



HAL
open science

Examining the Design Process for 3D Interactions in Performing Arts: A Spatial Augmented Reality Cyber-Opera Case Study

Cagan Arslan, Florent Berthaut

► **To cite this version:**

Cagan Arslan, Florent Berthaut. Examining the Design Process for 3D Interactions in Performing Arts: A Spatial Augmented Reality Cyber-Opera Case Study. IEEE Transactions on Visualization and Computer Graphics, In press. hal-04938727

HAL Id: hal-04938727

<https://hal.science/hal-04938727v1>

Submitted on 10 Feb 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Examining the Design Process for 3D Interactions in Performing Arts : A Spatial Augmented Reality Cyber-Opera Case Study

Cagan Arslan and Florent Berthaut 

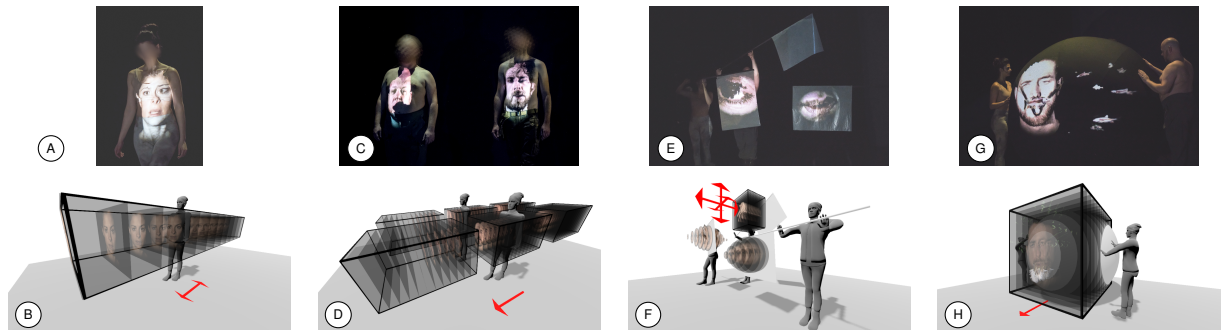


Fig. 1: Four 3D interaction sequences in Spatial Augmented Reality for the cyber-opera, with pictures from rehearsals on top and mockups showing virtual shapes, performers and 3D interactions on the bottom row : A-B) bi-directional audiovisual exploration of a video, C-D) unidirectional audiovisual playback of a dialog, E-F) 3D selection and manual playback of visual animations, G-H) 1D selection and automatic playback of a video. We show that the choices of gestures, virtual content and props were guided by the design criteria of expressiveness, efficiency, contextualisation and legibility.

Abstract— While 3D user interfaces are often designed with efficiency and accuracy in mind, artistic performances have their own very specific constraints and criteria for a successful interaction in mixed or virtual reality, which have yet to be fully understood. In this paper, we study the design of 3D interactions for a Spatial Augmented-Reality display in the context of a cyber-opera. We perform an analysis of design decisions taken during the multiple residencies and lab sessions, and we conduct a reflexive thematic analysis of interviews with the director and actors, which highlight how they integrated the story, the actors' performance, the audience experience and the technology. We identify four main criteria for designing 3D interactions in performing arts, namely their efficiency, expressiveness, contextualisation and legibility, and we derive guidelines for future research and creation.

Index Terms—performing arts, theatre, cyber-opera, spatial augmented reality, 3D interaction, extended reality

1 INTRODUCTION

The past decades have seen an increasing use of Extended Reality displays and 3D interaction in artistic performances (dance, theater, music ...) using technologies such as projection mapping (flat or with a fixed perspective), Pepper's ghost displays (e.g., optical combiners), Augmented Reality on large shared screens or mobile devices, and finally Mixed and Virtual Reality Head-Mounted Displays. The integration of these technologies entails design choices on the virtual content and interactions (gestures, techniques, devices). These choices can be very different from those made in other contexts such as industrial design, education, games or medical simulation, for which design guidelines exist [34].

However, while immersive performances have been regularly studied with a focus on audience experience [15, 36, 40] and technical constraints [24, 46, 52], little to our knowledge has been done on the design choices of 3D interactions. We believe that understanding the specifics of the performing arts domain in the design of 3D interaction is essential because it affects a growing number of users, both designers/artists and audiences, and can generate new challenges and

solutions for research in 3DUIs. Our goal is to provide insights and guidelines for both artists and researchers. To do so, in this paper, we study the integration of a spatial augmented reality interface in a cyber-opera. We observe the evolution of the design choices along two years of preparation, practice and performances, in particular looking at how the director combines feedback from the actors and other members of the company, SAR technological constraints, and story and characters constraints.

1.1 Related Work

The integration of interactive systems in performing arts, such as music, dance, theatre and opera, has long been studied.

1.1.1 Roles in artistic interactions

Public interfaces, especially in artistic performances and installations, involve many participants with various roles and trajectories between these roles [7]. For instance, Greuter *et al.* [24] define a number of roles in VR installations, covering the dimensions of agency and interest, including *passer-by*, *anticipator*, *orchestrator* and *appreciator*. The same work has been done on mixed-reality installations, for example by Fosh *et al.* [21] in an augmented sculpture garden. Some of these roles have been the focus of most research, namely those pertaining to the audience, who passively or actively attend performances, and to the performers, who rehearse, practice and play. Previous research has looked at ways to measure and understand the experience of a passive audience, using questionnaires and real-time feedback [3] but also physiological sensors [32], and at tools to enrich this experience during [15] and around performances [9]. Participatory performances have also been the focus of much research, in order to either provide

• Cagan Arslan is with CRISTAL, CNRS, Univ Lille. E-mail: cagan.arslan@univ-lille.fr.

• Florent Berthaut is with CRISTAL, CNRS, Univ Lille. E-mail: florent.berthaut@univ-lille.fr.

tools for actively engaging the audience [40, 50], or to understand this active experience [18]. The roles and trajectories of performers constitute the other frequently studied aspect of performances. Barkhuus and Rossito [4] study the process of rehearsing a play that integrates multiple technologies, providing insights on the trajectories of actors with respect to technology. Here we focus on the role of director, and their trajectories within the preparation of a performance. This aspect is sometimes tackled as a secondary aspect when studying actors [4], but we believe that it deserves a dedicated look in order to understand the main 3D interaction design decisions.

1.1.2 Interactive systems for performing arts

Opera has early on been a terrain for experimentation with interactive systems, with Ted Machover's *Valis* (1987), *The Brain Opera* [42], [37] or *Death and the Powers* [27]. These systems have been studied using in particular long term methodologies. For instance interactive dance performances have been thoroughly analysed by Latulipe *et al.* [31], looking at the temporal evolution of technology and highlighting how the technology and writing evolve over time [33]. Trajectories with the technology are further investigated by Bluff and Johnston [13] across time and multiple shows. They indicate how the technology and the relation between the technology and stakeholders (director, actors ...) evolve over time. Study of interactive systems has also produced design spaces for specific technologies, for example in the use of projection in theatre plays [47], or the strategies to cope with technological constraints and potential issues for mixed-reality performances [46]. Finally, much research has been conducted on the specific context and users of interactive systems in performances [41], in particular looking at the co-design between artists and designers. In the context of a theatre play, Tholander *et al.* [48] highlight how the "technology and narrative are co-constructed and shape each another" and how the roles of artist and designers intersect in the creation of the performance. Honauer and Hornecker [25] study the design process of interactive costumes and insist on the importance of involving artists as early as possible (during ideation) in the process of interactive systems for performances. Gonzalez *et al.* [23] provide guidelines for co-design in the context of dance performances. In this paper, we rather focus on understanding the criteria used for designing 3D interactions in order to inform future research and creation.

