Musical Drawing in 2D and 3D: Dimensions and Perspectives

Esther Gruy, Florent Berthaut

Musical Drawing in 2D and 3D: Dimensions and Perspectives

Esther Gruy
University of Lille
Lille, France
esther.gruy@univ-lille.fr

Florent Berthaut
University of Lille
Lille, France
florent.berthaut@univ-lille.fr

ABSTRACT

With advances in immersive displays and gesture tracking technologies, many novel interfaces for musical and visual expression have been developed, which often explore combinations of audio and visual productions. One category of such interfaces is musical drawing, in which artists simultaneously produce both visual and sonic content. Since it deals with two different sensory channels, one of the main problematic is to find the right balance between the sonic and the visual aspects of a project. In this article, we regroup the existing 2D and 3D musical drawing projects and we propose a set of dimensions that can be used to describe them, namely: Expression Orientation, Required Expertise, Collaboration, Visual Replay Value, Audio Replay Value, Modifications, Mapping Flexibility, Mapping Structure and Degree of Immersion. Using these dimensions, we analyse the design choices, discuss technological and technical constraints, and establish perspectives for future work.

CCS CONCEPTS

• **Applied computing** → **Sound and music computing**; *Performing arts*; • **Human-centered computing** → Interaction design theory, concepts and paradigms; *Mixed / augmented reality*.

KEYWORDS

Music, Sonification, 2D drawing, 3D drawing, Audio-visual interfaces, Human computer interaction

ACM Reference Format:
Esther Gruy and Florent Berthaut. 2024. Musical Drawing in 2D and 3D: Dimensions and Perspectives. In *Audio Mostly 2024 - Explorations in Sonic Cultures (AM '24), September 18–20, 2024, Milan, Italy*. ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3678299.3678317

1 INTRODUCTION

Audio-visual interfaces [12] have explored numerous ways of blending sound and image together, from projects such as the reacTable [18], a multi-user tabletop interface that mixes musical composition and visual programming, to the UPIC system, a musical composition tool using drawings to generate an audio output after being numerically processed. The latter can be considered as part of a subtype of audio-visual interfaces called musical drawing, that is described as the sonification of gestures, represented as strokes, inside a two or three-dimensional virtual environment. Musical drawing allows for diverse mappings between the sound, the visuals, and the gestures performed by a user, for example using the curvature, the colour or the speed of a stroke to affect its sonic properties such as its frequency, its tempo or its amplitude, to name a few. This choice of audiovisual mappings, which has long been an issue in the musical expression community [5], is therefore essential in musical drawing. Since sound is a major component of musical drawing, it makes it especially adapted to performances in front of an audience.

A specificity of musical drawing from classical drawing, or from a more conventional digital musical instrument (DMI), is that it can be solely focused on one medium or the other, meaning that the drawings can remain abstract, or that the audio output can lack a sense of cohesion. The choice of where to position a project between sound and visual is left to its designers and what they want to achieve with it. Some might use the drawings to benefit the sound (and vice versa), while others try to fuse the two together to create a complete audio-visual experience [12] that evolves over time and space. However, since sound and image work on different temporalities, with one being immediate and the other persistent, the optimal way to unite the two is unclear (and probably non-existent), and this is what makes musical drawing an interesting research field.

Another defining characteristic of musical drawing is its use of two sensory modalities together. This implies that, unlike in applications for visual artists or in acoustic instruments, the complexity has to be modulated, to stay below a certain threshold, as one cannot expect to handle a high complexity on both modalities at the same time. This can have a negative effect on the cognitive load of a user/performer, on top of being a challenge to use and, for an audience, to understand.

Previous research has, however, shown that multisensory input, during an interaction, can have positive effects on task performance [11], memory [13] and can increase a user's feeling of presence [11, 13]. Presence is defined as the feeling of *being there* inside a virtual environment. Since it deals with a subjective feeling associated with a positive outcome, it can be an important component when it comes to the design of expressive devices. It is, for instance, a dimension described in the Creativity Support Index (CSI) [10], a measure to assess how well a tool supports creativity, with it being referred as the sense of being fully immersed in a task.

Musical drawing, whether in 2D or 3D, remains a research subject that we believe has been not been studied thoroughly enough, although it deals with topics such as audio-visual interfaces, digital instruments, composition tools and performing arts. On a broader scale, it can be linked to research fields such as gestural interfaces or the analysis of audio-visual cross-modal perception.

AM '24, September 18–20, 2024, Milan, Italy

ACM ISBN 979-8-4007-0968-5/24/09
https://doi.org/10.1145/3678299.3678317

This review aims to provide a list of the 2D and 3D musical drawing projects that have been crafted over the years, while establishing several dimensions to classify them into. We believe that this dimension space makes it possible to analyse what has been investigated the most, why specific choices were made, and which areas could be interesting subjects for future research.

