EAI Endorsed Transactions
on Creative Technologies

Research Article **EAI.EU**

Design Strategies for a Hybrid Video Synthesizer

Sourya Sen[1,*], Nuno N. Correia[2], Koray Tahiroğlu[3]

[1]Independent Artist, Kolkata, India
[2]University of Greenwich, London, UK and ITI/LARSyS, Funchal, Portugal
[3]Aalto University, School of Arts, Design and Architecture, Helsinki, Finland

Abstract

INTRODUCTION: Technological advances have led to newer instruments and forms of audiovisual performance. Simultaneously, there has been a resurgence of analogue equipment in sound synthesis. However, such instruments for video synthesis are yet to be well established.

OBJECTIVES: The primary objective of this research is to propose design guidelines for designing hybrid video synthesizers and develop a working prototype.

METHODS: Through a study of historical and contemporary practises, design guidelines for a hybrid video synthesizer are proposed. A prototype instrument is then developed and used in live performances to justify the guidelines.

RESULTS: The hybrid approach towards audiovisual performance has some merit, however, it may only be relevant to a certain group of people already familiar with modular workflows.

CONCLUSION: The merits and demerits of our design proposal using the prototype instrument has been noted in this research. Future plans have also been discussed.

Received on 21 December 2020; accepted on 12 March 2021; published on 19 March 2021

Keywords: video synthesizer, eurorack, modular, audiovisual performance

doi:10.4108/eai.19-3-2021.169038

1. Introduction

Live audiovisual performance is an event of live manipulation of sound and image [1]. With rapid advances in technology over the last few decades, various strategies exist today towards composing and performing such material. These range from hardware analogue instruments, digital tools, commercially available tools, and artist-built tools, to name a few.

Parallelly, there has been a resurgence in analogue and analogue-like technology and their workflows, mainly for audio but also in the video realm. Most notably, the Eurorack format, pioneered by Dieter Doepfer in the mid 1990s has gained significant prominence [2]. According to Doepfer, the popularity of the modular platform helps performers to come up with "really distinctive sounds that couldn't have been generated in any other way". Visual performers have also been rediscovering analogue technology, both in the screen-based space, but also with other mediums such as vector synthesis and working with laser systems [3].

Digital technologies are known for their high fidelity reproduction and precision - but the imperfections of analogue technologies hold a certain attractiveness for artists [4]. This is particularly noticed in the hybrid

*Corresponding author. Email: ayruos@gmail.com

audio practises of today, combining analogue outboard gear with the convenience of digital audio workstations (DAW). While there has been a similar movement in the visual domain, the practise is far less prevalent. This digital-analogue divide that exists is of particular interest to the authors. The aim of our research is to identify design requirements for a hybrid video synthesizer and, based on that, to develop a proof of concept instrument.

2. Audiovisual Performance and Evolution of Video Synthesis

2.1. Audiovisual Performance

Audiovisual performance has a long history, going back to discussions on correlations between musical scale and colors by Greek philosophers, such as Aristotle and Pythagoras [5]. Early forms of audiovisual experiments have led to various disciplines in the recent past, namely, expanded cinema, live cinema, VJing and live audiovisual performances [1]. While the distinctions between them can be blurry, any intermedia work can be termed 'audiovisual' [1]. When it comes to artistic practices arising from the same, however, further distinctions can be made.

The Pythagorean concept of arranging colours in intervals similar to a system of organizing musical frequencies towards a practice of visual music as well as incorporating concepts of synaesthesia are often attributed to the development of an audiovisual practice [1] [5]. However, those concepts are not adequate towards defining the field. In reality, contemporary audiovisual production is an amalgamation of achievements in technology [1] [6]. Live audiovisual performance can be defined as an artistic practise incorporating live manipulations of image and sound which are time and media based as well as performative [1]. Furthermore, it does not comprise of a single style or technique but identifies common elements that identify such a practise [1]. Chion [7] makes the argument towards combining audiovisual elements to form a third "audiovisual element", which creates "added value". Grierson elaborates further, suggesting audiovisual works with added value as specific responses arising out of synthesis of sound and images and concludes:

> (...) the audiovisual work is more than a combination of its component parts. As such, the practise of audiovisual composition is not simply the production of audio and video. It is the process of composing audiovisual works which exploit added value. [8]

By creating a third audiovisual element, an audiovisual composition is more than a sum of their parts and exists beyond the audio or visual element in isolation. Elaborating further, Grierson [8] identifies the main strategies for such compositions as synchronization of audio and visual elements, audiovisual congruence, and binary opposition and provides a valuable framework for analysis of audiovisual compositions and is used to evaluate our approach later in this paper.

