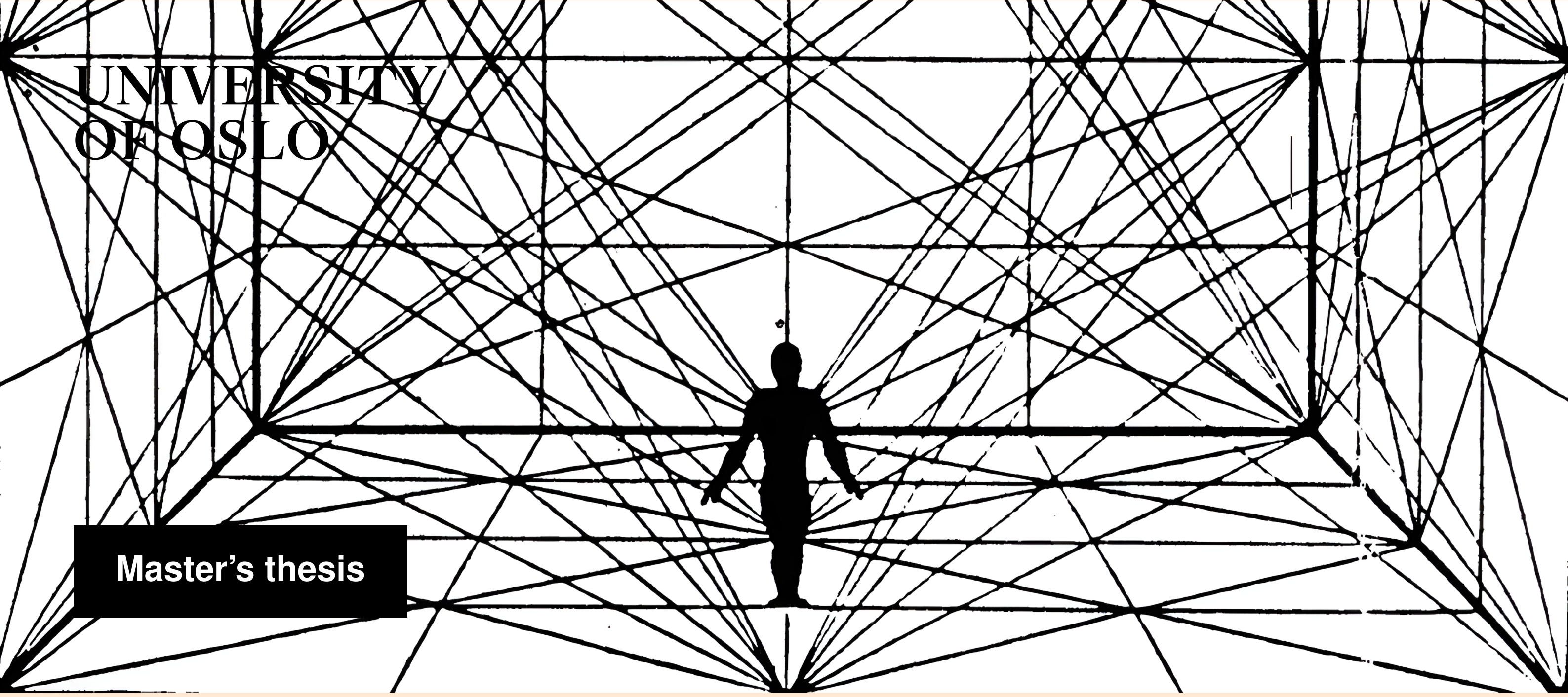
UNIVERSITY
OF OSLO



Master's thesis

The Shapeshifter

Co-constructing the body with optical, marker-based motion capture in live dance performance

Hugh Alexander von Arnim

Music, Communication and Technology
30 ECTS study points

Department of Musicology
Faculty of Humanities

Autumn 2023

Hugh Alexander von Arnim

The Shapeshifter

Co-constructing the body with optical, marker-based motion capture in live dance performance

Supervisor:
Tejaswinee Kelkar

Cover Image:
The Laws of Cubic Space by Oskar Schlemmer. From Schlemmer (1987)

Abstract

This thesis presents *The Shapeshifter*, a multi-modal system for interactive dance and a corresponding performance that employs an optical, marker-based motion capture system. *The Shapeshifter* is developed from an exploratory, research-creation approach to investigate the ways in which the use of motion capture in artistic contexts results in an environment in which the representation of the body is co-constructed. This thesis first outlines the co-construction model of motion capture in live dance and the theoretical methodology upon which this model is based. Thereafter, the iterative and participatory design of *The Shapeshifter* system and development of the performance, carried out with a research-creation collaborator are described. An evaluation of the system focused on the collaborator's experience of their body, their perception of its relation to the technological components of the system, the co-constructed representation of the body, and the system's latency form the final part of this thesis.

Acknowledgements

The work done in this thesis would not have been possible without the kind support and encouragement of the network of people surrounding me. The students and staff in the MCT programme, from whom I've learned so much, provided a stimulating and inspiring environment over the course of the previous years. Stefano Fasciani, the programme coordinator, facilitates this environment through offering an intense series of challenges while simultaneously enabling experimentation and the pursuit of individual interests and fields of study. My thesis supervisor, Tejaswinee Kelkar, went above and beyond in the amount of time she took and effort she made to offer feedback, read proofs, and shape nebulous and sprawling ideas into concrete concepts. Live Noven invested herself in the work, graciously offering so much of her time and so many of her ideas. My parents, Ingrid and Christian, provided emotional and financial support without which I would not have been able to pursue this programme. And most of all, I feel incredible gratitude that Isabel and Mia have gifted me with a warm, caring, and loving (and fun) home to return to at the end of each day.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In 2022, ABBA began their first concert tour in 42 years in support of their recently released album *Voyage*. Starting with a residency in a custom arena within the Queen Elizabeth Olympic Park in London which is scheduled to last until 2024, and followed by an upcoming global tour, each show features a performance of the band's greatest hits along with accompanying dance routines. The production engaged renowned choreographer Wayne McGregor to develop the show's dance components, who noted the difficulties in trying to get the members of ABBA to recapture the energy of their earlier years while emphasising the embodied nature of this challenge, stating that "it's been an amazing, kind of like technical and emotional challenge. How is it that you get artists who've not performed for a while back into their bodies, and enliven and confident in performing?" (ABBA XXI, 2022)

For the opening night of their residency, Agnetha, Benni, Björn, and Anni-Frid reunited in the Olympic Park. However, they did not take to the stage. They remained seated in the audience as they watched virtual representations of themselves as they appeared in the 1970's, so called *ABBA-tars*, perform the show along with a live band. Created with the use of a large-scale full-body and facial motion capture system, *Voyage* Tour producer Ludvig Andersson employed strong terms to elide the physical bodies of the members of ABBA with their *ABBA-tars*, claiming "that when you see this show it is not a version of, or a copy of, or four people pretending to be ABBA, it is actually them", going so far as to state that the technology captured "the soul of their beings" (ABBA Voyage, 2021). Although such rhetoric can be ascribed to the promotional nature of the interview, taking Andersson's comments in view of McGregor's points towards a striking implication; namely that the process of getting the members of ABBA "back into their bodies" involves constructing a new, virtual body for each member.

This example highlights the complex relationship between the physical body of the performer and its virtual representation. The central theme of this thesis revolves around this relationship, the process required to construct such representations, and how a performer perceives the relationship between their own body and the representation. The relationship between

Figure 1.1: ABBA back in their bodies. From 'The bigger picture: ABBA Voyage' (2022, p. 14-15)

the physical body and its medial representation obtained through the use of motion capture have already received critical examination, ranging from Ng and Bax's (2023) analysis of *Voyage* to inquiry into works dating back to the initial wave of live artistic productions which integrated motion capture technologies in the 1990's, exemplified in a series of articles examining the relationship between the physical and virtual bodies of dancer and choreographer Bill T. Jones in the sequence of works building from the original *Ghostcatching* (Barber, 2015; Dils, 2002; Goldman, 2003). However, Andersson's claim that the constructed bodies *are* ABBA holds a further implication which opens an additional line of inquiry for examining the relationship between the physical and virtual bodies. As noted by Plaete et al. (2022), a number of younger stand-ins also provided motion from which the virtual body was constructed. A team of engineers and animators were involved in the process of constructing the body from the motion data. The technologies involved implied ways of working, and imposed their own limitations on the body that could be constructed. If ABBA are back in their bodies, their embodied experiences are the result of a process of constructing the body from their motion which involves numerous actors, each of whom brings with them assumptions about what those bodies should be. The use of motion capture technologies results in a process in which the representation of the body is, in effect, co-constructed.

In this thesis, I aim to critically explore the co-construction of a virtual representation of the body through the use of optical, marker-based motion capture in live performance, specifically focusing on dance. Employing a research-creation methodology centred around the development of an interactive dance system and performance, the thesis is structured around the following three research questions:

RQ. 1 How does the use of a motion capture system co-construct a virtual representation of the body in live performance, and which assumptions about the body does it make?

RQ. 2 How can a multi-modal interactive system be iteratively designed from a perspective which foregrounds motion capture as co-constructing a representation of the body?

RQ. 3 How does a performer experience their body in relation to the technological components in performance with a system for interactive dance?

1.1 Contributions

In line with the above research questions, this thesis offers the following contributions:

1. The co-construction model, a theoretical framework which models the use of motion capture technologies to create a virtual representation of a physical body as a co-constructive process. In this thesis, I ground this co-construction model in the examples of *optical, marker-based MoCap* employed in *dance* works in a *live performance* context.
2. *The Shapeshifter*, a performance and multi-modal interactive system. This is documented through the following, which can be found at this link[1] and in the thesis appendices:
    (a) A video recording of a performance
    (b) The source code for the interactive system
    (c) Methodology for the evaluation of system motion to visualisation latency, adapting the methodology developed for latency measurements in head-mounted displays by Steed (2008).
    (d) The code used to perform computational analysis of the system latency
3. A qualitative analysis of the embodied experience of the project's collaborator of performing the work and their relationship to the technological components of the system from an embodied perspective.

[1]https://mct-master.github.io/masters-thesis/2023/12/12/hughav-the-shapeshifter.html

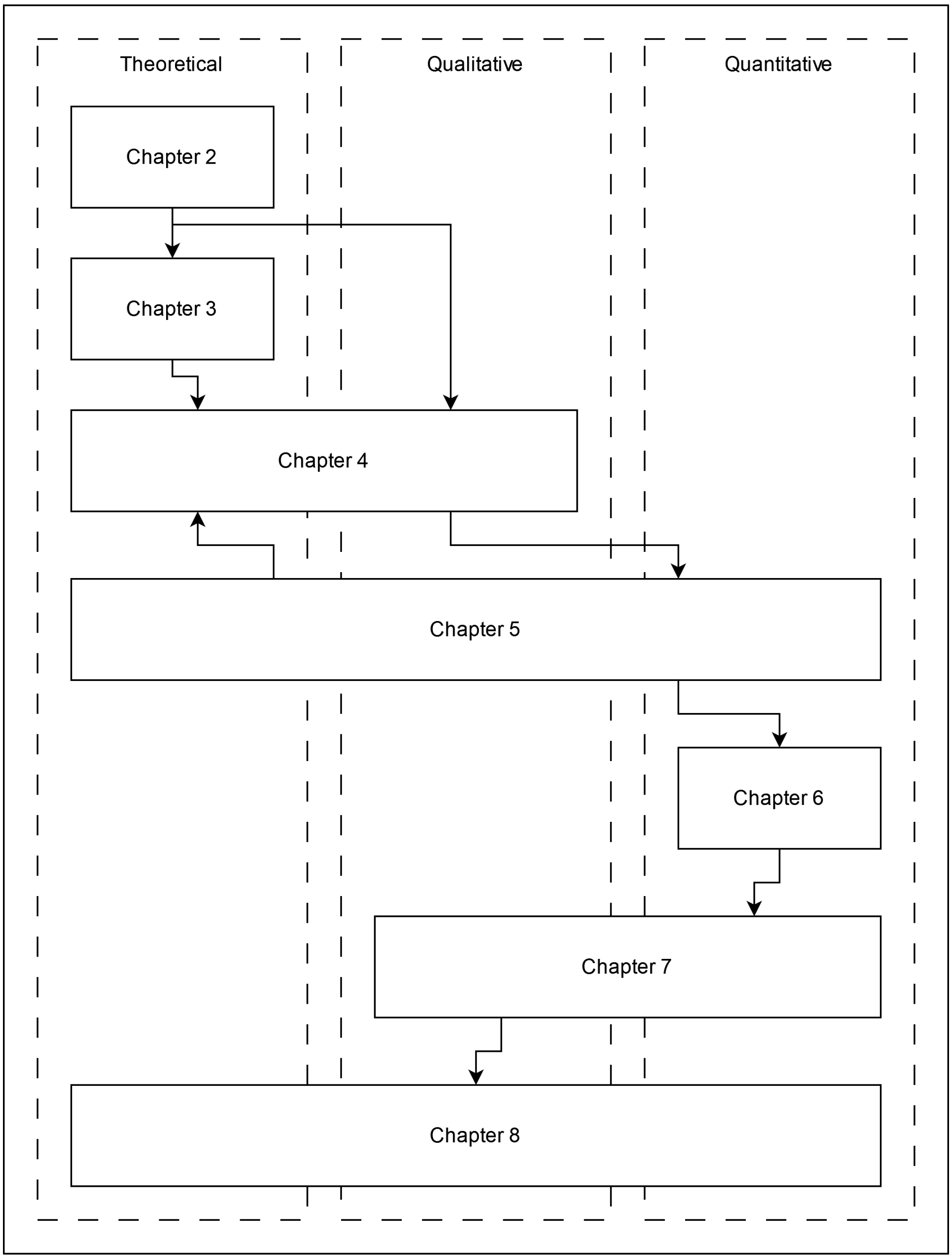


Figure 1.2: The structure of this thesis. Connections represent the flow of supporting ideas. The locations of these connections do not imply a theoretical, qualitative, or quantitative support, but rather imply a general support.