1.1.3 Extended Reality displays on stage

Originating from Antonin Artaud's "Le théâtre et son double" in 1936, the expression *Virtual Reality* is intrinsically linked to the illusion of objects and characters created in theatre performances. In fact, many theatre techniques, such as the Pepper's Ghost developed in the 19th century, were developed to create sensory illusions and, by relying for example on optical principles, prefigured and constitute the basis for modern Extended Reality displays now used on stage. Logically, digital immersive technologies such as head-mounted displays (HMD), stereoscopic displays have early on been employed to produce VR performances [19]. Extended Reality displays are now being used in all fields of performing arts, including theatre [22], opera [1], dance [36], and music [2] with a large variety of technologies and scenographies [52]. In this project, we focus on the use of Spatial Augmented Reality (SAR) [12, 39], meaning that the virtual content is projected on the physical surfaces of the stage and does not require performers or spectators to wear equipment such as HMDs or hold smartphones or tablets. SAR relies on projection on surfaces, such as walls, screens, floors, semi-transparent screens, fog or water screens [30], physical structures [13] but also on bodies [5]. It has been used in theatre plays [35, 38], in operas [54], for musical instruments [2, 51] and in dance performances [5], an example being the work of Klaus Obermeier¹. The displayed virtual content is either 2D, aligned with the physical surface it is projected on, 3D displayed with a specific perspective so that it works best from a sweet spot in the audience, or volumetric, *i.e.*, with slices of 3D content that appear when a physical surface moves through a virtual volume. For instance,

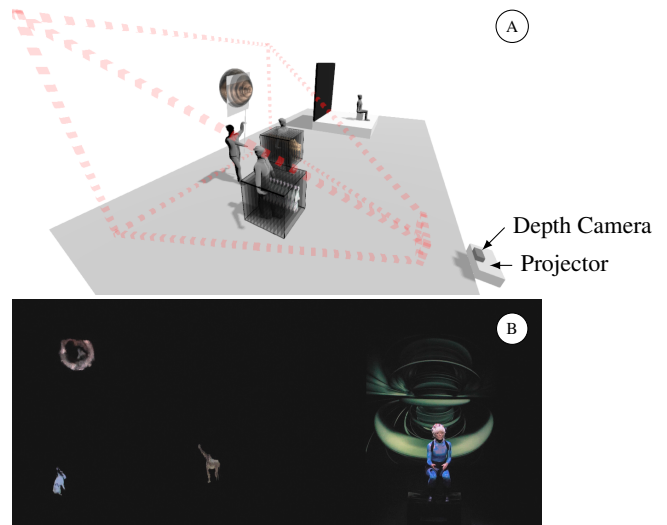


Fig. 2: The SAR display is used in the third act of the cyber-opera. Three shapes are revealed by the performers (two with their bodies, one with a flag) on the left part of the stage, while a fourth performer, the cyborg, seats in front of a large display. A) Schematic view from the left with, in dotted red, the volume in which virtual shapes can be placed, as the camera and projector frustum intersection. B) Front view from a performance.

Berthaut *et al.* [11] use depth cameras, projectors and an optical combiner to project slices of virtual objects in a shared space between stage and audience for musical performances.

1.2 Contribution

In this paper, we want to understand the process of designing 3D interactions in performing arts, through the study of a spatial augmented reality display in a cyber-opera. Our contribution is two-fold : 1) We propose a set of design criteria for 3D interaction in performing arts that can be used to inform creation and research 2) We provide a set of insights and guidelines to help this design process.

2 THE CYBER-OPERA

The cyber-opera *Terres Rares*² is created by the *Éolie Songe* theatre company and involves the collaboration between a director, a writer and a composer. There are four performers (actors / singers / acrobats) on stage and four live musicians (electronic drums, keyboard, flute, cello). The creation took place between 2019 and 2022, in a series of residences and was performed three times to the time of writing (in addition to partial performances after each residence).

2.1 Story and structure

The cyber-opera is set in the future. It recounts the expedition of three humans (one female, one male and one technologically-augmented male) and a cyborg, sent by a company to assess the presence of rare earth minerals in a distant (undisclosed) territory, after the company has lost track of an exploration robot, which fell down a rift. The cyber-opera is structured in three acts. The first act has a traditional operatic setting with limited use of technology on stage. The second act combines an operatic form with acrobatic sequences. At the end of act two, all human characters have either been killed, have died of frost or have disappeared, meaning that the actors playing these characters appear as ghosts or memories in the third act, an important context for the interaction design, as discussed further in Section 4.

¹<https://www.exile.at>

²<https://terrev.univ-lille.fr/>

2.2 Revealed Spatial Augmented Reality in the Third Act

In the third act, the opera takes a technological turn, as all events are now triggered by two AI entities, created by the fictional company to investigate the disappearance of the human characters, having retrieved the cyborg. The act starts with voices of the entities interrogating the cyborg and entering its memories. The entities then proceed to explore the memories looking for clues to incriminate one of the human characters for the fatal ending of the expedition, before finally creating a fake testimony for one character. While some memories are projected as video clips in the background above the stage, others are revealed interactively by the actors using our Spatial Augmented Reality (SAR) display and therefore appear at different positions on the left part of the stage as seen on Fig. 2, with the cyborg seated on the right side in front of a screen.

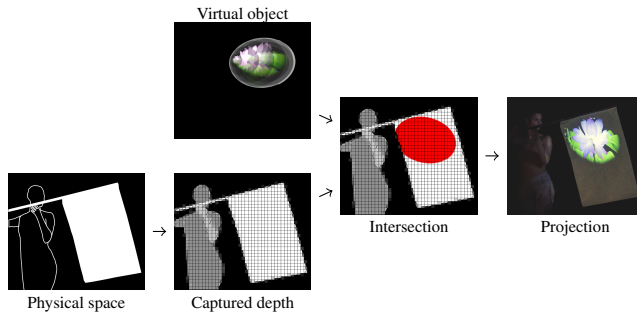


Fig. 3: Rendering pipeline of the SAR display. Here a sphere containing the frames of a video of a flower is revealed with a flag.

Our SAR technology relies on a common combination of projector and depth camera [28] and is implemented in the *Rivill* software³. The final chosen equipment are an Asus Xtion Live camera and a 1920x1080 (6000 lumens) laser projector. The display uses a volumetric approach of revealing/slicing virtual objects in the physical space, as explored notably by Cassineli *et al.* [17], Ben Guefrech *et al.* [6] or Berthaut *et al.* in an artistic application [11]. Virtual shapes (boxes, spheres, cylinders) are placed at various positions on stage. Choices for the primitive type, their placement and dimensions are discussed in detail in Section 3.2. These shapes are assigned an internal content, which can be a single 2D image, a single video, one or several sentences, separated frames of a video, a 3D texture and so on. When actors intersect a virtual shape with their body or props (captured by the depth camera), the intersection is reprojected in the physical space on the body or prop. They therefore "reveal" a slice of the shape, including the internal content and sometimes the outline of the surface.

Our "Revealed" SAR rendering pipeline is similar to the one described by Berthaut *et al.* [10]. Contrary to them, because the virtual objects are all primitives (boxes, spheres ...), we rely on signed distance functions to reduce the number of rendering passes to two. As shown in Figure 3, we first render the depth mesh captured from the depth camera to a texture, storing for each pixel the distance to the virtual camera. We then render the scene of virtual objects. For each fragment of the virtual shape, we test if the corresponding fragment of the depth mesh belongs is inside the transformed primitive volume. If it is the case, the fragment colour is selected depending on the virtual shape content. In the opera, we use either 3D textures comprising the frames of a video arranged along the Z axis, 3D textures with sentences arranged at multiple depths or 2D textures showing a video. The interactions with a virtual shape (*i.e.*, is it revealed or not, what is the position of the centre of the revealed part, what is the extent of the revealed part), can also be transmitted to an external software using the OpenSoundControl [49] protocol. In our case, this enables playing back the sound of a video when entering or moving through the shape using a PureData patch that implements granular synthesis.

³<https://gitlab.univ-lille.fr/mint/rivill>

2.3 SAR interaction sequences

During the SAR sequences, the two AI entities are exploring the cyborg's memories in order to find evidence to incriminate one of the characters for the accident. In the four sequences finally chosen at the end of the design process (which we describe in Section 3.2), the three actors are dressed in dark costumes (see Figure 5.A) which turn into a screen when they open their arms. Their interactions consist in revealing audiovisual content inside virtual shapes using either their bodies or props such as flags. In doing so, they have both informative, *i.e.*, showing the accessed memories to the audience, and expressive, *i.e.*, injecting life into these memories, goals. The four sequences are precisely timed, synchronised with the music and other video displays. Figure 1 provides schematic views of the virtual shapes used in these final sequences, together with how they appeared during the performance.