In his research on contemporary audiovisual performance practice, and on his own artistic work, Correia proposed the concept of Interactive Audiovisual Objects as a compositional strategy: modules integrating user interface with sound visualization [9]. He expands on Chion's concept of Audiovisual Contract by proposing an Interactive Audiovisual Contract: "a symbolic contract that the users enter into, agreeing to consider that sound, visualization and GUI (Graphical User Interface) form a single entity" (ibid.). Correia's concepts mainly relate to GUIs, but his concepts can also be applied to custom physical interfaces. Regarding interface design for audiovisual performances, Correia also developed the openFrameworks ofxAVUI software tool, where "flexibility of mappings between UI, sound and image, and ways to manage that flexibility, are fundamental qualities" [10]. This flexibility of mapping plays an important role in our discussion later.

The electronic analogue video synthesizer and computer-based approaches towards image generation and manipulation have historically contributed towards building visual tools for audiovisual performances which involve the features discussed above - strategies for creating added value as well as linear and non-linear mappings between audio and video. In the next section we shall discuss the role of the video synthesizer in particular and its historical significance to audiovisual practises.

2.2. The Role of Video Synthesizers

Collopy [11] argues that the earliest form of video synthesizers were modelled after music synthesizers of their time and this can be seen in the instruments built by Eric Siegel, Bill Etra and Steve Rutt, Dan Sandin, and others. Siegel defined his work as the "video equivalent of a music synthesizer" (as quoted in [12]), and Beck's synthesizer would incorporate "principles of control voltage and artificial signal production from audio synthesizers" [11]. The result of this influence meant that these video synthesizers could be controlled the same way as audio synthesizers – in a modular fashion with control voltages (CV), often compatible with each other. As Sandin states [12], "I just went through all the Moog modules... and said if you center their bandwidth to handle video and you do the right things with sync, what would they do?"

The term itself is a matter of debate. Dunn, writing on audio synthesizers, notes:

> (...) there is nothing synthetic about the sounds generated from this class of analog electronic instruments, and since they do not ‘synthesize’ other sounds, the term is more the result of a conceptual confusion emanating from the industrial nonsense about how these instruments ‘imitate’ traditional acoustic ones. [12]

Extending the concept to video, any electronic circuitry capable of producing video signals can be called a video synthesizer. The earliest forms of video synthesis involved feedback effects and direct manipulation of the video signals, and were being simultaneously developed and practised in the 1960s [11]. The true video synthesizer, for the purposes of a definition, would come from Eric Siegel in 1970, with his invention of the Electronic Video Synthesizer. In his own words, it was:

> (...) like the video equivalent of a music synthesizer, where you have a program board and you can start to set up a whole series of visual geometric happenings in color on the video signals—the screen—and this is designed for video composition. . . everything is composed right inside of the synthesizer. (as quoted in [12])

This paper refers to a video synthesizer in a similar fashion - a machine capable of producing standalone video signals without the need for any image input. While this may fulfil the definition of video synthesis, for a video synthesiser to become a successful instrument within an audiovisual setting, further details need to be discussed over and above the signal production mechanics.

For audiovisual composition using visual tools, the Experimental Television Centre (ETC)[1] identified the following key qualities – real time aspect to the creative process with an open ended architecture, indeterminacy in terms of outcome, interactivity with the tools, sound image synchronization, electronic generation of camera-less images and finally, merging within the placement of our body in real spaces [13]. These are important properties to keep in mind as video synthesis tools can and may exist in isolation, how they integrate within audiovisual environments and become an instrument is a key element of our discussion.

Parallel to the analogue world of video synthesis, digital advancements were also underway. On one hand there was the addition of digital circuitry in analogue synthesizers for added control [11], and the other, there was development of computer-based systems for graphics generation. The role of computers in the arts would only increase over time as they would become smaller, portable and more powerful and they play an important role in the artistic audiovisual performance world of today [14].

