Focal : An Eye-Tracking Musical Expression Controller

Stewart Greenhill
stewart.greenhill@gmail.com

Cathie Travers
accordion@cathietravers.com

ABSTRACT

We present *Focal*, an eye-tracking musical expression controller which allows hands-free control over audio effects and synthesis parameters during performance. A see-through head-mounted display projects virtual dials and switches into the visual field. The performer controls these with a single expression pedal, switching context by glancing at the object they wish to control. This simple interface allows for minimal physical disturbance to the instrumental musician, whilst enabling the control of many parameters otherwise only achievable with multiple foot pedalboards. We describe the development of the system, including the construction of the eye-tracking display, and the design of the musical interface. We also present a comparison of a performance between *Focal* and conventional controllers.

Author Keywords

Computer music, eye tracking, expression controller, augmented reality, user interface design

ACM Classification

H.5.2 [Information Interfaces and Presentation]: User Interfaces - input devices and strategies, interaction styles. H.5.5 [Information Interfaces and Presentation] Sound and Music Computing - methodologies and techniques, systems.

1. INTRODUCTION

This work focuses on electronic processing of the accordion, but the principles apply to any instrument. Electronic effects can enhance acoustic sounds, creating new timbres and sonic textures. Effects have parameters that the player may wish to modulate for artistic reasons. The challenge is to expressively control these parameters during performance.

The piano accordion has a three octave treble keyboard which is played with the right hand, and a 120-button bass keyboard which is played with the left. The treble and bass sounds come from independent reeds boxes driven by a shared bellows. The timbre of the sound can be varied by enabling various sets of reeds using the register switches above the treble and bass keyboards.

Playing the accordion occupies both hands so additional articulation usually involves pedals. Due to the shape and position of the instrument the accordionist often cannot see their feet, making it difficult to control more than one or two pedals. An accordion can weigh 10-15kg, so an added challenge is to maintain posture and balance while moving the feet between pedals. We aim to devise *a hands-free interface for expressive control which minimises the number of pedals.*

Eye trackers offer part of the solution by measuring where a person is looking. As a viewer interprets a scene the eyes move subconsciously in a series of quick jumps called *saccades* lasting around 30ms, interspersed with *fixations* of between 100 and 400ms. *Smooth pursuits* occur when the eyes track a moving object, but only *fixations* can be consciously controlled. Thus it is difficult to use gaze position alone for fine control without introducing filtering and lag. However, the eyes can very rapidly jump between objects, so gaze can reliably signal a choice between alternatives.

Most previous work on eye-controlled music has relied entirely on gaze position for the musical interface. Our novel approach is to separate control into two parts: *articulation* and *selection*. An expression pedal is used for articulation, and gaze selects which parameter to control. Parameters are shown graphically in a see-through head-mounted display, allowing free head and body movement and visual communication with other players and audience, and giving feedback about the state of all parameters. The movement of the eyes replaces the more costly movement of the feet between pedals.

2. EXPRESSION PEDALS

Pedals first appeared in the pipe organs of the 13th century. Pedalboards allowed organists to play simple bass lines, or to hold drone notes while continuing to play manuals with both hands. Expression pedals are situated above the pedalboard and are usually operated with the right foot, while the left foot plays the bass melody. Toe studs may also be used to enable quick stop changes. These can be seen in Figure 1 on either side of the expression pedals. This idea carried through to electronic organs, which often include momentary contact toe-switches on the volume pedal to trigger stop changes, note “glides”, or to toggle an automatic accompaniment.

Electronic effects became popular along with the electric guitar in the 1950s but these were typically incorporated into amplifiers. The first pedal-controlled effects appeared in the 1960s when transistors made it it possible to package an entire effect circuit and all controls inside the pedal. These pedals include a bypass foot-switch, and some effects such as “Wah” also have expression controls. This modularisation makes it easy to combine different effects but can lead to complexity, requiring a “tap-dance” routine when changing settings.

