# *Common Spaces:*
# Multi-Modal-Media Ecosystem
# for Live Performances

## Luís Leite

ESCOLA SUPERIOR DE MEDIA ARTES E DESIGN; UNIVERSITY OF PORTO;
INSTITUTO DE TELECOMUNICAÇÕES

## Rui Torres

FERNANDO PESSOA UNIVERSITY

## Luís Aly

UNIVERSITY OF PORTO

**ABSTRACT**

*Common Spaces* is an interface for real-time media convergence and live performance combining media, applications and devices. A multimodal media ecosystem was designed to respond to the requirement of a specific performance — how to mix multiple applications into a single environment. This collaborative environment provides a flexible interface for performers to negotiate, share, and mix media, applications, and devices. *Common Spaces* is a framework based on interoperability and data flow, a network of virtual wires connecting applications that "talk" to each other sharing resources through technologies such as OSC or Syphon. With this approach, media designers have the freedom to choose a set of applications and devices that best suit their needs and are not restricted to a unique environment. We have implemented and performed with this ecosystem in live events, demonstrating its feasibility. In our paper we describe the project's concept and methodology. In the proposed performance we will use the Digital Archive of Portuguese Experimental Literature (www.po-ex.net) as a framework, appropriating its database assets, remixing its contents, as well as the techniques and methods they imply, stimulating the understanding of the archive as variable and adaptable. These digital re-readings and re-codings of experimental poems further highlight the importance of the materialities of experimental writing, integrating self-awareness in the modes of exchanges between literature, music, animation, performance, and technology.

**KEYWORDS**

multimodality; live performance; media ecosystem; interoperability.

**RESUMO**

*Common Spaces* é uma interface para convergência de média em tempo real e performance ao vivo combinando média, aplicações e dispositivos. Um ecossistema multimodal foi projetado para responder ao requisito de um desempenho específico - como misturar várias aplicações num único ambiente. Esse ambiente colaborativo oferece uma interface flexível para artistas negociarem, compartilharem e misturarem média, aplicações e dispositivos. *Common Spaces* é um *framework* baseado na interoperabilidade e no fluxo de dados, uma

rede de fios virtuais conectando aplicativos que "conversam" uns com os ou-
tros compartilhando recursos por meio de tecnologias como OSC ou Syphon.
Com essa abordagem, designers dos média têm a liberdade de escolher um
conjunto de aplicações e dispositivos que melhor atendam às suas necessida-
des e não estão restritos a um ambiente único. Implementamos e realizamos
este ecossistema em performances ao vivo, demonstrando a sua viabilidade.
Neste artigo, descrevemos o conceito e a metodologia do projeto. Na apresen-
tação proposta iremos utilizar o Arquivo Digital da Literatura Experimental Por-
tuguesa (www.po-ex.net) como *framework*, apropriando os recursos desta
base de dados, remisturando os seus conteúdos, bem como as técnicas e mé-
todos que eles implicam, estimulando a compreensão do arquivo como variável
e adaptável. Essas releituras e recodificações digitais de poemas experimentais
destacam ainda a importância das materialidades da escrita experimental, in-
tegrando a autorreflexividade nos fluxos e trocas entre literatura, música, ani-
mação, performance e tecnologia.