1.2 Thesis Structure

This thesis is structured along three main threads: theoretical, qualitative, and quantitative. These are reflected in the contributions outlined in section 1.1. An overview of these threads in relation to the thesis chapters is presented in figure 1.2.

In chapter 2, I build the theoretical background for the subsequent chapters, outlining key concepts and preliminary work while simultaneously demarcating the thesis' scope. I present the co-construction model of motion capture in live dance performance, the main theoretical contribution of this thesis, in chapter 3, along with its methodological base. In chapter 4, I introduce *The Shapeshifter*, a performance and multi-modal interactive system, developed in view of the work presented in chapters 2 and 3 along with a collaborator from a research-creation approach. I introduce the collaborator, the performance space, and the concepts that we aimed to explore with *The Shapeshifter*, and additionally provide an overview of a performance from a functional perspective. In chapter 5, I outline the theoretical, qualitative, and quantitative methods that underlay the iterative and participatory design of the system and development of the performance, an in-depth interview that took place after a pilot performance, and an evaluation of the system motion to visualisation latency. I provide a systems-oriented description of *The Shapeshifter* in chapter 6, outlining the design and implementation of the system. In chapter 7, I present the results obtained through the work done in this thesis, followed by a discussion and contextualisation of these results, as well as direct answers to the thesis' research questions, in chapter 8.

Chapter 2

Background, Concepts, and Scope

In this chapter, I outline the theoretical concepts and background work that supports the subsequent chapters of this thesis, with this also serving to define the thesis scope. This includes preliminary work upon which much of the work done for this thesis is based.

2.1 Motion Capture

As defined by Nymoen (2013), at a fundamental level motion capture refers to "the use of a sensing technology to track and store movement" (p. 13). This definition is quite broad, enabling everything from, as noted by Nymoen, a pencil and paper to specifically designed technologies for capturing and representing motion to be delineated as a motion capture technology. Nymoen also clarifies the distinction between motion *tracking*, which comprises sensing and processing motion, and motion *capture*, wherein the processed data are also stored. Moving forward, I will use these terms when it is necessary to refer exclusively to contexts which meet Nymoen's definitions, but will use the abbreviation *MoCap* as an umbrella term when distinction is not required.

According to Nymoen's taxonomy, the most common sensing technologies employed to track motion are *acoustic*, *mechanical*, *magnetic*, *intertial*, and *optical*. It is the latter of these which I focus on in this thesis, specifically the variant of optical tracking which is *marker-based*, employing physical markers which are attached to the object to be tracked.

2.1.1 Optical, Marker-Based MoCap

The name of this method of MoCap concisely sums up how it works and what is involved, describing both the sensor and what is sensed. Optical refers to the use of cameras as the sensing technology, with these consisting of cameras that operate within the Infra-Red (IR) range of the electromagnetic spectrum. Marker-based refers to what is sensed by the cameras, namely some physical object placed within the environment or upon the human body for which the position is tracked by the camera system.

Optical markers are generally small, spherical balls that are coated with a highly IR reflective surface with an adhesive on one side. There are two main types of markers: active and passive. Active markers emit their own IR light and require a separate power supply, while passive markers reflect IR light originating from IR light emitters usually mounted on the cameras themselves.

A typical marker-based system uses multiple cameras. Once the camera system is set up, the capture volume, which is the physical space in which motion is to be captured, is calibrated using fixed marker distances, so that new objects can be accurately tracked. If the IR light that is either reflected (by passive markers) or emitted (by active markers) is captured by at least two cameras, the position of the marker within the capture volume can be triangulated after the raw camera data has had several processing steps applied in order to extract the position of the marker within the frame. The position of the marker within the capture volume is represented by a Cartesian coordinate relative to a user-defined origin and represents a dimensionless point in the capture space.

There are several considerations that must be kept in mind when using a marker-based system. Firstly, if a marker is occluded, that is hidden from the cameras view so that it is visible to fewer than two cameras, it can no longer be registered by the system. This means that the physical characteristics of the capture volume must be considered. For example, if there is a desire to capture within a specific section of the volume, any objects which occlude the cameras' line of sight must be removed, or at least placed in a position where at minimum two cameras can fully capture that section. The body of the performer can also cause occlusions, so even if an area is clear of objects there might be limitations placed upon motions that the performer can carry out. An example of this is floor work in dance, where any markers on the side of the performer's body which is against the floor might not be able to be captured.

Secondly, if several markers come into close proximity with one another, the system might be more imprecise. This is due to the system being unable to distinguish between the markers, especially if they come within the deviance of error of the system. This must be considered when multiple markers are to be place upon a small object.

Thirdly, any object that reflects or emits IR light will be registered as a marker by the system. In locations which contain many of either of these a lot of noise can be added. This noise can be in the form of missing markers, representations of markers that do not physically exist, or marker jumping.

When passive markers are used, individual markers do not possess a distinct identifier. In the case of motion capture, once a recording has been completed markers can be labelled and any gaps in the capture can be filled. Depending upon the amount of noise in the capture, this can potentially be quite a long and arduous process. If it is important that individual markers are consistently identifiable during the MoCap session (for example, in interactive performances which map a specific marker to

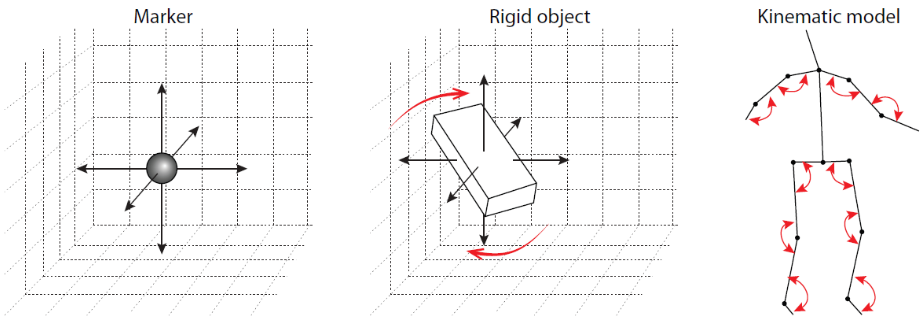


Figure 2.1: The process of constructing a kinematic model of the skeleton from a series of markers. From Nymoen (2013, p. 15)

a specific parameter), a method must be developed with which its identity can be preserved across any gaps.

A defining property of an optical, marker-based system is that the position of each marker is represented as part of a global coordinate system. This means that the position of each marker is not represented in relation to the position of other markers, but rather in relation to an origin that is defined during calibration of the capture volume. As a consequence, there is no inherent relationship recognised by the system between the motion of each individual marker that it captures. Instead, markers can be placed anywhere within the capture volume, and the relationship between the markers must be defined through a method chosen by the user. An advantage of this is that it is possible to capture a human body interacting with an inanimate object or multiple human bodies within a single capture, and the spatial relationship is preserved. The most common method of determining the relationships between these objects is through the definition of rigid bodies and the modelling of the human skeleton as a kinematic chain.

Since the positions of markers are represented as dimensionless points within the capture volume, several steps can be performed to obtain higher order properties. To obtain spatial dimensions and the orientation of an object, a constellation of markers can be defined as a rigid body. A rigid body is constructed with the assumption that the physical object that it represents is non-deformable, meaning that it does not change in shape, size, or internal structure when subject to external force. This implies that to define a rigid body, the constellation of markers must be in fixed positions, with the relative position and angles between each of the markers remaining consistent for the duration of the capture or period of tracking.

To model more complex objects, a series of rigid bodies can be joined together to form a kinematic chain, which represents this series as connected to one another by joints which have predetermined degrees of rotational and transformational freedom, as shown in figure 2.1. These are often organised hierarchically, with one rigid body serving as the root to which all other rigid bodies are chained. The modelling of kinematic chains can quickly become quite complex if there is a desire to move beyond simple models and a full discussion of this is beyond the scope of this thesis. However, I refer to Müller (2007b) for a more thorough description of this process in relation to the human body.

A common procedure is to define the hips as the root rigid body of the kinematic chain, with the two upper legs and the lower spine functioning as separate stems extending from the root. These models are often extremely simplified. For example, the spine is commonly modelled as consisting of either two or three connected rigid bodies in contrast to the 33 vertebrae found in the human spine.

Optical, marker-based motion capture has been applied across several fields in the arts and entertainment, most prominently in film and video games. However, there is a contingent of live dance works which employ the technology which I will overview. As the term *live* can signify several meanings, I will first outline the definition in use in this thesis.

2.1.2 Live Performance

Although this thesis is not directly concerned with the ontology of live versus reproduced performance, as I exclusively focus on the use of motion capture in a live performance context, the term *live* requires elucidation in order to clearly define the scope of works and contexts that I will discuss. I draw upon Auslander's (2023) line of argumentation that, in view of the increasing incursion of mediatisation into various facets of the arts and entertainment, live and mediatised performance do not form an opposing ontological binary defined by intrinsic qualities of the performance. Rather, they form two contexts of performance that define themselves in opposition to the other.[1] Drawing upon sociologist Irving Goffman's concept of social frames, Auslander argues that the live-ness of a performance is a construction, derived through the expectations and socio-cultural signifiers associated with the event. In effect, a performance becomes live when it is framed as such.

Carrying this definition of live performance into this thesis, I consider works which, to take two examples, are framed as ephemeral or which contrast mediatised or virtual aspects of performance with the spatio-temporally physically present to be live. Under such a definition of live-ness, the purview of this thesis extends to works not only involving interactive, real-time motion tracking, but also those involving pre-recorded motion captures. Although, as noted by Birringer (2004), interactive art requires a different understanding to non-interactive, multimedia performance, with the former grounded in an "aesthetics of process" (p. 168) in contrast to the latter's grounding in "aesthetics of spectacle" (p. 168), this thesis is not so much concerned with the process of interactive art but rather the process of co-constructing the body that occurs when a motion capture system is employed, regardless of whether this occurs prior to or during performance.

[1] Auslander also unambiguously states that this does not imply that the two terms share an ontology.

2.1.3 Motion Capture in Live Dance Performance

The Riverbed group, formed by Paul Kaiser and Shelley Eshkar, were responsible for several of the earliest forays into the integration of motion capture with dance (Dixon, 2007). With the group, Michael Girard and Susan Amkraut developed a kinematic modelling software named *Biped*, which formed the basis of Riverbed's collaborations with a number of prominent and influential dancers and choreographers. This software underlay their first collaboration with Merce Cunningham, an animated installation named *Hand-drawn Spaces*. Based upon this successful collaboration, Cunningham proposed to use the motion capture technology in a work that also involved live dance performers on stage. The result of this was a work titled *BIPED* in 1999, which was named after the modelling software that had been developed by Riverbed. The work featured dancers on stage, accompanied by projections of animated captures showcasing two or three dancers[2] executing a series of Cunningham's movement sequences onto a scrim. As reported by Abouaf (1999a), the process used to create the work involved a single afternoon of motion capture with the dancers. After processing the captured data, kinematic models were created which formed the basis for the following animation procedure. There were two main methods involved in the animations created from the kinematic models. The first was a rotoscoping technique, with hand drawn animations traced on top of the kinematic models by Kaiser and Eshkar. Abouaf describes these as "an expressive chalk skeleton against a black background" (p. 5).

The second method was the creation of a 3D model through mapping a spline to the kinematic model. Variations on this technique involved modifying the spline to represent more abstract forms. For the 3D animation, as noted by Dixon, much detail went into the modelling of kinematic effects, such as skin and tendon behaviour, and even "foot to ground collision response" (p. 188).

This method of mapping a kinematic model to a 3D animated model has proved influential to subsequent development of dance work involving the use of optical, marker-based MoCap. Dixon, in reference to an image of dancers in front of one of the animated figures featured in *BIPED*, notes that "*BIPED* images such as these have been so admired and reproduced that they have become archetypical of the digital dance and performance movement" (p. 193).

Following *BIPED*, Riverbed collaborated with Bill T. Jones on the installation *Ghostcatching*. In this work, several motion patterns performed by Jones were captured in a similar manner to those performed by the dancers in *BIPED*. The kinematic model created is mapped to representations meant to invoke "intertwinings of drawn strokes" (Jones et al., 1999). This was achieved by using the same systems that were involved in the production of *BIPED*, both in terms of the mapping of the kinematic model created from Jones' capture data to a series of splines, as well as the modelling of

[2] Abouaf (1999a) reports that the captures took place with two dancers, whereas Dixon (2007) reports three.

the skin and muscle behaviour (Baumgartner, 1999). After the premier of the work as an installation at The Cooper Union in New York, the work was later incorporated as a part Jones' *Breathing Show* tour. It sees multitudes of the animated figures spawning from each other and performing the patterns captured from Jones. These are accompanied by recitations recorded by Jones, ranging from song to spoken word.