2.3. To Digital And Back

Digital platforms and computers would also fundamentally change the landscape of electronic music. Modular synthesizers would lose favour for smaller and powerful digital synthesizers and eventually, computers [15]. The digital movement, while popular and seemingly all ubiquitous, would soon be challenged by artists looking for creative expressions within the perfections of the digital technologies [4]. A hacker movement on the one hand and analogue nostalgia on the other would eventually give rise to the modular Eurorack format, pioneered by Dieter Doepfer [16] [4]. While still relying on concepts of discrete modules and voltage control like the modular synthesizers of the past, this new format would both be smaller and portable, and also incorporate modern digital microcontrollers for a truly modern modular environment for sound synthesis. This is of course, in conjunction with the continued use of the computer, digital musical instruments (both hardware and software) and the digital audio workstation (DAW). The modern musical practice, therefore, adopts a modular and hybrid nature with the availability and use of the varied instruments and approaches which are easily available.

When it comes to video synthesis however, hybrid approaches are less sophisticated. Stemming from the hacker culture of the 90s, there are various DIY video synthesizers like the RGB3Trins and the CHA/V [4]. Companies like LZX Industries[2], are making commercially available analogue video synthesizers as Eurorack modules as well as semi modular instruments. To clarify, analogue here refers to the output signal as certain circuits used within these instruments do use digital integrated circuits in their design.

Digital technologies have been adopted in hardware video synthesizers but many of them lack modular control over parameters for integration in an audiovisual setting. A notable exception is the recently released Hypno by Sleepy Circuits which “digitally emulates and extends common analog video workflows that previously required whole systems to achieve … in a simple compact interface”[3] and might be the blueprint of future video synthesizers to come. As it is, it adheres very closely to the design guidelines proposed later in this paper. The Hypno might be the exception to the norm but the lack of a modular yet modern video

[1] http://www.experimentaltvcenter.org/

[2] https://lzxindustries.net/

[3] https://sleepycircuits.com/hypno

synthesis platform is important to note here considering the influence of modular audio synthesizers in the design of the pioneering video synthesizers.

Modular workflows in audiovisual composition and performances are not uncommon and either use the aforementioned modular instruments or are executed in the software realm. Max/MSP, Touchdesigner, Pure Data, etc are popular modular, patch based audio, visual and audiovisual composing and performance software. A recent trend of digital emulations of Eurorack modules and workflows are also increasingly popular with solutions like VCV Rack and Softube Modular. A recent project, the Urack [17] combines the Unity game engine with VCV Rack, a digital modular synthesizer format, to create a software modular audiovisual system.

3. Analysis of Contemporary Practises

As discussed earlier, audiovisual performance involves the creation of audiovisual works which exploit added value [8]. Many contemporary approaches exist with that same aim and a list of all such strategies will never be exhaustive. However, characteristics that define a flexible audiovisual instrument can still be identified. Franco et al. [18] argue that such an instrument should have the following features:

1. Real-time (improvisatory) performance capabilities for the creation of images and sound.
2. Compositional structures, event organization and modification.
3. Expressiveness.
4. Mapping flexibility between audio and visuals.
5. Modifiers and effects.
6. Learnability.

Levin [19] also proposes similar ideas, identifying the following characteristics:

1. The system makes possible the creation of dynamic imagery and sound, simultaneously in real time.
2. The system's results are inexhaustible and extremely variable.
3. The sonic and visual dimensions are malleable.
4. The system permits the performer to create or superimpose their own ideas over conventions and idioms of established visual languages.
5. The basic principles of operation are easy to deduce while sophisticated expressions are possible and mastery is elusive.

To summarize from the above, both Franco et al. [18] and Levin [19] discuss ideas of customization, flexibility and complexity that can be achieved through such an instrument. In practise, two distinct classes of instruments and platforms can be identified, 1. Usage of existing and commercial solutions towards building a performance, and, 2. Artist-developed tools and systems for performance.

When it comes to using existing tools and commercial solutions, many hardware and software systems are available, including visual jockey (VJ) tools. Usually, such an approach is easier to use with immediate outputs but come at cost of flexibility. They are either designed for certain fixed workflows and get difficult to work around, or the audio and visual tools are separate entities and are difficult to integrate with only rudimental mapping strategies available between them.

Artist-developed tools, either in software of hardware, tend to be more flexible and customizable as they are built from the ground up by the composer/performer according to their needs and tend to satisfy the properties noted by Franco et al. [18] and Levin [19]. However, this comes at the added expense of the development for the instrument design required for the composition process.