I.INTRODUCTION

This paper presents a framework designed for "transposing" spatial poetry into a multimodal media live performance. It attempts to explore the potential of interactive digital media within artistic contexts. *Common Spaces* was developed for an artistic performance articulating spatial poetry with the performative aspects of digital media. The performance "Untitled" was created by the art collective Retroescavadora, and was first presented at the bookstore Gato Vadio, as part of a series of interventions with the Digital Archive po-ex.net — "Arquivo vivo é Anarquivo!" [A Living Archive is an Anarchive!] —, and recently performed at ELO 2017 at Maus Hábitos, Porto. Our first concern was how to transpose spatial poetry into a multi-dimensional digital environment interrelating the performative, the visual, and the sound spaces in an expressive manner. Consequently, there was the need to search methods to combine and orchestrate these dimensions. Spatial poetry is described by Torres, Portela, and Sequeira as a "form of poetry based on intersemiotic processes in which various sign systems (visual, audible, verbal, kinetic, performative) and materialities (three dimensional, medial) are invoked and used in an expressive way" (2014). With this concept in mind we have imagined four distinct interaction scenarios, relating the performative space with the media spaces where texts were integrated and manipulated. To materialize this concept we have designed a media ecosystem providing a common ground for collaborative work supporting distinct creative applications. We have also developed a real-time work-flow environment based on the typical off-line media production where the resulting media from one application is used as the resource asset on a different one. The media designer can therefore decide the best tool to be used in each specific stage during the media production. However, live interactive media creation is typically based on a single environment, which might present limitations by the lack of features. Thus, we propose a framework based on interoperability, resource sharing, and media orchestration for real-time media mixing and generation. Setting up a collaborative environment based on diverse applications presents several challenges. Applications must be able to "talk" to each other and allow their resources to be shared without reducing the performance substantially. In the following sections we describe the conceptualization of the artistic performance "Untitled" (Fig. 1), as well its technological implementation.

**Figure 1**. Pictures from Untitled performance at Conde Duque in Madrid (2017).

This project further connects to Digital Humanities methodologies, acting on the intersection between computation and humanities, as well as focusing on the development of transferable tools and environments for collaborative work (Burdick *et al.* 2012). Digital media technologies extend human capabilities beyond analogue media (McLuhan, 1994). Emerging technologies are available to common users, promoting new social behaviors. They become the new media producers, as Lugmayr and Teras argue: "We transformed ourselves into a networked and knowledge based society" (2015). Big Data became an integral part of the human cultural behavior, but how can we deal with so much information? Data driven media, as stated by Lugmayr and Teras (2015), require a data-centric work-flow that depends on the content. They present a cross-disciplinary media centered approach towards the investigation of Big Data that focuses on the interplay between technologies, applications and media types. We extend this concept by presenting a flexible user interface to allow the performer to improvise with the data.



**Figure 2**. Dimension mapping diagram.

## II. CONCEPTUAL SYSTEM

In order to translate spatial poetry into a multidisciplinary collaborative environment that gathers physical and virtual spaces, we have developed the concept of *Common Spaces.* The common-space derives from the notion of common ground as the medium and the process of communication. It can be understood as a mutual understanding among interactors (Clark and Brennan, 1991) — as the iterative process of conversation for exchanging evidence between communicators —, as well as an interface. A successful exchange, however, demands coordination between both content and process. To coordinate the contents, both interactors must follow similar assumptions. They must establish a common ground based on mutual beliefs, or mutual knowledge. To coordinate the process they need real-time update and feedback. The interactor models the "meaning" in a collaborative interface based on successive approximations. As Brenda Laurel argues, the "interface becomes the arena for the performance" (1991). We have developed a conceptual system of coordinates in order to project the text into three spatial imaginary dimensions: visual, sound, and performative dimensions. The narrative was built based on four environmental spaces that represent a timeline (Fig. 2): from abstraction to object space, from text to physical space. Semiotic higher-level concepts were employed as references for the perceptors to interpret the meaning of the messages. For example, the word spacing in a printed text can be used as reference for a pattern of digital text. This system can be seen as an abstraction of the media model discussed by Lugmayr (2012), where smart ambient media contribute to the digital overlay over the real world. Therefore, we have searched for connections between real-world references and digital overlay perception, stimulating the perceptors' interpretation. Each space can be divided into the core principles of manifestation, experience, and physical/digital world. We have adopted a knowledge-based approach for the creation of each virtual environment providing clues for relating the virtual with the physical world. The narrative begins with the unformed text (abstract space), grows into objective representations (object space), proceeds with textual structures (text space), and evolves until its final physical appearance (physical space) (Fig. 3). Each stage is projected into the interconnected dimensions with distinct time-to-space mapping schemes: X – visual; Y – sound; Z – performative.