1.1 Why musical drawing?

One problematic which accompanied the arrival of digital instruments is their lack of expressiveness over musical controls. Physical (or acoustic) instruments necessitate precise movements and rigorous training to be played, and every variation has an impact on the produced sound. As for digital instruments, it is not necessarily the case. For example, with laptop performances, there is very little space for gestural expressiveness, and the reduced interactions with the instrument can make it harder for the audience to understand the gestures of the performer/musician [9, 30].

Drawing (or sketching) on the other hand relies on a set of gestures which are easily understandable by an individual. In the case of digital instruments, drawing can therefore be used as a way to increase the transparency of the instrument by a metaphor of a physical interaction [14]. Sketching gestures are also often viewed as providing expressiveness, including a tolerance to mistakes and providing a more organic feel to performances [37].

Because of the familiarity of the performed gestures, drawing is also a way for novices to experiment with music, providing a "low entry fee" [35] for such instruments [19, 28]. This is amplified by the fact that musical drawing creates *immediate feedback loop where drawing and playing sound is inseparable and intertwined* [31].

Furthermore, drawing in itself can be used in the production and the edition of sounds. In wavetable synthesis, for example, it can be a way to freehand soundwaves [29] and envelop curves [28].

Another point of interest for the combination of sound and visuals lies in the profusion of artistic work that explore sonification of visual content, e.g. turning images, drawing or animations into sound, to see the unusual sonic result that would come out of it. In the research field, one example is Daisyphone [8], a collaborative musical instrument with a circular structure, where several notes can be put inside a row. Each row is being played successively. This creates a sort of canvas for musical loops where users can freely experiment. The goal of Bryan-Kinns was to observe the behaviour of users when composing music in a remote collaborative environment. Drawing, either geometrical shapes or more elaborate forms, was a consistent behavioural pattern found within the participants. Some participants even completely discarded the musical aspect of the instrument and treated it as a drawing device. This shows the intuitive nature of drawing, and more globally, of visuals.

2 PRE-DIGITAL 2D MUSICAL DRAWING

Before the current computer systems and hardware, which allow for real-time sound processing and interaction in either 2D or 3D spaces, musical drawing relied on more physical interactions with a system. A few examples can be found in history.

One of them is Oramics [25], a system first designed in 1957 by Daphne Oram. It uses photo-electric sound synthesis, where users have to draw onto a set of ten synchronised 35mm clear film strips. The drawings and shapes create a mask for the light to pass through, and photo-electric cells, in turn, generate an electrical charge, which affects the audio output. It can control several sound parameters, such as frequency, timbre, duration, and so on.

As mentioned in the introduction, another example is the UPIC system [26], created by Iannis Xenakis, and dating back to 1977. The system is composed of a large digitising tablet and of a computer with a vector display. For it to work, a user draws waveforms, that are rendered and stored by the computer. Then, they can compose sounds by simply drawing onto the tablet, with the x-axis affecting the time and the y-axis the pitch. The drawings can be subjected to several algorithmic transformations to alter the sonic output, and compositions can be stretched in time from a few seconds to several hours. Early versions of the system did not allow for real-time sonic output, as data needed to be processed before producing sound, but that was later developed.

The ANS Synthesizer [20], created by Yevgeny Murzin in 1958, is similar to both UPIC and Oramics in the way it works. A user can draw onto a glass surface covered by non-drying black mastic, with the x-axis representing time and the y-axis representing pitch. The surface is then moved to a device, where a ray of light passes through the scraped-off part of the drawings and onto photocells, which controls amplifiers and bandpass filters. This instrument is polyphonic and can generate 720 microtones over 10 octaves.

3 2D DRAWING AS A MUSICAL TOOL

With the emergence of hardware and software tools, which can generate visual and sonic output in real-time, musical drawing moved to the digital space, starting with 2D graphical interfaces.

Among one of the first documented work using drawing to interact with an audio-visual interface, are two projects created during Levin's thesis [22]: *Yellowtail* and *Loom*.

In *Yellowtail*, drawings are animated into a loop as a way to make the resulting sound evolve over time. The system works with pattern-playback, a technique that uses the inverse spectrogram to synthesise sound from an image, to produce the audio output.

Loom dives into a similar approach. Sound is produced while a stroke is drawn: this way, the audio output depends on the stroke's current endpoint properties, like its thickness or its curvature. Once a stroke is completed, it progressively fades, then redraws itself with the exact same motion, as to create an audio-visual loop.

Other projects followed this idea of using drawing to produce sound. Different Strokes [37] is an interface for musical control which uses drawing to recreate a visual connection between the actions of a user and the audio output. This project is made for laptop performances, where everything is very automated, with the aim of letting the performer be creative on a live setting, just like they would be with acoustic instruments. This project works on a computer, either with a tablet or a mouse. The strokes colour, speed and shape influence the resulting sound, by linking them to pre-defined wavetables and using particles moving along the strokes to drive the sound.

Illusio [1] is a digital instrument which works using a guitar pedal and associates strokes to sonic loops, with the aim of letting users create their own control interface, for expressiveness and

musical exploration purposes. This project works on the principle of a "musical tree", where loops are ordered hierarchically, and modifications on a parent loop affects the children. It is made to be used with other instruments, to record, organise and modify the music on the sole basis of the drawings.

