Motus 4: new developments on a web-deployed software for composing real-time audiovisual interactions.

Augusto Piato Amstalden [1*], **Tiago Fernandes Tavares**[12]

[1]School of Electrical and Computer Engineering – University of Campinas
Av. Albert Einstein, 400, Campinas, SP

[2]Interdisciplinary Nucleus for Sound Studies
Rua da Reitoria, 82, Campinas, SP

a213364@dac.unicamp.br

Abstract. *Camera-based audiovisual interactions have been greatly used in art installations. For such, it is often necessary to use specialized hardware and software. This restricts audiovisual installations to areas that can both afford them financially and be physically accessed by the audience. In this work, we present the development of a web-deployed composition tool for audiovisual interactions that runs on the client side and does not require installing any additional software. The tool allows quickly configuring areas for movement capture, and customizing the corresponding audio feedback. Simultaneously, it provides visual feedback that can aid the audience to understand the experience. Consequently, the tool can be used to compose audiovisual interactions that reach a large audience via web. Moreover, its virtual and ubiquitous nature can foster a diversity of artistic explorations related to the remote characteristic of the relationship between audience and artwork.*

1. Introduction

Art installations are artwork that dialogue with the space around them. Interactive art installations have the additional characteristic of using the audience participation as part of its structure, that is, the artistic context only exists with the participation of an audience [1]. Audience participation has been present in a great amount of artwork throughout history, but only in the last few decades the interaction between the audience and the artwork has been mediated using sensors, actuators, and digital computers.

One important characteristic of digital computers is that they are able to emulate systems that are physically unfeasible. This characteristic is found, for example, in computer games, which commonly have physics engines that are unrealistic, yet foster highly engaging and immersive environments [2]. The link between the physical and the virtual worlds, provided by sensors and actuators, has enabled a diversity of interactive artwork, including augmented performances [3, 4], live audience participation [5], augmented musical instruments [6, 7], instrument development frameworks [8, 9], and art installations [10, 11].

Interactive art installations commonly require specialized equipment, including projectors, loudspeakers, the sensors themselves, computers, and a physical environment to host the installation. These elements can quickly become prohibitively costly, especially in economically disfavoured areas. Consequently, the access to the artistic richness brought by interactive artwork is socially and economically determined [12, 13].

Moreover, in 2020-2021 the world has suffered the COVID-19 pandemic. Within this period, the access to non-essential locations (including art galleries) has been greatly restricted. Another consequence of the COVID-19 crisis is that the Internet and social media have become more relevant as means to communicate and share information.

Such a worldwide transformation has lead to the development of several artwork exploiting virtuality and social isolation [14, 15]. Among those, there are ideas based on audio streams, interactive websites, and online streams of studio-based hardware. These ideas generate requirements for frameworks that allow using real-time client-side audiovisual elements, and a myriad of different frameworks have been employed for such.

This concept aligns with the ideas underlying the MOTUS [10], an audiovisual interactive installation that focuses on using affordable hardware for an engine that captures movement and converts it into musical and visual feedback. Cameras have been used in human-computer interaction for music for decades [3], and MOTUS employs this concept while deploying its software as a webpage, thus only requiring hardware that is becoming increasingly ubiquitous [16, 17]. Therefore, even though it is an art installation, it can be understood as a *virtual* one, as the audience does not have to be physically present in a particular place to participate on it.

In its original conceptualization, MOTUS was devised with intertwined technical, poetic, and aesthetic ideas. However, musicians using MOTUS raised the demand for ways to change the musical and visual structures behind it, that is, to use MOTUS as a music composition tool [10].

Here, we present the fourth iteration of MOTUS, namely MOTUS 4, deployed online at https://tiagoft.github.io/motus/. Unlike its predecessors, it does not carry an inherent poetic or aesthetic idea. Rather, it provides a set of tools that allow composers to quickly devise, experiment, and share interaction configurations.

*Supported by CNPq.