## III. DIMENSIONS

The three dimensions materialize the text-to-digital-space mapping, exploring the multi-modal-media inter-relations.

> (**X**) Visual dimension is projected onto two and three graphical planes. Visual spaces represent the distinct evolution stages of text (output), the animation in these spaces is used as input in other dimensions.

(**Y**) Sound dimension is characterized by reactive and performative sounds that are modeled by the actor in real-time and triggered by events from the virtual environment.

(**Z**) The performative dimension is characterized by the actor's expression while acting and the resulting performance in the virtual environment. The actor's voice and body movement as expressive means of manipulation (input) as well its representation (output) in physical/virtual spaces.

The actor interacts with the media environment with voice via microphone, with vision using a webcam, and with hands tracked by a Leap Motion device. Each interaction method presents distinct degrees of freedom.

## IV.ABSTRACT SPACE (GENESIS)

Abstract space represents the genesis of text and is characterized by a generative behavior. In this space all dimensions are interrelated: the visual dimension is generated according to features such as displacement, attraction, or oscillation; each feature is manipulated through the performative dimension generating sound.

(**X**) Visual: Particles lying on a 2D plane simulate living microorganisms with expressive behaviors. They are reorganized whenever the performer changes parameters, including their displacement through the attraction value.

(**Y**) Sound: A sound space that reflects the atomization of the acoustic phenomenon through density change or granular synthesis. Specific parameters are manipulated by signal processing techniques including filters, or resonators.

(**Z**) Performative: Fingers (touch). The performer models the visual dimension through interactive parameters presented in a multi-touch surface.
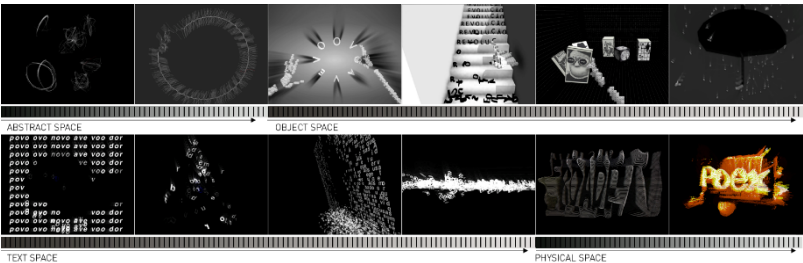


**Figure 3**. Visual dimensions: Abstract, object, text and physical spaces.

## V.OBJECT SPACE (PLAYGROUND)

In this space the performer plays with virtual interactive objects. It simulates a playground environment, similar to children playing with building blocks.

> **(X)** Visual: 3D dimensional text objects with physical properties are disposed in space for manipulation.

> **(Y)** Sound: Sounds are triggered when the performer's virtual hands touch the text-objects.

> **(Z)** Performative: The performer manipulates 3D text-objects as if they were toys in a playground searching for a meaning, embodying his interactions with the environment.

## VI.TEXT SPACE (STRUCTURE)

This space represents the structured text. Poems are presented in a 3D environment and deconstructed by the performer's hands. Each letter behaves as a particle with physical properties influenced by the environment reacting to the wind, gravity, and magnetic fields.

> **(X)** Visual: Textual structures from Portuguese experimental poems are presented, deconstructed and navigated in 3D planes.

> **(Y)** Sound: Voice interaction produces a futuristic "imaginary." Allied with sampling techniques, it attempts to recreate an Intonarumori (noise machine).

> **(Z)** Performative: Gestures transform the structure of the organized text, allowing the performer to destroy the structure, reorganizing its disposition in space.

## VII.PHYSICAL SPACE (RE-INTERPRETATION)

Printed spatial poems are the raw material for the manipulation and re-interpretation. The performer makes use of a webcam as an extension of his/her point of view. Fragments of captured text are then extruded from the paper plane as if they were a landscape of text.

