

ADVANCED MEDIA CONTROL THROUGH DRAWING: USING A GRAPHICS TABLET TO CONTROL COMPLEX AUDIO AND VIDEO DATA IN A LIVE CONTEXT

Steve Gibson & Justin Love

This paper demonstrates the results of the authors' Wacom tablet MIDI user interface. This application enables users' drawing actions on a graphics tablet to control audio and video parameters in real-time. The programming affords five degrees (x, y, pressure, x tilt, y tilt) of concurrent control for use in any audio or video software capable of receiving and processing MIDI data.

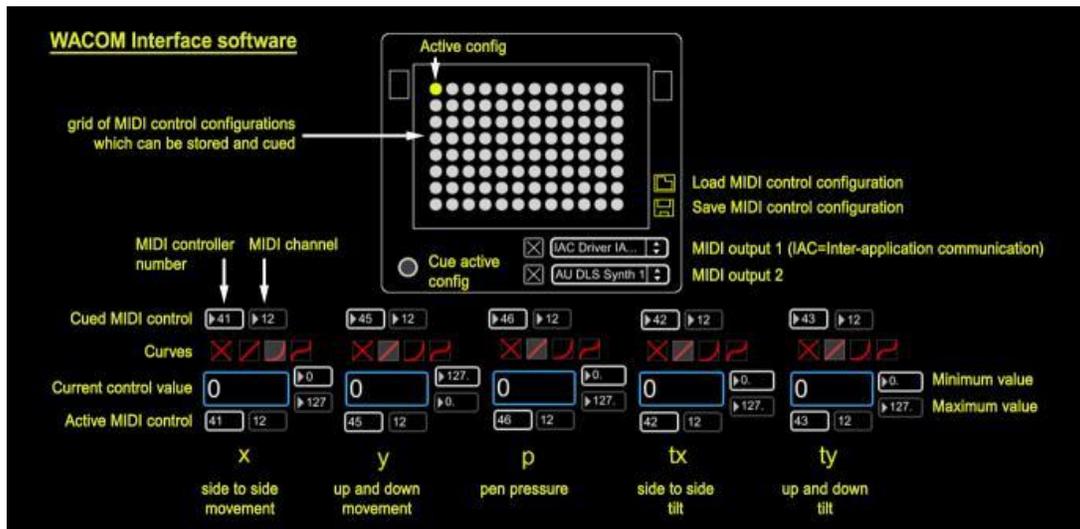


Fig. 1 - Wacom Interface with explanation, Justin Love and Steve Gibson, 2008-09.

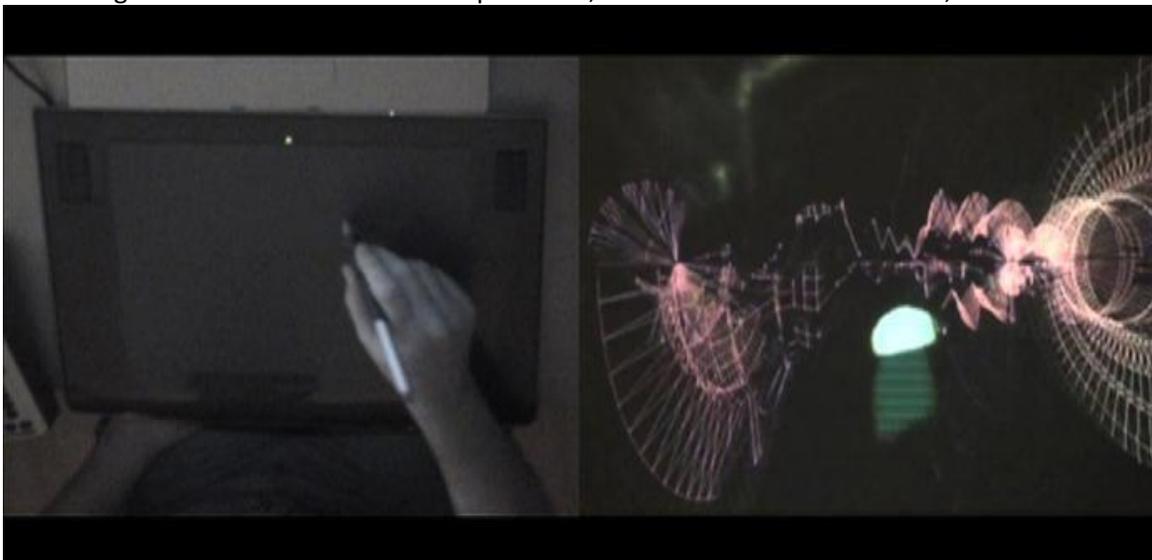


Fig. 2 – Wacom Tablet MIDI Demo. Live performance by Steve Gibson, January 2010. Video by Steve Gibson, 2010.