Magnusson's analysis of instrument design through the concept of affordances and constraints can be extended towards audiovisual platforms as well as purely visual tools. On the topic of commercial solutions versus artist developed tools, he writes,

> Problems with the former lie in the conceptual and compositional constraints imposed upon users by software tools that clearly define the scope of available musical expressions. It is for this reason that many musicians, determined to fight the fossilization of music into stylistic boxes, often choose to work with programming environments that allow for more extensive experimentation. However, problems here include the practically infinite expressive scope of the environment, sometimes resulting in a creative paralysis or in the frequent symptom of a musician-turned-engineer. [20]

The same phenomenon is also seen in purely visual performance systems, both in hardware and software. Striking the right balance between an extensive development framework and the constraints put into it often leads to more immediate output. To such an end, Magnusson [20] introduces the idea of mapping as design constraints. This is proposed through a model where the artist interacts with a controller using a mapping engine connected to the sound engine for performance and composition. Through this interaction and sonic and haptic feedback (and additionally, visual feedback), the musician can continue their process of composition and performance.

Regarding these control and interaction possibilities, it might not be too much to argue that our current digital musical instruments are likewise narrowly constrained. This is something we know, in many such artistic fields constraints can be a creative encouragement. Even if we could be able to provide "boundless space for musical expression", the design of computational tools account for particular types of thinking and particular modes of interaction [21]. It is possible to see the impact of these constraints in visual performance systems as well. Despite all the possible constraints, technology applied in mapping strategies between control interface and sound/video processing is an inherent part of the audiovisual performance systems and it opens up possibilities for performance practices to emerge through its technology design [22].

A video synthesizer or an audiovisual system can be analysed through a similar model. In contemporary practices, various controls are provided to the performer, which are mapped to various parameters of the visual engine for real time manipulation. Therefore, with regard to designing or developing an instrument enabling a performer to do the same, three things become important – the **control interface**, the **mapping engine** and the **visual engine**.

When it comes to mapping in relation to audiovisual performances, four approaches are possible [13]. These are namely, 1. Sound and image being derived from the same source, 2. Sound being used to modulate image, 3. Image being used to modulate sound, and finally, 4. Sound and image being modulated by a third source (ibid.). In the realm of composing and performing audiovisual works where one medium is not dependent on the other (i.e., not visualizing sound or sonifying image), therefore, the major interest lies in either deriving the sound and image from the same source or modulating sound and image by a third source.

4. Methodology

The research follows a practise-based methodology. Candy [23] defines practice-based research as an “investigation undertaken in order to gain new knowledge partly by means of practice and the outcomes of that practice” and this may be demonstrated through various outcomes, including creation of an artefact. Furthermore, practice-based research projects also need to be supplemented with “textual analysis or explanation to support its position and to demonstrate critical reflection” (ibid.).

To investigate our design proposal, contemporary performance strategies are analysed and design requirements for a hybrid instrument proposed based on that analysis. A proof of concept instrument is then developed with an iterative development process for practical analysis of the proposed design requirements. Initial stages of its development have been reported on the first author’s MA thesis [24] as well as a short conference paper [25]. The MA thesis focused on the overall design and development while the conference paper concentrated on the history of the video synthesizer leading to the design proposals. In this paper, we will reflect on the current state of the design of the synthesizer and discuss it in the light of the related work and literature presented above.

5. Proposed Design Strategies for a Hybrid Video Synthesizer

For developing a modern hybrid video synthesizer which would be physical, tactile and work in a modular fashion, based on the earlier discussions, the authors would like to propose the following strategies.

Firstly, it would have to be **compatible with the Eurorack format**. As discussed earlier, pioneering video synthesis platforms have had a rich history of inheriting from the modular sound synthesis platforms of their time. Following a similar strategy would have a few key advantages. First of all, with compatible voltage sources, it would be possible to drive both audio and video together. With a wide variety of Eurorack modules which are available, from sequencers to low frequency oscillators to function generators, it would allow for a wide variety of control voltages and modular routing possibilities, it would fulfil the characteristics as identified by Franco et al. [18] and Levin [19]. Additionally, these control voltages would be able to modulate both sound and image and when strategically used, correspond to the ideas stated by Grierson [8] towards exploiting added value. Finally, as the Eurorack standard is widely available and adopted by both artists as well as manufacturers/designers, it would leave us with developing only the visual side of an audiovisual performance system. Such a system would be infinitely customizable depending on the other modules which a particular performer would like to combine the video synthesizer with.

Secondly, it should be **both easy to use as well as extensively customizable** towards generating complexity. This would ideally lead to a digital system as even with different possible analogue control voltages, the core video synthesis engine in an analogue design would be immutable. Following the idea of mapping as design constraints [20], this instrument would: be able to map both physical inputs as well as control voltage inputs to visual parameters;

be immediately plug and play; and be infinitely customizable under the hood to fundamentally alter the core visual engines.

