

Audiovisual interactive artwork via web-deployed software:

Motus composes Homino-idea

Augusto Piato Amstalden, Tiago Fernandes Tavares, Anésio Azevedo
Costa Neto, Giovana Cardi Camarini

Universidade Estadual de Campinas; Instituto Federal de São Paulo | Brazil

Resumo: Muitas instalações de arte dependem de interações audiovisuais baseadas em câmeras, e isso geralmente requer hardware e software especializados. Consequentemente, as instalações audiovisuais são geralmente restritas a áreas mais ricas, nas quais o equipamento especializado pode ser adquirido e adequadamente hospedado. Em países com evidente desequilíbrio de renda ligado à localização, a restrição geográfica leva a uma restrição de audiência. Neste trabalho, apresentamos o desenvolvimento de uma ferramenta de composição para interações audiovisuais implantada na web no lado do cliente e não requer a instalação de software adicional. Simultaneamente, fornece feedback visual que pode ajudar o público a entender a experiência. Consequentemente, a ferramenta pode ser utilizada para compor interações audiovisuais que atingem um grande público via web. Exploramos ainda mais a ferramenta compondo a instalação audiovisual Homino-idea. Ela é inspirada nas interações entre humanos e o meio ambiente e pode ser exibida em locais de arte ou usada online.

Palavras-chave: instalação audiovisual interativa, hardware ubíquo, Processing, mata atlântica, paisagens sonoras

Abstract: Many art installations rely on camera-based audiovisual interactions, and this commonly requires specialized hardware and software. Consequently, audiovisual installations are usually restricted to wealthier areas, in which the specialized equipment can be afforded and properly hosted. In countries with an evident income unbalance linked to location, the geographic restriction leads to an audience restriction. 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. 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. We further explore the tool by composing the audiovisual installation Homino-idea. The installation is inspired by the interactions between humans and the environment, and can be either shown in art venues or used online.

Keywords: audiovisual interactive installation, ubiquitous hardware, Processing, atlantic forest, soundscapes

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 (MAMEDES et al., 2011). 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 (TAVARES; PAIVA, 2018). The link between the physical and the virtual worlds, provided by sensors and actuators, has enabled a diversity of interactive artwork, including augmented performances (WINKLER, 1998) (TAVARES et al., 2015), live audience participation (ARAÚJO et al., 2019), augmented musical instruments (TRAIL et al., 2012) (KELLER; GOMES; ALIEL, 2019), instrument development frameworks (CALEGARIO et al., 2019) (FIGUEIRO; SOARES; ROHDE, 2019), and art installations (TAVARES, 2015), (MAMEDES et al., 2011).

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 (VIEIRA; SCHIAVONI, 2020) (FIGUEIRO, 2019).

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 led to the development of several artworks exploiting virtuality and social isolation (NIME 2020, 2020) (MUTEK, 2021). 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 (TAVARES, 2015), 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 (WINKLER, 1998), and MOTUS employs this concept while deploying its software as a webpage, thus only requiring hardware that is becoming increasingly ubiquitous (TURNER, 2021) (ZUGARA, 2021). 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 (TAVARES, 2015).

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.

The remainder of this article is organized as follows. First, Section 1 brings forward non-artistic applications for the interaction concept discussed in this paper. Then, Section 2 discusses the development steps proposed in this work. The audiovisual installation Homino-idea is described in Section 3. After that, Section 4 brings further discussions on possible future applications and developments using MOTUS, and Section 5 concludes the paper.

1. Music Applications

MOTUS 4 is a highly customizable musical composing tool aimed at the video-based interaction with humans. It is inspired in making audiovisual interactive installations (MAMEDES, 2018). 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 (MOREIRA; JUSTI; MOREIRA, 2018) 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 (HOLANDA; BARBOZA; MEJIA, 2012). 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 on social media (WYZOW, 2020), like TikTok and Kwai. 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 (SHIPMAN, 2019), which aim to improve wellness in urban areas. The MOTUS 4 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.

2. Motus Development

2.1. 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 led to a difficult maintenance code, and quickly the MOTUS software became malfunctioning and obsolete.

In this implementation, we use another framework, p5.js, 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 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.