MIDI foot controllers are convenient for musicians who use digital effects or laptop-based sound processing. Figure

NIME'16, July 11-15, 2016, Griffith University, Brisbane, Australia.
.

Figure 1: Organ pedalboard with 32 bass keys. In the center are three three expression controllers: two pedals and one roller. Five toe studs an be seen on the right, and four on the left.

Figure 2: MIDI Foot Controller, including nine foot switches and two continuous controller pedals.

Figure 3: Expression pedal from an Excelsior Digisyzer (c1980) hybrid accordion showing dual toe-switches.

2 shows a pedalboard which includes nine digital switches, and two expression pedals. MIDI controllers can be easily bound to parameters of effects or synthesizers in music software such as Ableton Live. Bank switches allow the performer to activate preset pedal bindings while playing.

Pedal manufacturers have largely overlooked the innovation of the organ toe switch, which allows events to be triggered without moving the pedal or moving the foot to another pedal. Commercial toe switch equipped expression pedals have a latching switch mounted *underneath* the pedal, meaning that to trigger the switch the player must completely depress the pedal, losing the current pedal position. For this work we have resurrected and MIDI-enabled some pedals from discarded electronic organs. Figure 3 shows a similar type of pedal, equipped with dual toe-switches. Note that the switches are mounted at the sides of the pedal, not underneath it.

3. EYE TRACKING

Eye trackers work by processing images of the eye to compute the coordinates of a person's gaze. The eye is illuminated with infrared (IR) light in the range 780 to 880nm which is invisible to the human eye and therefore does not distract the user. The main task of the image processing is to accurately locate the pupil in the eye image. When the light source is close to the optical axis of the lens the pupils appear bright, as they do in "red-eye" flash photographs. When the light source is off-axis the pupil appears dark. The reflection of the light source from the cornea is called the "glint". The location of the pupil relative to the glint can be used to normalise the gaze vector with respect to the light source. Using two light sources creates a second glint which allows a tracker to adjust for a degree of head movement [2].

Two types of eye tracker are distinguished by the location of the eye camera. *Remote* trackers have the camera fixed in the environment, usually on a display or object of attention. *Head mounted* trackers are worn as part of glasses or a helmet. *Remote* trackers calibrate gaze relative to a fixed surface such as a computer screen. *Head mounted* trackers may also include a second "world" camera which captures the scene from the user's point of view, and can be calibrated to register gaze in the image coordinates of this camera.

A major application of commercial eye trackers is to understand how people attend to different visual environments, for marketing research or usability studies. These systems can range from US$10K to beyond $20K which puts them beyond the reach of most musicians. Several low cost devices have recently appeared, in the range of US$100 to $200 aimed at consumer applications such as gaming. The Tobii eyeX (`http://tobii.com`) and the Eye Tribe (`http://theeyetribe.com`) are remote trackers designed to be used in front of a computer display and can be characterised by their *operating distance* and the size of their *head box*, the volume within which calibration is accurate. Both trackers have a head box of 40x30cm at 65cm. Range is limited by camera resolution and lighting: 45–100cm (Tobii), and 45–75cm (Eye Tribe).

We evaluated the Eye Tribe tracker, which unlike the Tobii unit includes a cross-platform SDK. However we found the size of the head box too restrictive for musical performance, even when seated in front of the tracker. Fortunately, there are many open source software and hardware designs available for developing custom eye trackers.

We chose to implement a tracker based on Pupil [7, 8], an open-source platform for ego-centric eye tracking. Pupil is cross-platform (Linux, Mac, Windows) and is actively supported by Pupil Labs (`http://pupil-labs.com`) who maintain and develop the software and also sell pre-built hardware in several configurations. They also maintain a DIY reference design using commercial webcams for which 3D printed parts can be ordered.