In the years since, further works have explored the possibilities afforded through mapping a kinematic model of the skeleton to an animated figure in dance involving both motion capture and motion tracking. These inculde a re-envisioning of *Ghostcatching* in 2010 as *After Ghostcatching* (Barber, 2015), several works undertaken by Downie (2005) which included collaborations with Merce Cunningham and Trisha Brown for which, critical of design approaches grounded in a perspective of mapping, he developed an agential approach towards kinematic modelling of the skeleton from marker positions, Vincs and McCormick's (2010) use of the model to drive representation outwards from the body of the dancer in a stereoscopic projection with a real-time system, Meador et al.'s (2004) integration of the animated figure with multi-image projection, Strutt's (2022) telematic project, developed during the COVID-19 lockdown in the United Kingdom, for which the kinematic model is streamed in real-time over the internet and animated in a second location, and McCormick et al.'s (2014) use of machine learning techniques to learn a dancer's gestural vocabulary and construct a further animated figure placed in juxtaposition with an animation constructed from the dancer.

There are, of course, several dance works which employ optical, marker-based MoCap that do not attempt to kinematically model the skeleton. For example, the work *Lucidity* (James et al., 2006) uses a custom-built tracking engine to trace the position of a dancer as cloud of points from which higher level features such as dancer proximity and groupings are extracted. Limb motion was modelled with statistical methods relating to the point cloud. Vincs and McCormick (2010) make use of tracing the position of the dancer's hands in an alternative performance. Landry and Jeon (2017) map features extracted from marker data to sonification parameters.

Although widely used across a number of fields for the purpose of motion capture, optical, marker based methods have not experienced significant adoption for the purposes of real-time motion tracking in live performance, with other methods of tracking motion such as inertial systems being preferred (Bevilacqua et al., 2011). This is reflected, for example, in James et al. (2006), stating that they aim to move away from the use of optical, marker-based methods of motion tracking for future work on *Lucidity*, due to the intrusive and opaque nature of the system. I connect this desire for non-intrusiveness and transparency to Sterne's (2001) concept of mediate auscultation, which I outline in the following section.

2.2 Mediate Auscultation and Construction of the Body

In this thesis, I aim to construct a theoretical lens through which virtual representations of a physical body or body part can be examined. As the extent to which a body is represented can be nebulous, I relate this to the use of a MoCap system in a way that implies a sense of what Sterne (2001) terms mediate auscultation. Taking the stethoscope and medical listening as an example, Sterne describes mediate auscultation as a form of hearing mediated through the employment of a sensing device in which the sensor apparatus is "erased from consciousness" (p. 123) and which leads to a situation in which "the tool stands in for a whole process from which it erases itself" (p. 123). Although specifically focusing on audible modalities (hence auscultation), I transfer the concept to the employment of MoCap which is framed in a manner which removes the materiality of the sensor from the conceptual frame of the work. In other words, in contexts in which the raw sensor data obtained from the MoCap system is processed into a form of motion data which elides this processing with the object to be tracked itself. Here, I limit this to cases in which the object to be tracked is the human body.

Practically, this can take the form of framing the work through rhetoric that elides the body of the performer with a virtual representation of the body that comes into being through MoCap, such as is the case with the *ABBA-tars* described in chapter 1 or the artists' statement on *Ghostcatching*, where Jones et al. (1999) write that "the body of Jones *is* multiplied into many dancers" (p. 1, *emphasis added*). However, this can also present within processing steps that imply a conceptual erasure of the MoCap system's sensing apparatus. On a relatively straightforward level, this can take place in cases such as the positioning of a marker on the hand and conceptualising the position of the marker extracted from the raw sensor data *as* the position of the hand, both in terms of framing the processing (i.e. labelling the motion data as *hand* or applying constraints based upon what a hand "can do") as well as the performer's approach towards the technology (are they moving their hand or moving their hand in order to move the marker).

It is also important to address that the terminology of *construction* in relation to the MoCap system can be read as implying that the technology itself is displaying a degree of agency in the process of emergence of the virtual body. Although investigation into the agency of technological systems is a developing direction of study in several fields concerned with the intersection of technology and the arts,[3] a discussion of the presence of agency in relation to MoCap systems is beyond the scope of this thesis. Instead, when I apply the terminology of *creation* in relation

[3]For example, there is a growing interest in technological agency as an alternative approach to mapping in the field of interactive music systems, especially in relation to the application of machine learning techniques and feedback systems, see Eldridge et al. (2021), Erdem et al. (2022) and Magnusson et al. (2022) for several recent examples.

to technological systems, I am employing it in contexts relating to the way in which their affordances and limitations, and their shaping of data contributes to fashioning the emerging virtual body beyond the intentions of the human actors involved in the process.

2.3 Research-Creation

The work done in this thesis comprises a project following Stévance and Lacasse's (2018) research-creation. This is a broad methodological approach in dialogue specifically with previous discussion on artistic and practice-oriented research. In contrast to these perspectives,[4] with research-creation Stévance and Lacasse offer a methodological approach for the integration of artistic practice and academic research more expansive in scope. Noting that this is a broad, overarching concept, its scope encompasses not only the various forms of artistic research but also non-artistic practices that result in artefacts such as technological innovation. They offer a definition comprising six key concepts:

> Research-Creation is understood as (1) an *approach* applied to (3) an individual or multiple-agent (2) *project* combining (4) research *methods* and creative *practices* within a dynamic frame of (5) causal *interaction* (that is, each having a direct influence on the other), leading to both (6) scholarly and artefactual *productions* (be they artistic or otherwise).
>
> (Stévance and Lacasse, 2018, p. 123)

Their approach embraces the employment of any methodology within the creative and research process, so long as these inform and influence one another. They stress that this interaction between the creative and research process must be causal in nature, in that the products of the research process would not otherwise exist without the simultaneous undertaking of a creative process, or vice versa.

Moreover, Stévance and Lacasse emphasise that the focus on the project over the individual working on the project leads to the research-creation approach being well-suited to collaborative endeavours. They make certain to differentiate between *cooperation* and *collaboration*, emphasising that cooperation is the unsystematic joining of two parties in order to work, mostly independently, towards achieving a result but that collaboration involves a dynamic process of a group voluntarily working together to achieve a common goal. In the case of this thesis, the collaborator on the project is a dancer and physical theatre practitioner from the local area. I will introduce the collaborator and define the roles that we assumed in the project in chapter 4, when the collaborator becomes more present in the work. A description of the specific research methods employed in this thesis can be found in chapter 5.

[4]See Arlander (2009), Barton (2017), Borgdorff (2006), Candy (2006), Finley (2007), Frayling (1993), Hansen (2017), Klein (2017) and Ladd (1979).

2.4 Preliminary Work

To aid in contextualisation of the work carried out for this thesis I will briefly present *Reconfigurations*, a previously developed interactive system for dance, as well as an approach we termed *Motion Pointillism* that we began to develop during work on *Reconfigurations* and that was further refined during work on *The Shapeshifter*. Both of these inform the work that is documented in the following chapters. As this work took place in a collaborative context with my thesis supervisor, I switch to employing the plural while describing this work.

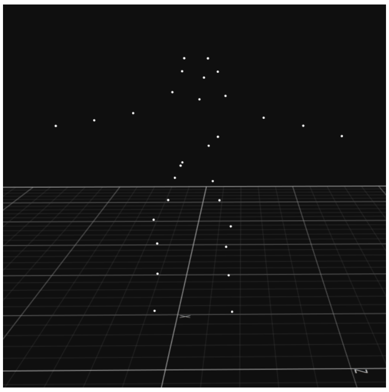
2.4.1 *Reconfigurations*

Reconfigurations originated from a desire to interrogate how the body was represented by the MoCap system installed in the laboratory at the Department of Musicology called the *Portal*. Having previously worked with this system, as well as other optical systems, for work on motion analysis projects, we started to think about how the MoCap software visually presents the human form as a construct of the motion data that was captured. This likewise led to the recognition of several limitations that systems imposed upon how the body could be represented by the software, both in terms of hard limitations in terms of the capabilities of the software, as well as soft limitations imposed by the software's design.

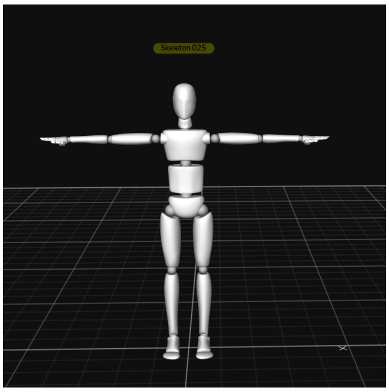
At a fundamental level, the MoCap system, *Optitrack Motive*, presents the position of markers captured by the system. These are visualised within a 3D representation of the capture volume as small, coloured spheres, as shown in figure 2.2a. Motive allows the organisation of a collection of markers as either a rigid body or a kinematically modelled skeleton. The skeleton can be represented as rigging one of several meshes such as a mannequin, as shown in figure 2.2b. Alternatively, the skeleton itself can be represented by a series of lines, which join the spheres of the joints along the path of the skeleton's bones, shown in figure 2.2c. A rigid body is similarly represented, with a series of lines demarcating the boundary of the object between the spheres of the markers, shown in figure 2.2d.

While this representation of the kinematic model is recognisable as the figure of a person, there is a sense of the uncanny to this representation of the human body. The body is reduced to a series of points and reconstructed through joining these points with a series of pre-determined connecting lines. Importantly these connecting lines are also fixed in terms of their properties relative to the body that they are portraying, freezing it in this uncanny form.

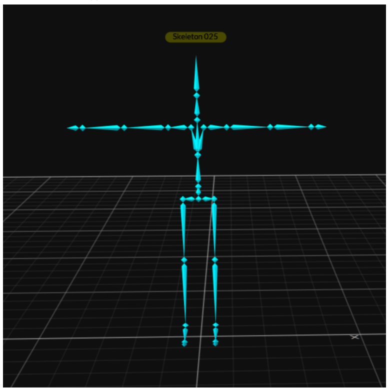
We began to consider how we could unfreeze this form and remove the constraints imposed by the modelling process. What if the markers were not assumed to be in fixed positions on the body, so that the modelling process didn't break down when a marker is moved? What if the connecting lines were malleable, not presuming to reconstruct a part of the body in the capture? What if we didn't conceptualise a single human body as the boundary of the modelling process, enabling the encompassing



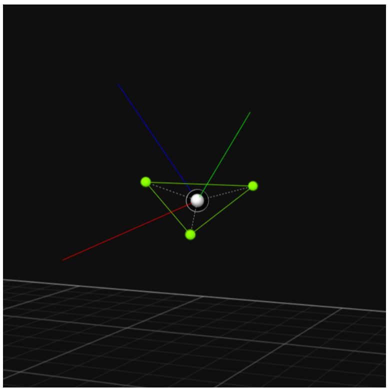
(a) The position of individual markers visualised as small, coloured spheres.



(b) A mesh in the form of a mannequin rigged with a kinematic model of the skeleton.



(c) The skeleton visualised as a series of lines representing bones and spheres representing joints.



(d) An individual rigid body, also visualised with spheres and lines.

Figure 2.2: Various visualisations of MoCap data in *Optitrack Motive*. Most major optical, marker-based software manufacturers employ similar principles for visualisation, even if specifics may vary.

of inanimate objects, parts of the environment, and even a second physical body as part of the construction of the form?

Unfortunately, the Motive software is quite obstinate with its modelling process, being especially inflexible when it comes to reconfiguring these in real-time. There is no manual way to create connections between markers. Rigid bodies, once defined, stop being tracked if a marker moves outside of the margin of error (usually a couple of millimeters). A skeleton model is picked from a list of presets, each requiring the wearing of a specific marker set, and then the parts of the body that should be modelled are fixed. A custom skeleton can be defined, but this requires creating a custom XML file, something that is not possible either post-facto of a recording or in real-time. Moreover, it is quite complex to do even without these constraints and still implies the wearing of markers in specific positions.

We therefore decided to use Motive to stream marker positions and create our own software that would enable this malleability. Using the

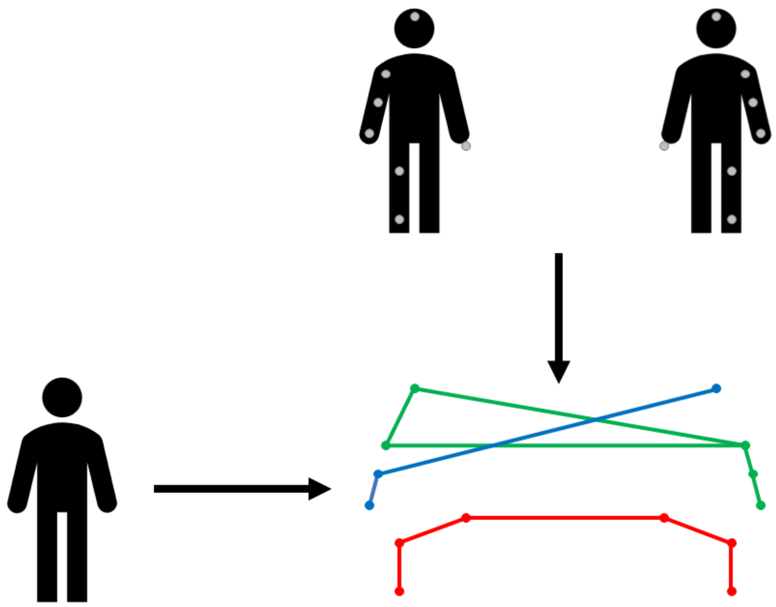


Figure 2.3: The two dancers position the markers in the performance space, either through wearing them, holding them in their hands or positioning them somewhere in space. The third performer uses the software to group the markers into bodies in real-time and draw malleable bones between them. The resulting figures are then displayed on a video wall in front of the performers and audience.

programming language Python, we in effect recreated a simplified version of the MoCap software display. Each marker is represented in a pseudo-3D virtual capture volume. Depth is simulated through altering the size of each circle. However, we added an element of interactivity, which functions as a reconfigurable modelling process. Markers can be assigned to a "body" on the fly, with each body represented by a different colour. Markers belonging to the same body can be joined together with a line representing a bone by clicking on one marker and drawing to another. However, instead of these lines being fixed in position and length, they follow the markers to which they are connected, changing in length and relative position. Markers can also be disabled, removing the marker and any connecting bones from the display. We also decided to try and work with occurrences such as occlusions which are generally treated as errors in the system. When a marker is occluded, or otherwise not recognised by the system, its representation remains frozen in place. This means that a marker can be purposefully covered to hold it in position while the rest of the body moves to a different position.