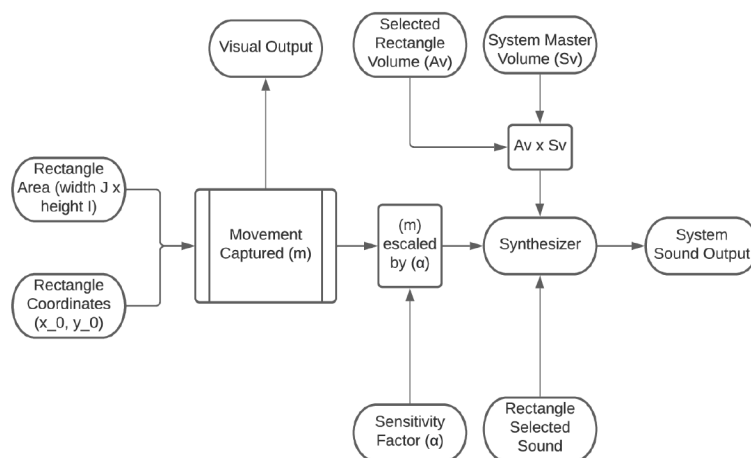
The application was deployed online using Github pages¹. This is especially interesting because it facilitates continuous updates and uses the high up-time provided by the service. Next, we discuss the new features implemented for this development stage.

2.2. 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.

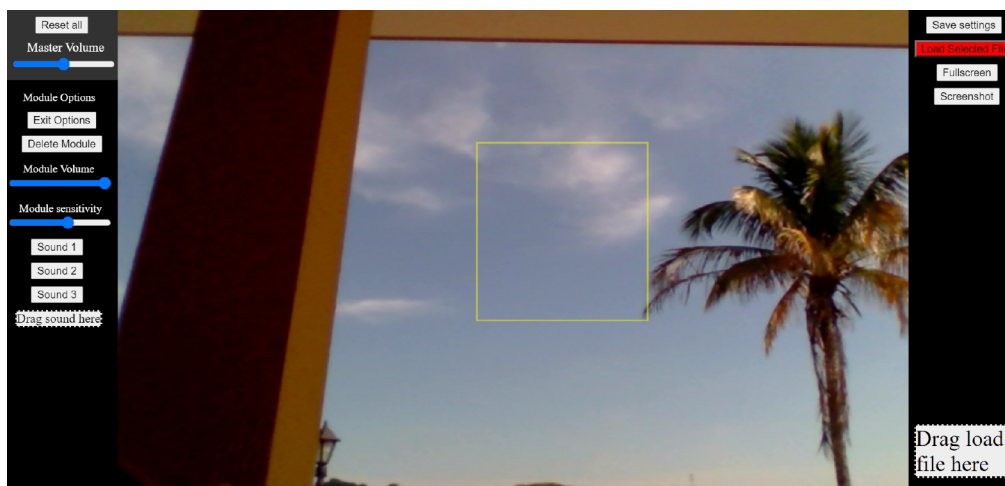
¹ <https://tiagoft.github.io/motus/>

FIGURE 1 – Block diagram for MOTUS' sound engine. Movement is captured from pre-defined rectangles within the screen and the amount of movement is used as a control parameter to the sound synthesizer.



MOTUS 4 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:

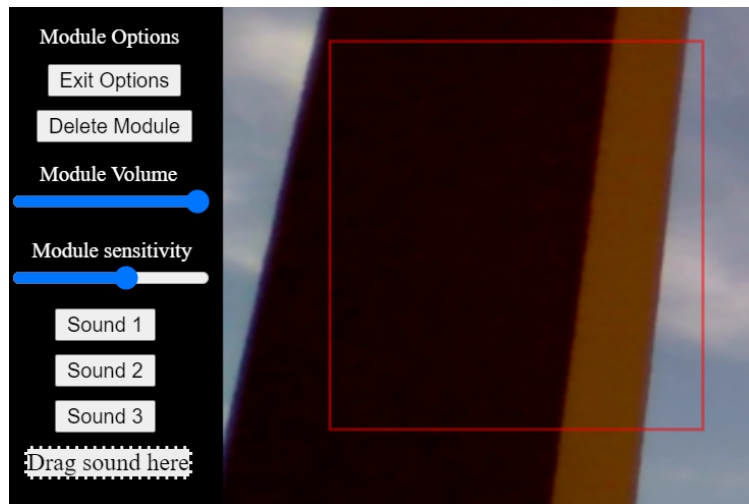
EQUATION 1 – Equation used by MOTUS' movement caption engine.