The Pupil software processes up to three video streams (a world camera, and one or two eye cameras). The pupil is located in the eye cameras based on the center-surround feature proposed by Swirski [11]. This feature is particularly effective in highly off-axis eye images. With an error threshold of 5 pixels, Pupil's detection rate on Swirski's test data-set approaches 90%, compared to only 40% for ITU Gaze Tracker, and less than 20% for Starburst [7]. Performance on regular eye images is better since the viewing angle is more favourable.

4. EYE CONTROLLED MUSIC

Hornof [4] reviews previous works on eye-controlled musical performance. The first is "Intuitive Ocusonics" (1999) which uses Steim's "Big Eye" as simple analyser for eye video. This software allows sounds to be triggered when an object (the pupil) moves through predefined "hot zones" in the image. "Oculog" (2007) uses computer vision to track the position of the pupil, and the raw (x,y) coordinates are fed into a tone generator to control the note frequency (x) an velocity (y). The performer responds to objects in the environment, and the dynamics of the eye movement (eg. saccades, jitter) contribute to the sonic texture.

"EyeMusic 0.9b" (2004, [3]) is the first work to precisely register the user's gaze relative to a display and uses a commercial tracker. The performer views stimulus images on a display, and the software identifies fixations, and uses horizontal and vertical position to control the pitch of separate voices. As part of the performance the gaze cursor is shown on the original images. In a later work "EyeMusic 1.0" (2007, [5]) the display includes moving objects (bouncing balls) which trigger sounds when they interact with the user's gaze point.

"Eye Harp" (2011, [14]) aims to reproduce some aspects of traditional musical instruments, allowing the performer to control rhythm and melody. It includes a matrix-based step sequencer, and also allows tunes to be played by choosing notes and chords with eye fixations.

"Eye.Breathe.Music" (2010, [1]) is a musical instrument that uses gaze to control pitch, and breath pressure for dynamics. A headset includes camera, a pressure sensor and a 12x3 array of LEDs. Horizontal position of the LED corresponds to the note of a chromatic scale, and vertical position indicates an octave offset.

These previous works can be divided into three groups.

1. "Ocusonics" and "Oculog" use raw pupil position to control a generative musical process. There is no display, and the performer responds to visual stimuli in the environment.

2. In "EyeMusic" sound is generated by interactions between the performer's gaze and displayed stimuli (eg. bouncing balls).

3. "Eye Harp" and "Eye.Breathe.Music" include a musical control surface through which the performer can select particular notes, as they do in a traditional musical instrument.

Our system differs from these works in several ways.

Firstly, *Focal* it is not a musical instrument but instead augments traditional instruments. It is used in conventional performance and will only be successful if the improved expressiveness outweighs any constraints imposed by the system.