Finally, it would need to be **compatible with contemporary display technologies** to easily integrate with modern studios and performance venues.

6. Design and Development of a Prototype Instrument

A prototype instrument following these strategies has been built, around a Raspberry Pi with custom hardware and software. The hardware developed integrates the Raspberry Pi computer in an Eurorack environment and the software runs the synthesis engines, programmed using the openFrameworks[4] creative coding platform. It fulfills the proposed design requirements as mentioned above. The details are furnished in the next sections, broken up into Eurorack integration, balancing simplicity and complexity, and adherence to modern standards.

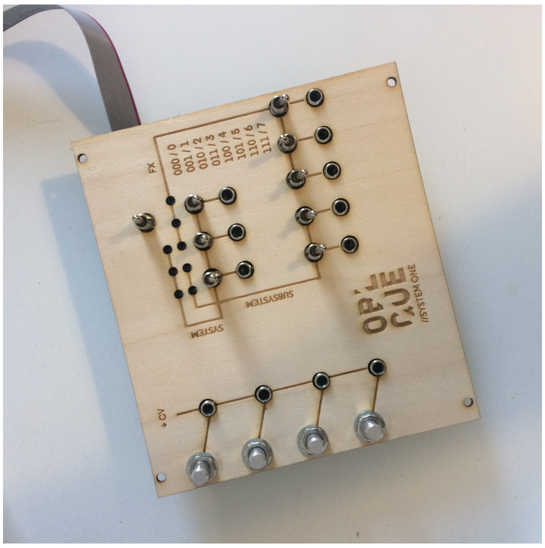

Figure 1. The Front Panel of the Prototype Instrument

6.1. Eurorack Integration

When integrating with Eurorack control voltages, there are two different kinds of voltages. The first of these are continuous voltages, which can be anywhere within the Eurorack voltage standards in the form of low frequency oscillators (LFOs), envelope and function generators, and finally, stepped voltage sequences. These are generally used as modulators in an Eurorack system. The second kind of voltages are triggers and gates, the former being momentary gate signals, which are a binary on/off state and used to trigger and sequence events.

Our design (see Figure 1) has four bipolar CV inputs on the left and eight gate toggles in two rows for both continuous voltage inputs as well as gate/trigger sequences. While the former is mapped to parameters of a particular visual engine, the latter allows for sequencing between different engines as well as additional modifiers. All the CV and gate inputs are controllable from the front panel but are overridden by a voltage source once a cable is plugged in. In case of the continuous voltage inputs, the panel control acts as an attenuator when a cable is present. A detailed explanation of the circuitry and the controls are provided in the Github repository of the project[5].

6.2. Balancing Simplicity and Complexity

Keeping in mind ease of use, the video synthesizer, once plugged in, immediately has a video output. Using the panel controls or using external CV sources, they can also be immediately manipulated and mapped with various degrees of control. The software environment has however been carefully designed to give many levels of customizable options to the performer.

At the very core level, the hardware interfacing layer is kept separate from the visual engine. Therefore, in the most extreme case, the performer is free to build their own visual engine and program it in whichever way they choose. However, one does not need to go that low level. Currently, the visual engines are loaded as banks of 'systems' and 'subsystems' which are selected via the gate inputs or toggle switches. Each of these 'systems' are GLES (openGL for embedded systems) fragment shader files which are accessible as text files to be easily modifiable without needing to recompile the entire application. Additionally, the CV parameters are also exposed in the same shader files and can be remapped at will. Decoupling the hardware interface and the visual engine and also implementing a shader based visual engine has been deliberate decisions to keep the software stack as customizable as possible while maintaining a plug and play video synthesizer.

Additionally, the shaders have been developed with a cross-platform syntax. A desktop version of the visual synthesis engine was also developed to freely experiment with shaders and their mappings on a desktop environment without needing to access the Raspberry Pi over other methods.

6.3. Adherence to Modern Technological Standards

One of the reasons for choosing the Raspberry Pi was its default HDMI output, a modern standard for

[4] https://openframeworks.cc/

[5] https://github.com/sourya-sen/rpi_Synth

connecting to displays. As modern computers tend to have HDMI (or similar digital) outputs, performance venues and projectors increasingly have HDMI as one of the default options for their display format. Alongside capture cards and HDMI switchers and mixers, an HDMI signal is more easily integrated into performance systems unlike analog displays which have inherent issues of capturing, reproducing and projecting in todays day and age.