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

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.

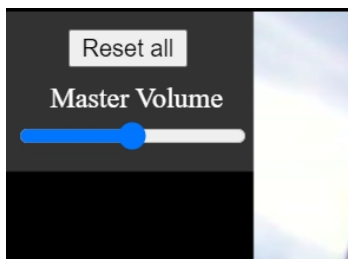
Also, we modified the audio/music synthesis algorithm so that it can be freely changed to allow artistic explorations. In the current version, users can change the sound sample that is linked to each rectangle, as well as its volume. The composer can use one of the predefined sound samples or upload 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 be seen in Figure 3.

FIGURE 3 – Screenshot crop showing MOTUS' rectangle settings configuration buttons.



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 4 – Screenshot crop showing MOTUS' general settings configuration buttons.



Furthermore, new minor components were 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.

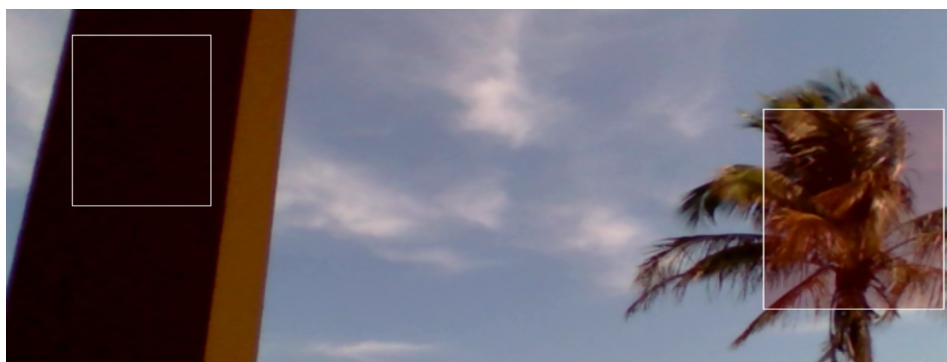
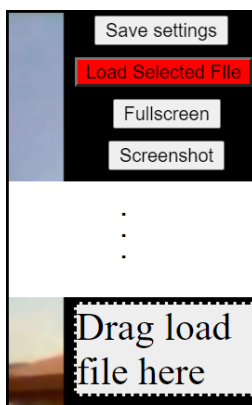


FIGURE 6 – Screenshot crop showing MOTUS' save and load settings configuration buttons and a drag'n'drop area that allows users to upload their own previously saved files.



3. Composing Homino-idea

Homino-idea was devised and constructed from a diversity of aesthetic and technical reflections. In this section, we bring forward these reflections. We start by discussing how Homino-idea provides an embodied listening experience based on the Atlantic Forest soundscapes. Then, we proceed to discuss the technical changes made on MOTUS 4 to fully implement Homino-idea's interactive art installation.

3.1. Embodied Soundscapes of the Atlantic forest

Unprecedented media access to local natural soundscapes around the world via streaming and field recordings is drawing people's attention to conservation efforts. As ecoacousticians have shown us, the more biophonic variations a soundscape has, the more complicated it sounds. Henceforth, from rainforests to savannas, each natural environment around the world has its own internal logic and coherent system.

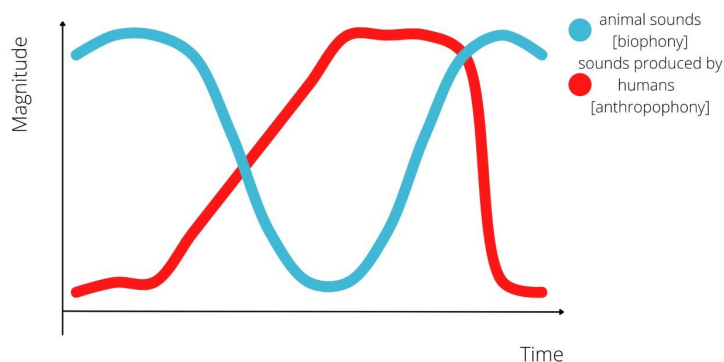
Homino-idea is a virtual sound art installation that invites the audience to an embodied listening experience via the sound of the Atlantic Forest. This South American forest stretches along the Atlantic coast of Brazil from northeast to southeast. This forest's fauna typically hides from humans, and, for this reason, they can only be heard in locus if the human visitor remains quiet.