Introduction

The paper considers the connection between drawing technique and media control both generally and specifically, postulating that dynamic drawing in a live context creates a performance mode not dissimilar to performing on a musical instrument or conducting with a baton. The use of a dynamic and physical real-time media interface re-inserts body actions into live media performance in a compelling manner. Performers can learn to “draw/play” the graphics tablet as a musical and visual “instrument”, creating a new and uniquely idiomatic form of electronic drawing.

Electronic drawing is generally (though not exclusively) limited to pictorial or representational drawing with a pen and graphics tablet in order to produce still images or frames for animations in software such as Adobe Flash. Given the authors’ background in media art, physical computing and transmedia applications, we were more interested in the idea of repurposing the graphics tablet as media control device. We were particularly interested in the act of electronic drawing as a means of manipulating sound and video rather than as an output to an actual drawing, whether on-screen or for print.

Other 3D “tablet-like” devices have been used for manipulating sound, including the popular Kaoss Pad from Korg. [1] These devices are generally very limited in size: the Kaoss Mini KP has the following dimensions: 4.17 in (W) x 5.08 in (D). In the context of live performance this size does not allow for dramatic gesture.

The notion of the dramatic gesture was very important for us since we wanted to use a tablet as the only visible performance element. It was imperative that the audience be able to identify performer gesture with perceptible results in the media used (live sound and video in our case). To achieve this aim a larger area for drawing would be required. Finally, another limitation of other 3D controllers such as the KAOSS pad is that they are ‘only’ 3D. In order to achieve complex control of the media elements through our drawing actions, we wanted to be able to use more than three parameters at a time.

For the above reasons we decided to use a graphics tablet as a control device for gestural performance, and in particular we selected the Intuos 3 from Wacom (now superseded by the Intuos 4. [2] The Intuos 3 is 9 inches x 12 inches in dimension and has five degrees of control. This large surface combined with five possible control parameters made it an ideal choice. Also the fact that one uses a pen with a graphics tablet makes it more useful as a performance instrument, since the pen is a visible object that audience members can identify in much the same way as they would a conductor’s baton (though the audio-visual results we aspired to were quite different from those of an orchestra).

Background: Live music performance as a visual medium

Our interest in repurposing a device such as a graphics tablet for media control stems from the authors’ dissatisfaction with certain modes of electronic music performance. Live music has traditionally been experienced as a partially visual medium, with the visible actions of the performer holding the attention of the listener: “Making music involves not only the communication of musical sounds but is also characterized by a continuously changing and meaningful use of facial expressions, body movements, and hand gestures. Until the late nineteenth century, music performances were almost always experienced as audio-visually integrated activities.” [3]

With the advent of electro-acoustic “tape-based” music in academic electronic music and laptop performance in popular electronic music, the visual reference of the performer became “un-stuck” from the resulting sound produced in a live performance scenario. The relatively long history of tape-based electro-acoustic music has attempted to deal with the lack of performance spectacle by ever-increasing complex “diffusions” of the sound in space. This often involves ever-increasing numbers of speakers placed strategically throughout the room. For example, the BEAST “diffusion” system at Birmingham University has “over 100 discrete channels” of sound available. [4]

When used effectively this can create a sort of architectural soundscape, in which the listener follows the sound around the speakers and thereby intuits an image of sound as an object in space, albeit virtually. For the most part though, it is hard to avoid the conclusion that the use of ever-increasing numbers of speakers is a rather desperate ploy to obfuscate the fact that the audience does not easily respond to music in the absence of the visual spectacle of the performer. In short, in the majority of cases, the tape-music concert is one that generally does not satisfy the audience need for performance complexity. The genuinely live performance event has possibilities for variation, expression, errors, and communication with the listeners/viewers that are lacking in purely tape-based performances.

Similarly, pure laptop-based popular electronic music faces a similar crisis. When the performer is hidden behind the screen, it is simply impossible to recognize or even infer what he or she might be doing. As with tape-based music, the visual aspect in laptop performance is reduced to a virtually-inferred spectacle at best. (It should be said that in laptop performance there is at least the presence of an actual performer, and though it may be difficult to ascertain his or her actions, at least he or she is generally doing ‘something’ live).