The remainder of this article is organized as follows. First, Section 2 brings forward non-artistic applications for the interaction concept discussed in this paper. Then, Section 3 discusses the development steps proposed in this work. After that, Section 4 brings further discussions on possible future applications and developments using MOTUS, and Section 5 concludes the paper.

2. Music Applications

MOTUS is a highly customizable musical composing tool aimed at the video-based interaction with humans. It is inspired in making audiovisual interactive installations [1]. However, this same technology can be used in a diversity of other applications, as we discuss below.

Music has shown to have beneficial effects in Alzheimer patients [18] when used alongside regular therapies. These effects include higher engagement and happiness during treatment, and increased motor coordination and memory strength. Also, rheumatologic disease therapies can use feedback to improve control in daily activities, such as dancing or walking [19]. These two aspects are combined in MOTUS, which provides both a motion-guided feedback and a musical response. Thus, its sound and movement capture engines are potentially useful in these contexts.

Also, nowadays there has been an increase in the usage of video and music sharing social media [20], like TikTok [21] and Kwai [22]. We anticipate that MOTUS' underlying ideas can be used in interesting, valuable filters to transform home-made recordings. Then, the transformed audio-powered videos can be shared with online communities for entertainment, content, or advertisement purposes.

Finally, we highlight that many locations are currently leaning towards sustainable development and becoming smart cities. This transformation calls not only for technical uses of data and sensing, but also specific types of interactive, permanent art exhibitions [23], which aim to improve wellness in urban areas. The MOTUS software is an important step towards making this type of art available even in less wealthy areas, as it uses ubiquitous hardware and it can foster a myriad of aesthetic ideas to be explored by local artists.

All of these applications have guided several transformations to the MOTUS' backend. These transformations aim at facilitating the customization, both by artists and developers. We discuss these technological decisions in the next section.

3. Motus Development

3.1. Technology stack

The first version of MOTUS was implemented using JavaScript vanilla with a Webkit framework, which provides functionalities regarding using the webcam and the client computer's audio system. This framework has particularities for each webbrowser, and, due to security reasons, its updates frequently change the API. This lead to a difficult maintenance code, and quickly the MOTUS software became malfunctioning and obsolete.

In this implementation, we use another framework, p5.js [24], which provides similar functionalities as those initially provided by Webkit. P5.js, however, has a more stable API, which is maintained by the Processing Foundation [25] and caters to a large community of artists, programmers, and practitioners. P5.js also provides community-based support and a vast collection of example-codes and libraries.

Although p5.js allows for a greater code standardization, there are some browser-specific behaviors that must be accounted for. Different browsers might have diverse security concerns regarding the use of some features like video-cameras, microphones, audio reproduction. However, the current implementation of MOTUS, as far as our tests comprised, is not harmed by these particularities hence the system can be used in any major browser.

The application was deployed online using Github pages (https://tiagoft.github.io/motus/). This is especially interesting because it facilitates continuous updates and uses the high up-time provided by the service. Next, we discuss the new features implemented for this development stage.

3.2. New features and variations

The update in MOTUS described here implements new features that arise from decoupling the image processing and audio synthesis ideas from the aesthetic proposals. As seen in Figure 1, the new proposal does not provide an aesthetic proposal of its own; rather, it simply models the flow of information through elements that can be further changed as it becomes necessary.

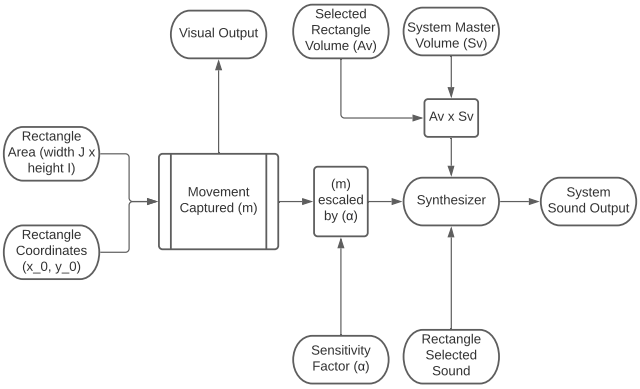

Figure 1: Block diagram for MOTUS' sound engine. Movement is captured from predefined rectangles within the screen and the amount of movement is used as a control parameter to the sound synthesizer.