Using this software, we developed an improvisatory dance work for three performers. Two MoCap performers improvise dance phrases within a performance area, employing up to 30 markers which they are free to

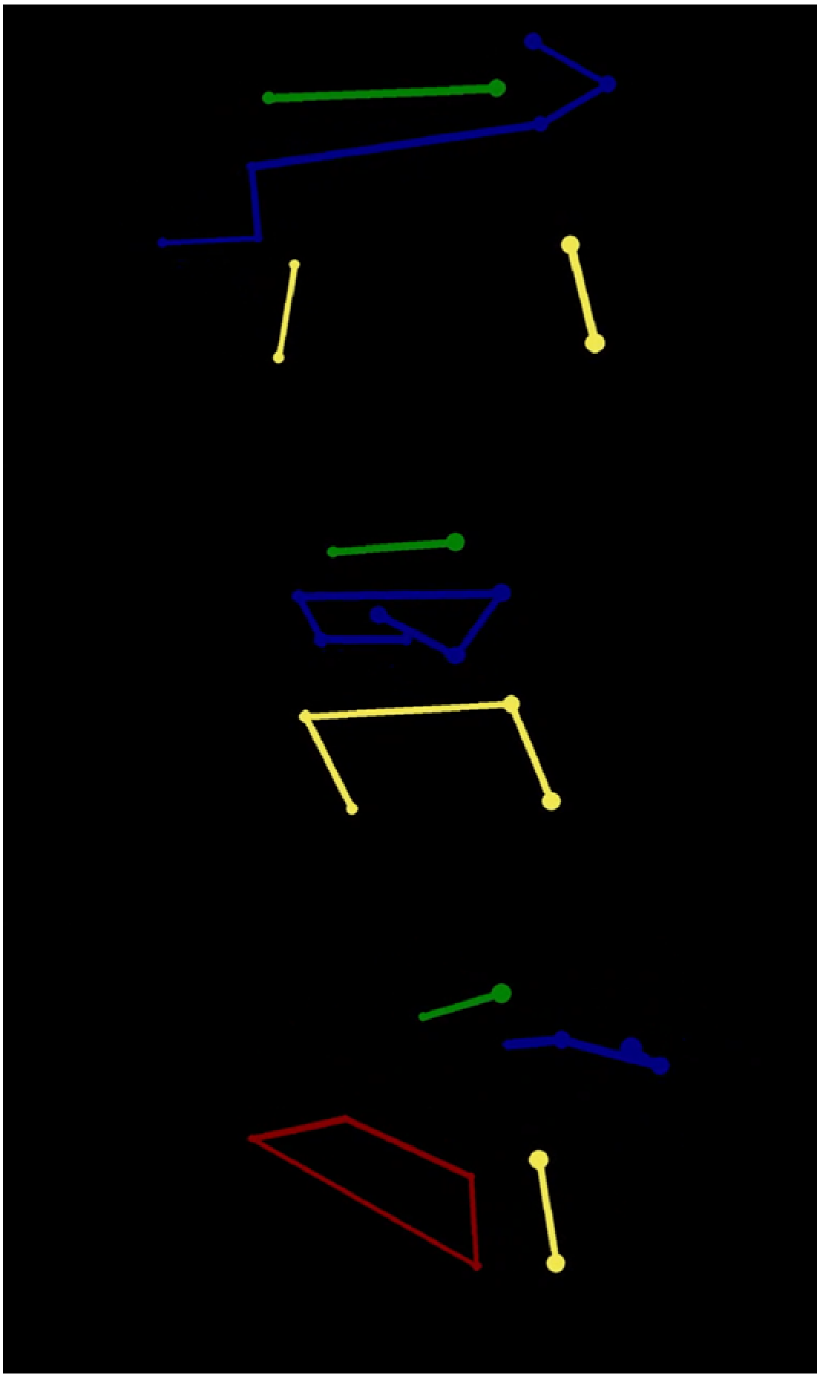


Figure 2.4: Three stills of the visualisation display from a performance of *Reconfigurations*. Each dancer was wearing markers along one side of their body divided across the lateral plane, resembling the positioning of markers required for kinematic modelling on one half of the body. Initially, this was used to create a combined body for the dancer, with the dancers coordinating their motion to move the shared representations. As the performance progressed, parts of the representation started to split off into more abstract shapes. Markers were eventually removed from the body and placed upon inanimate objects.

move and place wherever they wish at any point during the performance. The third performer uses a digital interface to control the software to reconfigure and connect these markers into up to five bodies. There is no limitation to the markers that can be connected within a body, meaning that connections can be built across both moving performers, and even incorporate inanimate objects to which they attach a marker. Likewise, there is no limitation in the other direction. Each performer can place their markers in positions corresponding to a marker set used for kinematic modelling, and these can be designated a single body and connected to resemble the way in which the modelled skeleton is represented in the MoCap software. Figure 2.3 shows how this functions from the performers' perspective. Instead of being pre-defined at the outset of a performance, the model of the body is configured and reconfigured during the performance. Facing the performers is a video wall, displaying the bodies as a mirror of the physical performance space. We also created a sound synthesis engine to which motion parameters are mapped to generate musical material across a performance. In figure 2.4 several stills from a performance are visible.

2.4.2 Motion Pointillism

While working on *Reconfigurations*, the concept of MoCap as a co-constructive process began to develop. From a perspective grounded in the concept's foundations, we outlined an approach towards working with MoCap termed *Motion Pointillism*.

We started by returning to considering what it is that we fundamentally aim to track and capture when using optical, marker-based MoCap, namely the position of dimensionless points within the capture volume that represent the position of the markers in physical space. Most MoCap software visualises these in a similar manner to *Optitrack Motive*, with small, unconnected spheres of identical colour. On captures of people wearing full body marker sets that are required for kinematic modelling, it is quite possible to recognise a human form from these representations before any modelling has taken place. This recognition can even take place at quite a personal level. For example, Jeannie Steele, one of the dancers who performed motion captures for *BIPED*, reported that she was able to discern herself from the collection of motion patterns that were recorded for the piece when observing the raw position data before any modelling had occurred (Abouaf, 1999b). However, although the human form is recognisable, if markers are removed from the representation one by one, slowly the human form that is visible dissolves. The work done in shaping the MoCap data into a representation of the human form is being performed in part by those viewing the representation.

This thought stands at the centre of the approach. We do not conceptualise these points as referring to the motion captured from a human body which can then be mapped to the output modalities of a system, whether through methods such as the rigging of an animation to a modelled skeleton or treating sensor data as representative of a performer's input gesture. We instead envision these points as referring to the point within the physical capture volume at which the markers are positioned. We then view the development of their referentiality as part of the performance, taking place within the perception of those performing and viewing the work, leveraging the ambiguity and tension inherent in the implication that the marker is referring to an object within the physical capture volume but the uncertainty of what exactly this is. It is up to the viewer to "join the dots" so to speak. We liken this to the pointillist movement in the visual arts, in which painters worked with points of individual colours and allowed these to blend in the perception of the viewer. Here it is points in space that refer only to the motion of a marker, that blend to construct a form in the viewer's perception. Crucially, we see this as a collaborative approach, involving the performer, the audience, the performance system designer, as well as the MoCap system itself.

We developed five guidelines that can be considered when aiming to work from this approach. These are:

Guideline 1 The MoCap system should be as opaque as possible.

Everyone involved in a performance should be highly aware of the fact that any visual that is viewed on a screen is not a representation of an object in the physical space. Rather, what can be seen is a visualisation of a position in the physical space mirrored in a virtual space. To emphasise this point, no attempts should be made to conceal aspects of the MoCap system, neither markers nor cameras. This is connected to:

Guideline 2 The MoCap system cannot provide errors.

Marker occlusions, confusions, and noise are major reasons for either abandoning working with the system or using a method such as kinematic modelling which provides a way of counteracting these phenomena. We see these as an opportunity for the MoCap system itself to contribute towards the construction of forms. These can be purposefully worked into performance. For example, a performer can cover a marker to either remove it from the system or to hold it in a fixed position.

Guideline 3 The MoCap component of the performance should only work with points, but how those points are presented is open.

An interactive performance system built on top of the MoCap system should only be provided with the coordinates of each marker to work with. However, how each coordinate is presented within the virtual mirror of the physical space is at the discretion of those who create the performance. They can be connected to each other or to a separate point in the capture space, have transforms applied to them, and be represented by any object. What is important, however, is that any work done with the motion data does not assume that this motion originates from any specific source within the physical capture volume.

Guideline 4 Markers can be placed anywhere at any time.

Markers do not have a set location that must be maintained throughout the performance. The performer can attach them to their body, hold them in their hand, place them on objects, or drop them on the ground. The performer is encouraged to change the locations of markers throughout the entirety of the performance. The body is in focus, but it is not a boundary.

Guideline 5 The performer must be able to see visualised forms.

In many multimedia works involving MoCap in real-time, visual representations created from the MoCap process are projected behind the performer onstage. As we view the creation of the form as part of the work itself, and this as a collaborative process that takes place in the perception of all present, it is vital that performer also be able to take part in this process and view the configurations that can extend from their motion.

2.5 Summary

In this chapter I outlined several key concepts that support the following chapters. I provided an overview of optical, marker-based MoCap and its employment in live dance performance, defining live as a contextual phenomenon that encompasses uses of MoCap in both interactive systems as well as for the playback of previously recorded captures. I provided a summary of the concept of mediate auscultation, a conceptualisation of a sensing device that elides the device with that which is being sensed, and related this to the idea of constructing the body through the use of MoCap. Through outlining these concepts I additionally provided an outline of the scope of this thesis relating to the key terms in the research aim presented in chapter 1. I outlined research-creation as the broad methodological approach employed in this thesis, and concluded by outlining two pieces of preliminary work that inform the following chapters.

Chapter 3

The Co-Construction Model

In this chapter, I present the co-construction model of MoCap, the main theoretical contribution of this thesis. This model underlies the development and evaluation of *The Shapeshifter*, the interactive system and performance described in subsequent chapters. I begin this chapter by presenting the model, which comprises a complementary approach to ways of thinking about MoCap from a systems-oriented perspective. Thereafter I outline the methodological underpinnings of the co-construction model. This takes the form of considering structures along which assumptions about the body are embedded in a representation of the body, taking the use of kinematic modelling with optical, marker-based MoCap as an example. I frame this through a theoretical grounding in a conceptualisation of the body taken from disability studies and relate this to the concept of mediate auscultation.

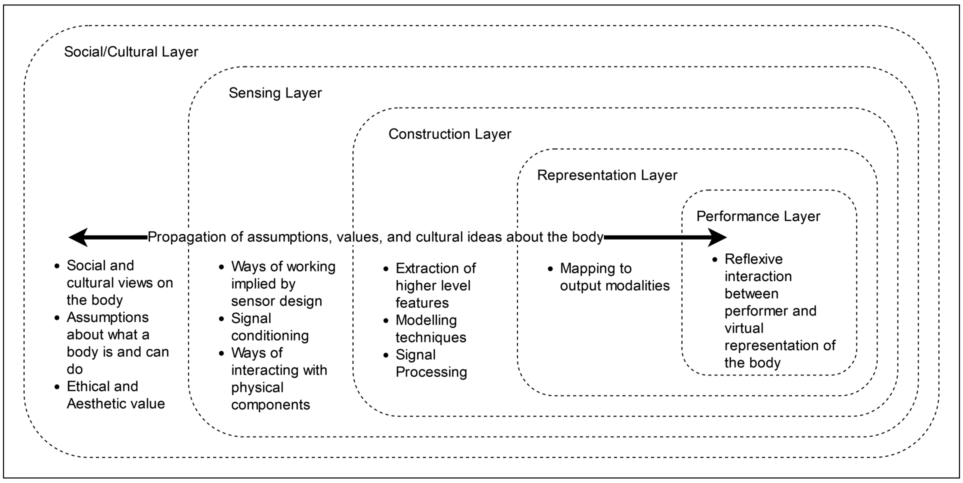


Figure 3.1: The co-construction model of MoCap in live dance.

3.1 MoCap as Co-Constructive Process

Figure 3.1 presents a model of MoCap in live dance as a co-constructive process. Rather than tracing flows of information through systems, the co-construction model traces vectors through which assumptions, values, and cultural ideas about the body are embedded in and co-construct a virtual representation of the body that interacts with the physical body of the performer during live performance.

The model depicts this occurring across five layers. These should not be understood as distinct and separated, but rather as nested within one another. They depict the increasing concretisation of the form of the virtual body, until this is realised in performance in relation to the physical body of the performer. Crucially, the flow of the embedding of assumptions and values is depicted as unidirectional, starting in the outer-most layer and flowing in towards the performance, but rather shows the propagation of assumptions throughout the entirety of the model.[1]

The outer-most layer depicts the wider social and cultural environment in which the performance is occurring, encapsulating the rest of the process and imbuing it with a system of values relating to the depiction of the body. Within this layer there are three layers relating to the technological systems of the MoCap process, namely the sensing layer, the construction layer, and the reproduction layer. [2]

The sensing layer encapsulates contributions to the construction propagating from the design of the sensing components. This can cover, for example, the requirement to apply signal conditioning methods in order to obtain usable data, or aspects of the physical design that imply certain ways of working, such as the inability of optical, marker-based MoCap to capture occluded markers. This also covers the presentation of physical sensors themselves to the performer, which imply certain ways of working or physical requirements for interaction.