The introduction of physically expressive performance aspects into electronic music has increased in the past ten to fifteen years. With improvements in computer speeds and the growing number of gesturally expressive media control devices available, electronic music has become increasingly engaging within a genuine performance context. The laptop is still employed live in most cases, but at least it is being controlled by a performer with some other device that the audience can relate to as an instrument. In this regard the live electronic performance medium (with a physically present and “active” performer) bears some resemblance to the drawing medium, in that deliberate and unconscious gestures in time form a basis for the artist input in both mediums. In both drawing and live performance errors are allowed (or are, at the least, quite inevitable). Meandering, testing, chance, and happy accidents in improvisational or semi-improvisational mediums such as live electronic music or DJing, are quite relatable to doodling, dreaming and sketching in the field of drawing.

Background: The VJ as live performer

It is worth noting that the VJ faces a similar problem to that of the laptop performer: how do his or her actions relate to the music and visual performance and how does the audience perceive those actions as in any way corresponding to the audio-visual results? VJing generally relies heavily on digital signal processing (DSP), in which the VJ software detects aspects of the incoming audio (e.g. hard attack beats coming from a kick drum) and consequently applies effects, cross-fades, or clip edits based on that information. This normally creates an obvious connection between the audio performance and the resulting visual world, at least on the rhythmic level.

While these DSP tools provide a solid automatic connection between the audio and video in a VJ performance, there remains the presence of the VJ performer: as with laptop music performance, in the absence of any noticeable interface beside the laptop, the performer's actions are somewhat mystifying to the audience.

For the above reasons we decided that the graphics tablet served as a potent tool for uniting the DJ (or electronic musician) and the VJ under a single interface. In addition by using the large surface of the tablet as an interaction device we created a dynamic performance instrument that the audience could relate to in a very physical manner.

Initial research

Our initial research into programming options for using MIDI with a Wacom tablet uncovered some pre-made applications, but in general we were unsatisfied with both the usability and "look and feel" of these solutions. For example, one such application that we considered was the "Tablet 2 MIDI" interface developed by Livelab in Denmark. [5] This program allows complex mapping of MIDI data to multiple parts of the graphics tablet. The essential problem with this model is that it uses a complex menu-driven user-interface that is not intuitive. In addition the application is Windows-only, and given the fact that the authors (and much of the digital media community) are Mac users, we determined that this "off-the-shelf" solution was not ultimately viable. If we wanted a solution that would meet our needs and the needs of a more general digital media community, it would be necessary to develop our own alternative application.

For this reason we decided to look at creating our own interface using Max from Cycling 74. [6] Max has an application programming interface (API) that allows users to create their own external objects. The extensibility and modularity of Max has resulted in a large community of developers that create and share their custom Max external objects. To implement our interface we incorporated a Max external object designed by Olaf Matthes that outputs data produced by the Wacom Tablet. [7] In the Max patch, the data from the Wacom tablet is routed and mapped to a series of user-definable MIDI messages that can then be used to produce audio and video manipulations and transformations.

Interface design

It was essential that the interface that we created be entirely contained in one window, with no need for submenus or even normal file menus. Fig. 1 illustrates our basic interface design. Side-to-side (x), up-and-down (y), pressure (p), tilt x (tx), and tilt y (ty) can each be mapped to any of 127 possible MIDI control change parameters on 16 different MIDI channels. In addition, control curves can be applied to each parameter. For example, a linear control curve produces a consistent rate of change for a given parameter, whereas an exponential control curve causes a parameter to change slowly at first and then increase rapidly towards the end. The MIDI output can be routed to two different MIDI ports simultaneously. The grid at the top allows the user to save up to 96 different MIDI control change setups in one file.

Connecting with audio and video

To connect our application with other MIDI software, we used Interapplication MIDI. This allowed us to instantly map tablet actions to an effect or parameter in the audio or video software. Naturally only five

tablet parameters can be used to control video or audio data at the same time, but it is possible to route the same MIDI control data to multiple functions in the video or audio applications. With the grid of 96 storable configurations, once the user gets tired of the same configuration a new one can be loaded instantaneously. It is possible to cue a configuration (similar to how a DJ would cue a new track) and press one button to immediately activate the new setup.

The applications we use with our tablet audio-visual interaction are Ableton Live (for audio) and Modul8 (for video). Both of these programs are ideal for live audio-visual performance as they have extensive MIDI support and are built for non-linear performance situations in which the user may want to change, modify or apply effects to audio-visual materials instantaneously.