MOTUS still relies on acquiring movement intensities in rectangles using a webcam video stream within

a web browser, as seen in Figure 2. This concept is important because it allows the software to be deployed via web and to be used with a hardware setup (webcam, laptop speakers, Internet connection) that is common for most computer music enthusiasts.

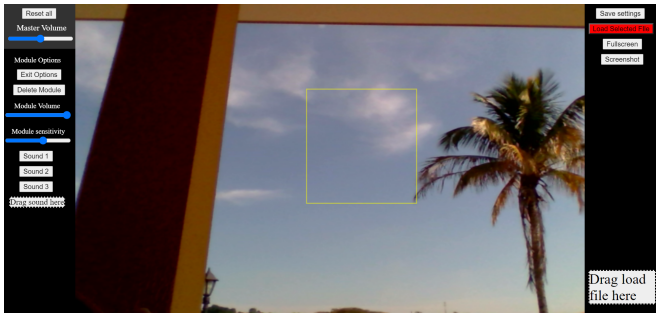

Figure 2: Screenshot of MOTUS' interface. On the left there are buttons to change general settings and rectangle's settings. On the right there are buttons related to the save and load engine, and to the screen settings.

However, its behavior was changed so composers can create, configure, and delete them. Each rectangle has an area (A), which is defined by its width J and height I, and is placed in the screen at coordinates (x_0, y_0). The movement within each area is calculated within each $I \times J$ rectangle at each frame t using the grayscale pixel values $p[i, j, t]$, as:

$$m = \frac{1}{IJ} \sum_{i=0}^{I-1} \sum_{j=0}^{J-1} |p[x_0+i, y_0+j, t] - p[x_0+i, y_0+j, t-1]|. \quad (1)$$

The movement amount m is highly impacted by characteristics of the webcam and by environmental factors such as ambient lights. Also, because of the averaging factor $\frac{1}{IJ}$, larger rectangles tend to reject more delicate movements. These factors are accounted for by allowing the user to configure a sensitivity factor α that scales m as desired or needed by the composer.

Also, we modified the audio/music synthesis algorithm so that it can be freely changed as to allow artistic explorations. In the current version, users can change the sound sample that is linked to each rectangle, as well as its volume. The composer can use one of the predefined sound samples or upload a sample of their own. The uploaded sample is stored in-browser, that is, it is not stored in any cloud servers, which avoids server liability in any copyright issues that might appear. These features can e seen in Figure 3.

A global setting for the master volume was also added, as seen in Figure 4, allowing a faster on-site configuration. All of these modifications allow the program to be used as a composition tool. This idea contrasts with the earlier 2015 *MOTUS*, in which the whole system was strongly linked to particular aesthetic choices such as the sound production behavior and the motion capture design.

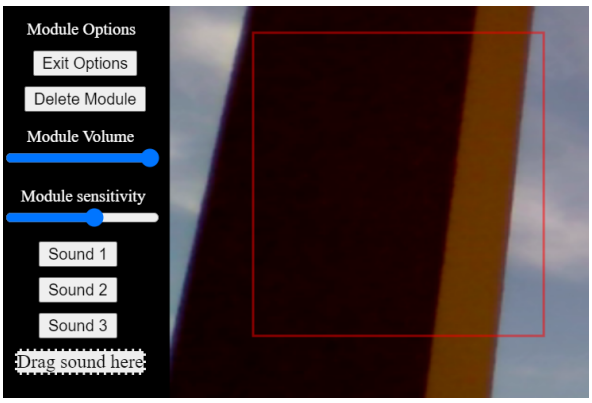

Figure 3: Screenshot crop showing MOTUS' rectangle settings configuration buttons.