7. Discussion

As mentioned above, the design of the instrument followed an iterative process. This also included performances at various festivals like LPM 2018, Ääniaalto 2018 and Eyemyth 2019 among others (see Figure 2). Throughout the development cycle, various modifications have taken place with learnings from each stage of development and performances taken into account.

7.1. Patching Strategies for Audiovisual Performance

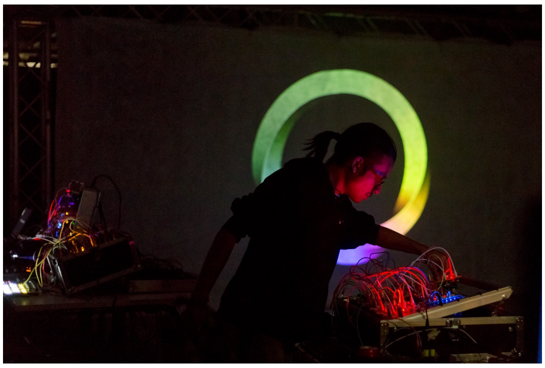

Figure 2. The first author performing with the prototype synthesizer at Ääniaalto 2018

The currently developed instrument in an audiovisual setting very clearly adheres to audiovisual composition where image and audio are both modulated from a third control voltage source [13]. Patching amplitude envelopes to the audio source as well as a parameter in the video synth creates immediate synchronicity towards building added value [8]. Patching gate sequences to 'system' and 'subsystem' in time with audio intervals also creates distinct sound-image events and compositional structures [18]. For example, gate sequences clocking the main sequence can be patched in parallel to change the 'system' or 'subsystem' parameters. Within a patch, the envelop driving the amplitude can be used to also change any CV-addressable parameter of a shader.

One drawback has been noticed in the current design scheme however. There is currently no way to route any sort of visual parameter to the sound generation module – ie, the image cannot modulate the sound in any way currently, although the converse is possible [13]. A system like this would be more tightly integrated if it were possible to control audio features, for example, the amount of reverb, depending on the overall brightness of the image, akin to how an envelope follower works in the audio domain.

7.2. Ease of Use and Complexity in Practise

The design of the software keeping in mind ease of use but allowing for extensive customization has already been described above. Based on the features defined by Franco et al. [18], the instrument allows for real time improvisatory performance capabilities; can work with compositional structures and event organization; can be expressive via the means of control voltages; and offers flexibility towards mapping between image and sound.

With the availability of certain modifiers, it is also possible to further add effects, and with a simple panel control structure comprising a few knobs and switches, is not too difficult to learn as well. Therefore, there are enough strategies implemented towards attaining audiovisual congruency or binary opposition, features discussed by Grierson [8] towards "exploiting added value", with clear design choices made towards affordances and constraints in tune with Magnusson's discussion [8], offering a mapping engine with tactile and visual feedback.

A fair amount of the evaluation above is based on the assumption that the user already has an understanding of modular synthesizers. As the first author realised while presenting the instrument at an international festival, this means that the instrument can appear confusing to an audience who do not have such understanding. Without having other modules offering control voltages and audio sources, the video synthesizer only makes up for one half of an expressive audiovisual performance platform and the importance of the Eurorack ecosystem cannot be stressed enough.

Finally, with the modular design of the hardware and software, further complexity can be derived and the entire functionality can be altered if the performer wishes to do so. This can range from customizing shaders to integrating additional hardware, like cameras, to the Raspberry Pi. For such an endeavour, the performer would need to be familiar with certain paradigms - however, this is not absolutely necessary as the instrument currently has all the core functionality, including a set of shaders to immediately start using within an audiovisual context.

7.3. Working in the Eurorack Format

Even though Eurorack has a defined standard, it is common to see different manufacturers working within their own ranges of voltage tolerance. For example, most Make Noise[6] modules expect unipolar control voltages, whereas most Mutable Instruments[7] modules expect bipolar control voltages instead. Therefore, the Eurorack standard is quite fragmented in nature. While there are various modules designed towards attenuating and offsetting voltages to use in such a situation, many times there is a fair amount of guesswork involved with regards to what a module is outputting or expecting.

This led us towards adding additional code to the *ControlVoltage* class later in software, as we had not anticipated this when starting development of the design of the instrument. If we were starting out again or redesigning the hardware extensively, we would approach this issue from a hardware point of view but adding attenuverters at each CV input – a feature that is commonly seen in physically larger Eurorack modules. Attenuverters, unlike attenuators, allow for scaling and offsetting the control voltage inputs on the panel itself and can be extremely useful when working with a wide range of CV inputs, specially if they follow different voltage standards. This does add in more complexity to the hardware design and additionally may take up extra space on the panel as well and raise the manufacturing cost, but is something to deliberate when designing modules within the Eurorack format.