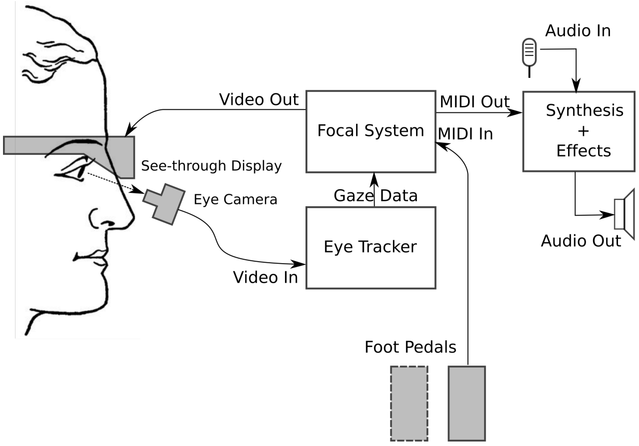

Figure 4: Overview of Focal system

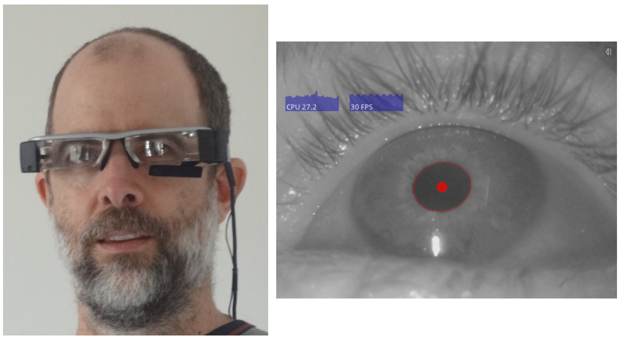

Figure 5: See-through head mounted display with eye tracking camera (left). Sample infra-red eye image (right).

Secondly, with the exception of "Eye.Breathe.Music" previous systems limit movement due to the head-box of a remote tracker, or the need to be correctly positioned in front of the display. We use a head-mounted eye-tracking display which is less restrictive.

Thirdly, although *Focal* generates MIDI expression data, this is not derived from gaze. Gaze defines only the *context of the expression*, or *which* parameter is controlled. The articulation comes from a separate expression controller such as a pedal. For the performer, both the instrument and the expression sensors are familiar. Glancing at an object to indicate intent is also a familiar non-verbal communication, making system comparatively easy to learn.

5. DESIGN AND CONSTRUCTION

Figure 4 shows an overview of our design. The user wears an eye-tracking head-mounted display which shows a graphical interface over their field of view, but does not obscure normal vision. The eye tracker computes gaze position in display coordinates. This is fed into the main *Focal* process which manages the user interface and MIDI event mapping.

The *display* is an EPSON Moverio BT-200AV see-through mobile viewer. These augmented reality "smart glasses" include two miniature 960x540 pixel LCDs, with optics that project the display over the user's field of vision. The image width is 23 degrees of arc, equivalent to an 80 inch display at 5m, or a 13 inch laptop screen at arms length. The glasses run Android 4.0.4 on a 1.2GHz OMAP processor,

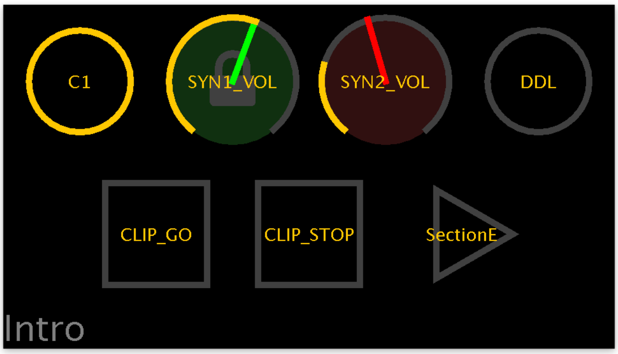

Figure 6: *Focal* interface elements for performance of *Elegy #2* (Travers, 2012). Section 6 describes the control types, and the corresponding signal processing is detailed in Section 7

and include sensors such as a GPS, accelerometer, gyroscope and camera. A separate adapter can wirelessly send HDMI video to the glasses. The accuracy of the tracker is about 0.5 degrees, so the tracking resolution is 1/46 of the display width, or about 20 pixels.

Figure 5 (left) shows the Moverio headset with an eye tracking camera mounted below the left eye. The camera is a Pupil Labs 30Hz IR camera which includes IR illumination for dark pupil tracking. This is attached to the Moverio using a custom 3D printed mount which gives three rotational degrees of freedom for aligning the camera to the eye. Figure 5 (right) shows a sample image from the camera, as displayed in the user interface of the Pupil tracker. The computed pupil location is shown as a red ellipse, and a dot indicates the centre position. Note that this image is flipped because the camera is upside-down when positioned for the left eye.

The *expression pedal* was constructed by adding a MIDI circuit to the volume pedal of a discarded electronic organ. This has a potentiometer attached to the main pedal, and a toe switch mounted on the left side of the pedal. The two sensors are mapped to separate MIDI continuous controllers. The controller also supports a second expression pedal with up to two toe switches, such as the Digisyzer pedal (Figure 3).