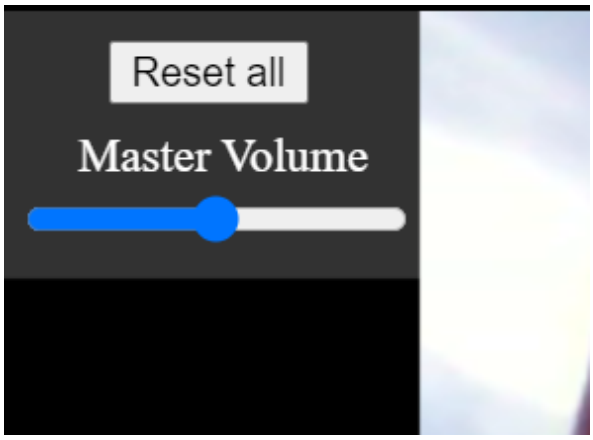

Figure 4: Screenshot crop showing MOTUS' general settings configuration buttons.

Furthermore, new minor components where added to improve usability, like a full-screen mode, a "reset all" functionality (Figure 4) that removes all rectangles, and a red flashing color within each rectangle that lights up proportionally to the calculated movement m serving as visual feedback, as seen in Figure 5. Lastly, we added a save and load option, so that users can export/import their creations as JSON files, allowing them to share their interaction compositions or save them for later use (Figure 6).

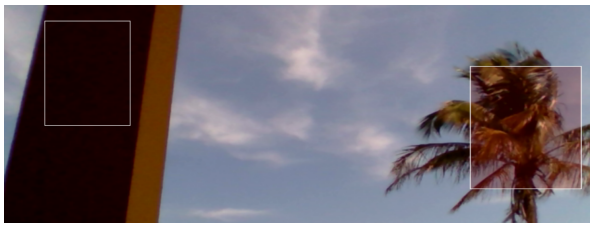

Figure 5: Screenshot crop showing MOTUS' visual feedback based on movement. On the left there is a rectangle without movement, on the right there is a rectangle with captured movement emitting visual feedback.

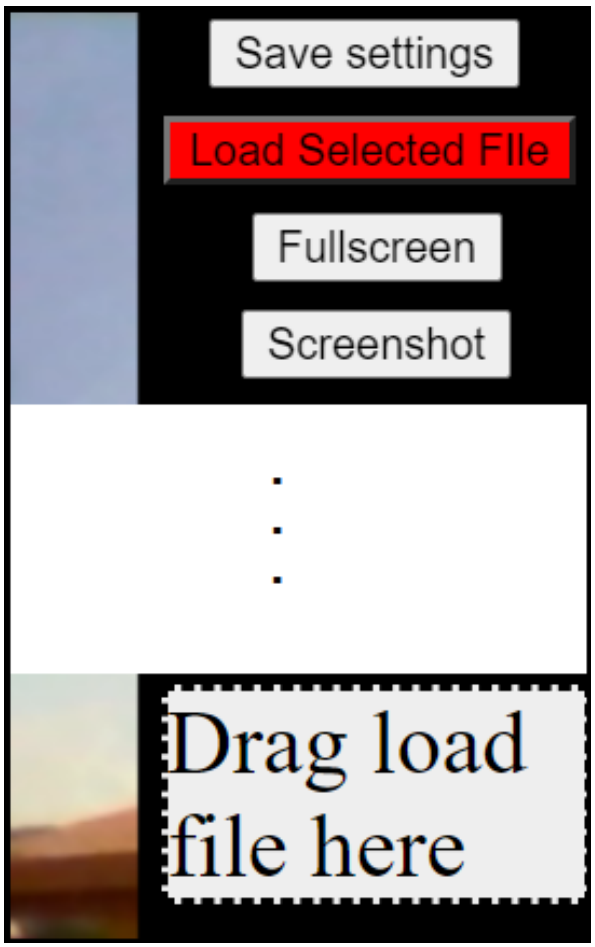

Figure 6: Screenshot crop showing MOTUS' save and load settings configuration buttons and a drag'n'drop area that allows users to upload their own previously saved files.