Monologue : This sequence illustrates the first step in the recreation of a character from the cyborg's recorded data from audiovisual fragments. It consists of the actress interacting with a single long virtual box ($W=100\text{cm}$, $H=70\text{cm}$, $D=500\text{cm}$), as seen in Figure 1.B, starting at 1 meter above the stage and 1 meter from the front. It contains frames from a video of her character. What the audience see and hear is the face of the female character, animated by the back and forth movements of the actor, making a series of vocal noises (stutters, muffled vowels, consonants, grunts ...), as shown in Figure 1.A. The actor improvises, stopping to maintain a vowel or consonant or oscillating between two sounds, to highlight the process of rebuilding a legible memory. The whole sequence lasts 55 seconds, then the shape disappears.

Dialog : The second sequence illustrates the reconstruction of a more complex scene, *i.e.*, a dialogue between characters. Each of the two male actors interacts with 3 aligned virtual boxes (see Figures 1.C and 1.D). All boxes have the same dimensions ($W=80\text{cm}$, $H=64\text{cm}$, $D=120\text{cm}$), they are separated by a gap of around 30cm, they are placed at a height of 150cm and start at around 1 meter from the front of the stage. The two groups of three shapes are separated by a gap of 50cm. Note that the faces of the characters do not fill the full width of the shapes so that the horizontal gaps are visually bigger. The boxes contain frames from videos of sentences pronounced by the actors' respective characters. What the audience see and hear is a dialogue between the two male characters, where each face alternatively pronounces a clearer, although strangely articulated sentence. The sequence lasts 1 minute and 20 seconds.

Shapes : The third sequence depicts memories of various living organisms emerging spontaneously from the cyborg's data, despite the efforts of the AI entities to control the search for evidence. As shown in Figure 1.E and 1.F, during the sequence, virtual shapes appear at seven positions on stage : three at the front and two at the back at the same height of around 1.5 meters and two other at the back at a height of around 2.5 meters. For each position, three shapes are displayed successively, either spheres elongated across the Z axis or long boxes. They contain frames of short videos of body parts (hand, arm, eye, back), animals (pidgin, eagle, fawn, giraffe, rabbit, elephant, lion, fish) and plants (moving flowers, moving branches, time-lapse of plants growing). All shapes are 60cm wide, 50cm high and 150cm deep. What the audience see is the apparition of these animations at various positions above the stage, revealed by the actors' bodies and by a flag. The sequence lasts 2 minutes.

Speech : The fourth and last sequence illustrate how the AI entities finally choose to generate a fake speech from one of the characters in which he takes responsibility for the accident, turning it into sabotage in defence of the environment. A large inflatable balloon (approximately 1 meter in radius) is brought down from the ceiling to around 1 meter above the ground, as shown in Figure 1.H, and intersects multiple virtual boxes. What the audience sees and hear is a first a video of the character's face emerging from a black background, and speaking with either visual glitches or incrustations of plants, birds or fish evolving around his head, as shown in Figure 1.G. Finally, the balloon stabilises as the AI entities have finished generating the speech. A new virtual box appears, replacing the four previous ones. It contains the generated video, which is played back with the synchronised sound. At the end

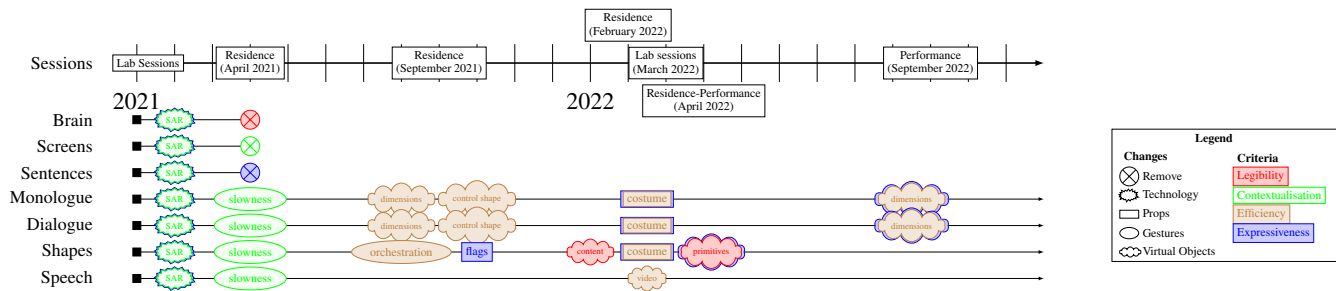


Fig. 4: Timeline of work sessions and design process with the main (not all), changes and design criteria for each of the interaction sequences.

of the video, the box disappears, leading to the final song of the cyborg. The balloon here is manipulated by a technician so that the content seems to appear on stage without any human intervention. This final sequence lasts 2 minutes and twenty seconds.

3 FIELD STUDY

Our goal is to understand how 3D interactions in the four sequences described in the previous section were designed, in particular how virtual content, gestures and interactions were chosen. To do so, we used a performance-led "in the wild" research approach, by following the tests, rehearsals and performances of the cyber-opera *Terres Rares* for more than two years. Rogers *et al.* [45] define research in the wild as *involving designing, prototyping, implementing and studying technology in situ, outside of the laboratory, and with the goal of articulating how it affects people's everyday experiences and appropriation practices*. Benford *et al.* [8] define it as *a set of explorations concerned with collaborating with artists to create cultural experiences that are deployed "for real", ... , to be experienced by the public who should see them as cultural artefacts rather than laboratory prototypes*. To synthesise, the aim of research in the wild is to study the relationship between a technological idea and its real world application.

3.1 Data collection

The design process went through many iterations and testing across both lab sessions and three main residences, depicted in Figure 4. Lab sessions involved the director, accompanied by either one or several technicians, dressmaker or prop master. During the residences, tests of the sequences were performed most of the time with the actors, but sometimes also with assistants. There were four main residences which took place in April 2021, September 2021, February 2022 and April 2022. The first three residences were followed by public presentations of at least the third act, and the fourth one was directly followed by the opening and the second full public representation of the cyber-opera. Another public performance took place in September 2022, and others are planned for 2025. Before and between these, 4 lab sessions took place : one in October 2019, two in March 2022 and one in April 2022.

During these residences and lab sessions, we documented the successive discussions, tests, and design choices through text notes, videos and pictures. In addition to the design logs, we conducted two interviews with the director, one in April 2022 and one in June 2022, and two interviews with each of the three actors during the April 2022 residence.

3.2 Analysis of the design process

After having assembled all notes from the residences and lab sessions, we extracted all design decisions by looking at four dimensions of the design of 3D user interfaces : tracking and display technology, physical props, gestures and virtual objects.

In addition to the four finally chosen sequences, three more were tested during the residences and discarded, as shown in Figure 4. The first relied on two performers moving through a 3D texture of the MRI scanner of a brain using a large mobile fabric screen. The second involved the manipulation of several semi-transparent fabric screens placed in a line so that the content revealed by the first screen was

duplicated on the screens behind. They were used to display the same content as the monologue sequence. The last removed sequence was composed of two lines of three shapes placed behind each other, similar to the dialogue sequence, but each containing half of a sentence. Two actors, one for each line, used paddle shaped props to display the sentences, moving backward and forward and therefore recomposing full sentences with changing meaning.

3.2.1 Choice of technology

The first design choice pertained to the selection of the SAR technology, instead of other mixed-reality displays such as headsets or 3D screens, and happened before and during the first lab sessions. The director favoured a technology that: 1) allowed him to place virtual content freely on stage without the constraints of a fixed screen, in order to adapt to the scenography; 2) did not require the audience to wear headsets or glasses, because of the large expected attendance of the cyber-opera; 3) enabled some form of expressiveness and embodiment in the way the virtual content was accessed, here by being revealed by the actors' bodies and gestures. During lab sessions, the depth sensing and projection hardware was tested in order to ensure both the legibility of the virtual content and a sufficiently large performance volume. Discussions with the team highlighted the need for a simple and stable technology which would be easy to replace in case of hardware failure and easy to set up, especially without the presence of the researchers. The first step was to adapt the projection volume to the interaction space, with a trade-off between the pixel density of the projection and the available volume for placing virtual objects. During lab sessions we started by experimenting with short throw projectors but quickly switched to narrower angle projector in the first residence to preserve the pixel density at the distance of 6 meters from the projector where some of the shapes started. This however meant that the volume for placing content was reduced in width. A laser projector was finally chosen that matched as closely as possible the field-of-view of the chosen depth camera and provided sufficient brightness. The choice of depth camera was based mainly on depth sensing quality. It was decided to rely on pattern-coded structured light, which produced a lower level of noise than stereo depth cameras, and therefore a more stable mesh for computing intersections with the virtual content. Although the resolution of the chosen depth camera (640x480) does not ensure a very accurate sampling, especially at long distances, this was not deemed an issue because large surfaces were used to reveal the content, *i.e.*, bodies and flags. Finally, we were expecting issues with latency of the system (30fps sensing, 60fps display), namely visible misalignment between props and projection, but the performers' movements were purposely slow in order to evoke ghosts / memories of their respective characters. The chosen technology therefore aligned with the context of the performance.