The construction layer covers the moulding of the sensor data into forms that represent the body. This can occur, for example, through methods such as signal processing techniques that shape the input data into certain forms, the extraction of higher level features that imply the body's shape or ability. This can take the form of more or less explicit construction, for example, through the application of skeletal or muscular modelling or through the assumption that gestures reliant upon a specific body part will

[1] This should not be taken as an indication that the way in which the body is constructed and represented in a single dance work alters social and cultural views on the body, but rather, on aggregate, the depiction of bodies in cultural works can structure social and cultural attitudes towards the body as much as the inverse.

[2] I recognise that there is a large gap in scope between a wider social and cultural environment and a dance performance involving the use of MoCap. The inclusion of this layer is not intended to function as a method to pinpoint precise sites within this environment where assumptions and values are located, and this layer could potentially be excluded from the model. However, I believe encapsulating the inner layers within this layer serves the function of maintaining a sense of clarity of where many assumptions and values originate, and thus aids in encouraging consideration of which assumptions various actors may be bringing.

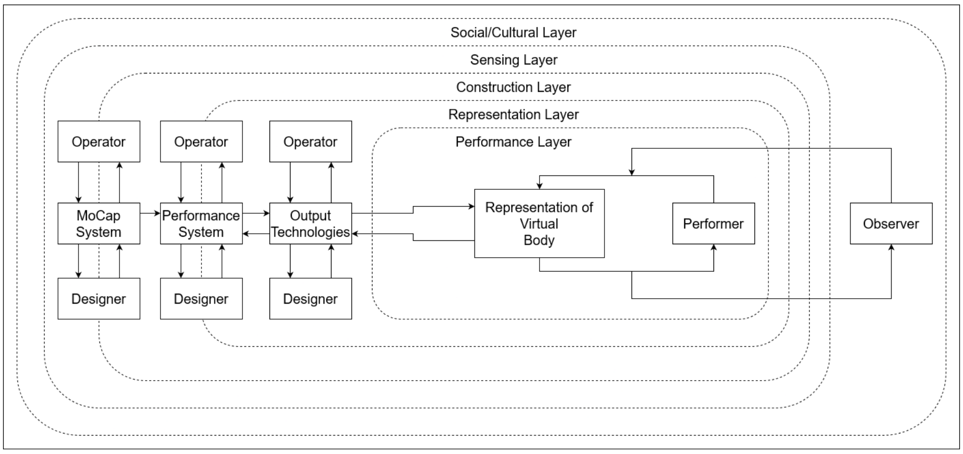


Figure 3.2: The co-construction model applied to a performance of *Reconfigurations*.

be provided as input.

In the reproduction layer, the constructed body is mapped to an output of the technological systems in one or several modalities. The result of this is a representation of the virtual body that is present in the performance alongside the physical body of the performer.

Within the inner-most performance layer, the performer interacts with the virtual representation of the body. Through their performance, they shape the form of the virtual body and, likewise, the form of the body shapes their performance. This can take place explicitly, for example within the context of interactive performance, or implicitly, as is the case for performance using reproduced motion capture wherein the performer's physical body is placed into relationship with the capture.

Applying the model to a performance involves identifying sites in which assumptions and values are embedded. This comprises identifying technological systems and components, the designers of the systems, and those who interact with, operate, or observe the systems to construct the representation of the body and determining the flows of embedded assumptions and values. An example of this is shown in figure 3.2, which maps these for a performance of *Reconfigurations*. For example, this shows the MoCap system which is employed to track the position data of markers, with the design of *how* it extracts the position data forcing certain ways of working and conceptualisation of the body. This places it as overlapping the sensing layer and the lower levels of the construction layer. We designed the components of the *Reconfigurations* system on top of this. As we were working with marker position data, this means that the *Reconfigurations* system played a large role in construction of data into a representation of the body through designating the relationships between marker positions. The system then presented these constructed representations in sonic and visual modalities over various output technologies, which in turn shape the representations. The performer interacts with the representation in the performance layer, into which the assumptions and values embedded in the system in the sensing, construction, and representation layers propagate. This process takes place in a wider social and cultural environment from which these assumptions and values emerge. The observer is situated within this environment, interpreting the representation of the body in line with the assumptions and values which they take from the social and cultural environment. This figure could be reproduced on increasing layers of granularity in relation to system components in order to identify these assumptions and values, and how they actualise within the system. Systematising these sites and flows enables targeted analysis of the assumptions and values brought by an actor or embedded within a system or system component, as well as aiding in design through considering components which can be accessed by designers and operators of systems, possible assumptions and values that they embed, and critical self-reflection on behalf of the actors who interact with these technological systems and components.

There are several things to note about this model. First, the designation of systems and components within a performance can be quite fluid and can cover multiple layers or be encapsulated in a single system that occupies space across the sensing, construction, reproduction, and performance layers. Likewise, systems can be conceptualised in a variety of ways at varying levels of granularity. Therefore, the purpose of the co-construction model is not to create a static and comprehensive map of the technologies and actors involved in a performance, but rather to offer a framework with which the embedding of assumptions and values can be systematically assessed. Second, when referring to performance systems, I do not imply solely interactive systems. With this I include all technological systems employed to reproduce the constructed body, from projectors to speakers to screens. A dancer performing in front of an animated figure created through motion capture is a distinctly different performance context depending upon whether the figure is displayed on a small television or projected onto a screen several meters tall, which can, for example, result in the embedding of assumptions and values about the height of a body. Third, the designation of actors in the example above as performer, observer, designer, and operator do not refer to individuals, but rather to roles. It might be that several of these roles are fulfilled by a single person, or indeed that a single person fulfils all roles. In such cases, the role that they are assuming can also shape the assumptions that they embed in the representation in different ways at different times.[3]

[3]This is well demonstrated through an anecdotal discussion that I had with a dancer/musician duo who worked with optical, marker-based motion tracking for a project. They employed a kinematic modelling procedure to construct a skeleton from the dancer which was then rigged to animate a mesh. However, they explained that the environment in which they were working resulted in a lot of high-frequency noise in the marker positions. Instead of applying filtering to the marker data in order to remove this noise, they instead aimed to create a performance in which the constant jitters were worked into the performance. In terms of the co-construction model, instead of attempting to remove artefacts originating in the sensing layer, which would be considered a limitation of the system when constructing a model of the body extending from social and cultural ideas about the body (namely that bodies don't constantly jitter and body parts do not suddenly

Fourth, the model does not imply a temporal element. The operations occurring in the sensing, construction, and reproduction layers might take place prior to or during the performance, or be split across these temporal frames. For example, in many situations in which reproduced captures are employed, the operations of the sensing and construction layers are situated prior to the performance, and operations on the reproduction layer take place both prior to and during performance. Moreover, the operations of the performance layer are doubled, occurring once during the capture sessions and again during the performance, each shaping the body in subtly different manners.

The co-construction model is intended to serve as a supplement to more systems-oriented models of MoCap that trace information flows through systems, such as those found in fields which interact with live dance and MoCap such as interactive dance (Camurri et al., 2016; Tragtenberg et al., 2019), animation (Bodenheimer et al., 1997; Gleicher, 1999; Kitagawa & Windsor, 2008), and interactive music (Birnbaum, 2007; Fasciani, 2014; Marshall, 2009; Miranda & Wanderley, 2006), with an overview of these models and perspectives, and their relationships with mediate auscultation and the co-construction model available in appendix A. It offers a step towards a way of systematically considering where assumptions about the body enter the performance, and whose assumptions they are. Conversely, this can also aid in system design, specifically in terms of thinking about where to situate the various systems, and whether to employ various functionalities and components in view of the assumptions that they embed. In the following section, I outline the methodology underlying the co-construction model through identifying several mechanisms that can shape the form of the co-constructed representation of the body.

3.2 Mechanisms that Shape the Virtual Body

The range of forms that the human body can take is remarkably heterogeneous. However, the use of MoCap to represent the body in live performance requires some proxy quantity to be sensed and from this, information about the physical body is extrapolated to construct its virtual representation. This can take forms ranging from the construction of a kinematic model of the skeleton to rig an animation or designing interactions and mapping these to an output in an interactive system. Therefore, in order to construct a body from these proxy quantities a number of assumptions must be made about the physical body, the form it takes, and the actions that it can perform. This becomes especially pertinent when a system involving the use of MoCap is intended for generalised use, in which a range of heterogeneous bodies may come into contact with a system that employs a singular model of the body constructed around these assumptions. However, as noted by Hayles (1999), specifically with regard to the body's

change position), the artefacts were carried through and integrated into the construction and reproduction layers. These actualised in the representation of the body with which the performer interacts in the performance layer.

relationship with technology, there is a broad gulf between *the body* as a generalised form and an individual's sense of embodiment. Hayles frames this dichotomy in terms of normativity and social and discursive concepts:

> Embodiment differs from the concept of the body in that the body is always normative relative to some set of criteria. To explore how the body is constructed within Renaissance medical discourse, for example, is to investigate the normative assumptions used to constitute a particular kind of social and discursive concept ... In contrast to the body, embodiment is contextual, enmeshed within the specifics of place, time, physiology, and culture ... Embodiment never coincides exactly with "the body," however that normalized concept is understood.
>
> (Hayles, 1999, p. 196)

It is from this generalised *body* that designers of technology must draw their assumptions, at least in part. Designers of systems that are employed in artistic practice are, in effect, embedding a cultural reproduction of the body within the representations of the technology, which, as argued by Balsamo (1995) from a feminist perspective, results in an upholding of binaries and value systems: "body parts are objectified and invested with cultural significance. In turn, this fragmentation is articulated to a culturally determined 'system of differences' that not only attributes value to different bodies, but 'processes' these bodies according to traditional, dualistic gendered 'natures'" (p. 234).

One such "system of difference" is the concept of normalcy, developed in a context of studies of the body as a construct of historically informed discourse in disability studies in the 1990's. A key point of consideration was the function through which a wide range of heterogeneous bodies which have few commonalities wind up marked as disabled.[4] Although historically disability had been viewed, for example, in religious terms as a result of divine affliction or afflatus (Straus, 2011), in the 19th and 20th century disability came to be framed in medical terms, locating the site of disability solely within the body (Scully, 2008). This framing of the term results in a view in which "[n]egative experiences encountered by a person with disability, including discomforts and lack of freedoms and capabilities, are all, on this model, due solely to the disability's character as a biological issue - a problem to be dealt with medically" (Hall, 2018, pp. 38-9).

Davis (1995) traces the development of this view on disability as alongside the development of statistical methods to assess physical qualities of populations, exemplified by the 19th century French statistician

[4]It is important to note the distinction between impairment and disability. As noted by Straus (2011), an impairment is an "underlying biological or medical condition" in contrast to "the meanings conferred on impairment by social and cultural construction" that form the disability (p.4). Sterne (2021) offers an account of impairment, including discussion of its relationship to disability, highlighting the context dependent nature of the relationship, in particular through his concept of the *normalised impairment*.

Adolphe Quetelet's concept of the *homme moyen*, a figure representing the average human which Davis describes Quetelet as stating represents the "average of all human attributes in a given country" (p. 26). It is against this figure of the average human against which extremes of the distribution can be defined, and those possessing bodies that fall into these extremes come to be seen as deviant.[5] As Davis clarifies, such thinking does not just affect those marked as disabled, but rather determines wider social thinking about the body, representing "a social process that intimately involves everyone who has a body and lives in the world of the senses" (p. 2).

This process also steers cultural domains, with Davis taking the nude in art as a case study. He writes:

> For example, if I ask you to think about the nude in art, chances are good that you will visualize a specific kind of body. Chances are remarkably good that the body will be female, white, and not visibly impaired. Few readers would imagine an Asian woman or a woman of color, even fewer a nude using a wheelchair. The reasons for such visualized assumptions are complex, involving further assumptions about beauty, about idealization, about sexuality, about gender, and so on. Intricately placed in that web of assumptions is a power move, I would call it, to fix the body as entire, intact, whole.
>
> (Davis, 1995, p. 11)

This concept of normal body in cultural domains is further explicated by Garland-Thomson (1997) in her figure of the *normate*. Garland-Thomson examines social and cultural representations of bodies marked as disabled, and argues that disability arises out of a socially layered exclusionary discourse, "not so much a property of bodies as a product of cultural rules about what bodies should be or do" (p. 6). In view of this, Garland-Thomson defines the figure of the *normate*, "the figure outlined by the array of deviant others whose marked bodies shore up the *normate's* boundaries ... a very narrowly defined profile that describes only a minority of actual people" (p. 8). Reynolds (2019) builds upon this figure, framing his analysis in relation to ability, describing the *normate* as "the tain of the mirror of ableism ... the invisible mechanism that allows slippage from being to being-able" (p. 244). He notes that the

[5]Davis also traces these developments with the concurrent development of the eugenics movement, through which the charting of distributions of physical characteristics can be advanced to quartile-based rankings. As Davis notes, "[t]he rather amazing fact is that almost all the early statisticians had one thing in common: they were eugenicists. ... An important consequence of the idea of the norm is that it divides the total population into standard and nonstandard subpopulations. The next step in conceiving of the population as norm and non-norm is for the state to attempt to norm the nonstandard" (p. 30). It should also be noted that when measuring a population along multiple dimensions, basing decisions around the idea of the average person as the exemplar of all dimensions is not the best course of action, as exemplified by the US airforce's attempt to define the average anthropometry of their pilots in the 1950's (Daniels, 1952).

influence of the *normate* is not limited to those marked as disabled, but also for instance a "job candidate ... picked over another because they are perceived to be more attractive, conflating cultural ideals of beauty with labor-related abilities" (p. 244). He relates these social forces to an embodied perspective, describing how "the *normate*, ever furnishing normative measures, reigns over the scale, scope, and content of ability expectations, it shapes everyone's experience of embodiment" (p. 255).