We use the open-source *Pupil Capture* software for tracking the eye camera. Our only modification is the addition of a new calibration plugin which registers gaze in screen coordinates. There is no world camera, so the stock calibration procedures cannot be used.

The *Focal* software implements the user interface and MIDI mapping. It consists of roughly 2K lines of Java code, and receives eye positions from Pupil Capture via the ZeroMQ protocol. Eye positions are analysed for fixations using the I-DT dispersion algorithm [10]. The software manages asynchronous events from the MIDI interface, eye tracker, and GUI. Active rendering is used to achieve a constant frame rate of 30 frames per second. The system outputs control information in the form of MIDI continuous controller (CC) and note on/off messages.

6. THE FOCAL USER INTERFACE

The *Focal* system produces no sound but is designed to be used with an external effects processor or music software such as Ableton Live. The user defines "virtual" MIDI controls which remap the MIDI pedal events when activated via gaze. Some have independent state (*continuous* and *toggle* controls) and some do not (*note* and *discrete* controls). In this section "control" means a "virtual" MIDI control; physical MIDI controllers will be named "toe switch" or "expression pedal." Figure 6 shows the main interface elements. Note that the display uses simple high-contrast elements because it will be projected over the user's field of view.

Continuous controls appear as a dial with the current value indicated by the "needle" position and a coloured arc. When the user looks at a dial, the needle changes colour and the dial enlarges slightly to show that it is active. When the dial value is *synchronised* with the pedal state, the needle becomes green and any change to the value is transmitted immediately as a MIDI CC message (eg. SYN1_VOL control in Figure 6). When the pedal is in a different state the dial is *unsynchronised* and the pedal value is shown by a red needle (eg. SYN2_VOL control). In this state the pedal value is displayed but not transmitted. The user adjusts the pedal position (needle) to match the dial value which is always indicated by the coloured arc. Once the values match, the dial becomes synchronised and begins transmitting its value. This "sync" gesture is quite quick and easy since the needle position shows which direction the pedal should be moved.

Toggle controls appear as a solid circle which is coloured when on, and grey when off. Tapping the toe switch toggles the state, which is transmitted as a MIDI CC message with value 0 (off) or 127 (on). Toggle controls may optionally be *grouped*, so that enabling one control disables all other controls in a group. This is similar to "radio buttons," except that it is also possible for all buttons to be off.

A *note control* appears as a solid square, and transmits a MIDI note on/off event when the toe switch is pressed/released. Note controls can be used to trigger sounds which have been mapped to keys of a sampler instrument. A *discrete control* is a continuous control with only two values (0 and 127). These are operated like notes using the toe switch, but instead send MIDI CC messages.

To simplify the display, controls may be grouped into *scenes*. This reduces clutter, making it easier for the player to find a control on the screen. Fewer controls also mean more space per control which is helpful since larger controls are less sensitive to random noise in the gaze position.

A *scene control* appears as a right-facing triangle and triggers a change to the named scene, functioning like a hyperlink in a web page. Scenes can be arranged in a graph to match the musical structure of a piece. The user can nominate a MIDI CC controller which transmits the scene index when a new scene is activated. This can be linked to the *scene selector* in Ableton Live to activate a Live scene, or to a *chain selector* in a Live rack to control instrument patches or effects for each scene.

A *keys control* pops up a scene containing 12 note controls in a chromatic scale. This can be used to play simple melodies. A *preset control* activates a previously saved system state, restoring the values of all controls. A *macro control* executes a programmed sequence of actions, which can include MIDI note and control events and timed delays.