We named the installation "Homino-idea" inspired by the fact that we, *Homo sapiens*, belong to the genetic branch *Hominoidea*, a primate superfamily from which other giant primates also descend. Species in this group share many common features in their auditory systems (STOESSEL et al., 2016), and generally perceive and interact with auditory cues from the environment. "Homino-idea" invites listeners to revisit their hominid roots, transcend the noise wall of modern cities, and listen to the soundscapes of the Atlantic Forest.

The creative core of Homino-idea is the interaction between sounds produced by humans and biophony (sounds produced by natural species in soundscapes). The sounds produced by humans were designed with a digital synthesizer, and we refer to them as "drones". The biophony primarily used field recordings of frogs (*anurans*), howler monkeys (especially *Alouatta guariba*), and a diversity of insects and small mammals to create the soundscape.

The embodied interaction in the installation relies on the movement detection algorithm used in MOTUS 4. As shown in Figure 7, when the camera captures movements, the drone sounds are triggered, blurring the forest soundscape. When there is no movement, the drones swell rapidly, and the natural soundscape slowly subsides.

FIGURE 7 – Illustrative diagram showing how the interaction between the triggered sounds (drones/anthropophony) and the natural soundscapes changes over time. The vertical axis, the magnitude, can be described as the energy released by each sound source.



This dynamic between the drone sounds and the natural soundscape is based on the *Acoustic Niche Hypothesis*, which states that animal voices are organized to avoid acoustic conflicts with other species. This hypothesis also supports the idea that animal species avoid singing when a persistent and disarming human sound source outshines their vocalization. This means that Homino-Idea emulates the acoustic disturbance caused by human sounds in a particular soundscape and its consequences for the health of the habitat to raise awareness for environmental conservation.

The installation virtually expels the audience from their urban homes and invites them to an auditory and physical experience inspired by the Atlantic Forest's sounds. Because it is deployed using MOTUS 4 (web software and ubiquitous hardware), it can be accessed from anywhere, without requiring physically going to a cultural center. Also, it requires the audience to remain self-aware of their own movements (as to avoid disturbing the virtual forest), thus becoming a physical experience, that is, providing a participatory, embodied listening experience related to the Atlantic Forest.

3.2. Technical changes

Motus 4 was developed within the p5.js framework, based on the Processing environment. Both the framework and the environment gather a large group of active users and developers, which facilitates finding solutions to problems that might appear. Also, the engine was programmed and developed using the model-view-controller (MVC) technique (FOOTE, 2022).

This technique combined with the used framework made MOTUS 4 a concisely partitioned code. Which means that any change can be made only modifying a single portion of the project. For example, the audio syntheses engine can be entirely modified by only editing a controller module called "HI_Synthesizer".

The first main modification applied into MOTUS 4 was precisely on its sound engine. The default behavior changed a single sound sample's characteristics, whereas Homino-idea requires a contrast between two sound samples.

The modified engine still uses the movement captured seen in Equation 1, although in a different way. The movement m is now used to determine the amplitude of the drone, as seen in Equation 2. The values of d_weight are weights added to the filter to either smoothen or increase the movement's effect, depending on the value of m .

EQUATION 2 – Equation used by Homino-idea's drone's synthesizer.

$$drone = (d_weight * m) + ((1 - d_weight) * drone).$$

This means that the drone's sound is directly controlled by the audience movement, mimicking the sound disturbance a human would create whilst visiting the Atlantic Forest. Figure 7 shows how fluid the drone's sound can be, as long as the audience's movement is equally fluid. This translates perfectly to the idea of total control over the disturbance caused in the environment.

The anthropophony sound behaves differently, it's not affected directly by the movement, but by the drone's states (defined by its amplitude). This means that the forest sound will act as a state machine, either rising, falling or maintaining itself, depending on the drone's amplitude value. This shows how the audience isn't directly controlling the forest behavior, but merely affecting it.

As such, the movement signal is filtered, delayed, and non-linearly mapped to define the volume of biophony sounds. For such, we tested many different functions, and, although we chose one, we could not find any objective reason that made one mapping strategy “better” than the other. The parameters were manually adjusted so that the audience could quickly listen to the forest again after a long period of moving. Although this does not correspond to the real-life forest, we understand that taking too long to respond can make the audience believe that the system is unresponsive, thus harming the experience.