3.2.2 Choice of physical props

Some decisions on props were made to align them with the context of the story. Because their characters have disappeared at this point of the performance, the actors are supposed to be very discreet, barely visible, so that the focus of the audience is on the memories that they reveal. However, they should remain visible as human "avatars" who

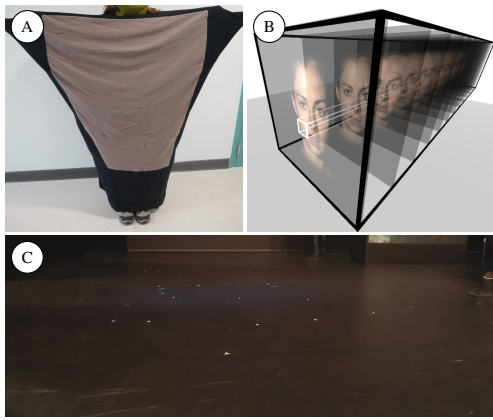


Fig. 5: Techniques developed to improve the efficiency in interactions : A) Costume designed to maximise the visibility of content while maintaining expressiveness. B) Smaller virtual control shape (in white) added inside the content shape (in black) in order to limit sound variations due to inaccuracy in depth sensing. C) Pieces of tape on stage that mark out either the centre or extremities of shapes.

interactively replay the memories. Actors were first using their bodies, either naked or with a skin-tight tops. Tests were then made with large pieces of fabric held in front of them. Because these hid the bodies too much, it was finally decided to create costumes, shown in Figure 5.A that would preserve a visible human shape. These are black on one side, so that actors can remain almost invisible when they are not interacting, and they transform into a large projection surface when the arms are extended. Similarly, the large black balloon used in the final sequence to display the fake speech was chosen in order to switch from human/organic manipulations conveying memories to a more abstract synthetic sequence, giving less importance to the actors. Other decisions were made to increase the expressiveness of the sequences. In the *shapes* sequence, a light and slightly translucent material was chosen for the flags in order to provide the sensation of volume and the possibility of "poetic" manipulations by the performers, as seen in Figure 1.E. For the lower virtual objects, however, costumes were then preferred as they increase the dynamic range of movements. The choice of costumes over flags was also made for efficiency reasons. In fact, flags were too inaccurate for the selection of the various frames of videos along the Z axis, because of their lack of rigidity. The same reason guided the selection of the size and material of the various physical props, so that they matched the dimensions of the virtual content on the horizontal and vertical axes.

3.2.3 Choice of gestures

Many decisions regarding performers' gestures were made to ensure the efficiency of access and display of the virtual content. For instance, part of the movements became more orchestrated and written over time, notably in order to make collaborative interactions more fluid and to avoid occlusions of one performer's content by another performer. Performers would then rehearse for some of the sequences when and where they should enter the shapes. Interaction feedback and guides were also designed to increase efficiency. When projections happened on the costumes, it was difficult for performers to perceive them visually. They could rely on the auditory feedback from the played back sound in the monologue and dialogue sequences to understand where they were inside or between shapes. However, to remember the limits of the shapes, it was decided to use pairs of tape marks placed on the floor, identifying the beginning and end of the shapes. Some of them can be seen in Figure 5.C. Finally, in the *speech* sequence, the manipulations of the balloon by the three actors were replaced by a mechanism to lower the balloon and move it back and forth more efficiently. Other decisions were made to add expressiveness to the gestures. Performers were encouraged to avoid gestures and interactions that were too didactic,

and rather to find more playful, poetic gestures, for example playing with their breathing to alter the sound and visuals on the dialogue and monologue sequences. The lack of expressiveness in interaction was one of the reasons why the text sequence, deemed too didactic, was abandoned. The context in the story was also a prominent aspect of choices in gestures. It was decided to favour slow movements throughout the entire act, in order to reinforce the feeling that the actors play ghosts or avatars of their characters. The director also insisted on wanting the actors to find a reason within their character for the way they moved and interacted with the virtual content, in particular in the *shapes* sequence, leaving some amount of improvisation. Finally, the absence of actors, and therefore gestures, in the last sequence emphasises the artificial aspect of the content, purely generated by the AI agents.

3.2.4 Choice of virtual objects

The first criterion for virtual objects was to make their manipulation more efficient, *i.e.*, to ensure that the content was correctly presented to the audience. Their dimensions, in particular their depth, were therefore often increased, leaving black frames at the beginning and end of the content, to make sure that performers did not miss important frames when entering or leaving. The spatial placement of virtual objects was also tweaked so as to avoid occlusions between performers during interactions. In the *speech* sequence, the content of the four boxes was changed from static frames that had to be scrolled through, to videos playing automatically, because the mechanism used to control the balloon was not accurate enough to parse the content precisely. Fade-ins and fade-outs were added to help actors anticipate automated changes of content. The control of sound playback in the *dialogue* and *monologue* sequences was made more efficient by using the intersection with a narrow hidden control shape instead of the large visual boxes to limit variability in depth from the full body surface, as shown in Figure 5.B. Interactions were also made more clearly discernible by the audience, for instance by discarding content that did not feature variations, *e.g.*, between frames of a video, that were sufficiently strong. In the *shapes* sequence, in order to avoid blurry content due to too many frames being revealed over a short distance, video length was reduced to only one or a few seconds. Some choices also depended on a trade-off between legibility and expressiveness in interaction. In the *shapes* sequence, the primitive (box or sphere) and appearance of virtual objects was chosen to either produce a softer entering or movement in the shape, using a sphere primitive and an alpha gradient (see the eye in Figure 1.E), or to obtain a sharp entry that preserves all the frames in the animation, using a box primitive without gradient. In addition to selecting the videos, text and images from the story, the content was also made to match other displays presented on stage, and therefore the context, such as a brain scan animation behind the cyborg, and their placement was altered to match the lighting chosen for the act.

3.3 Thematic Analysis of Interviews with the Director

We conducted a thematic analysis of the two interviews we had with the director in pursuit of understanding the key points of the design process from their point of view. We adopted a reflexive approach to thematic analysis [14] to be able to interpret the data in the light of what we have witnessed during the residences. After transcribing the two interviews, the two authors coded separately then grouped the codes into themes without searching for a consensus of meaning [44]. We often interpreted the underlying meaning of the content to reflect the connections to what happened during the residences. The themes were then discussed and refined. The process resulted in 8,475 words of transcribed interviews, 118 codes and finally the five themes described below.

3.3.1 How the intangibility of the shapes fits the context

There is a strong coupling between the technological concepts and the world the cyber-opera tries to portray. From the first stages of the production, technologies "*should make sense in relation to the artistic practice that is implemented on the stage*". In the context of *Terres*

Rares, taking place in a time when humanity is at its dystopic end, augmented reality is a convenient tool to underline the fragility of the humanity. "Connected to the dramaturgy of the moment, virtual shapes are there to talk about the fragility of humanity in an admirable way." The connection between the virtual objects and the fragility of humanity is established through intangibility. Virtual objects are intangible; they are in a sense ephemeral, short-lived. They appear one moment and disappear the next moment; "the spectral, I am there, I am not there, ghostly side of the shapes" represent the humanity occupying only a minimal fraction of earth's history.