4. Discussion

In this work, we present a composition software based on a previous audiovisual installation [10]. The composition software untangles the aesthetic and poetic ideas from the technical one, thus allowing quickly composing and customizing a diversity of interactions. It does not require installing specialized software or understanding programming languages, which favors an easier learning process.

As such, we anticipate that composing interactions can be useful in several contexts, many of them diverse from art itself. Table 1 shows some possible applications that can be directly fostered by each of the newly added features.

We highlight that the creative use of each of the features can diverge from our initial idea. As an example, the customizable sound selection feature clearly fits the development of social media pieces and interactive installations, but can play a major role in therapy, as the sound can be changed to enhance the patients engagement.

An important aspect of this composition tool is that it is entirely deployed using the Internet, thus users only have to use a regular web browser. This is important because it prevents composers from having to install specialized software or learn complicated frameworks. Also, it allows the audience to access artwork without necessarily having to physically go to a specific location. Although not mandatory (as the artist can simply develop an installation and run it on a physical site), the virtuality of this interaction poses a vast field for developments and explorations.

Features	Applications
Save and Load	Social Media
Customizable Area Selection	Therapeutic Installations
Customizable Sound Selection	Social Media & Interactive Exhibitions
Visual Feedback	Therapeutic Installations
Modifiable Sound Synthesizer	Therapeutic Installations & Interactive Exhibitions
Sensitivity Modification	Therapeutic Installations

Table 1: New features added to MOTUS and their aimed applications.

Virtuality in artwork has become especially important during the COVID-19 crisis, which calls for social isolation, hence the use of telecommunication technologies has become necessary. Hence, Motus 4 comes as a timely tool to foster artistic and technical explorations during the pandemics.

5. Conclusion

This paper discusses new technical developments to the MOTUS engine, previously used in an audiovisual interactive art installation. It has been entirely ported to the P5.js framework, which facilitates its development. Also, we have added several new features, as requested by users in a previous study [10], from which we highlight a music composing tool that allows customizing the mapping between movements and audio.

The workflow of our tool comprises of defining rectangles in a videocamera stream from which motion is captured, and then mapping this motion to control customizable sound production agents. This allows to compose motion-based interactions that can be used for audiovisual installations, both virtual and on-site. The whole system executes within a web browser (`https://tiagoft.github.io/motus/`), hence it is not necessary to download and install any additional software in the host computer.

In the future, we will aim at collaborating with potential users of the engine to find new requirements. Also, we will evaluate the impact of using MOTUS and its underlying ideas in several contexts, including art installations, physical therapy, social media, and other possible applications.

References

[1] Clayton Mamedes. Participação em instalações interativas (in portuguese). *Revista Interdisciplinar Internacional de Artes Visuais*, 5(1), 2018.

[2] Raul Paiva and Tiago Fernandes Tavares. Impact of algorithmic composition on player immersion in computer games: A case study using markov chains. *Revista Música Hodie*, 18(1):61–73, 2018.

[3] Todd Winkler. Motion-sensing music: Artistic and technical challenges in two works for dance. In *Proceedings of the International Computer Music Conference*, 01 1998.

[4] Tiago Fernandes Tavares, Gabriel Rimoldi, Vânia Eger Pontes, and Jônatas Manzoli. Cooperative musical creation using kinect, wiimote, epoc and microphones: a case study with mindsounds. In *Sound and Music Computing 2015*, 2015.

[5] João Araújo, Avner de Paulo, Igino Silva Junior, Flávio Schiavoni, Mauro César Canito, and Rômulo Costa. A technical approach of the audience participation in the performance 'o chaos das 5'. In *Anais do XVII Simpósio Brasileiro de Computação Musical*, pages 28–34, Porto Alegre, RS, Brasil, 2019. SBC.