Another important feature that was modified for Homino-idea was MOTUS' visual-feedback engine. While in MOTUS the visual feedback was made by the translucent color red appearing over areas where movement was captured, Homino-idea's feedback was made by flashing random black squares over the area. This effect was an aesthetic choice, one that we found to be able to give the notion of “glitches”, and disturbances of the ambiance. Figure 8 shows an example of this visual effect.

FIGURE 8 – Screenshot showing Homino-idea's new visual feedback engine. To start the feedback a short stimulus was given by a user that then left the scene. Throughout the image random black squares can be seen, indicating that movement was detected.



Even after a large customization process to incorporate the sound dynamics of Homino-idea, the fluidity of Motus 4 has been largely used in the project. This makes Homino-idea an extremely flexible arrangement, which in turn reflects the core of the project's ideas and inspirations. Homino-idea is hosted online at <https://hominoidea.motus.art.br/>

4. Discussion

In this work, we present a composition software based on a previous audiovisual installation (TAVARES, 2015). 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.

TABLE 1 – New features added to MOTUS and their aimed applications.

Feature	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

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 patient's 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.

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.

MOTUS 4's application in Homino-idea brings forth the applicability of ubiquitous hardware and software for audiovisual installations with a strong artistic component. Moreover, its development highlights how easily new features can be added to MOTUS 4.

5. Conclusion

This paper discusses new technical developments to the MOTUS 4 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 (TAVARES, 2015), from which we highlight a music composing tool that allows customizing the mapping between movements and audio.

The workflow of our tool consists of defining rectangles in a video-camera 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, hence it is not necessary to download and install any additional software in the host computer.

We used the MOTUS 4 engine in “Homino-idea”, a movement-reactive audiovisual installation based on the dynamics between humans and animals in the soundscape of the Atlantic Forest. We developed the installation using web-based software and ubiquitous hardware so that it can be experienced by audiences who do not reside in cultural centers. “Homino-idea” aims to provide an embodied listening experience of the soundscape, taking audiences from their homes into a virtual forest and raising awareness of the human impact on animal extinction and habitat destruction.

REFERENCES

ARAÚJO, João; DE PAULO, Avner; JUNIOR, Igino Silva; SCHIAVONI, Flávio; CANITO, César; COSTA, Rômulo. A technical approach of the audience participation in the performance 'O Chaos das 5'. In: *Anais do XVII Simpósio Brasileiro de Computação Musical*. São João del-Rei: SBC, 2019. 28-34.

CALEGARIO, Filipe; TRAGTENBERG, João; CABRAL, Giordano; RAMALHO, Geber. Batebit controller: Popularizing digital musical instruments' technical development process. In *Anais do XVII Simpósio Brasileiro de Computação Musical*. Porto Alegre, RS, SBC, pages 217–218, 2019.

FIGUEIRÓ, Cristiano Severo. *Desobediência sonora: selos de música experimental e suas tecnologias de sustentabilidade*. EDUFBA, 2019.

FIGUEIRÓ, Cristiano; SOARES, Guilherme; ROHDE, Bruno. Esmeril - an interactive audio player and composition system for collaborative experimental music netlabels. In *NIME 2019*, 2019.

FOOTE, Brian. *Model-view-controller*. Available at: <<http://www.laputan.org/pub/papers/POSA-MVC>>. Accessed on: 1 mar. 2022.

HOLANDA, Thiago; BARBOZA, Paulo José Mote; MEJIA, Dayana Priscila Maia. *Tratamento fisioterapêutico em pacientes acometidos por Alzheimer*. Referência bibliográfica. 2012.

MAMEDES, Clayton Rosa; GARCIA, Denise H. L.; FORNARI, José; MANZOLLI, Jônatas. Abstrações: An audiovisual installation based on motion recognizing. In *2011 4th International Congress on Image and Signal Processing*. Volume 2, pages 586–589, 2011.

MAMEDES, Clayton. *Participação em instalações interativas (in portuguese)*. Revista Interdisciplinar Internacional de Artes Visuais, 5(1): UNESPAR, 2018.

MOREIRA, Shirlene Vianna; JUSTI, Francis Ricardo dos Reis; MOREIRA, Marcos. *Intervenção musical pode melhorar a memória em pacientes com doença de Alzheimer? Uma revisão sistemática*. 2018.