Different Strokes and Illusio both target a musically experienced audience, unlike Articulated Paint [19], which is a musical interface strictly designed for novices. The objective is to let users experiment musical expression easily, without having to spend years learning an instrument. The project works with a tablet acting as a canvas and a paintbrush to drive the interactions. There are two modes available, with one being focused on expressive gestures, and the other using a more conventional approach with a notational system. For both of them, the movements of the brush and the pressure applied on the canvas influence the audio output.

The project providing the most visually diverse results might be Sounding Brush [31] and its seven different brushes available. Sounding Brush is an instrument that works on a tablet and puts the visual elements of the system on the same level as the soundscape. The brushes offer a large panel of visual to audio, audio to visual and gestural to audio and visual mappings, which permits users to construct elaborate sounds while focusing on the drawings.

The last project presented in this section is m-shaper [21]. In addition to the drawing and musical aspects it provides, there is also a shape generation component to the system, which adds to the visual content. A user sketches on a Wacom tablet, that is able to get information such as the tilt angle and orientation of the pen, the pressure applied on the tablet and the speed at which they are drawing. These parameters, in addition to the position of a stroke, drive the audio output (instrument, note and octave) and the shape generation. There is also a possibility for users to change the mappings between the gestures/visuals and the sound, and to define whether sound is produced while drawing or when a stroke is done.

All these projects do share elements in common while adding different perspectives to the research field. For example, Different Strokes, Illusio and Articulated Paint, in contrary to the others, choose to put the musical aspect above the drawings, but their goals, functioning and targeted audience remain vastly different. It is the same with Levin's projects and Sounding Brush, which while blending drawing and sound together, have thought out their mappings and visual choices differently. Even when sharing a similar basis of musical drawing in 2D, since there is not a right solution on how to construct a prototype, the possibilities are vast, and still largely under explored.

4 FROM 2D TO 3D

With the recent evolution of 3D display technologies (in Virtual or Mixed Reality, now combined as Extended Reality, using headsets or other immersive displays) and their increasing availability, 3D drawing has opened possibilities to explore, and so has 3D musical drawing.

For instance, Musical Brush [34] explored creativity with musical drawing in several 3D contexts. This project uses a smartphone and its motion characteristics to generate the visual and audio outputs according to a user's gestures.

"Drawing Sound" [27] is a multi-user environment in Mixed Reality using a HoloLens and bare-handed interactions to draw. Here, the space is delimited into six vertical layers of sound. Similarly to Different Strokes, particles follow a stroke on a loop and produce sound according to the height of its current position. When two strokes intersect, an audio-visual effect is produced, defined by the layer in which the intersection happens. This can produce sounds such as piano, synths, bells and so on.

Finally, there is *Reflets* [6], a Mixed Reality environment for musical performances, that works using reflecting transparent panels, projectors and depth cameras: performers and audience members can reveal 3D virtual content placed into the scene by physically intersecting with it. In one of the use-cases of this project, *SoundPaths*, a user, with a WiiRemote and a position tracker, draws 3D sound paths that are revealed by the spectators. The performer controls the sound by intersecting a 3D stroke which plays an audio loop starting at the revealed position.

These projects, as far as we know, are the only 3D musical drawing systems that currently exist. This leaves a lot of room for experimentation, especially considering the possibilities 3D technologies offer.

4.1 The challenges of 3D drawing

A few things can explain the actual lack of projects focusing on 3D musical drawing. Most of them are related to the way 3D drawing technologies work and the issues surrounding it.

Firstly, going from 2D to 3D increases the degrees of freedom in gestural interaction, which has a significant effect on the complexity of the task and the cognitive load of the user. One has to take into consideration that this is just regarding the drawing aspect of musical drawing, and that associating sound can add more complexity to that, depending on the mappings and the overall design of the system. In this aspect, 2D drawing can seem more approachable.

Then, there are the problems regarding stereoscopic displays such has Extended Reality (XR) Head-Mounted Displays (HMD), frequently used for 3D drawing. Due to an inaccurate rendering of depth cues, they may cause the depth perception to be faulty (due to the vergence-accommodation conflict), and consequently lead to difficulties when precisely interacting with 3D content [2, 3]. Other optical see-through HMDs, such as the HoloLens, possess a very narrow field of view of the device, which can prevent users from having sufficient visual feedback on the drawing that they are producing, especially with large drawings close to them. Systems like Spatial Augmented Reality (SAR), which use the physical space and therefore do not possess the same perception issues as other technologies, are currently not used as tools for 3D drawing, except for *Reflets*, because the constraints they pose on displaying full drawings have not been tackled by research.

Another aspect comes from the absence of a physical surface. Even when accompanied by vibrotactile feedback when drawing is activated, 3D drawing remains in mid-air, and therefore lacks contact with a surface that provides primary passive haptic and auditory feedback. The interaction with this surface also benefits musical drawing because it allows expert users to rely on familiar controlled dimensions of contact, pressure and tip orientation on

paper or on a tablet. This familiarity and expertise is lost in 3D drawing.