> **(X)** Visual: The text is captured from the paper plane and reinterpreted through digital processes including video oscillators that control, offset and "z-displace" the video signal to achieve a 3D rasterized image.

> **(Y)** Sound: Glitch and malfunction aesthetics creating a sonic space with micro-failures. The glitch effect is achieved by continuous sine wave signal triggered whenever a percentage of white area is recognized.

 **(Z)** Performative: The performer operates a camera, trying to capture textual fragments inscribed on sheets of paper; his/her voice contributes to the transformation of text into another dimension. The volume of his/her voice is mapped to the "z-displace" extruding the images.

## VIII. MULTIMODAL INTERACTION

The multimodal interaction schemes present distinct ways to interpret and manipulate the text-space (Fig. 4a):

**Voice** (spoken text) – The human voice transmits emotions, and it is an expressive resource of textual interpretation. The microphone captures the voice, which is then segmented, processed and mapped to specific functions.

**Vision** (text selection) – Our vision allows us to read texts in many different ways. We are free to select where and how to read the text. The performer operates a web camera to select his/her point of view of the printed text. The image is then rasterized and extruded with "z-displace" based on the Rutt/Etra video synthesizer.

**Hands** (text manipulation) – We use our hands for writing, for manipulating objects or for speaking with gestures. We employed two different types of hand-based interfaces: a multi-touch device for control and a Leap Motion device for expressive manipulation. The performer can manipulate text-based objects with both hands, can write words with fingers and produce gestures representing words.
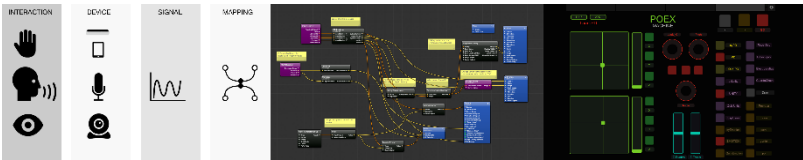


**Figure 4**. (a) Multi-modal diagram; (b) Quartz Composer patch; (c) iPad TouchOSC Interface.

## IX. COMMON SPACES FRAMEWORK

The *Common Spaces* framework provides the interface for mapping multiple inputs to multiple outputs in a collaborative environment. This multimodal and multimedia collaborative framework is characterized by 4 steps that describe the data flow process:

a)   **Input**: Performers interact with the system using hands, voice, and vision (point of view). Interaction is captured through the input device and transformed into digital signals.

b)  **Process**: Digital signal processing is applied to reduce the noise and make the data more useful; the segmentation procedure also helps to use this data in a meaningful way.

c)  **Map**: Mapping makes the interaction data available through all applications and provides the semantics for assigning the input action to the output behavior in the visual and sound dimension.

d)  **Output**: Each application shares its output, which is then mixed or combined into one or more outputs, or routed as input for another process.

## X. SHARING RESOURCES

Resource sharing is relevant in the conceptualization of a generic collaborative environment. Performers can continue working with their elected applications and share specific functionalities with the rest of the team. Today's computer performance lead to the development of local media sharing techniques that take advantage of dedicated hardware. Audio and video signals are routed through applications by virtual wires in a similar way as the patch cords transport the signals in the old analog synthesizers. Frameworks such as Syphon for Macintosh (Butterworth and Marini, 2010; Ingalls and Place, s.d.; Jarvis, 2014) or Spout (Butterworth and Marini, 2010; Jarvis, 2014) for Windows provide video and image sharing technology working directly on the computer graphics card for optimal performance. On the other hand, the most common audio sharing technology can be found in frameworks such as JACK (Butterworth and Marini, 2010; Davis and Letz, s.d.; Jarvis, 2014) for multi-platforms or Soundflower for Macintosh (Davis and Letz, s.d.; Ingalls and Place, s.d.). There are different approaches for exchanging control data, but the most popular is based on Open Sound Control (OSC) (Schemeder, Freed, and Wesser, 2010). This network protocol can control different types of media, applications and equipment and is widely supported. We combined these frameworks in a unique environment for media data flow.