Viewing the use of MoCap in performance in light of this, it becomes possible to identify a process through which the representation of the body through the technology is imbued with values based upon assumptions made long before work on the project begins. The technologies used do not exist in a vacuum and in many cases are appropriated for use in the arts from other domains. This is especially the case for optical, marker-based MoCap, which is heavily linked to clinical gait analysis, a domain grounded in a medicalised view of the body (Downie, 2005).[6] This intersects with the concept of mediate ausculation, which Sterne (2001) directly links to medical sensing devices.

Framed through a discussion of the appropriation of biosensors designed as medical sensing devices for artistic purposes, Naccarato and MacCallum (2017) enter into dialogue with Sterne (2001) and the concept of mediate auscultation. They argue that the concept of mediated sensing implies in turn that an un-mediated sensing must also exist, and that this overlooks the fact that all sensing, even unaided by technological apparatus, is to some extent mediated. The sensor cannot be removed from the causal chain of perception and viewing it as such can lead to the masking of ethical and aesthetic values which are imbued into design of software and hardware.

Moreover, they frame their argument in terms of using sensors as control devices for interactive systems and argue that sensors employed in this manner imply a rigid causality and representation between the body being sensed and its form in the resulting media. This requires an "empirical conception" of what the body part being sensed is and what it can do, with the authors writing that

> In control based interaction, be it with biosensors or motion tracking, comparable assumptions regarding what bodies (or body parts, or bodily processes, or bodily gestures) are, and therefore can do, form the ethical basis from which aesthetic mappings are designed.
>
> (Naccarato and MacCallum, 2017, p. 6)

This also manifests in the design choices made by the designer(s) of the software systems involved in the work. As Naccarato and MacCallum note, the chain of causal links between the sensor and its medial representation

[6] Downie notes the relationship of the MoCap system with clinical applications to express frustration that this leads to a situation in which the system provides either extremely accurate tracking data or no tracking data at all.

(in more systems terms, the process of mapping the system's input to its output) consists of multiple steps of processing, including normalisation of the data provided by the sensor and the removal of noise. They argue that these steps themselves comprise aesthetic and ethical value judgements and provide a first-hand example of a choreographer working on a motion tracking system that was being designed for people with disabilities. In a session with a participant with Parkinson's disease, the system was unable to track the specific movements that the participant was making due to his specific style of movement. The choreographer kept requesting that the participant stand still in order to calibrate the system. As the participant was unable to do this, the choreographer described the participant as having broken the system. Linking this incident to the aesthetic and ethical values of the choreographer, Naccarato and MacCallum write that

> It was apparent, and not surprising, that the reference point of the system was situated in the body of the choreographer who seemed to view the non-normative movements of otherly-abled individuals as problematic. Importantly, the choreographer and software designers could not have created a system based on preset mappings of cause and effect between movement and media that was available to all bodily configurations and capabilities. In collaborations that involve movement and computing, the movement of the performer does not only point inward to the body as a self-contained entity; rather, movement hints at relationships between the body of the human performer with other elements in the given context, including the hardware and software being used for computation, as well as the designers of the interactive system. Software and hardware systems necessarily reflect and propagate the ethical and æsthetic value systems of their designers, whether intentionally or not.
>
> (Naccarato and MacCallum, 2017, p. 5)

It is possible to outline several material examples of the imbuing of aesthetic and ethical values taking the example of kinematic modelling in optical, marker-based MoCap, with the commonly used technique presenting the construction of a representation of the body at its most explicit. As outlined in section 2.1.1, defining a model of a kinematic chain involves designating a constellation of markers as a non-deformable rigid body and linking these together through a series of joints which determine the degrees of freedom between each rigid body. This requires the designers of software that defines a kinematic model to answer two difficult questions: what form does a human body take and what does and does not count as human motion? These two questions intersect with two requirements that are aimed for in the development of the models, namely that they are generalisable, that is that they can be used by more than one person, and that they represent a simplified model of human kinematics.

For kinematic modelling to function correctly, markers must be positioned on the body in accordance with predetermined locations so that the rigid bodies required for the model are correctly defined. In light of the drive towards generalisability, designers tend towards pre-defining the location of the markers which form the rigid bodies in advance, and in effect must determine the form of the bodies whose kinematic chain can be modelled with their system. In many cases, if the entire marker set that is required for the model is not present or individual markers are not positioned with the correct spatial relationship to one another (within a margin of error), the modelling process will not function correctly, or in some cases, is incapable of functioning at all.[7] As a result, the system is rendered unusable by those, for example, who do not possess a body part that is required by the model.

The requirement to create a simplified model of the kinematic chain of the human body is likewise shaped by social and cultural conceptions of the body. To take one example, the spine is often modelled as two to four rigid bodies connected by joints with either one or two degrees of rotational freedom. These rigid bodies are often modelled as forming a direct line between the pelvis and the skull. Such a model does not account for differences in spinal shape, such as found, for example in people with scoliosis (Schmid et al., 2016). Although models have been developed that can reproduce the spine in more detail,[8] these more complex models require more complex marker sets to function and have not found widespread use in live dance involving MoCap. The larger number of markers required increases the visibility of the MoCap system (as well as increasing the points at which occlusions and noise can occur), which negatively intersects with the drive towards transparency from the MoCap system. Moreover, even though such models provide a closer approximation of the kinematics of the human body, they are nonetheless approximations, and therefore still require decisions on behalf of the designer in terms of how the kinematic chain should be modelled.

As a cumulative effect of these factors, assumptions start to form about the types of bodies for which kinematic modelling is intended. This is namely a body which possesses all of the body parts required by the model, can reach poses and perform motion patterns which are recognisable to the model, and have a body who's nuances of form and motion can be represented through the simplifications required. This is not to say that bodies which do not meet these requirements cannot be modelled, but that the range of models widely available in commercial MoCap systems must be modified and adapted or a new model must be created from scratch. This is not a simple process, often hidden behind a barrier of knowledge of kinematics, mathematical representation, and computer programming. This also applies to those who possess a body which does

[7] As is the case, for example, for the skeletal modelling algorithm found in *Optitrack Motive* (OptiTrack, 2022).

[8] For example, the IfB-marker set developed at ETH Zürich makes use of a large number of markers positioned on the back in the location of individual vertebrae and can be used to model the spine of a person with scoliosis (Zemp et al., 2014).

fit the mould for the generalised models if they wish to adapt a model to create a representation that they think better fits their own conception of themselves, for example by concealing a part of their body. The result of this is the development of specific ways of working with the software based around these ways of conceptualising the body. These propagate through any other technological components involved in the MoCap, for example in determining the types of cameras used, their placement, and processing applied to minimise occlusions and marker noise, as well as out into the physical environment and the physical bodies involved.

This is a single example based upon a single technique. However, each component involved in the MoCap process introduces its own assumptions about the body, from sensing systems to systems used to reproduce the captured or tracked data which propagate outwards, intersect, and contribute in a complex manner to the construction of the virtual body. Moreover, all actors involved in the process, from the technologies' designers, to the users of the system, to performers who are placed in relationship with the virtual representation, carry their assumptions and contribute to the form of the virtual body.

3.3 Summary

In this chapter, I presented the co-construction model of MoCap in live dance. This is the main theoretical contribution of this thesis, and is intended to serve as a framework for systematically considering actors, sites, and technologies where assumptions and values are embedded into representations of the body constructed through MoCap. This model is intended to serve as a supplement to more systems-oriented perspectives, such as those outlined from the fields of animation, interactive music, and interactive dance. This model was developed along a methodological structure building from a disability studies perspective on the body, taking the use of kinematic modelling through optical, marker-based MoCap as an example. In the following chapter I introduce *The Shapeshifter*, a performance involving an interactive system designed employing the framework provided by this model.

Chapter 4

The Shapeshifter

In this chapter I present *The Shapeshifter*, an interactive performance that was developed within the research-creation context of this thesis. This is centred around an interactive system that builds upon the work done for *Reconfigurations* within a framework informed by the co-construction model presented in chapter 3.

This chapter begins with an introduction of the collaborator in the research-creation project, describing their background, interests and skills, and role within the project, and a description of the performance space in which development took place. Accordingly, moving forward I switch to the plural *we* when describing elements in which they were involved.

After this, I detail the concepts behind the performance that we developed at the projects outset, as well as the guidelines for development and the technologies, modalities, and ideas that we wished to explore. Following this, I present an overview of a performance.

4.1 The Research-Creation Collaborator

The collaborator in the research-creation project is a performer who has experience and background in a variety of performance artforms, ranging from acting, to dance, and music. This is reflected in their primary interest in physical theatre, a form of theatre focused on the body with roots in both avante-garde theatre and dance (Sanchez-Colberg, 1996). They additionally have an interest in scenography and these two interests have provided them with a perspective focused on the relationship between the physicality of the body and how this relates to the space in which the body is located. They also work with vocalisation, relating this to their focus on the physicality of the body. They primarily describe themselves as an actor, however they emphasise the fluidity in descriptive terms as being more useful to those being provided with a description of their work. As a result, context dependent, they describe themselves as an actor, dancer, performer, or performance artist. They stress, however, that the application of varying descriptors has no influence on the way in which they perform.

They currently primarily work as part of an ensemble at a local theatre which specialises in physical theatre. We did not have contact prior to the start of this project in January 2023, and were connected by my thesis supervisor. Prior to this project, they had little prior experience of working with interactive systems and no experience of working with MoCap. Their expectations of MoCap were therefore primarily shaped by popular depictions of the technology, mainly its use for performance capture for blockbuster films. Although they understood that this project would likely be quite different to those contexts, their expectations of the technology were still mainly shaped by performance capture and animation use-cases, namely that there would be an animated figure rigged with data that they provide. They also saw their connection to the project as heavily relating to their use of the voice in their work.

4.2 The Performance Space

As I will describe in section 5.1, the methodology employed in the development of *The Shapeshifter* performance and design of the interactive system took a non-generalised approach. While this implies that the system is being designed for a specific performer for a specific performance, this also applies to the physical space in which the work is performed. The aim is not to design a system that can be easily and quickly transported between performance spaces, but which fully utilises and leverages the space for which it is being designed. In addition, the physical properties of this space play a role in interaction with the MoCap system, as well as determine aspects such as the physical location of the performer and audience and the position of objects which draw their focus.

The performance space for *The Shapeshifter* is the *MCT Portal*, shown in figure 4.1, located at the Department of Musicology at the University of Oslo.[1] The *Portal* is a laboratory designed to enable the artistic and research-focused exploration of physical-virtual communication with a specific focus on music. To this end, it is outfitted with the following equipment:

1. A *Midas M32* mixing console[2]
2. A main stereo speaker array consisting of:
    (a) Two *Genelec 8030C* speakers[3]
    (b) A *Genelec 7050B* subwoofer [4]
3. A spatial audio array consisting of eight *Genelec 8040B* speakers[5]
4. An *OptiTrack* optical MoCap system with eight Flex 13 cameras[6]

[1] https://www.hf.uio.no/imv/english/about/rooms-and-equipment/mct-portal/
[2] https://www.midasconsoles.com/product.html?modelCode=P0B3I
[3] https://www.genelec.com/8030c
[4] https://www.genelec.com/previous-models/7050b
[5] https://www.genelec.com/8040b
[6] https://optitrack.com/cameras/flex-13/

Figure 4.1: The *MCT Portal* prepared for a performance.

5. A video wall consisting of four *NEC MultiSync X555UNV* arranged in a two × two grid[7]

A floor plan of the *Portal* along with the relative position of the installed technological systems is visible in figure 4.2.

The *Portal* possess dimensions of 10.3 m in length, 4.9 m in width, and 2.8 m in height. A control booth of size 1.9 m × 3.2 m is situated at one end of the room, meaning that possible performance space comprises 8.4 m × 4.9 m.

The spatial audio system utilises an octagonal arrangement of eight speakers, each positioned at a radial distance of 2 m from the origin. The speakers are located at 45° intervals around the circumference of the array, with a 22.5° offset from a reference line pointing directly at the video wall of 0°. The origin is situated 4.5 m from the video wall and 2.3 m from the south wall, and the array is mounted on a truss at a height of 2 m. This allows for unobstructed motion through the performance space. However, this means that the speakers are positioned above head height.