Finally, the current research on 3D drawing is not especially focused on expressiveness, but rather on more technical constraints related to the issues mentioned above. Many articles dive into the beautification of strokes, visual guidance, input and interaction techniques, and so on. This might partially explain the lack of 3D musical drawing projects, adding that researchers might find 2D more suited to their objectives, and that 3D equipment are still rather new and/or expensive.

5 A DIMENSION SPACE FOR 2D AND 3D MUSICAL DRAWING

A dimension space is a way to better understand and communicate design considerations, raise possible issues and explore different design options, according to the properties of a device or system [15]. It is a part of the Design Space Analysis approach [23], although without explicitly using the *Question, Option, Criteria* method [7]. Since the projects described in the sections 3 and 4 have a high design variability from one another, a dimension space can be a way to look into the design choices in detail, find common trends, and uncover some under explored areas.

Hence, after reviewing the different 2D and 3D projects, we established several dimensions as an attempt to classify them and get a global vision of their designs. However, due to the occasional lack or information and/or the unclear explanations given inside the articles in regards to the dimensions, some elements inside this classification could be inaccurate or open to debate.

All the projects, classified in alphabetical order, can be found in Table 1. The description of the different dimensions are explained in the following subsections.

5.1 Expression Orientation

The first dimension that we identified is the Expression Orientation (EO). It defines whether a project is more drawing oriented, audio oriented, or tries to blend both in a sort of middle ground. It can therefore take three values: Sound, Drawing and Both. The EO is determined according to the project's main research focus. To be in the Both category, a project would have to put the same emphasis on the audio and the visual side of the application. The main difference would be that when focusing specifically on sound, or specifically on drawing, the other component's role is set to be an interaction tool for the music or the drawing.

5.2 Required Expertise

The Required Expertise (RE) defines whether a project is adapted to a specific level of user expertise, according to the previously determined Expression Orientation. If the EO is Sound, then the corresponding expertise would be musician; if it is Drawing, the expertise would be visual artist; if it is Both, then it would correspond to a blend of the above. This dimension can take four values: Novice, Advanced, Expert and All. If the required expertise is not disclosed in the article, we can argue that it is not relevant, and therefore that the project is adapted to all expertise. This dimension also draws inspiration from the dimension space proposed by Birnbaum *et al.* [7].

5.3 Collaboration

The Collaboration (Collab) dimension analyses whether a project supports collaborative work with two ore more participants. It can take two values: Yes (the project supports collaborative work) and No (the project does not support it). It is important as collaborative instruments may allow for assigning roles to different users, *e.g.,* one more focused on the visual aspect and the other on the sonic one, or as they may enable more complex visual or sonic results. With the advent of shared immersive virtual environments, 3D drawing applications may also enable a shared performance between distant users which can collaborate in creating a shared artistic content.

5.4 Visual Replay Value

The Visual Replay Value (VRV) focuses on the way the drawing side of the application is designed, by looking at how the strokes behave inside the environment. Once completed, they can be static (*i.e.,* remain visible), fade out, fade and then redraw themselves, or move on a looping animation. The values are therefore Static, Fade, Loop and Animated. There can be several values for one project, since some of them have different functionalities (for example linked to different brushes), which can imply different drawing behaviours. Controlling this behaviour might also be part of the instrument expressive capabilities.

5.5 Audio Replay Value

The Audio Replay Value (ARV) is similar to the VRV, except that it is focused on the musical side of the application and how the audio output is designed. It can be immediate, meaning that the sound is not replayable, either manually by the user or automatically inside a loop, whether it fades or is continuous; it can be manually replayable by a user, or can repeat itself inside a loop without the need for a user's action. Finally, the sound can evolve over time on its own, meaning that the user does not interact with the audio output for it to change, leading to a higher tolerance to interruption [24]. The four values are therefore Immediate, Manual, Loop and Evolving. Similarly to the VRV, a project can fit into one or more of those categories.

5.6 Modifications

The Modification (Modif) dimension dives into the possibility for a user to modify the work they have done while interacting with the system. It is mostly linked to the drawing side of the application with, for example, the possibility of removing a stroke, changing its shape, erasing a part of it or moving it inside the environment, although musically this could involve applying sonic transformations to the resulting sounds. The possible values are either Yes (some modifications can be made) or No (once drawn, the result is final). This dimension does not go further as to establish a modification spectrum, since what is supported depends on how a project was designed, and they are not necessarily comparable with one another.

5.7 Mapping Structure

The Mapping Structure (MS) focuses on how the audio-visual mappings are designed: whether they are from visual to sound, from sound to visual, or from gestures to both sound and visual. To be

more precise, it deals with which component actually drives the other one. For example, if the drawing parameters (like colour, transparency or stroke weight) affect the audio output, or on the opposite, if it is the audio parameters (like grain, tempo or reverberation) which change how the visuals look. Therefore, the possible value are V->S, S->V and G->VS. It is possible for a project to fit into one or more categories, since it can have, for example, both gestures and visuals driving the other components.