7.4. Advantages and Disadvantages of Our Approach Against Other Solutions

Modern video synthesiser tools and modular approaches exist in the contemporary realm but our approach has a few key differences. From the standpoint of artist-developed tools versus commercial solutions [20], our approach takes into consideration the advantages and disadvantages of both the strategies and tries to design a platform which is both easy to use and simple to get up and running like commercial solutions but keeps low level access for extensive modification. A critical feature of developing our instrument was making it as platform-like as possible and not for one-off use.

There are a few projects that exist in the Eurorack realm for video synthesis. Ours differ from the LZX series of products by using digital standards and an HDMI signal compared to analogue video outputs of the LZX instruments. Additionally, a key difference between both the LZX series and the Hypno is the open nature of the software stack in our implementation allowing it to be customized in terms of functionality, something the analogue video synthesis tools as well as the Hypno do not offer.

Complying to the Eurorack standards have both pros and cons, however. The advantages have been noted above, like having flexible patching strategies for improvisation and indeterminacy, but there are certain drawbacks as well. The video synthesizer is as powerful as the Eurorack modules around it allow it to be and is suited towards performers already working within the Eurorack ecosystem. This makes it a rather niche instrument within the already niche subject of audiovisual performance. Additionally, the discrepancy in the varying voltage standards within Eurorack has already been noted. This creates added difficulty while designing for, and working with, other Eurorack modules. This also relates to the earlier comment on the instrument being only as powerful as the Eurorack modules used in conjunction with the video synthesizer.

Finally, for our proof of concept instrument, we chose a Raspberry Pi for the various advantages it offers – it is a small Linux computer with general purpose input output ports for integrating custom hardware and also offers a native HDMI output. However, the graphics performance of the Raspberry Pi left much to be desired and certain compromises had to be made. This mainly included capping the maximum frames per second of the image generation at 25 frames per second and even then being restricted to simpler shaders for realtime performances. Advanced techniques involving complex geometries or raymarching could not be executed on the Raspbery Pi at a decent frame rate. Looking at modern devices like smartphones, tablets and even smartwatches, it is obvious that high performance graphics capabilities can be integrated in small form factors and hopefully with our instrument as well as the design guidelines discussed above, other more powerful instruments would soon exist, either as finished products or platforms to build off of.

8. Conclusion and Future Work

In this paper, we have proposed a design for a modern hybrid video synthesizer format for the purposes of live audiovisual performance. We have taken inspiration from the pioneering video synthesis platforms and their design based on contemporary modular synthesizers. Additionally, we have also analysed the different features that make up a flexible audiovisual performance instrument and the different approaches towards it. Through this, our proposal boils down to three key aspects: compatibility with the Eurorack format, balance of ease of use and customizability, and adherence to modern technological standards.

Our motivation came from the fact that there were none such instruments fitting our needs and by designing a proof of concept instrument, we present further

[6] https://www.makenoisemusic.com/
[7] https://mutable-instruments.net/

discussion on this topic. By using our instrument for live performances with various degrees of success, we can claim that following the proposed design strategies does lead to a robust audiovisual performance instrument with the characteristics discussed in this paper, some of the disadvantages withstanding. From a technical standpoint, the video synthesiser has successfully identified the video projection system at each venue and ran crash free for the entire duration of the performances. For future work, we aim to refine our instrument further, addressing the drawbacks that have been mentioned in the paper.

In the meanwhile, we have released our design open source and wish to formally present it to the artistic world at large and gather further user feedback. It would be useful to gather feedback from other artists, as well as to investigate how they implement this instrument within their own practices and Eurorack systems. By releasing the project open source, we also hope to see it being adopted and modified by other artists for their own custom solutions.

Acknowledgements

The authors wish to thank the staff and students of the Department of Media, Aalto School of Arts, Finland for their encouragement, counsel and facilities for the realization of this project.

References

[1] Ana Carvalho and Cornelia Lund. *The audiovisual breakthrough*. Collin & Maierski Print GbR, 2015.

[2] Ableton. Interview with modular synthesis guru Dieter Doepfer, February 2015. URL https://www.ableton.com/en/blog/dieter-doepfer-completing-the-circuit/.