In normal interaction, the toe switch triggers a single discrete event such as switching a control from an off-state to an on-state, and the expression pedal position drives the value of continuous controls. Pressing the toe switch on a continuous control triggers a special "lock" event. This causes the expression control to remain *locked* to that control independent of the gaze focus until a second lock event is triggered. If triggered on the same control, the pedal is *unlocked* and control again follows gaze in real-time. If triggered on a different control, the lock transfers immediately to the new control.

Locking is useful for maintaining expression while working with other digital controls (toggles, notes, keys). Locking also allows the simultaneous operation of two pedals. The system maintains separate locks for each pedal based on which toe switch was pressed. When activated via locking or gaze focus, background colour indicates which pedal is bound to the control: green for right and red for left. In this way, the user can simultaneously control up to three devices: two locked continuous controls and one digital control. For example, the user could play a simple bass line whilst modulating two synthesis parameters.

7. EVALUATION

Our system is under evaluation by professional accordionist and composer Cathie Travers (`http://cathietravers.com`). For her piece *Elegy #2 (2012, for acoustic accordion and electronics)* we compared the set up and performance with and without *Focal*.

The piece runs for 6 minutes and uses the following digital effects hosted in Ableton Live: delay, synthesizer 1, synthesizer 2, clip launcher. Pedals used in the original setup include a Roland FC-300 MIDI foot controller (see Figure 2), two passive expression pedals (controlled by the FC-300), and a Keith McMillen 12-Step Controller. These are arranged around a saddle chair, which allows the performer to swivel to reach each device. Functions mapped to the pedals include: start and stop a low C drone on both synthesizers (C1), separately control the volume of each synthesizer (SYN1_VOL, SYN2_VOL), start and stop a pre-recorded audio clip (CLIP_GO, CLIP_STOP), enable and disable a digital delay (DDL), play a bass line from a C minor scale, using 8 of the 12 notes from one octave. There are a total of 12 digital switches and 2 continuous controls. The piece requires retriggering the clip while it is still playing, so includes separate start and stop buttons, rather than a single toggle. The *Focal* controls were grouped into two scenes, with the scale notes in the second scene and all other controls in the first; the delay control appears in both scenes. The first scene is shown in Figure 6. Note that in a live concert there is a third continuous control pedal for the master volume output from Ableton Live, which must be balanced against the amplified live accordion volume level.

A video of the original performance is available online [13]. At the time of writing we are preparing a comparison video. This and other demos and updates on the project are available online: `http://stewartgreenhill.com/articles/focal/`.

During this evaluation we discovered several missing features which have since been added. First is the "latching note" which is a note that works like a *toggle*. In the example, note C1 starts when pressed, and stops when pressed again, rather than on release of the first press. Second is the "mono note group", a group of latching notes where the activation of a note cancels the activation of other notes. Optionally, in a "legato" note group, the note-off event of the previous note is deferred until after the note-on event of the next note.

We note a degree of mental effort associated with the eye tracking, though it is unclear yet whether this is more or less than a regular pedal system. It is still necessary to learn through practice the correct sequence of events, but to focus gaze we must also consciously resist the natural tendency to glance at distracting stimuli. When focus is tied to gaze we must not accidentally lose control mid-gesture, and have added rules to prevent a control that is currently changing from losing focus. While we allow explicit focus control for continuous controls via locking, we currently don't have this option for digital controls, although it could be added using a second toe-switch. In practice the exact behaviour may depend on user preference, as it does for window focus in some operating systems (click-to-focus versus focus-follows-mouse). Currently, display latency and calibration drift are confounding factors (see 7.1) which distract from the natural intent of the interface, and the situation will be clearer once these are resolved.

As discussed in Section 1, we used organ pedals because commercial pedals do not include toe switches. Compared to expression pedals such as the FC-300, organ pedals have a greater range of travel (around 30 degrees) and are "looser" meaning that they may sometimes move slightly when operating the toe switch. A "stiffer" mechanism is preferred, so we are investigating two options: tightening the mechanism on the organ pedals, or fitting toe switches to commercial expression pedals, either by adding a microswitch arrangement, or by using a pressure sensor.