5.8 Mapping Flexibility

The Mapping Flexibility (MF) defines if a user has the possibility to change the mappings between the gestures, the audio output and the visuals, or if they are already defined and considered final. The values are either Yes (they can be changed) or No (they are already implemented and offer no possibility of change). For instance, even if a project offers several different mappings between the audio and the drawings and the user can switch between them, if they cannot be edited, it will be classified as No.

5.9 Degree of Immersion

Finally, the last dimension is the Degree of Immersion (DoI). Immersion can be defined as *a description of a technology* that *describes the extent to which the computer displays are capable of delivering an inclusive, extensive, surrounding and vivid illusion of reality to the senses of a human participant* [32]. Consequently, the Degree of Immersion of a system or device is linked to its technological capacity to provide an environment where the physical reality disappears (inclusive), the sensory modalities are included (extensive), the virtual environment has a large field of view (surrounding) and the device or system provides a rich and qualitative content (vivid) [32]. Here, the DoI is a continuous value, that goes from Low (with a simple computer screen or a tablet), to Mid (with Mixed Reality displays such as Spatial Augmented Reality to see-through Head-Mounted Displays), to High (with Virtual Reality Head-Mounted Displays). The dimension in which the system works (2D or 3D) is given within parenthesis.

6 DISCUSSION

There are several elements we can take from Table 1 to analyse further. The next subsections are dedicated to this.

6.1 Perspectives on musical drawing

One of the first noticeable thing is the absence of collaborative features in the projects, except for Drawing Sound and *Reflets*. Music is a medium well adapted to collaboration (groups, orchestras), and since drawing is natural and easy to understand (especially when it comes to communicating), it would not be too far-off to imagine the possibility of adding collaborative features in some of the projects. From a technological standpoint, all the devices presented could support it. Musical drawing could be a great opportunity for collaborative sessions between users, especially in a shared virtual environment that would then benefit from a strong level of co-presence [33], no matter their levels of expertise, as well as for performances with several people contributing to the audio-visual output, like it is the case with *Reflets*.

Also, all the projects have a low Degree of Immersion, this time again with the exception of Drawing Sound and *Reflets*, which are in Mixed Reality. This is explained by the fact that, in 2D, projects rely on either tablets or computers to function. A high DoI would no be a relevant feature, unless the drawings are in 3D, since a HMD is not needed to draw in 2D. Hence, new 3D musical drawing projects could explore higher Degrees of Immersion.

As for the mappings, two aspects can be discussed. Firstly, the majority of projects use gestures (speed, pressure, movement in space...) to drive the audio and/or the visual output. This can be linked to a desire for expressiveness that is often associated with gestures and body movement (for example with playing an acoustic instrument). However, more complex mappings strategies, *e.g.,* with feedback between visual and sonic parameters, should probably be explored in order to increase expressive capabilities and enable more expert control [17]. Secondly, there is generally no possibility for a user to change the mappings, with the exception of m-shaper. This might be because the audio-visual mappings are not easy to establish, and that could be rather confusing for someone, who has not designed the system or has no knowledge about sound parameters, to do it themselves. Having a bit of freedom could, however, be interesting for a user, since expressiveness is involved. This could make them feel more comfortable using the device in regard to their own sensitivity and sensibility.

Regarding the Required Expertise, the projects are either adapted to all expertise levels, or there is a gap from beginner to expert without anything specifically made for the individuals in between. This can be explained by the fact that the projects for novices are made for people who want to experience music without having to learn an instrument, and the projects for experts are made to be used along with their technical skills, and would not be useful to someone without enough experience to actually master them. This leads to advanced players staying in an in-between with nothing crafted for their needs. Again here, this leaves room for the design of instruments which would allow for a learning curve that, for example, simplifies one of the output modalities in order to help users gain expertise in the other, with mappings that evolve over time.

Focusing on the Visual Replay Value, in most projects, the drawings are completely static (except for Levin's work and a few Sounding Brush brushes). We believe that it could be interesting to try and find more ways to make the drawings match the temporality and/or natural variation of the sound. It could take the form of shrinking or twisting strokes to make them fade, shaking the whole drawing along with the tempo, or using sounds or gestures to temporarily animate the strokes.

For the Audio Replay Value, the projects work with either immediate sounds (continuous or with fading effects) or loops. Apart from *Reflets*, there are no options to manually replay a sound, which is something that could be explored and allow for creativity in musical composition. Also, only one project uses the Evolving-type sound, while that could be something interesting to experiment on, as it could also drive the visuals along.