The timing of the appearance of virtual shapes is strongly connected to the separation between the context of different acts. The virtual objects do not intervene until the third act, where "we are not in a theatrical narration as we had in act one, nor at an operatic level as in act two, we are in a kind of path much more uncertain for the spectator himself who makes his imaginary journey." At that uncertain point, through the fragility of the humanity and the intangibility of the virtual shapes, "a completely poetic universe is created".

3.3.2 Integrating the technology efficiently on stage

The story of *Terres Rares* takes place in an era of humanity where everything is automated. The technologies on stage are chosen carefully to reflect this mechanical side of that world. For the directing staff, "there was this idea that act three is triggered by a button and that it is the machines that deploy act three." On stage there were multiple technologies; not only SAR but also some frequently used in contemporary performing arts, such as computer-assisted music or computer-assisted lighting. Thanks to the accumulation of the technologies, "the show is a machine, and it's a machine that runs the show". As a result, the relation of the actors to the stage changes; they are bound to the rules and the constraints of the machine. The synchronisation of all elements of the third act are achieved by sound cues, but the actors are also guided by the images. "The actors are animated by the images, they know when to go by the sounds."

On the other hand, the machine is bound to the physical constraints of the environment. As there are boundaries of the stage, there are boundaries of the SAR space, which means "the virtual objects have their own limit that is part of the scene". For instance, "a virtual volume makes it possible to exit the projection at the front and [] at the back." Boundaries are imposed by the technological limits of the depth camera and the projector because "there is less definition in the back of the stage and a clean cut of the virtual frame in the front of it." Considering the touring nature of spectacles, boundaries need to be adapted to different venues; the size of the stage and the distance to the spectators are the determining parameters of a clean rendering of the virtual content. "We could use this device in a room of 100, 300, 400, 500 people. Beyond 500, perhaps we need to enlarge the props and enlarge the images".

3.3.3 A trade-off between legibility and expressiveness

Next to the constraints of the venue, the properties of the physical medium on which the images are projected (the *revealer*) also play an important role on the legibility of the content. As the size of the revealed part of the virtual object is limited to the size of the physical medium, a larger virtual object is more legible, but it requires a larger *revealer*, that is why "we invented these flying squirrel costumes made it possible to enlarge the projection surface to specify the sharpness and to increase the luminosity." However, a larger virtual object is more luminous and the reflected light mildly illuminates the surrounding area. This is a problem when the actor needs to be hidden. "When we look at a more luminous surface, we see the user, for example Marie-Yvettes's shoes, and it spoils everything." In addition, a larger *revealer* is harder to manipulate; it slows down the interactions and the transitions between virtual objects. Wearable *revealers* can respond to the trade-off between the size and the manoeuvrability by morphing the projection surface according to the needs of the scene. The flying squirrel costumes "would be an object bigger than the body. We can play with that, we can close it halfway, we can disable the bottom part, we can work with it by opening and closing."

The flexibility in the use of *revealers* comes to support the dynamism of the scene. If the interactions are too ordered, the scene is deemed to be too didactic. For instance, "putting a rigid sign to make a shape or text appear is very legible for the spectator, but it is not very interesting." If the content is too easily revealed, virtual objects are indistinguishable from a simple projection, regardless of the underlying technology. For example, "the curtains are too didactic. We have to find a more fluid, more playful way to make the shapes appear" Actors' movements between virtual objects contribute to the dynamism of the scene as much as the form of the "revealers". "the flying squirrel costumes offer a possibility to the actor's body to move between several shapes and to play in a dynamic way." It's up to the actor to appropriate the technology to balance the legibility and the expressiveness; "the actor must tame the shape and be able to use it intelligibly".

3.3.4 The director as intermediary between the story, the technology, the actors and the audience

As the themes above demonstrate, for the cyber-opera to be successful, one needs to ensure the coherence between the context of the spectacle and the technology, to make sure that the constraints of the stage and the SAR space match and to guarantee the legibility of content before the opera is shown to the public. The director plays a crucial role in putting the pieces together, by taking multiple responsibilities during the design process.

Firstly, the director is the main technical referent after the researchers. They are the person who will try "to understand the limits of the device, to get used to it, to try to exceed them as much as possible with the people who implement it to find the best results." They will try to push the limits of the technology to explore the possibilities in expressiveness. The director uses their knowledge on the limits of the technology to guide the actors. Actors being used as *revealers* in certain conditions, the director is "in the best position to see what worked and what didn't". Especially, when interacting with virtual objects "intended to be audible and legible by the public", actors need detailed feedback on how to respond to cues and move their bodies. In our case, the director decided to "put marks on the ground to direct their movements". However, the limits imposed on the actors by the director does not mean that actors should not have liberty to explore. At some point, the director realised that they were "being a little too coercive and that", and they "needed to give a little freedom." Once the actors become autonomous, the director needs to validate "what the spectators of the first row and the back row see". In the end, putting the director at the intersection of everything is not an uncommon approach. According to them, "from the moment it's an art project, you need someone who controls the entire process if you want to have a higher artistic quality." In the case of a cyber-opera, this may be more apparent because of the novelties in the technology and story-telling. Still, the director does not consider themselves as the ultimate decision maker as "it is the collaboration that led us to a certain number of tests, to a certain number of conclusions, to refuse some things and to accept others."

3.3.5 Co-appropriation by director and researchers

While the researcher and the director are exploring new ideas, they "are in the process of creating a new vocabulary", which means that if they find themselves in a similar situation, they could reference the previous steps to find a common ground. This "vocabulary of intelligibility" also applies to the other members of the project, such as the actors. For instance, "we experimented differently with each virtual shape" using this common vocabulary, and once the actors appropriated the vocabulary, "they knew what they could or cannot do and manage on their own, like grown-ups." At that "state of mutual understanding and trust", the team may "tackle any other project without any problem".

When the researcher is added to the "classic process of theatrical experimentation", the director and the researcher depend on each other to appropriate and to accelerate the new working method. For the director, this involves being "more and more aware of the possibilities the machine offers." For the researcher, it is mainly about "sensing more and more what the director is looking for". The major acceleration arrives when both parties manage to work separately and combine the

result of their individual work. In later experiments, when *"we got together, it worked with very little adjustments, it means that there is a predictability from both the researcher and the artist."*

3.4 Thematic Analysis of Interviews with Actors

We conducted semi-structured interviews with the three actors who interacted with the SAR system, during the second residence, both during rehearsals and after a first public representation. We asked them about their experience with the system. We followed a similar procedure than with the director, resulting in 3,883 words, 55 codes and finally the four following themes that allow us to understand how they perceived the 3D interaction. This analysis is more concise than that of the director and serves to refine the previous findings.

The SAR technology fits the context of the opera : Actors commented on how the technology fitted the context of the story, *i.e.*, how it allowed for *"revealing mental and sensible realities not visible to the eyes"*, for *"accessing a dream or a thought that becomes reality"*. In doing so, they also acknowledge the consistency between the chosen SAR system, which *"strangely makes the image both dreamed and tangible"* and the displayed content that in the story originates from memories and evocative images.

Expressiveness comes from the dialog between the body and the virtual content : Actors highlighted that it is their interaction with the technology that *"make things very sensitive"*, that *"makes this universe exist"*. More specifically they perceive a dialog between their body and the images, where *"the image constrains the body"* and at the same time where *"the image and sound follow the rhythm of my improvised corporeal rhythm"*. To them there is something *"extremely sensual that can be developed, something very concrete"*. The system therefore constitutes an (musical) instrument for the body, *"a type of harp really. There is something with all these potential strings, as if we had slices with which we can play, moving forward and backward"*.

Learning allows going from technical usage to the expression of intention : Actors strongly insisted on the importance of practicing with the system in order to reach an expressive use. To them the interaction is at first *"very technical, and then you get over the technique"*. One of them first felt *"like I was serving the technique, so that it could have been me or someone else playing"*, meaning that they did not feel they could express themselves. But *"eventually it becomes like playing"* and with *"training we can appropriate it"*. Indeed after the performance they felt *"less awkward, less looking for something, it is smoother"*, also pointing out that they were *"probably more accurate on their intentions"*.

Legibility needs to be balanced with context and expressiveness : Finally, actors provided hints on legibility of the interaction and content for the audience. For instance, they expressed concerns on the quality of props (including clothes) to make the image more legible. They also insisted on the trade-off between legibility and the context, in particular in the case of *"the playback dialog that should be understandable but still possess the strangeness of a flow that it is not natural"*, in other words that it should not be *"too perfect, otherwise it would not be interesting"*.