[6] S. Trail, M. Dean, T. F. Tavares, G. Odowichuk, P. Driessen, A. W. Schloss, and G. Tzanetakis. Non-invasive sensing and gesture control for pitched percussion hyper-instruments using the kinect. In *New Interfaces for Musical Expression - NIME 2012*, Ann Arbour, Michigan, U.S.A., May 2012.

[7] Damián Keller, Cláudio Gomes, and Luzilei Aliel. The handy metaphor: Bimanual, touchless interaction for the internet of musical things. *Journal of New Music Research*, 48(4):385–396, 2019.

[8] Filipe Calegario, João Tragtenberg, Giordano Cabral, and Geber Ramalho. Batebit controller: Popularizing digital musical instruments' technical development process. In *Anais do XVII Simpósio Brasileiro de Computação Musical*, pages 217–218, Porto Alegre, RS, Brasil, 2019. SBC.

[9] Cristiano Figueiró, Guilherme Soares, and Bruno Rohde. Esmeril - an interactive audio player and composition system for collaborative experimental music netlabels. In *NIME*, 2019.

[10] Tiago Fernandes Tavares. An interactive audio-visual installation using ubiquitous hardware and web-based software deployment. *PeerJ Computer Science*, 1:e5, 5 2015.

[11] Clayton Rosa Mamedes, Denise H. L. Garcia, José Fornari, and Jônatas Manzolli. Abstrações: An audiovisual installation based on motion recognizing. In *2011 4th International Congress on Image and Signal Processing*, volume 2, pages 586–589, 2011.

[12] Rômulo Vieira and Flávio Schiavoni. Fliperama: An affordable arduino based midi controller. In *Proceedings of Nime 2020*, 2020.

[13] Cristiano Severo Figueiró. *Desobediência sonora: selos de música experimental e suas tecnologias de sustentabilidade*. EDUFBA, 2019.

[14] NIME 2020. "installations" accessed jul. 15, 2021. https://nime2020.bcu.ac.uk/installations/.

[15] Mutek. "mutek virtual" accessed jul. 15, 2021. https://virtual.mutek.org/en/mutek.

[16] Ash Turner. "how many smartphones are in the world?" accessed jun. 23, 2021. https://www.bankmycell.com/blog/how-many-phones-are-in-the-world.

[17] zugara.com. "webcam penetration rates adoption" accessed jun. 23, 2021. http://zugara.com/webcam-penetration-rates-adoption.

[18] Shirlene Vianna Moreira and Francis Ricardo dos Reis Justi and Marcos Moreira. Intervenção musical pode melhorar a memória em pacientes com doença de alzheimer? uma revisão sistemática. 2018.

[19] Thiago Holanda and Paulo José Moté Barboza and Dayana Priscila Maia Mejia. Tratamento fisioterapêutico em pacientes acometidos por alzheimer: Referência bibliográfica. 2012.

[20] Wyzowl. "why video is exploding on social media in 2021" accessed jul. 14, 2021. https://www.wyzowl.com/video-social-media-2020/.

[21] Zhang Yiming. "homepage" accessed jul. 14, 2021. https://www.tiktok.com/pt-BR/.

[22] JOYO TECHNOLOGY PTE. LTD. "about us" accessed jul. 14, 2021. https://www.kwai.com/about/.

[23] Heather Shipman. Smart art for smart cities. In Miroslav Mateev and Panikkos Poutziouris, editors, *Creative Business and Social Innovations for a Sustainable Future*, pages 251–253, Cham, 2019. Springer International Publishing.

[24] Processing Foundation. "overview. a short introduction to the processing software and projects from the community." accessed jun. 07, 2021. https://www.processing.org/overview/.

[25] Processing Foundation. "homepage" accessed jun. 07, 2021. https://www.processingfoundation.org/.