for music playing such as eye gaze only, gaze + switch, and gaze + partner mode.

Author Keywords
Accessibility; Eye tracking; Gaze input; Motor impairments; Multimodal interaction; Musical interfaces.

ACM Classification Keywords
K.4.2. Computers and Society: Social issues – assistive technologies for persons with disabilities.

ASSETS '19, October 28–30, 2019, Pittsburgh, PA, USA

ACM ISBN 978-1-4503-6676-2/19/10.
https://doi.org/10.1145/3308561.3354603

INTRODUCTION
In gaze-based musical interfaces, it is important to balance having large-enough targets to accommodate for tracker error [4] with interactions that can be completed in a timely manner according to the gaze event that is being detected (e.g., saccades, dwell). These different gaze events can constrain the rate of playable notes when used to trigger musical events, posing a challenge to control rhythm. Another challenge in generating computer music with the eyes is that gaze input for voluntary active selection is counter-intuitive with the user's natural tendency to use eye gaze for inspection [11,17]. Exploring an instrument is a natural part of playing music, so having gaze-responsive regions in a graphical UI can cause involuntary activation of visual targets (the "Midas Touch" problem). In addition, gaze interaction alone does not allow for multiple notes to be selected at a time, unless multiple notes are already grouped in one "virtual key." The lack of expressive mechanisms in gaze can limit what a user can express with a digital instrument using gaze alone.

To address these challenges, prior work has focused on designing musical interfaces that combine gaze plus other complementary input modes such as switches and facial gestures [1,2,19]. These interfaces have used gaze for selection while using the additional mode for playing and modifying the musical output. Other approaches [2,3,14,15] have designed virtual on-screen keys with different configurations to avoid the Midas Touch problem and facilitate music playing with just the eyes. We present Dueto, an interface that explores novel multi-modal gaze-based interactions for accessible live music playing.

Different to prior work and to extend it [2,7], Dueto presents an eye-gesture feature for nuanced harmony and melody creation and a collaborative multi-player mode. In this work, we describe how Dueto can support musical expression for people with motor disabilities.

DUETO

Dueto was developed using the Universal Windows Platform and the Gaze Interaction Library on Windows 10. We used a commercial eye tracker, the Tobii PCEye Mini. Dueto supports melody and harmony creation through multiple input modalities. The bottom part of Dueto's UI is dedicated for harmony creation using triad chords composed of the root note, the third note, and the fifth note, organized in what we call a harmony ladder. As shown in Figure 1, Dueto uses a pre-selected key chosen by the user (in this example, C major) from which the harmony ladder is based. Having a constraint like this one allows the harmony portion to be easier to explore as playing any key on the harmony ladder would sound good, giving the user freedom to explore different chord combinations. It is important to note that the user can turn the chords on/off based on their preference: they can choose to interact with the harmony ladder to play individual notes (chords off) or use the ladder to create chords (chords on).

A melody is a succession of notes forming a distinctive sequence. On the upper section of the interface (Figure 1, left), we have a scrollable piano keyboard. For users who have had some previous musical exposure, the spatial layout of the piano keys can serve as a reference. The keys of the piano are spread out to create rest areas for the gaze, in order to mitigate Midas Touch issues. Each key acts as a button. In order to play a melody, a player must activate the key. There are different ways to do this depending on the selected interaction modality. The piano keys scroll to change octaves when the user activates the arrow buttons.

With the goal of designing gaze UIs to manage trade-offs of user abilities and costs of using certain input modalities, we explore interactions with alternative modality combinations. The idea behind Dueto is not to substitute one mode with another mode, but to provide different alternatives to carrying out the same task. Throughout the design process we used the guidelines provided by Orio et al. [12] to make Dueto have 1) learnability, 2) explorability, and 3) feature controllability. We designed different iterations based on feedback from two musicians, one of whom is a daily eye-tracker user. With this in mind, in addition to the modality options, Dueto allows the user to choose how to visualize their gaze position in different ways: as a vector made of the trace of gaze points from the last 30ms ("show vector" button) or as a gaze point cursor on the screen, when toggling the "show cursor" button.