MUTEK. “Mutek virtual”. Available at: <<https://virtual.mutek.org/en/mutek>>. Accessed on: 15 jul. 2021.

NIME 2020. “Installations”. Available at: <<https://nime2020.bcu.ac.uk/installations/>>. Accessed on: 15 jul. 2021.

PAIVA, Raul; TAVARES, Tiago Fernandes. *Impact of Algorithmic Composition on Player Immersion in Computer Games: A Case Study Using Markov Chains*. Revista Música Hodie, 18, 1, 61-73, 2018.

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

STOESSEL, Alexander; DAVID, Romain; GUNZ, Philipp; OBERHOFFNER, Tobias; SPOOR, Fred; HUBLIN, Jean-Jacques. *Morphology and function of neanderthal and modern human ear ossicles*. Proceedings of the National Academy of Sciences: 113:201605881, 09, 2016.

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

TAVARES, Tiago Fernandes; RIMOLDI, Gabriel; PONTES, Vânia Eger; MANZOLI, Jônatas. Cooperative Musical Creation Using Kinect, WiiMote, Epoc and Microphones: a case study with MinDSounds. In: *Sound and Music Computing 2015*, 2015.

TRAIL, Shawn; DEAN, Michael; TAVARES, Tiago Fernandes; ODOWICHUK, Gabrielle; DRIESSEN, Peter; SCHLOSS, W. Andrew; TZANETAKIS, George. Non-invasive sensing and gesture control for pitched percussion hyper-instruments using the Kinect. In: *New Interfaces for Musical Expression - NIME 2012*. Ann Arbor, Michigan, U.S.A., 2012.

TURNER, Ash. "How many smartphones are in the world?". Available at: <<https://www.bankmycell.com/blog/how-many-phones-are-in-the-world>>. Accessed on: 23 jun. 2021.

VIEIRA, Rômulo; SCHIAVONI, Flávio. Fliperama: An affordable arduino based midi controller. In *Proceedings of NIME 2020*, 2020.

WINKLER, Todd. Motion-sensing Music: Artistic and Technical Challenges In Two Works For Dance. In: *Proceedings of the International Computer Music Conference*, 1998/01.

WYZOWL. "Why video is exploding on social media in 2021". Available at: <<https://www.wyzowl.com/video-social-media-2020/>>. Accessed on: 14 jul. 2021.

ZUGARA. "Webcam penetration rates adoption". Available at: <<http://zugara.com/webcam-penetration-rates-adoption>>. Accessed on: 23 jun. 2021.

ABOUT THE AUTHORS

Augusto Piato Amstalden is both a Computer Engineering and Sound Engineering student at University of Campinas (UNICAMP). Although his work in audiovisual interactions is fairly new, he is determined to learn new skills and improve existing ones to maximize his knowledge about this work area. Fascinated by music for a long time, he has found a friendly face in applying his programming skills in a vast range of musically-related software developments. ORCID: <https://orcid.org/0000-0002-2248-5673>. E-mail: a213364@dac.unicamp.br

Tiago Fernandes Tavares has a BsC in Computer Engineering, a MsC and a PhD in Electrical Engineering, all issued by Unicamp. His research comprised digital signal processing, machine learning and models for human-computer interaction. ORCID: <https://orcid.org/0000-0002-9104-613X>. E-mail: tavares@dca.fee.unicamp.br

Anésio Azevedo Costa Neto was born in Ituiutaba, Minas Gerais, Brazil. He holds a Bachelor's degree in Philosophy from the Universidade Federal de Uberlândia (UFU), a Master's degree in Fine Arts in Contemporary Art and a Ph.D. in Extended Perceptual Fields at the Universidade de Brasília (UnB). His research interests include sound art, audiovisual media and technology, and the transformation of organic sounds recorded in the Brazilian Cerrado. ORCID: <https://orcid.org/0000-0003-4616-2243>. E-mail: Anesio.azevedo@ifsp.edu.br

Giovana Cardi Camarini was born in Votuporanga, São Paulo, Brazil. She is currently graduating in information systems at Instituto Federal de São Paulo - Campus Votuporanga. ORCID: <https://orcid.org/0000-0003-1516-6903>. E-mail: giovanacardicamarini@gmail.com