The tablet as a live performance device

With the possibility of seemingly endless mappings of tablet data to corresponding audio and video effects we were acutely aware that some sort of logical mapping structure would have to be established in order to allow the interface to be used effectively in a performance situation. This harkens back to the concern that we had at the outset: we wanted to be sure that tablet actions had observable results so that audience members would be able to clearly identify the relationship between the live drawing and the performed audio-visual elements.

In the past Steve Gibson has worked on several projects which explore 3D interfaces and the connections between audio and video mappings and 3D control. The most important of these projects was *Virtual DJ*, in which the user is given control over an audio environment and lighting effects by moving in 3D space. [8] In short this project used standard control configurations between user actions and results in the audio-visual system. For example, raising the hands would generally play an upwards melody and with each new note a light would change colour. This allowed the user to identify that their motions were having observable and repeatable effects in both the audio and visual realms. In essence this simulated the effects of synaesthesia, a condition in which persons can often see colours in response to particular sounds, tones or musical notes: “Synesthesia is an involuntary joining in which the real information of one sense is accompanied by a perception in another sense.” [9]

For our tablet interactions we thought of the tablet interface as a miniaturised version of the multi-dimensional spatial interface used in *Virtual DJ*. Thus logical interactions could be inferred by testing user actions with system results. For instance, drawing on the y-plane on the tablet (up-down) could logically map to audio volume or low-pass filter (which would have a similar effect to volume, without completely removing the sound at the bottom of the tablet) and image opacity. Therefore we built a limited series of controls to be employed by the user on the tablet and held to these controls throughout our performance. The videos at the following two links show how three of the tablet parameters (x, y and p) were mapped to audio effects:

http://www.telebody.ws/TRACEY/Tablet_demo1.mov

http://www.telebody.ws/TRACEY/Tablet_demo2.mov

Conclusions

In summation we have found that the Wacom graphics tablet is a powerful device for controlling live audio and video in a performance situation. The act of live drawing, though here removed from its traditional reference to a produced “drawing” (either on-screen or in print), is one that allows for dramatic gesture in a way that pressing keys on a computer keyboard or moving a mouse could never hope to achieve. In addition the fact that the graphics tablet can unite five degrees of control over live audio and video makes it an ideal tool to consolidate the roles of the DJ and the VJ under one control interface. Finally we present a live example of the tablet used in a performance situation with one of the authors controlling live audio and video with his drawing actions (Fig. 2): http://www.telebody.ws/TRACEY/Tablet_Demo_Split.mov.

The Wacom MIDI software can be downloaded at: <http://www.roguescience.org/files/wacom2MIDI.zip>.

Learn how to setup the Wacom MIDI software to use with your Wacom tablet: http://www.telebody.ws/TRACEY/Wacom_MIDI_Setup_Demo.mov.

Note

A longer version of this paper was originally published by TRACEY as part of their Drawing and Technology issue: <http://www.lboro.ac.uk/departments/sota/tracey/dat/gibson-love.html>

TRACEY'S Copyright Notice

In submitting material to Tracey, contributors thereby grant permission for it to be published only on the Tracey website. Contributors retain copyright of their material and may, of course, use it elsewhere after publication in Tracey.

References and Notes:

1. Korg Kaoss Pad official Web Site, <http://www.korg.com/product.aspx?&pd=278> (accessed April 18, 2011).
2. Wacom Intuos 4 official Web Site, <http://www.wacom.com/intuos/> (accessed April 18, 2011).
3. P. Graham, F.A. Russo and W.F. Thompson, “Seeing music performance: Visual influences on perception and experience,” *Semiotica* 156–1/4 (2005): 177–201.
4. Birmingham ElectroAcoustic Sound Theatre official Website, <http://www.beast.bham.ac.uk> (accessed April 19, 2011).
5. Live Lab Tablet 2 MIDI official Web Site, <http://www.livelab.dk/tablet2midi.php> (accessed April 19, 2011).
6. Cycling 74 official Web Site, <http://cycling74.com/> (accessed April 19, 2011).
7. nullmedium official Web Site, “Olaf Matthes’ external objects for Max/MSP,” <http://www.akustische-kunst.org/maxmsp/> (accessed April 19, 2011). Note: no longer functioning.
8. Steve Gibson’s official Web Site, “Virtual DJ,” <http://www.telebody.ws/VirtualDJ/> (accessed April 19, 2011).
9. R. Cytowic, *Synesthesia: A Union of the Senses* (New York: Springer-Verlag, 1989), 1.