Gaze Only: If a user decides to use Dueto in gaze only mode, they can play a melody by either interacting with the harmony ladder or the scrollable piano keys. Dwell time of 100ms is used for active selection in this mode. To enable a gaze-only player to play a melody and harmony (chords) together, we include the functionality of recording a sequence using the harmony ladder and replaying it so that the user can play a melody on top of it. To access the looping functionality, the user can use the buttons located on the bottom left side of the interface. A user can play minor and major chords using eye-gestures. Major chords and minor chords follow their own unique pattern. In order to play a chord, a user must look at the root note of the chord and then complete the triad gesture *(Figure 1, right).*

Gaze + Switch: In the gaze + switch mode a player can play a note by looking at it and then using a switch to activate it. This two-step process mitigates the Midas Touch problem, enabling users to explore the interface without accidentally triggering notes. Playing a note by using gaze for fast key selection while using a tapping or pressing motion to control a switch can create nuanced rhythms and enable better tempo control than activating notes with gaze alone.

Gaze + Partner: The gaze + partner mode enables two players to interact with Dueto by having one player perform gaze-only interactions and having a second musician using multi-touch input. To enable multi-touch input, we used a Makey-Makey [20] and created a small capacitive touch keyboard with 8 keys. The recommended interaction is to split melody and harmony between players, similar to performing a four-hands piano piece. In this scenario, the gaze-only player can perform eye-gesture chords, while the second player can directly interact with the piano keys. The gaze-only player controls the piano scrolling, selecting what area of the piano maps to the multi-touch keys.

DISCUSSION AND CONCLUSION

Dueto introduces new elements yet to be evaluated in the space of gaze-based music playing, including an eye gesture harmony interaction, a partner modality, and the impact of enabling different types of gaze visualizations while playing. We predict that having an eye gesture-based harmony will allow users to play with their own style, as gestures are not performed the same way and at the same rate by every user. This will give the experience of playing Dueto a sense of explorability and nuance. We also look forward to evaluating the impact of having a partner as an additional input. This gaze + partner modality will require coordination and communication. Designing musical UIs in the context of diversity of abilities requires the exploration of multiple and alternative input modes. Understanding how to optimize each input modality can inspire new musical interactions that can in the end enrich the space of gaze-based HCI, while also making music creation accessible to people with varied physical abilities.

ACKNOWLEDGEMENTS

We thank the MSR Enable and Ability teams as well as Jeremy Best for his feedback as a musician and experienced eye-tracker user and Jonathan Brachthäuser for his supportive guidance on musical composition and theory.