Projects	EO	RE	Collab	VRV	ARV	Modif	MF	MS	DoI
Articulated Paint	Sound	Novice	No	Static	Immediate	No	No	G->VS	Low (2D)
Different Strokes	Sound	Expert	No	Static	Loop	Yes	No	G->VS, S->V	Low (2D)
Drawing Sound	Both	All	Yes	Static	Loop	Yes	No	G->VS	Mid (3D)
Illusio	Sound	Expert	No	Static	Loop	Yes	Yes	V->S	Low (2D)
Loom	Both	All	No	Loop	Loop	No	No	G->VS	Low (2D)
m-shaper	Both	All	No	Static	Immediate	Yes	Yes	G->VS	Low (2D)
Musical Brush	Both	Novice	No	Static	Immediate	No	No	G->VS	Low (3D)
Reflets	Sound	All	Yes	Static	Manual, Loop	No	No	G->VS	Mid (3D)
Sounding Brush	Both	All	No	Static, Fade, Animated	Immediate, Evolving	No	No	V->S, S->V, G->VS	Low (2D)
Yellowtail	Both	All	No	Animated	Immediate	No	No	V->S	Low (2D)

Table 1: Musical drawing projects classified according to different dimensions: Expression Orientation (EO), Required Expertise (RE), Collaboration (Collab), Visual Replay Value (VRV), Audio Replay Value (ARV), Modifications (Modif), Mapping Flexibility (MF), Mapping Structure (MS) and Degree of Immersion (DoI)

6.2 Expressiveness

There is one element that could explain the absence of projects where the Expression Orientation is Drawing. If music and drawing are not both expressively balanced in the design of a project, then one becomes a simple tool for the other. This is the case for projects such as Different Strokes, Illusio or Articulated Paint: the expressive focus is solely put on the musical aspect, and the drawing part of the device is simply a mean of interaction. If this was to be reversed, and the Expression Orientation was focused on the drawing, then the sound would be treated as a tool. Not putting expressive considerations into the sound design would most likely take the musical aspect out of the project, and thus not be considered musical drawing. We can, for instance, take the example of Xu *et al.* [36]: they applied sonification to a sketching system, but with the intention of using it as a guide for narrow fields of view in AR. Both drawing and sound are mixed together, but without expressive intent.

This leads to design considerations linked to the Expression Orientation of a project. If the EO is Sound, this creates a composition tool or a digital instrument; if the EO is Drawing, it produces an audio guide for drawing trajectories; if it is Both, it gives an audio-visual performative tool/instrument.

It might be possible to have a musical drawing project where the Expression Orientation is Drawing, but where the sonic output would help the visual expression, *e.g.,* by guiding 3D drawing gestures in the case of disrupted depth perception. In this instance, looking into audio-visual cross-modal congruence mixed with more "musical" sounds could be an interesting avenue for research.

6.3 What 3D technologies change

3D drawing offers significant changes from its 2D counterpart, along with its problems (as stated in section 4.1). However, it leaves a lot of room for experimentation.

One of the main advantage of 2D drawing, apart from being less demanding from a cognitive point of view, is that it does not need a lot of space to function. Most of the projects here relied on tablets or computers, which are small, static devices. Going from 2D to 3D implies a greater need for space, which is not suited for all types of projects: some like Illusio or Different Strokes use drawing purely as a tool combined with other devices or instruments and would not find any benefit in switching to 3D. It would make the job of the performer harder, since they would need to interact with other devices or instruments while simultaneously dealing with the 3D technology. On the other hand, the opportunity of being immersed in a 3D virtual drawing opens new perspectives in terms of gestural interaction but also in terms of the navigation and manipulation of the produced drawings, which can take advantage of existing 3D interaction techniques [4].

Another important aspect is the interaction device used for musical drawing, and in particular the haptic feedback that it provides. While in 2D drawing, haptic feedback (at least passive from pen to surface contact) is usually inherent to the interaction with the 2D screen or tablet, in 3D drawing, haptic feedback can be non-existent or very limited, *e.g.,* with vibrotactile feedback only. This leaves room for the creation of dedicated interaction devices that would allow musical drawing in 3D the same richness in feedback as one has in 2D drawing, *e.g.,* by reintroducing some components of 2D physical drawing within a 3D virtual space. This could imply combining mid-air movements traditionally associated with 3D drawing with more physical interactions above or on a surface. Potential solutions for this are hand-attached or hand-held devices that would provide a freely moving physical surface, or even encounter-type haptics. Our first answer to this is *MagneTip* [16], which combines physical interaction with a magnetic pen and a surface coil, with mid-air drawing in extended-reality.

The added dimension from 2D to 3D provides opportunities for additional audio to visual, visual to audio, and gestural to audio and visual mappings. Also, it enables the use of the whole body rather than being limited to the arms and the hands. This provides

a performer with much more freedom to actually occupy the space, and to be physically expressive.

There is also an advantage of 3D musical drawing from an audience point of view. 2D musical drawing, since it uses small devices and interacts on a flat surface, implies restricted gestures, that can be very fine and hard to decipher if not up-close. This can prevent the audience from understanding the actions of the performer [9], especially if this is their only way to view the actions that are taking place (*i.e.,* if they are not broadcast, for example using a projection on a screen). 3D technologies, and the space it provides, would allow for bigger gestures, which would be more understandable for an audience, even without having access to the content of the virtual environment. Furthermore, the 3D navigation within a virtual drawing opens novel scenographical opportunities, where the audience is transported inside the drawing, to focus on various elements during a performance.