The cameras for the *OptiTrack* system are likewise mounted upon the truss at the positions specified in figure 4.2. Due to the large number of items mounted upon the ceiling truss, as well as due to the truss itself, markers located above a height of approximately 2 m are often occluded and not reliably captured by the system. In addition, there is an area extending approximately 2 m in front of the video wall in which markers

[7]https://www.sharpnecdisplays.eu/p/dc/en/products/details/rp/x555unv.xhtml

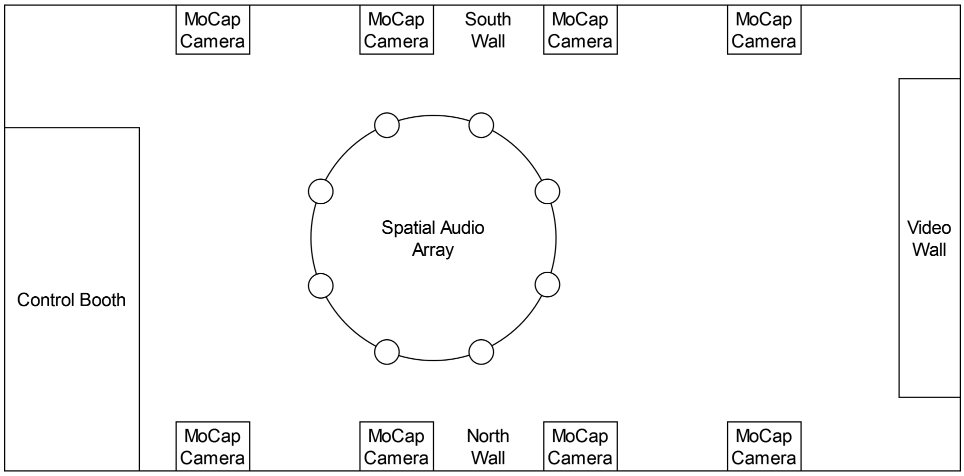


Figure 4.2: A floor plan of the *MCT Portal*. Included are also the relative position of the permanently installed technologies.

are not reliably captured by two cameras, especially if they are occluded from the view of the majority of the cameras by an object such as a human body.

For performance, the entire area of the *Portal* in front of the control booth was cleared and a large dance mat of 1.6 m × 7.5 m is placed along the centre of the room. This aids in reducing marker noise. A rectangle of approximately 1 m × 1.5 m is marked around the origin of the spatial array as the area for the audience. Audience members are free to stand or to sit. This has the capacity for four to six audience members depending upon configuration.

The north wall of the *Portal* consists of a row of windows. Black, molton stage curtains are drawn across these when the MoCap system is in use. In addition, the rooms lights, which are on a dimmer switch, are dimmed to their lowest possible level.

The technical setup of the *Portal* was radically altered in July 2023. As the system component of *The Shapeshifter* was designed specifically for the room in the state described above, this precluded the use of the system until a number of components and modules had been altered. This resulted in an additional limitation placed upon the work done in this thesis, which is discussed in more detail in section 8.3.

4.3 The Performance Concept

The concept for *The Shapeshifter* grew out of work done during development on *Reconfigurations*. *Reconfigurations* explicitly emphasised the collaborative element of constructing a body between multiple performers and the role that the observer plays in the co-construction of the body. With *The Shapeshifter*, we aimed to create a work for a single performer in order to explore the reflexive relationship between the performance and the technology in the process of co-constructing the body. We aimed to place the per-

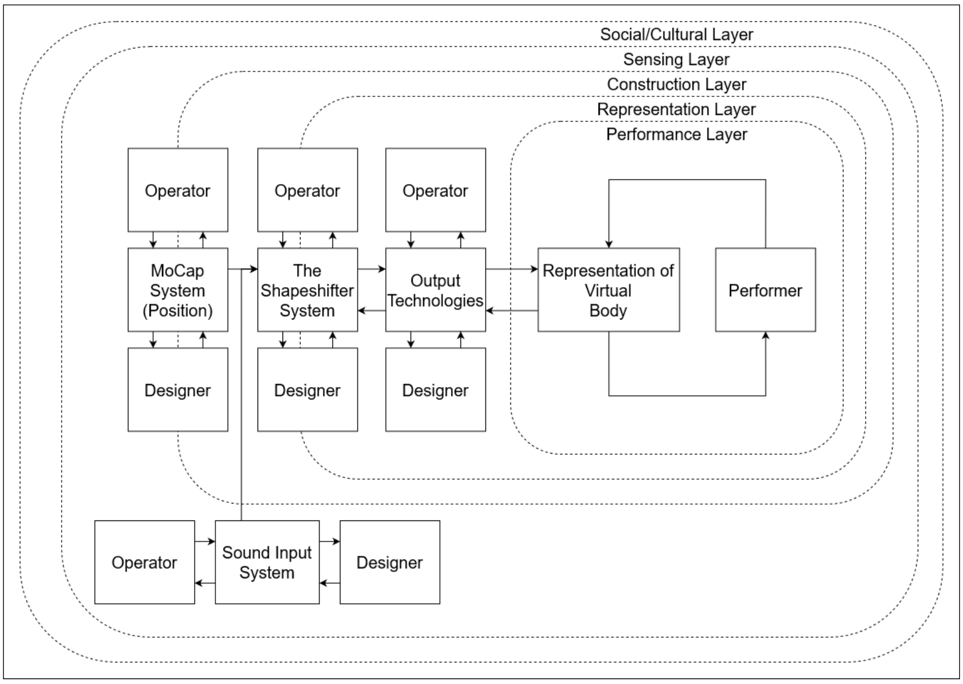


Figure 4.3: A high-level overview of *The Shapeshifter* system in view of the co-construction model.

former into a relationship in which they are cognisant of the ways in which the technological components of the performance are shaping the forms of the body that are constructed, at minimum within the representation layer, and additionally foreground how this feeds back into the performer's motion.

As we would again be working with the *Motive* MoCap system installed in the Portal, and therefore the processing of raw sensor data into motion data would be closed to us, we primarily focused upon the higher levels of the construction layer and the representation and performance layers through the development of a performance system that would sit on top of the MoCap system. We conceptualised the relationship between these elements and ourselves during a performance, applying the co-construction model, as shown in figure 4.3.

Applying the guideline for *Motion Pointillism* as described in section 2.4.2, in a similar manner to *Reconfigurations* we decided to solely work with the position data provided by the MoCap system. Any definition of relationships between the dimensionless points of position would be defined within the performance system. In this way, we would position the performance system as encompassing a part of the construction layer, as well as the representation layer.

We settled upon several concepts that we wished to explore with *The Shapeshifter*. Firstly, within the performance system, and again at minimum within the representation layer, we wanted to explore multiple instances of the relationship between the performer and the virtual body. However, we didn't want these to be static throughout the performance and instead actualise in various, unpredictable forms.

Additionally, we aimed to explore the relationship between the improvisatory and the choreographed. We framed this in terms of conceptualisation of the markers used by the performer and the tendency to elide these as stand-ins for parts of the body. Through this lens, constraints and limitations placed upon marker positioning are directly transferred as constraints applied to the body of the performer. We related this to ideas of objects as a means of constructing constraints in dance performance. A short review of these ideas is available in appendix B. In view of this, we aimed to create a context within *The Shapeshifter* in which it is explicitly markers that must follow a choreography, thus through extension applying this choreography to the virtual body being constructed. We aimed to contrast this with an improvisatory nature to the performer's motion. If the markers are conceptualised as stand-ins for parts of the performer's own body, however, the performer then likewise becomes a body following the pre-determined choreography.

These ideas were also reflected in our aims for the visualisation provided by the performance system. Prior to the start of the development process, the collaborator watched several videos of the *Reconfigurations* system in use and we experimented in a recreation of the system which rendered the visualisations in a fully 3D environment. The collaborator noted that this reminded them of works by Oskar Schlemmer, a central figure in the Bauhaus movement who worked across a variety of disciplines including dance, in particular *Pole Dance*, a work in which a dancer is affixed with long white poles, which not only extend the body but also create imbalance and restrict its motion (Lahusen, 1986). Building upon this, we aimed to take inspiration from Schlemmer's thinking in the design of the performance system component and how this relates to its role in the construction of the body. Schlemmer's work relates the human body to space specifically through the use of costuming. It is with costuming that the human form can be transformed into the *Tänzermensch* (man as dancer). This transformation was illustrated by Schlemmer through four methods of relating the body to space, shown in figure 4.4. A short review of Schlemmer's ideas on costuming and its function to transform the human to the *Tänzermensch*, as well as technological re-imaginings of his work, is available in appendix C.

In light of this, we aimed to view the visualisation component of the performance system as a form of costuming for the virtual body that was being co-constructed within *The Shapeshifter* through thinking in terms of the methods outlined in figure 4.4. Working with the guidelines provided by the *Motion Pointillism* approach, we would provide each of the instances of the virtual body that we aimed to construct with a distinct geometric shape for each point, a distinct colour, and moreover, a distinct form of motion through limiting degrees of freedom of the representation of individual markers' motion. This would, in effect, restrict the way in which the constructed body could move and would allow us the opportunity to examine how this materialised in the way in which the performer experienced their own motion and body.

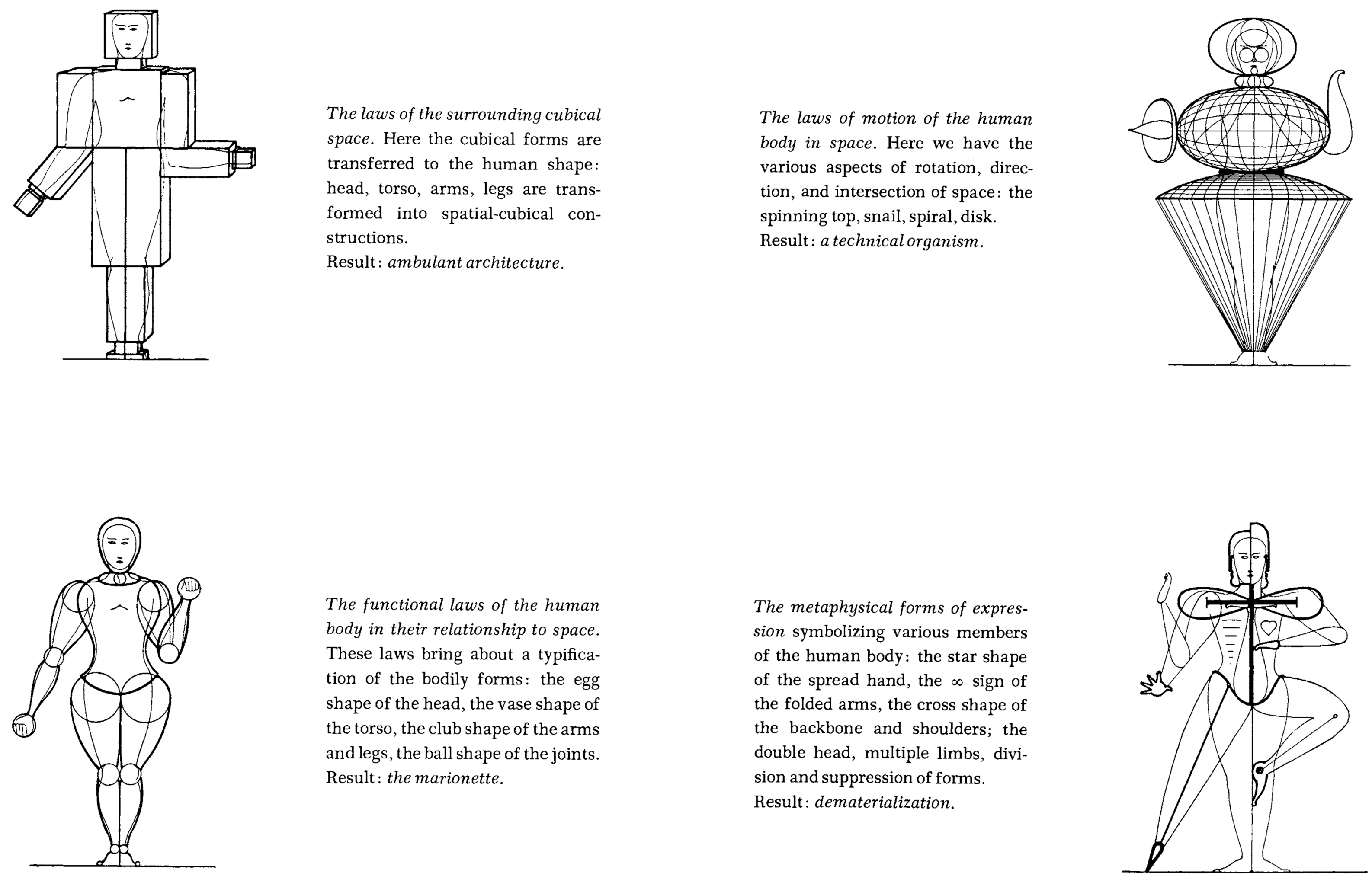


Figure 4.4: The methods fundamental to the transformation of the human body into the *Tänzermensch*. From Schlemmer (1987).

Finally, we also wished to expand the modalities upon which the performance centred from just motion and image. With the collaborator's interest in vocalisation, we aimed to include both the performer's voice and a sound output modality from the system in the performance.

4.4 An Overview of a Performance

In this section I provide a description of a pilot performance of *The Shapeshifter* that took place in May 2023 at the University of Oslo.

The Shapeshifter is an improvisatory dance work for a single performer. Prior to the performance, the performer positions up to 30 position markers wherever they please, either on the body, attached to other objects, or placed within the environment. During the performance, they are also free to reposition these whenever and wherever they wish. A performance consists of nine phases, during each of which the performer improvises a motion pattern and accompanying vocalisations, shown in figure 4.5. To trigger the end of a phase, each of the position markers must be located within a corresponding space in the physical performance area. Each phase presents a different visualisation style both for the virtual representation of the position marker and any connections drawn between markers as well as distinct limitations on how the visualisation of each marker can move. At the end of the nine phases, the cycle begins again. During the second run-through of the phases, the performer's vocalisations for each phase from the previous run-through are looped within the corresponding phase in the current run-through. Starting in the third run-through of the motion phases, the representations of the markers and connections and their motion limitations begin to shift, interpolating between combinations of the representations of all nine phases, shown in figure 4.6. The interpolation is based upon several factors, relating to the similarity of the performer's motion and vocalisations to the motion patterns and vocalisations performed in the previous run-throughs. Likewise, the looped vocalisations begin to twist and distort away from the original recordings. When the performer vocalises during the current run-through, both the audible and visual representations are pulled back into their original state from the first run-through. As the number of repetitions increases, it becomes increasingly difficult for the performer to purposefully control the representations, building to a climax in the seventh and final run-through of the nine motion patterns.