One advantage of the *Focal* system is the reduced floor real-estate devoted to pedals. Currently, a dedicated laptop is used for the eye tracking and user interface, which partially counteracts the reduction in gear due to the replaced pedals. Efforts are underway to make the eye tracker run on the glasses, which would result in a more portable solution. Another advantage is a reduced time and motion required to switch between different pedal functions. This should be more significant for more complex pieces, but in practice it may depend on how well a piece can be structured into simple scenes. This is a topic for future study.

7.1 Technical Issues

This section outlines issues for future development.

Currently the main limitation of our system is the robustness of the eye tracking. Calibration is generally good, but small changes in the camera position can occur with head movement, which reduces accuracy over time. Typically, the system must be recalibrated after about 10 minutes, which limits its use in musical performance. Nevertheless, freedom of movement is better than is offered by remote trackers. Some camera movement is due to the design of the Moverio headset, which has soft adjustable nose pads that occasionally flex with head motion. This design may be partly to allow the Moverio to be worn over prescription spectacles, and when used in this way stability can be improved by fixing the Moverio to the spectacles frame. Replacing the nose pads with conventional spectacles pads would also improve stability. Indeed the upcoming BT-300 Moverio offers a similar "improved nose design".

We can devise ways of stabilising the headset, but a current limitation is the assumption that the camera is fixed to the head. This is partly addressed with the 0.7.4 release of the Pupil software (March 2016), which includes a new 3D eye model based on the work of Swirski [12]. The 3D model uses pupil motion to infer the location of the eyeball with respect to the camera. This information can be used to compensate for camera movement over time. In addition, the 3D model is more robust to occlusions of the pupil in the image due to reflections, eyelashes, or eyelids. Future releases of the Pupil platform are expected to officially support 3D calibration and gaze mapping for head mounted displays, which should give dramatically better stability than is presently achieved by the system.

HDMI video from the host computer is sent to the Moverio display from a special wireless adapter, leading to two issues. Firstly, we found the connection process to be very erratic, though once a connection is established it seems robust. However, wireless is subject to interference which may be outside control in a new performance environment, so a cable connection is preferred. Secondly, there is a display lag of up to 250ms in the video streaming which makes the interface seem sluggish. A solution would be to implement the *Focal* interface on the Moverio's Android processor, which would also allow the use of the built-in sensors (accelerometer, gyro) for expression control. Also running the the eye tracking on the Moverio would result in a self-contained, robust and very lightweight solution. A USB MIDI cable connection to the foot pedals (or other expression controllers) would allow the system to avoid using wireless.

At the time of writing there are no eye-tracking head-mounted displays on the market, and EPSON's Moverio is one of the few optical see-through displays available. Google Glass was withdrawn from production in January 2015, but Google was awarded a patent in April 2015 for a heads up display with an eye tracking camera [9] which suggests a possible future product. Similar displays have been demonstrated using free-form prisms [6]. Eye trackers are being incorporated in virtual reality (VR) headsets for gaming (eg. Fove, and Occulus Rift), but the application for augmented reality (AR) is less clear. Eye tracking VR headsets could be much more robust since they are enclosed units and not susceptible to external lighting. However, musicians are unlikely to accept the isolation of a VR headset where there is no visual communication with the outside world.

We have not yet examined the response of the system under stage lighting. The eye camera uses infrared (IR) light to image the eye and may be affected by bright or uneven IR lighting. IR is produced by tungsten halogens which are often used for spot lighting. However, LED stage lighting is becoming common and contains almost no infrared. Very bright light (eg. sunlight) makes it difficult to see the display, and for this the Moverio includes optional shades which could also be used to block ambient IR.