7 CONCLUSION

In this article, we reviewed the existing 2D and 3D musical drawing projects in order to classify them, according to eight dimensions: Expression Orientation, Required Expertise, Collaboration, Visual Replay Value, Audio Replay Value, Modifications, Mapping Flexibility, Mapping Structure and Degree of Immersion.

This classification, and the resulting analysis, led to conversations around the design of the projects and brought considerations about the usability, expressiveness, technical and logical constraints, as well as possible additions one can introduce to this research field.

Therefore, we hope that this article opens new perspectives for future work in 2D and/or 3D musical drawing.

ACKNOWLEDGMENTS

The authors would like to thank the researchers included in this review for their work.

REFERENCES

[1] Jerônimo Barbosa, Filipe Calegario, Veronica Teichrieb, Geber Ramalho, and Giordano Cabral. 2013. Illusio: A drawing-based digital music instrument. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. 499–502.

[2] Mayra Donaji Barrera Machuca and Wolfgang Stuerzlinger. 2019. The Effect of Stereo Display Deficiencies on Virtual Hand Pointing. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3290605.3300437

[3] Anil Ufuk Batmaz, Mayra Donaji Barrera Machuca, Duc Minh Pham, and Wolfgang Stuerzlinger. 2019. Do Head-Mounted Display Stereo Deficiencies Affect 3D Pointing Tasks in AR and VR?. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 585–592. https://doi.org/10.1109/VR.2019.8797975

[4] Florent Berthaut. 2020. 3D interaction techniques for musical expression. *Journal of New Music Research* (2020). https://doi.org/10.1080/09298215.2019.1706584

[5] Florent Berthaut, Myriam Desainte-Catherine, and Martin Hachet. 2010. Combining audiovisual mappings for 3d musical interaction. In *International Computer Music Conference*. p–100.

[6] Florent Berthaut, Diego Martinez, Martin Hachet, and Sriram Subramanian. 2015. Reflets: Combining and Revealing Spaces for Musical Performances. In *Proceedings of the International Conference on New Interfaces for Musical Expression* (Baton Rouge, Louisiana, USA) *(NIME 2015)*. The School of Music and the Center for Computation and Technology (CCT), Louisiana State University, Baton Rouge, Louisiana, USA, 116–120.

[7] David Birnbaum, Rebecca Fiebrink, Joseph Malloch, and Marcelo M. Wanderley. 2017. *2005: Towards a Dimension Space for Musical Devices*. Springer International Publishing, Cham, 211–222. https://doi.org/10.1007/978-3-319-47214-0_14

[8] N. Bryan-Kinns. 2004. Daisyphone: the design and impact of a novel environment for remote group music improvisation. In *Proceedings of the 5th Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques* (Cambridge, MA, USA) *(DIS '04)*. Association for Computing Machinery, New York, NY, USA, 135–144. https://doi.org/10.1145/1013115.1013135

[9] Olivier Capra, Florent Berthaut, and Laurent Grisoni. 2020. Have a seat on stage: Restoring trust with spectator experience augmentation techniques. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference*. 695–707.

[10] Erin Cherry and Celine Latulipe. 2014. Quantifying the creativity support of digital tools through the creativity support index. *ACM Transactions on Computer-Human Interaction (TOCHI)* 21, 4 (2014), 1–25. https://doi.org/10.1145/2617588

[11] Natalia Cooper, Ferdinando Milella, Iain Cant, Carlo Pinto, Mark White, and Georg Meyer. 2015. The effects of multisensory cues on the sense of presence and task performance in a virtual reality environment. (2015).

[12] Nuno N Correia and Atau Tanaka. 2017. AVUI: Designing a toolkit for audiovisual interfaces. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 1093–1104.

[13] H.Q. Dinh, N. Walker, L.F. Hodges, Chang Song, and A. Kobayashi. 1999. Evaluating the importance of multi-sensory input on memory and the sense of presence in virtual environments. In *Proceedings IEEE Virtual Reality (Cat. No. 99CB36316)*. IEEE, 222–228. https://doi.org/10.1109/VR.1999.756955

[14] Sidney Fels, Ashley Gadd, and Axel Mulder. 2002. Mapping transparency through metaphor: towards more expressive musical instruments. *Organised Sound* 7, 2 (2002), 109–126.

[15] T. C. Nicholas Graham, Leon A. Watts, Gaëlle Calvary, Joëlle Coutaz, Emmanuel Dubois, and Laurence Nigay. 2000. A dimension space for the design of interactive systems within their physical environments. In *Proceedings of the 3rd Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques* (New York City, New York, USA) *(DIS '00)*. Association for Computing Machinery, New York, NY, USA, 406–416. https://doi.org/10.1145/347642.347799

[16] Esther Gruy and Florent Berthaut. 2024. MagneTip: Reintroducing a Physical Interaction Loop for 3D Musical Drawing in Extended Reality. In *New Interfaces for Musical Expression*. Utrecht (Netherland), Netherlands. https://hal.science/hal-04649511

[17] Andy Hunt, Marcelo M Wanderley, and Matthew Paradis. 2003. The importance of parameter mapping in electronic instrument design. *Journal of New Music Research* 32, 4 (2003), 429–440.