4 FINDINGS

From our analyses, we define a set of interaction design criteria and we provide guidelines future performances.

4.1 Criteria for designing 3D interactions

In most 3D interaction contexts [34], the prevailing design criteria remain the efficiency of the design, in particular the time required to accomplish a given task, and the error rate. Other common criteria pertain to the user experience, judged through the lenses of usability or presence. In performing arts, often the goal of an interactive system lies rather in the interaction itself, *i.e.*, the succession of gestures and responses from the system, rather than in a final result, and involves both the user and the observers of the interaction. In the case of *Terres Rares*, the task in all sequences is to select and manipulate content so that it is visible for the audience. From a 3D User Interfaces perspective, they can be classified as 3D selection and manipulation techniques.

The selection part here corresponds to entering a virtual shape with one's body or prop, and the manipulation of moving through the slices in order to control what is visually and sonically displayed by each volume/shape.

From the analysis of the design process, we believe that four main criteria were used to assess how well the chosen interactions match the task, namely : contextualisation, expressiveness, legibility and efficiency. These criteria are far from being independent, so that the director is constantly adjusting the design in order to find the correct balance and to give the appropriate weight to each of them, depending on the experience that he wants to create for the audience. Therefore, it is essential to consider the criteria not as goals to reach at all times, unlike efficiency in other application fields, but rather as parameters that the director adjusts, managing tensions between them differently in each sequence. Figure 4 depicts when the major design decisions were taken for each of the interaction sequences.

4.1.1 Contextualisation

The main criterion for design appears to be the contextualisation of the interaction. In fact, all elements of the interaction are at various stages of the design process confronted to their relevancy with respect to 1) the moment of the story at which they are used 2) the roles of the performers' interaction 3) the placement of the content on stage, relative to the scenography. This criterion is less prominent in other fields such as industrial design or scientific visualisations, where the context might constrain some aspects of the interaction, *e.g.*, choice of familiar gestures or representations. Contextualisation of the interaction can be facilitated by associating the researchers throughout the writing and rehearsing process. If it is not the case, it can happen, including in some of the authors' own experience, that the interactive devices are either replaced by non-interactive content or simply discarded. It also benefits from multiple rehearsals of the full performance, which reveal how the interaction fits in the moment of the story.

4.1.2 Expressiveness

The second criterion is expressiveness, which corresponds to the *player freedom* defined by Jordà [29], that is the controlled flexibility, or amount of variability, in the execution of the task. In our case, it was ensured by choosing costumes, props and virtual objects that allowed performers to obtain a variety of results when controlling the temporal and spatial components of their exploration of the audiovisual memories. Expressiveness is in that sense strongly linked with the notion of appropriation [53]. Performers are in fact explicitly asked to create a personal interpretation in their interaction, both reflecting the context of the story and their role, but also improvising from a vocabulary of gestures that they build over time. Expressiveness can be facilitated by using props that increase the range and vocabulary of gestures, in our case flags and costumes, or by choosing content that offer more variability when manipulated (*e.g.*, combining shape variation and frame selection in the *shapes* sequence), increasing the gestures-to-result mapping complexity [26]. The search for expressiveness however needs to be balanced against other criteria such as efficiency. In the *dialog* sequence for example, a more efficient way of showing the content would be to directly play the video at the normal rate when actors enter the shape, while allowing them to control the play head might lead to inaccuracies but also gives them more flexibility. Expressiveness might also sometimes conflict with legibility, such as for gestures or content manipulations that are too subtle to be perceived by the audience. The director therefore has to choose a trade-off between an interaction that is either more didactic, *e.g.*, with a limited set of gestures, or more expressive. Finally, designing for expressiveness requires leaving room and time for appropriation by the performers and for adjusting the design. In the case of the cyber-opera *Terres Rares*, frequent changes in the sequences (even the day before the first performance), and the complexity of the opera (number of acts and sequences that had to be rehearsed) meant that actors had less time to explore the interactions, which in turn limited the opportunities to refine the design.

4.1.3 Legibility

Legibility refers to how clear the relation between the performers actions and the resulting visible (and audible) effect is. For instance, in the *shapes* sequence, it was decided to simplify both the content (simpler motions in the successive frames) and props (costumes instead of flags with less stable surfaces) in order to make the control of the animations clearer for the audience. Legibility is linked to the notions of transparency [20] or attributed agency [15] in the sense that it should provide a clear understanding of what the performers are doing and how it influences the perceived result. In the classification defined by Reeves *et al.* [43], improving legibility implies increasing the visibility of both actions and effects. As explained above, legibility must be weighted against expressiveness, which might be degraded if techniques are designed to be "transparent" for the audience, with simple one-to-one mappings between gestures and effects [26]. It should also preserve the overall atmosphere at a certain point in the story, which might require hidden actions or effects. The search for legibility highlights the importance of the director, whose trajectory successively have them take the roles of designer, performer and also audience member. This last role, also taken at various times by other members of the theatre company, helps anticipate how the audience members will perceive the interaction. However, it implies the combination of per-sequence rehearsals, and full rehearsals during which the interaction is seen as part of the full scenography.

4.1.4 Efficiency

The last criteria that we see guiding the design choices is the efficiency of interactions, that is how fast and correctly the desired result can be reached. In our case, the result can be displaying a particular image at a defined position on stage or playing back an intelligible sentence. Efficiency can be evaluated in an objective manner when looking at if the interactions were performed in the allocated time and with the expected result. In our case, the third act was precisely timed, with a synchronisation between the music and other displays on stage, so that for example the actors had to make sure that all sentences in the *dialogue* sequence were played before the end of the sequence, or that all shapes in the *shapes* sequence were shown at least once. Efficiency is mostly in conflict with the *expressiveness* criteria, which might require variability in controls and might lead to errors, and the *contextualisation* criteria, which can call for gestures, props or content that slow down or make the interaction artificially more difficult. Efficiency can be improved by reducing the degrees of freedom in gestures [34] or by reducing the risk of errors in the manipulations.

4.2 Insights and guidelines for artists and researchers

Based on the criteria defined above and our experience of the many residences and lab sessions, we propose the following insights and guidelines that can help both researchers and artists during the design of 3D Interactions for performing arts.

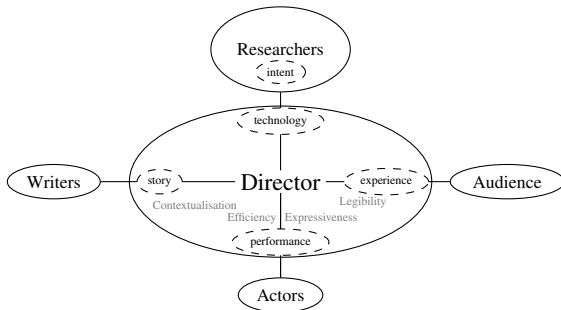


Fig. 6: The director internalises the other roles in order to make design decisions based on associated criteria. Symmetrically, researchers end up internalising the artistic intent.

4.2.1 Roles and temporality in the co-design process

Figure 4 depicts the moments when the major design decisions were made, which leads us to two observations : 1) all design criteria are used at all stages of the creation 2) large changes can also happen at any moment, such as changing costumes, especially when the director is aware of the capabilities of the technology, for instance changes in the dimensions and positions of the dialog and monologue were performed hours before the third full public performance. As discussed above and shown in Figure 6, the design choices made the by director are informed by their "trajectories between" and "internalisation of" the roles of performer, writer and audience member. This is particularly visible during residences, as the director spatially alternates between the stage and various positions in the audience in order to ensure the legibility and contextualisation of the interactions, and to illustrate expressive gestures. Regarding their relation with researchers, our experience confirms the results of previous studies on the importance of early rehearsals and tests [25] thanks to which the technology and writing can evolve together over time [33]. Finally, our collaboration with the director seems to have been reinforced by the *explication of both artistic and scientific goals of the project*. We therefore recommend that the goals of artists and researchers are clearly stated from the beginning of the collaboration (and regularly recalled). This will ensure that roles are understood by all. It will also facilitate the appropriation of the technology and research by the artists and of the artistic intent by the researchers, as shown in Figure 6, leading to researchers anticipating design choices and artists integrating constraints.