8. CONCLUSIONS

We presented the design of *Focal*, a novel musical expression controller which incorporates eye tracking. The major innovation is to use gaze to control to routing of expression control data between virtual control devices which are rendered in a see-through head mounted display. This allows the control of many parameters otherwise only achievable with multiple foot pedalboards. Although our system uses pedals for accordion performers, the principle can be applied to both conventional control sources (eg. breath pressure) and novel control sources which might be incorporated in hyper-instruments. The system could also be useful for musicians with movement disabilities. Audio input could be replaced with a synthesis process, and a puff/sip pressure sensor could be used for expression control.

We built several parts of the system due to the lack of off-the-shelf solutions. This includes the see-through eye-tracking head mounted display, and the toe-switch enabled expression pedals. Although necessary for the design these are not the focus of our work. We hope that in the future robust commercial display systems will become available. Big names like Google have ventured into this space, but a clear market is yet to emerge.

Our initial evaluation with accordions suggests that the system could provide a viable alternative to conventional controllers. Some advantages and disadvantages have been identified, and we expect that more will be discovered with future investigation.

9. REFERENCES

[1] S. Bailey, A. Scott, H. Wright, I. Symonds, and K. Ng. Eye. breathe. music: creating music through minimal movement. In *Proceedings of the 2010 international conference on Electronic Visualisation and the Arts*, pages 254–258. British Computer Society, 2010.

[2] D. W. Hansen and Q. Ji. In the eye of the beholder: A survey of models for eyes and gaze. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 32(3):478–500, 2010.

[3] A. Hornof and L. Sato. Eyemusic: making music with the eyes. In *Proceedings of the 2004 conference on New Interfaces for Musical Expression (NIME)*, pages 185–188. ACM, 2004.

[4] A. J. Hornof. The prospects for eye-controlled musical performance. In *Proceedings of the 2014 conference on New Interfaces for Musical Expression (NIME)*, pages 461–466. ACM, 2014.

[5] A. J. Hornof, T. Rogers, and T. Halverson. Eyemusic: Performing live music and multimedia compositions with eye movements. In *Proceedings of the 2007 conference on New Interfaces for Musical Expression (NIME)*, pages 299–300. ACM, 2007.

[6] H. Hua, X. Hu, and C. Gao. A high-resolution optical see-through head-mounted display with eyetracking capability. *Optics express*, 21(25):30993–30998, 2013.

[7] M. Kassner, W. Patera, and A. Bulling. Pupil: An open source platform for pervasive eye tracking and mobile gaze-based interaction. `http://arxiv.org/abs/1405.0006`, April 2014.

[8] M. Kassner, W. Patera, and A. Bulling. Pupil: An open source platform for pervasive eye tracking and mobile gaze-based interaction. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*, UbiComp '14 Adjunct, pages 1151–1160, New York, NY, USA, 2014. ACM.

[9] H. S. Raffle and C.-J. Wang. Heads up display. US Patent(US 9001030 B2), April 2015.

[10] D. D. Salvucci and J. H. Goldberg. Identifying fixations and saccades in eye-tracking protocols. In *Proceedings of the 2000 symposium on Eye tracking research & applications*, pages 71–78. ACM, 2000.

[11] L. Świrski, A. Bulling, and N. Dodgson. Robust real-time pupil tracking in highly off-axis images. In *Proceedings of the Symposium on Eye Tracking Research and Applications*, pages 173–176. ACM, 2012.

[12] L. Swirski and N. Dodgson. A fully-automatic, temporal approach to single camera, glint-free 3d eye model fitting. In *Proceedings of the 2013 International Workshop on Pervasive Eye Tracking and Mobile Eye-Based Interaction (PETMEI)*, 2013.

[13] C. Travers. ELEGY #2 - Cathie Travers / accordion, Ableton Live, FC300, 12-Step. `https://www.youtube.com/watch?v=9UvIq-fvxYI`, 2012.

[14] Z. Vamvakousis and R. Ramirez. The eyeharp: An eye-tracking-based musical instrument. In *8th Sound and Music Computing Conference*, 2011.