## XI. ECOSYSTEM AND DATA-FLOW ORCHESTRATION

Our ecosystem is based on a set of applications and devices that are linked together by protocols that provide virtual connectivity (Fig. 4b). Two different types of data flow are used in our system: the media flow and the control flow. Contents such as video, images and audio are shared via the media flow. On the other hand, user interaction and control data are sent via the control flow. We adopted the OSC and MIDI protocols for data control and synchronization, while

Syphon and Soundflower were used as media sharing channels. The selected applications support those protocols and were divided into: a) control and media mixing, b) media generators, c) hybrid.

a)   The control and mixing applications were used to: orchestrate the media; control the cues; handle the mapping; media process and mixing. (Qlab, Osculator)

b)   Media generators were employed to respond to specific requirements of each space during the narrative: to generate sound, image, animation, or capture and process the video image. (Pure data, Unity3D, eMotion)

c)   Hybrid applications were employed to provide both functionalities: to generate and control the media. (Kyma, Quartz-Composer)

The performers operate a simplified graphical user interface (GUI) implemented on iPads using TouchOSC. We developed a custom interface for each space providing access to specific functionalities (Fig. 4c). The performer is able to orchestrate the media by pressing just one button, switching from one output to a different one or to jump a cue in the script receiving a visual feedback on the device. The script can instruct the system to close one application, and open a different one while remapping the video output. To facilitate the connection between devices and services we have used the Remote Strings Protocol (RCP). It provides auto-discovery and auto-negotiation features. It is available for Unity through the open source plug-in "Stringless" (Leite, 2016a). It is compatible with Openframeworks through ofxRemtoteUI, developed by Oriol Ferrer Mesià to control variables from a remote User Interface (UI) (Mesià, 2013). It can also be managed through "Pull The Strings" (Leite, 2016b), which is a marionette-programming environment. The abstract space was built with CHDH Egregore software for Pure Data (AKA, 2015). We have made adaptations to map hand gestures to sound parameters through Leap Motion, and to support custom MIDI messaging as well to support Syphon. The Text space was made with the digital creative tool "eMotion", by Adrien M. and Claire B (2013). eMotion is an experimental physically based animation system for real-time performance. The Physical space was developed inside Quartz Composer using the v002 Rutt Etra plug-in (Marini and Butterworth, 2008), which simulates the voltage-controlled video special effects systems and video synthesizers, originally developed by Steve Rutt and Bill Etra in the early 1970s.
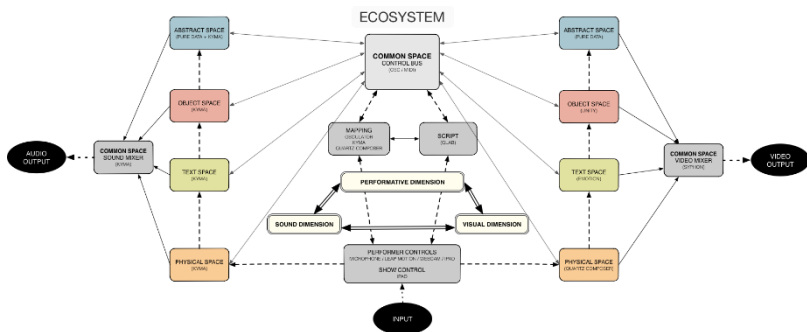
**Figure 5**. Multi-modal-media ecosystem diagram.