4.2.2 Specifics of Revealed Spatial Augmented Reality

Although we believe that these criteria and guidelines can be generalised to other 3D user interfaces on stage, some particularities of our display technology ("Revealed" SAR), which appeared during the interviews, might influence the director's design process. The first is the strong limitation, or even absence, of visual feedback, especially when performers are revealing content using the costumes, *i.e.*, directly on their bodies. Compared to 3D interfaces displayed on a stereoscopic screen or with a HMD, this absence of visual feedback led to specific design choices for efficiency, *i.e.*, resorting to tape on the floor and providing margin for error to enter and leave shapes. Relying on slices of virtual content instead of a full perspective rendering of virtual shapes also put constraints on designing for legibility. For instance, because the audience can not see the whole shape within which performers are moving, the changes in the displayed slices of content need to be strong enough so that the effect of the actors' actions are clear. On the other hand, the technology allows for placing content freely on stage, as long as the performers are able to reach it and reveal it, compared to other SAR systems such as projection mapping. This freedom means that the director is able to use the full space of the stage, and that the expressiveness can benefit from large scale mid-air movements. We therefore advocate for the use of this specific technology, which : 1) although similar to more standard on-body projections, offers more flexibility in the placement of and interaction with virtual shapes on stage; 2) can easily be combined with 2 dimensional projection mapping, stereoscopic displays and HMDs.

4.2.3 Design for legibility and expressiveness

3D interactions should be designed with controllable legibility for observers. This can mean providing parameters to adjust the visual feedback, the amplitude and nature of gestures, the size of physical props in order to make both actions and effects more easily perceivable and understandable by the audience. One option is to design following a level-of-detail approach as proposed by Capra *et al.* [16] in the case of musical performances. During residences, these levels of detail can then be easily selected by the director in order to estimate their impact on the audience. During performances, they can either be fixed or even changed dynamically as a way to switch between magical and expressive interactions [43].

The degree of freedom in interaction should also be quickly adjustable. This is important in order to both ensure that important content is always displayed to the audience at the correct time and place, but

also that there is room for interpretation for the actors. As mentioned before, efficiency and expressiveness are sometimes in conflict but not necessarily, so that instead of a continuum, multiple parameters can be defined which affect one or both of them. Efficiency can be increased by integrating disambiguation techniques, reducing the degrees of freedom or relying on one-to-one mappings between gestures and visual/sonic output. Expressiveness can be increased by (re)integrating ambiguity (noise, autonomous variations in content, ...) in interaction as highlighted by the actors in their interviews, relying on multidimensional controls and many-to-many mappings. These settings can be changed during residences to evolve with the performers learning the system and in relation with their appreciation of how they are able to achieve the goals and still be able to express their intentions (as pointed out during interviews with actors).

5 CONCLUSION, LIMITATIONS AND FUTURE WORK

In this paper, we studied the design process of 3D interactions in the cyber-opera *Terres Rares* across more than two years of residences and lab sessions. Our results suggest that design decisions were made according to four main criteria namely contextualisation, expressiveness, legibility and efficiency, and that the director carefully weigh them in order to select gestures, props, virtual content and technological components. We believe that these criteria can be found in other performing arts that employ mixed or virtual-reality displays, especially if they involve some amount of storytelling. Our results also highlight the importance of the role of the director, whose trajectory allows them to internalise other roles involved, including the researchers. Our approach suffers from some limitations : 1) the thematic analysis process was performed only by the two authors and might have benefited from a longer process of merging a refining codes, from longer interviews with actors and from interviews with audience members; 2) the guidelines and insights that we propose should be confirmed through comparisons with other performances relying on 3D interactions with a variety of display technologies in order to differentiate shared and specific design decisions; 3) Among future work, we believe that the use of our specific SAR technology and its associated absence of visual feedback in the context of performing arts raises important questions regarding guiding techniques for actors and visual augmentations for the audience [15]. We also want to investigate design methodologies for expressive 3D interaction techniques that would rely on and refine our proposed design criteria.

ACKNOWLEDGMENTS

The authors would like to thank the *Éolie Songe* theatre company and the team behind the *Terres Rares* cyber-opera. This work was supported by the "Fédération de Recherche Sciences et Cultures du Visuel, CNRS 2052", by the Région Hauts-de-France through the "TerRev" Stimule funding, by the "Equipex+ Continuum" and by the IRCICA (UAR-3380).

REFERENCES

- [1] Composing an interactive virtual opera: the virtualis project. *Leonardo*, 35(3), 2002. 2
- [2] C. Arslan, F. Berthaut, A. Beuchey, P. Cambourian, and A. Paté. Vibrating shapes : Design and evolution of a spatial augmented reality interface for actuated instruments. In *NIME 2022*, jun 16 2022. <https://nime.pubpub.org/pub/xy1df39n>. 2
- [3] S. Astrid Bin, F. Morreale, N. Bryan-Kinns, and A. P. McPherson. In-the-moment and beyond: Combining post-hoc and real-time data for the study of audience perception of electronic music performance. In *Human-Computer Interaction-INTERACT 2017: 16th IFIP TC 13 International Conference, Mumbai, India, September 25–29, 2017, Proceedings, Part I 16*, pp. 263–281. Springer, 2017. 1
- [4] L. Barkhuus and C. Rossitto. Acting with technology: rehearsing for mixed-media live performances. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pp. 864–875, 2016. 2
- [5] A. Barnett. The dancing body as a screen: Synchronizing projected motion graphics onto the human form in contemporary dance. *Computers in Entertainment (CIE)*, 7(1):1–32, 2009. 2
- [6] F. Ben Guefrech, F. Berthaut, P. Plénacoste, Y. Peter, and L. Grisoni. Revealable Volume Displays: 3D Exploration of Mixed-Reality Public Exhibitions. In *IEEE Conference on Virtual Reality and 3D User Interfaces*. Lisbonne (Virtual), Portugal, Mar. 2021. Honorable mention. doi: 10.1109/VR50410.2021.00023 3
- [7] S. Benford, G. Giannachi, B. Koleva, and T. Rodden. From interaction to trajectories: designing coherent journeys through user experiences. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 709–718, 2009. 1
- [8] S. Benford, C. Greenhalgh, A. Crabtree, M. Flintham, B. Walker, J. Marshall, B. Koleva, S. Rennick Egglestone, G. Giannachi, M. Adams, et al. Performance-led research in the wild. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 20(3):1–22, 2013. 4
- [9] S. Benford, C. Greenhalgh, A. Hazzard, A. Chamberlain, M. Kallionpää, D. M. Weigl, K. R. Page, and M. Lin. Designing the audience journey through repeated experiences. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–12, 2018. 1
- [10] F. Berthaut, C. Arslan, and L. Grisoni. Revgest: Augmenting Gestural Musical Instruments with Revealed Virtual Objects. In *International Conference on New Interfaces for Musical Expression*. Copenhagen, Denmark, May 2017. 3
- [11] F. Berthaut, D. Martinez Plasencia, M. Hachet, and S. Subramanian. Reflets: Combining and Revealing Spaces for Musical Performances. In *Proceedings of NIME*, May 2015. 2, 3
- [12] O. Bimber and R. Raskar. *Spatial Augmented Reality Merging Real and Virtual Worlds*. 08 2005. doi: 10.1201/b10624 2
- [13] A. Bluff and A. Johnston. Devising interactive theatre: trajectories of production with complex bespoke technologies. In *Proceedings of the 2019 on Designing Interactive Systems Conference*, pp. 279–289, 2019. 2
- [14] V. Braun and V. Clarke. Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health*, 11(4):589–597, 2019. 5
- [15] O. Capra, F. Berthaut, and L. Grisoni. Have a SEAT on Stage : Restoring Trust with Spectator Experience Augmentation Techniques. In *Designing Interactive Systems (DIS)*. Eindhoven, Netherlands, July 2020. doi: 10.1145/3357236.3395492 1, 8, 9
- [16] O. Capra, F. Berthaut, and L. Grisoni. Levels of Detail in Visual Augmentations for the Novice and Expert Audiences. *Computer Music Journal*, 2021. 8
- [17] A. Cassinelli and M. Ishikawa. Volume slicing display. In *ACM SIGGRAPH ASIA 2009 Art Gallery: Emerging Technologies*, 1 pages, pp. 88–88, 2009. doi: 10.1145/1665137.1665207 3
- [18] T. Cerratto-Pargman, C. Rossitto, and L. Barkhuus. Understanding audience participation in an interactive theater performance. In *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational*, pp. 608–617, 2014. 2
- [19] S. Dixon. A history of virtual reality in performance. *International Journal of Performance Arts & Digital Media*, 2(1), 2006. 2
- [20] S. Fels, A. Gadd, and A. Mulder. Mapping transparency through metaphor: towards more expressive musical instruments. *Organised Sound*, 7(2):109–126, 2002. 8
- [21] L. Fosh, S. Benford, S. Reeves, B. Koleva, and P. Brundell. See me, feel me, touch me, hear me: trajectories and interpretation in a sculpture garden. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 149–158, 2013. 1
- [22] D. Gochfeld, C. Brenner, K. Layng, S. Herscher, C. DeFanti, M. Olko, D. Shinn, S. Riggs, C. Fernández-Vara, and K. Perlin. *Holojam in Wonderland : Immersive Mixed Reality Theater*. *Leonardo*, 51(4):362–367, Aug. 2018. doi: 10.1162/leon_a_01644 2
- [23] B. Gonzalez, E. Carroll, and C. Latulipe. Dance-inspired technology, technology-inspired dance. In *Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design*, pp. 398–407, 2012. 2
- [24] S. Greuter, F. Mueller, and T. Hoang. Designing public vr installations:. In *Designing Interactive Systems Conference, DIS '22*, 15 pages, p. 792–806. Association for Computing Machinery, New York, NY, USA, 2022. doi: 10.1145/3532106.3533454 1
- [25] M. Honauer and E. Hornecker. Challenges for creating and staging interactive costumes for the theatre stage. In *Proceedings of the 2015 ACM SIGCHI Conference on Creativity and Cognition*, pp. 13–22, 2015. 2, 8
- [26] A. Hunt, M. M. Wanderley, and M. Paradis. The importance of parameter mapping in electronic instrument design. *Journal of New Music Research*, 32(4):429–440, 2003. 7, 8