[18] Sergi Jordà, Martin Kaltenbrunner, Günter Geiger, and Ross Bencina. 2005. The reactable. In *ICMC*. Citeseer.

[19] André Knörig, Boris Müller, and Reto Wettach. 2007. Articulated paint: musical expression for non-musicians. In *Proceedings of the 7th International Conference on New Interfaces for Musical Expression* (New York, New York) *(NIME '07)*. Association for Computing Machinery, New York, NY, USA, 384–385. https://doi.org/10.1145/1279740.1279826

[20] Stanislav Kreichi. 1995. The ANS synthesizer: Composing on a photoelectronic instrument. *Leonardo* (1995), 59–62.

[21] Doo-Young Kwon. 2015. m-shaper: A Sketch Drawing System for Musical Shape Generation. *Journal of the Korea Academia-Industrial cooperation Society* 16, 2 (2015), 1381–1387.

[22] Golan Levin. 2000. *Painterly interfaces for audiovisual performance*. Ph. D. Dissertation. Massachusetts Institute of Technology.

[23] Allan Maclean and Diane McKerlie. 1995. *Design space analysis and use representations*. John Wiley & Sons, Inc., USA, 183–207. https://doi.org/10.5555/209227.209245

[24] Joseph Malloch and Marcelo M Wanderley. 2017. Embodied cognition and digital musical instruments: design and performance. In *The Routledge companion to embodied music interaction*. Routledge, 438–447.

[25] Peter Manning. 2012. The oramics machine: From vision to reality. *Organised Sound* 17, 2 (2012), 137–147.

[26] Gérard Marino, Marie-Hélène Serra, and Jean-Michel Raczinski. 1993. The UPIC system: Origins and innovations. *Perspectives of New Music* (1993), 258–269.

[27] Ryu Nakagawa, Ryo Komatsubara, Taku Ota, and Takahisa Mitsumori. 2018. Drawing Sound in MR Space: A Multi-User Audiovisual Experience in Mixed Reality Space. In *Proceedings of the Virtual Reality International Conference - Laval Virtual* (Laval, France) *(VRIC '18)*. Association for Computing Machinery, New York, NY, USA, Article 20, 4 pages. https://doi.org/10.1145/3234253.3234309

[28] Kyosuke Nakanishi, Paul Haimes, Tetsuaki Baba, and Kumiko Kushiyama. 2016. NAKANISYNTH: An Intuitive Freehand Drawing Waveform Synthesiser Application for iOS Devices.. In *NIME*. 143–145.

[29] Gerard Roma and Anna Xambó. 2008. A Tabletop Waveform Editor for Live Performance.. In *New Interfaces for Musical Expression*. 249–252.

[30] W Andrew Schloss. 2003. Using contemporary technology in live performance: The dilemma of the performer. *Journal of New Music Research* 32, 3 (2003), 239–242.

[31] Sourya Sen, Koray Tahiroğlu, and Julia Lohmann. 2020. Sounding Brush: A Tablet based Musical Instrument for Drawing and Mark Making. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, Romain Michon and Franziska Schroeder (Eds.). Birmingham City University, Birmingham, UK, 331–336. https://doi.org/10.5281/zenodo.4813398

[32] Mel Slater and Sylvia Wilbur. 1997. A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence: Teleoperators & Virtual Environments* 6, 6 (1997), 603–616.

[33] Anthony Steed and Ralph Schroeder. 2015. Collaboration in immersive and non-immersive virtual environments. *Immersed in media: Telepresence theory, measurement & technology* (2015), 263–282.

[34] Rafael Valer, Rodrigo Schramm, and Luciana Nedel. 2020. Musical Brush: Exploring Creativity in an AR-based Tool Combining Music and Drawing Generation. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, 634–635. https://doi.org/10.1109/VRW50115.2020.00168

[35] David Wessel and Matthew Wright. 2002. Problems and prospects for intimate musical control of computers. *Computer music journal* 26, 3 (2002), 11–22.

[36] Haonan Xu, Fei Lyu, Jin Huang, and Huawei Tu. 2022. Applying sonification to sketching in the air with mobile ar devices. *IEEE Transactions on Human-Machine Systems* 52, 6 (2022), 1352–1363.

[37] Mark Zadel and Gary Scavone. 2006. Different strokes: a prototype software system for laptop performance and improvisation. In *Proceedings of the 2006 Conference on New Interfaces for Musical Expression* (Paris, France) *(NIME '06)*. IRCAM — Centre Pompidou, Paris, FRA, 168–171. https://doi.org/10.5555/1142215.1142255