[3] Derek Holzer. *Vector synthesis: a media archaeological investigation into sound-modulated light*. Self Published, 2019.

[4] Joseph Hyde. The new analogue: media archaeology as creative practice in 21st century audiovisual art. In Andrew Knight-Hill, editor, *Sound and image: aesthetics and practices*, pages 188–205. Focal Press, 2020. ISBN 978-0-367-27146-6.

[5] William Moritz. The dream of color music, and machines that made it possible. *Animation World Magazine*, 2(1), 1997.

[6] Luísa Ribas. *The Nature of Sound-Image Relations in Digital Interactive Systems*. PhD thesis, Universidade do Porto Faculdade de Belas Artes, Porto, 2011.

[7] Michel Chion. *Audio-Vision: Sound on Screen*. Columbia University Press, New York, 1994.

[8] Mick Grierson. *Audiovisual Composition*. PhD thesis, University of Kent, Kent, 2005. URL http://www.strangeloop.co.uk/Dr.%20M.Grierson%20-%20Audiovisual%20Composition%20Thesis.pdf.

[9] Nuno N. Correia. *Interactive audiovisual objects*. Aalto University, 2013.

[10] Nuno N. Correia and Atau Tanaka. AVUI: Designing a Toolkit for Audiovisual Interfaces. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pages 1093–1104, New York, NY, USA, 2017. ACM. ISBN 978-1-4503-4655-9. doi:10.1145/3025453.3026042. URL http://doi.acm.org/10.1145/3025453.3026042.

[11] Peter Sachs Collopy. Video Synthesizers: From Analog Computing to Digital Art. *IEEE Annals of the History of Computing*, 36(4):74–86, October 2014. ISSN 1058-6180. doi:10.1109/MAHC.2014.62.

[12] David Dunn. *Eigenwelt Der Apparate-Welt Pioneers of Electronic Art*. The Vasulkas, 1992.

[13] Hank Rudolph. ETC's System. In Kathy High, Sherry Miller, and Mona Jimenez, editors, *The Emergence of Video Processing Tools*, pages 473–84. Intellect Ltd, 2014.

[14] Chris Salter. *Entangled: technology and the transformation of performance*. MIT Press, 2010.

[15] Frank Trocco and T. J. Pinch. *Analog Days: The Invention and Impact of the Moog Synthesizer*. Harvard University Press, June 2009. ISBN 978-0-674-04216-2. Google-Books-ID: CoUs2SSvG4EC.

[16] Robert Fantinatto. *I Dream of Wires*. Waveshaper Media, 2014.

[17] Matthew Hughes and Andrew Johnston. URack: Modular audio-visual composition with Unity and VCV Rack. *New Interfaces for Musical Expression*, 2020.

[18] Enrique Franco, Niall J L Griffith, and Mikael Fernström. Issues for Designing a flexible expressive audiovisual system for real-time performance & composition. *New Interfaces for Musica Expression*, 2004.

[19] Golan Levin. *Painterly Interfaces for Audiovisual Performance*. PhD thesis, M.Sc. Dissertation. Massachusetts Institute of Technology, Cambridge, MA, 2000. URL http://acg.media.mit.edu/people/golan/thesis/thesis300.pdf.

[20] Thor Magnusson. Designing Constraints: Composing and Performing with Digital Musical Systems. *Computer Music Journal*, 34(4):62–73, December 2010. ISSN 0148-9267, 1531-5169. doi:10.1162/COMJ_a_00026. URL http://www.mitpressjournals.org/doi/10.1162/COMJ_a_00026.

[21] Andrew McPherson and Koray Tahiroğlu. Idiomatic Patterns and Aesthetic Influence in Computer Music Languages. *Organised Sound: an international journal of music and technology*, 2019.

[22] Koray Tahiroğlu, Thor Magnusson, Adam Parkinson, Iris Garrelfs, and Atau Tanaka. Digital musical instruments as probes: how computation changes the mode-of-being of musical instruments. *Organised Sound*, 2019.

[23] Linda Candy. Practice Based Research: A Guide. *CCS Report*, V1.0, 2006. URL http://www.creativityandcognition.com/wp-content/uploads/2011/04/PBR-Guide-1.1-2006.pdf.

[24] Sourya Sen. *Developing a hybrid video synthesiser for audiovisual performance and composition*. Aalto University, 2019.

[25] Sourya Sen and Nuno N. Correia. Design strategies for a hybrid video synthesizer. In *Proc. SIIDS 2020 Sound, Image and Interaction Design Symposium*. SIIDS, 2020.