- [27] E. Jessop, P. A. Torpey, and B. Bloomberg. Music and technology in death and the powers. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pp. 349–354. Oslo, Norway, 2011. doi: 10.5281/zenodo.1178051 2
- [28] B. Jones, R. Sodhi, M. Murdock, R. Mehra, H. Benko, A. Wilson, E. Ofek, B. MacIntyre, N. Raghuvanshi, and L. Shapira. Roomalive: Magical experiences enabled by scalable, adaptive projector-camera units. In *Proceedings of the 27th annual ACM symposium on User interface software and technology*, pp. 637–644, 2014. 3
- [29] S. Jordà. Digital instruments and players: part i—efficiency and apprenticeship. In *Proceedings of the 2004 conference on New interfaces for musical expression*, pp. 59–63, 2004. 7
- [30] B. Koleva, I. Taylor, S. Benford, M. Fraser, C. Greenhalgh, H. Schnädelbach, D. Vom Lehn, C. Heath, J. Row-Farr, and M. Adams. Orchestrating a mixed reality performance. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 38–45, 2001. 2
- [31] C. Latulipe, E. A. Carroll, and D. Lottridge. Evaluating longitudinal projects combining technology with temporal arts. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1835–1844, 2011. 2
- [32] C. Latulipe, E. A. Carroll, and D. Lottridge. Love, hate, arousal and engagement: exploring audience responses to performing arts. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 1845–1854, 2011. 1
- [33] C. Latulipe, D. Wilson, S. Huskey, B. Gonzalez, and M. Word. Temporal integration of interactive technology in dance: creative process impacts. In *Proceedings of the 8th ACM Conference on Creativity and Cognition*, pp. 107–116, 2011. 2, 8
- [34] J. J. LaViola Jr, E. Kruijff, R. P. McMahan, D. Bowman, and I. P. Poupyrev. *3D user interfaces: theory and practice*. Addison-Wesley Professional, 2017. 1, 7, 8
- [35] J. Lee, Y. Kim, M.-H. Heo, D. Kim, and B.-S. Shin. Real-time projection-based augmented reality system for dynamic objects in the performing arts. *Symmetry*, 7(1):182–192, 2015. 2
- [36] J. Lomet, R. Gaugne, and V. Gouranton. Could you relax in an artistic co-creative virtual reality experience? In *ICAT-EGVE, joint international conference of the the 32nd International Conference on Artificial Reality and Telexistence & the 27th Eurographics Symposium on Virtual Environments.*, pp. 1–9, 2022. 1, 2
- [37] T. Machover. *Hyperinstruments: A Progress Report, 1987-1991*. MIT Media Laboratory, 1992. 2
- [38] M. R. Marner, S. Haren, M. Gardiner, and B. H. Thomas. Exploring interactivity and augmented reality in theater: A case study of half real. In *2012 IEEE International Symposium on Mixed and Augmented Reality-Arts, Media, and Humanities (ISMAR-AMH)*, pp. 81–86. IEEE, 2012. 2
- [39] M. R. Marner, R. T. Smith, J. A. Walsh, and B. H. Thomas. Spatial user interfaces for large-scale projector-based augmented reality. *IEEE computer graphics and applications*, 34(6):74–82, 2014. 2
- [40] D. Mazzanti, V. Zappi, D. G. Caldwell, and A. Brogni. Augmented stage for participatory performances. In *NIME*, pp. 29–34, 2014. 1, 2
- [41] M. J. Nicholas, S. C. Daffara, and E. Paulos. Expanding the design space for technology-mediated theatre experiences. In *Designing Interactive Systems Conference 2021*, pp. 2026–2038, 2021. 2
- [42] M. Orth. Interface to architecture: integrating technology into the environment in the brain opera. In *Proceedings of the 2nd conference on Designing interactive systems: processes, practices, methods, and techniques*, pp. 265–275, 1997. 2
- [43] S. Reeves, S. Benford, C. O’Malley, and M. Fraser. Designing the spectator experience. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 741–750, 2005. 8
- [44] K. A. R. Richards and M. A. Hemphill. A practical guide to collaborative qualitative data analysis. *Journal of Teaching in Physical Education*, 37(2):225–231, 2018. 5
- [45] Y. Rogers and P. Marshall. Research in the wild. *Synthesis Lectures on Human-Centered Informatics*, 10(3):i–97, 2017. 4
- [46] A. Rostami and D. McMillan. The normal natural troubles of virtual reality in mixed-reality performances. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pp. 1–22, 2022. 1, 2
- [47] D. Z. Saltz. Live media: Interactive technology and theatre. *Theatre Topics*, 11(2):107–130, 2001. 2
- [48] J. Tholander, C. Rossitto, A. Rostami, Y. Ishiguro, T. Miyaki, and J. Reki-moto. Design in action: Unpacking the artists’ role in performance-led research. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–13, 2021. 2
- [49] M. Wright, A. Freed, et al. Open soundcontrol: A new protocol for communicating with sound synthesizers. In *JCMC*, 1997. 3
- [50] Y. Wu, L. Zhang, N. Bryan-Kinns, and M. Barthet. Open symphony: Creative participation for audiences of live music performances. *IEEE MultiMedia*, 24(1):48–62, 2017. 2
- [51] X. Xiao, B. Tome, and H. Ishii. Andante: Walking figures on the piano keyboard to visualize musical motion. In *NIME*, pp. 629–632. Cambridge, MA, 2014. 2
- [52] V. Zappi, F. Berthaut, and D. Mazzanti. From the Lab to the Stage: Practical Considerations on Designing Performances with Immersive Virtual Musical Instruments. In *Sonic Interactions In Virtual Environments*. Springer VS, 2022. 1, 2
- [53] V. Zappi and A. P. McPherson. Dimensionality and appropriation in digital musical instrument design. In *NIME*, vol. 14, pp. 455–460. Citeseer, 2014. 7
- [54] Y. Zhang, Y. Shen, W. Zhang, Z. Zhu, and P. Ma. Interactive spatial augmented reality system for chinese opera. In *ACM SIGGRAPH 2019 Posters*, pp. 1–2. 2019. 2