## XII. CONCLUSION

The *Common Spaces* framework was designed for transposing spatial poetry into a digital media live performance (Fig. 5). Several environments were created to represent the text in distinct media spaces using a multi-sensorial approach. Creative and collaborative performances require a media ecosystem that provides interoperability in a flexible environment for improvisation based on real-time media manipulation. *Common Spaces* provides a model for live data orchestration among applications and devices. It is based on multimodal and multimedia sharing, taking advantage of the available resources. This ecosystem provides a methodology for collaborative creative work. A custom graphical user interface can be customized using applications such as TouchOSC for mobile devices that accommodate specific functionalities for each performer. We believe that a generic and transparent interoperability protocol can replace OSC to support collaborative work incorporating video, audio, and control messages. Our system was successfully employed in a live performance showing its feasibility. This model can be useful for other live performances.

**REFERENCES**

AKA (2015). 30 Jun. 2017. CHDH Egregore. http://wwwchdhnet/egregore_sourcephp.
BURDICK, Anne, Johanna Drucker, Peter Lunenfeld, Todd Presner, and Jeffrey Schnapp (2012). *Digital Humanities.* MIT Press, Cambridge, MA.
BUTTERWORTH, Tom, and Anton Marini (2010). Syphon. Simplified BSD software license. 30 Jun. 2017. http://syphon.v002.info

CLARK, Herbert H and Susan E Brennan (1991). "Grounding in communication." *Perspectives on socially shared cognition.* Eds. L B Resnick, J M Levine and S D Teasley. American Psychological Association, Washington, DC, US, 127-149. 30 Jun. 2017. https://doi.org/doi:10.1037/10096-006

DAVIS, Paul and Stéphane Letz (s.d.) JACK – Audio Connection Kit. BSD license. 30 Jun. 2017. http://www.jackaudio.org

INGALLS, Matt and Tim Place (s.d.). Soundflower. The MIT License. 30 Jun. 2017. https://github.com/RogueAmoeba/Soundflower-Original

JARVIS, Lynn (2014). Spout. Simplified BSD licence. 30 Jun. 2017. http://spout.zeal.co

LAUREL, Brenda (1991). *Computers as Theatre.* Boston, MA: Addison-Wesley Longman.

LEITE, Luís (2016a). Stringless – Remote Control for Unity3D. 30 Jun. 2017. https://github.com/grifu/StringlessUnity

–––––––––– (2016b). Pull The Strings. 30 Jun. 2017. https://githubcom/grifu/Pull-The-Strings.

LUGMAYR, Artur and Marko Teras (2015). "Immersive Interactive Technologies in Digital Humanities: A Review and Basic Concepts." *ImmersiveME '15 Proceedings of the 3rd International Workshop on Immersive Media Experiences.* 31-36. 30 Jun. 2017. https://doi.org/10.1145/2814347.2814354

LUGMAYR, Artur. (2012). "Connecting the Real World with the Digital Overlay with Smart Ambient Media--applying Peirce's Categories in the Context of Ambient Media." Multimedia Tools Appl. 58.2: 385-398. 30 Jun. 2017. https://doi.org/10.1007/s11042-010-0671-3

M, Adrien and Claire B (2013). eMotion. 30 Jun. 2017. http://wwwam-cbnet.

MARINI, Anton and Tom Butterworth (2008). v002 Rutt Etra. 30 Jun. 2017. http://v002.info/plugins/v002-rutt-etra/

MCLUHAN, Marshall (1994). *Understanding Media: The extensions of man.* Cambridge, MA: The MIT Press.

MESIÀ, Oriol Ferrer (2013). ofxRemoteUI. 30 Jun. 2017. https://githubcom/armadillu/ofxRemoteUI.

SCHEMEDER, Andrew, Adrian Freed, and David Wesser (2010). Best Practices for Open Sound Control. 1-10.

TORRES, Rui, Manuel Portela, and Maria do Carmo Castelo Branco Sequeira (2014). "Methodological Rationale for the Taxonomy of the PO.EX Digital Archive." *New Literary Hybrids in the Age of Multimedia Expression: Crossing borders, crossing genres.* Ed. Marcel Cornis-Pope. Amsterdam and New York: John Benjamins Publishing Company. 42-55.