REFERENCES

1. Sam Bailey, Adam Scott, Harry Wright, Ian Symonds, and Kia Ng. 2010. Eye . Breathe . Music : creating music through minimal movement. In Proceedings of the 2010 international conference on Electronic Visualisation and the Arts (EVA'10), pp. 254-258. BCS Learning & Development Ltd., Swindon, UK, 254-258.
2. Nicola Davanzo and Marco Porta. 2018. Playing Music with the Eyes Through an Isomorphic Interface. In Proceedings of the Workshop on Communication by Gaze Interaction (COGAIN '18). ACM, New York, NY, USA, Article 5, 5 pages. https://doi.org/10.1145/3206343.3206350
3. Carlos H. Morimoto, Antonio Diaz-Tula, José A. T. Leyva, and Carlos E. L. Elmadjian. 2015. EyeJam : A Gaze-Controlled Musical Interface. In Proceedings of the 14th Brazilian Symposium on Human Factors in Computing Systems (IHC '15), ACM, New York, NY, USA, Article 37, 9 pages. https://doi.org/10.1145/3148456.3148493
4. Anna Maria Feit, Shane Williams, Arturo Toledo, Ann Paradiso, Shaun Kane, and Meredith Ringel Morris. 2017. Toward Everyday Gaze Input : Accuracy and Precision of Eye Tracking and Implications for Design. In Proc. CHI Conference on Human Factors in Computing Systems (CHI '17). ACM. https://doi.org/10.1145/3025453.3025599
5. James D. Foley, Victor L. Wallace, and Peggy Chan. 1984. Human Factors of Computer Graphics Interaction Techniques. *IEEE Computer Graphics and Applications* 4, 11: 13–48. https://doi.org/10.1109/MCG.1984.6429355
6. Stewart Greenhill and Cathie Travers. 2016. Focal : An Eye-Tracking Musical Expression Controller. In Proceedings of the 2016 Conference on New Interfaces for Musical Expression (NIME '16). 230–235.
7. Anthony J Hornof. 2014. The Prospects For Eye-Controlled Musical Performance. In Proceedings of the 2014 Conference on New Interfaces for Musical Expression (NIME '14). 461–466.
8. Anthony Hornof, Jeffrey Stolet, and Tim Halverson. 2008. Bringing to Life the Musical Properties of the Eyes. University of Oregon Department of Compuer and Information Science, Technical Report 08. 2008; 5.
9. Ishan Chatterjee, Robert Xiao, and Chris Harrison. 2015. Gaze + Gesture : Expressive, Precise and Targeted Free-space Interactions. In Proceedings of the 2015 ACM on International Conference on Multimodal Interaction (ICMI '15). ACM, New York, NY, USA, 131-138. https://doi.org/10.1145/2818346.2820752
10. Robert J.K. Jacob. 1991. The Use of Eye Movements in Human-Computer Interaction Techniques : What You Look At is What You Get. ACM Transactions on Information Systems, Vol. 9, No. 3: 152–169. http://dx.doi.org/10.1145/123078.128728
11. Manu Kumar, Andreas Paepcke, and Terry Winograd. 2007. EyePoint : Practical Pointing and Selection Using Gaze and Keyboard. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '07). ACM, New York, NY, USA, 421-430. https://doi.org/10.1145/1240624.1240692
12. Nicola Orio, Norbert Schnell, and Marcelo Wanderley. 2001. Input devices for musical expression: borrowing tools from HCI. In Proc. NIME: 1–4. https://doi.org/10.1162/014892602320582981
13. Vijay Rajanna and Tracy Hammond. 2018. A Fitts ' Law Evaluation of Gaze Input on Large Displays Compared to Touch and Mouse Inputs. In Proceedings of the Workshop on Communication by Gaze Interaction (COGAIN '18). ACM, New York, NY, USA, Article 8, 5 pages. https://doi.org/10.1145/3206343.3206348
14. Zacharias Vamvakousis and Rafael Ramirez. 2012. Temporal Control In the EyeHarp Gaze-Controlled Musical Interface.In Proceedings of the 2012 Conference on New Interfaces for Musical Expression (NIME '12).
15. Stephen Vickers and Matthew Smalley. EyeGuitar : Making Rhythm Based Music Video Games Accessible Using Only Eye Movements. In Proceedings of the 7th International Conference on Advances in Computer Entertainment Technology (ACE '10). ACM, New York, NY, USA, 36-39. https://doi.org/10.1145/1971630.1971641
16. Jacob O Wobbrock and Shaun K Kane. 2011. Ability-Based Design : Concept , Principles and Examples. ACM Transactions on Accessible Computing (TACCESS). Article 9 (April 2011), 27 pages. http://dx.doi.org/10.1145/1952383.1952384
17. Shumin Zhai, Carlos Morimoto, and Steven Ihde. 1999. Manual And Gaze Input Cascaded (MAGIC) Pointing. In Proceedings of the SIGCHI conference on Human Factors in Computing Systems (CHI '99). ACM, New York, NY, USA, 246-253. http://dx.doi.org/10.1145/302979.303053
18. Xbox adaptive controller. Hardware product available at https://www.xbox.com/en-US/xbox-one/accessories/controllers/xbox-adaptive-controller
19. Eye Conductor. Software description available at https://abledata.acl.gov/product/eye-conductor
20. Makey-Makey. Hardware and firmware available at https://makeymakey.com