A performance takes place with the performer facing a video wall which mirrors the physical capture volume with a virtual capture volume. The audience is also positioned within the performance space, with the performer moving around the audience. To support the idea of the audience being within the performance space, the looped vocalisations are played back over a spatial audio system in two manners. The first is an underlying sound bed that slowly envelops the performance area over the course of the performance. The second positions each vocalisation at the position of the performer at the time that the vocalisation was recorded.

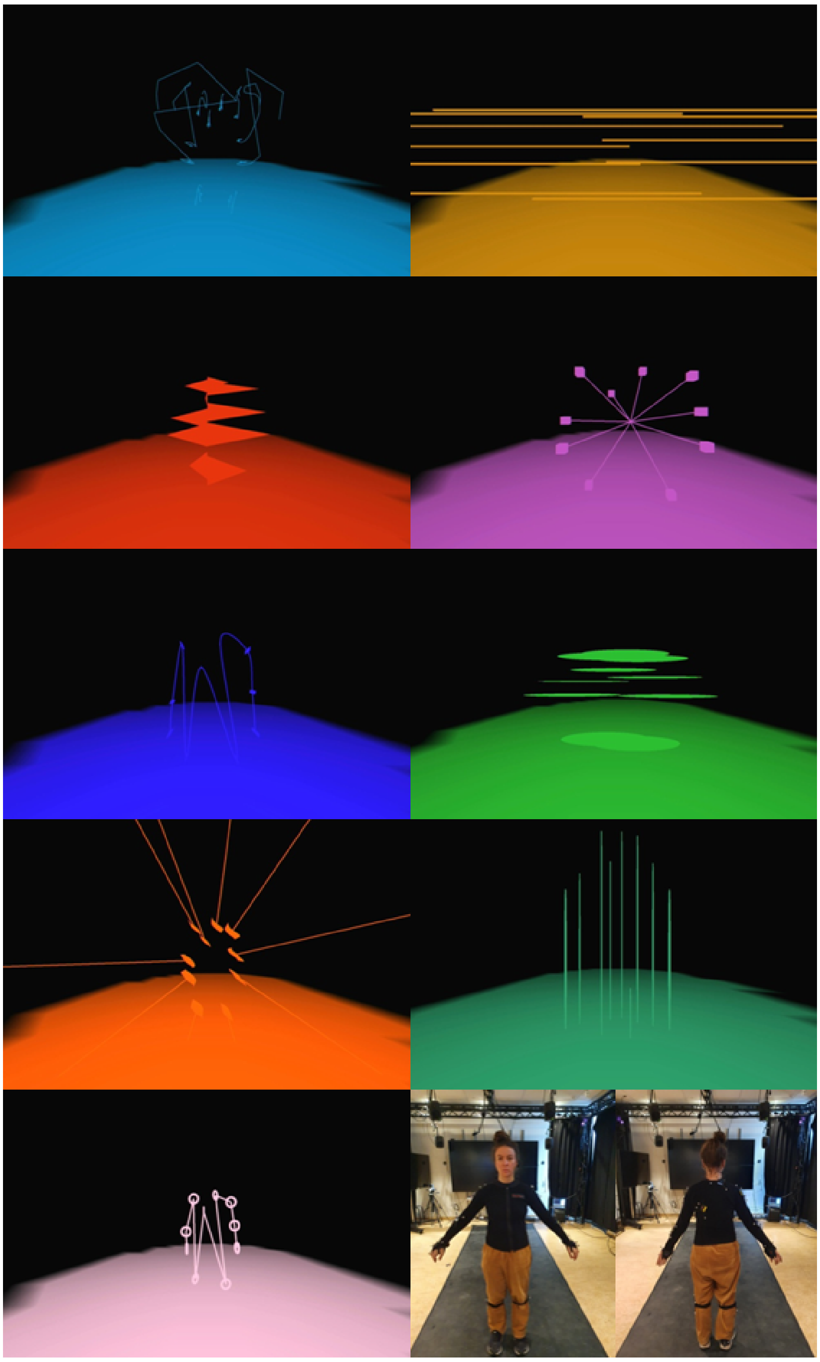


Figure 4.5: The initial visualisations of each of the nine motion phases for *The Shapeshifter*. These comprise a distinct style for representing each marker, sometimes including connections to other markers or points in space. These connections are not just calculated spatially, for instance the first phase creates a spline that passes through the positions where the marker was located across the previous few seconds of the performance. The final image shows the collaborator with the locations of the markers, all of which were positioned on their body.

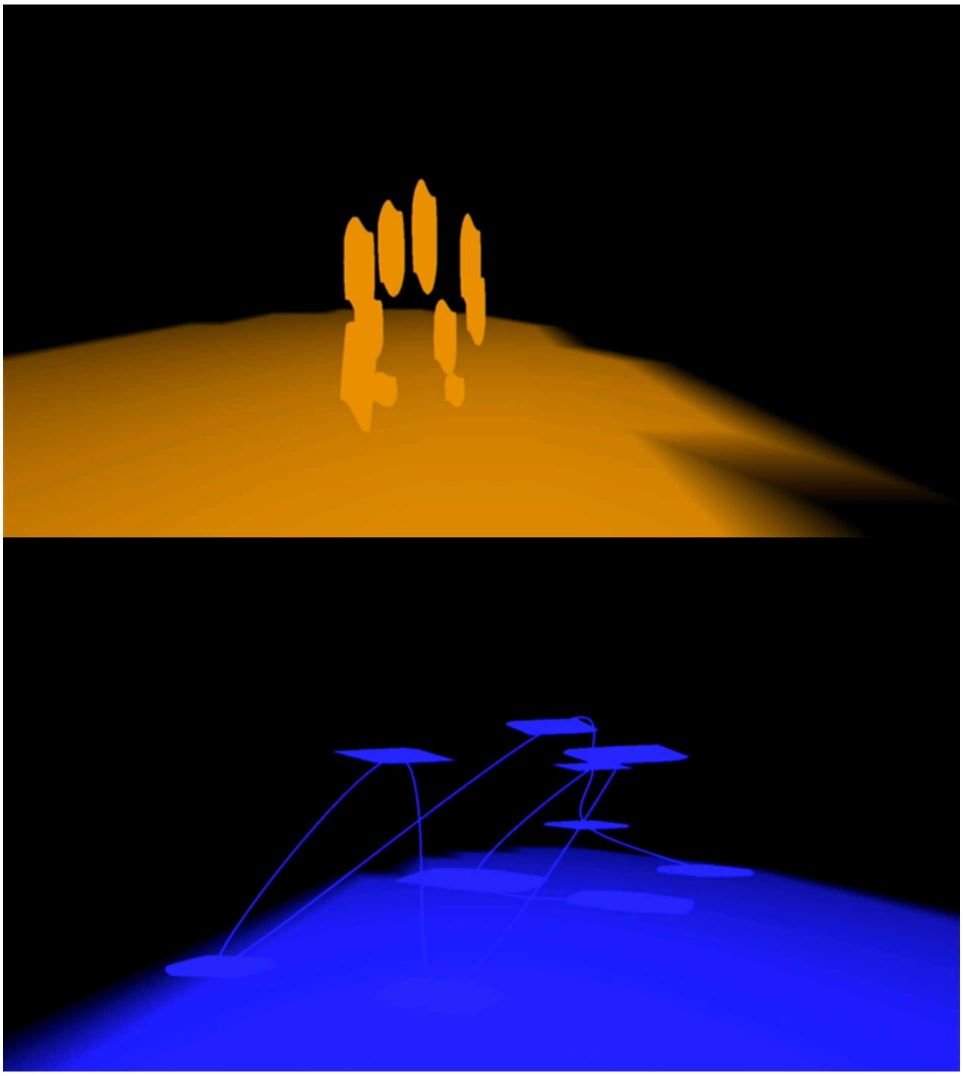


Figure 4.6: Two examples of the representations shifting to interpolate between the initial visualisation styles. This is based upon a combination of the motion patterns and vocalisations performed by the performer over the course of the performance. The amount of interpolation is also reactive to the performer's voice in real-time, providing an effect of a constantly shifting form. The interpolation also applies to the motion patterns of visualisation and parameters of the looped sounds.

The playback shifts between the two techniques based upon whether the performer is currently vocalising. The first technique is used to playback the looped vocalisations with the parameters of the current interpolation applied. The second, to reproduce the vocalisations in their original state.

4.5 Summary

In this chapter I provided an overview of *The Shapeshifter*, a performance and interactive system that forms the basis of the rest of the thesis. I firstly introduced the collaborator in the research-creation project, a performer who specialises and has a strong interest in physical theatre, and the performance space. After this, I introduced the concepts that we wished to explore with the work, namely the application of the guidelines for *Motion Pointillism* introduced in section 2.4.2, the idea of the motion capture marker as an object of constraint for the performer, and the grounding of the performance system and visualisation design in Oskar Schlemmer's ideas on costuming as a method to transform the human body into the *Tänzermensch*. I related these ideas to the co-construction model. Finally, I provided an overview of a performance of *The Shapeshifter* in its state for a pilot performance that took place in May 2023.

Chapter 5

Methods

In this chapter I outline the methods applied to the design of the *The Shapeshifter*, an interview on the collaborator's experience that took place after the pilot performance, and the evaluation of the system's motion to visual latency.

5.1 Design Methodology

Development of *The Shapeshifter*, both the system and the performance, was structured around two closely related methods, iterative design and participatory design.

Iterative design is a method that involves a series of design cycles or iterations. It originated in the first wave of Human-Computer Interaction (HCI), which focused on "human control of computational mechanisms[, . . .] concerned with the best way to design input and output affordances to facilitate effective human/machine couplings" (Gunkel, 2018, p. 12). A critical aspect of the original iterative process is that users rigorously evaluate each iteration, and the results inform future iterations (Nielsen, 1993). This emphasis on user evaluation highlights the close relationship between iterative design and user-centered design, for which, as noted by Gould and Lewis (1985), iterative design is one of fundamental principle. This presupposes a focus on the user and the task.

This focus stands in contrast to the current third wave of HCI, which is "concerned not with the capabilities or operations of the two interacting components—the human user and the computational artifact—but with the phenomenon of the relationship that is situated between them" (Gunkel, 2018, p. 13). This difference in epistemology is encapsulated by Kaye (2009) within the terms *task-focused* and *experience-focused HCI*. He frames his distinction of these terms in relation to human practice, noting the difference in ontology between these two approaches towards what human practice consists of. He notes that task-focused approaches see human practice as the sum of component parts. This rationalist approach not only assumes that these constituent parts are in and of themselves rational, but also that those observing can understand them in a similar manner. In contrast, the experience-focused approach sees human practice as complex and indivisible, therefore requiring a holistic approach to its understanding. Kaye additionally notes the differences in the practices that determine directions of research and/or creation. Under task-based approaches, these practices usually stem from the identification of a problem that requires a solution.

The use of iterative HCI methodologies that reframe the process of evaluation away from delineating a problem and attempting to solve it have found much ground in fields concerned with interactive performance, a specific context in which it can be especially difficult to frame the development of a system in terms of problems to solve. This is especially pertinent when the system to be developed is intended for a specific performance, rather than a generalised performance tool to be integrated into artistic performance by a wide range of artists. For example, Fdili Alaoui (2019), drawing on Bardzell's (2010) feminist HCI and Blythe's (2017) depiction of HCI as a narrative centred on *overcoming the monster* of a problem or challenge, frames her description of the iterative development of the interactive dance work *SKIN* not in terms of the solutions to the problems of the dancers and choreographer that the interactive components solve, but rather in terms of "the trade-offs, the decisions, the tensions and the negotiations that emerged from integrating technology in art" (p. 1197), explicitly framing her work as anti-solutionist. She calls for an embrace of the messiness that is inherent in artistic practice within wider HCI methods and practices.

Closely related to iterative design is the participatory design approach. As outlined by Hansen et al. (2019), in reference to Bjerknes et al. (1987), the participatory design approach is rooted in a political ideology of "democracy, empowerment, skilfulness and quality of process and product" (p. 2). Making reference to Simonsen and Robertson (2013), Hansen et al. characterise participatory design as a process of "investigating, understanding, reflecting upon, establishing, developing and supporting mutual learning between multiple participants in a design process" (p.2).

Analysing participatory design through a lens of program theory, Hansen et al. identify five main activities that take place within the process of participatory design.[1] These are field studies, workshops, prototyping, infrastructuring, and evaluation. Hansen et al. note, however, that a key distinction of participatory design in relation to alternative approaches such as co-design, contextual design or user-centred design, is that these activities result not just in the production of design artefacts but also result in intangible products such as new knowledge, procedures, and organisational arrangements. In light of this, participatory design fits well with the wider research-creation context framing the broader project.

Participatory design has been applied in a context of interactive dance, for example, by Brenton et al. (2014), Fdili Alaoui et al. (2013) and Landry

[1]As noted by Hansen et al., in program theory a common approach is to identify the input, process, and effects of a specific program. They go on to further subdivide the process component of the program into mechanisms, "the general underlying principles that generate effects" (p. 3), and activities, "the particular way or the medium through which the mechanism is brought into action" (p. 3).

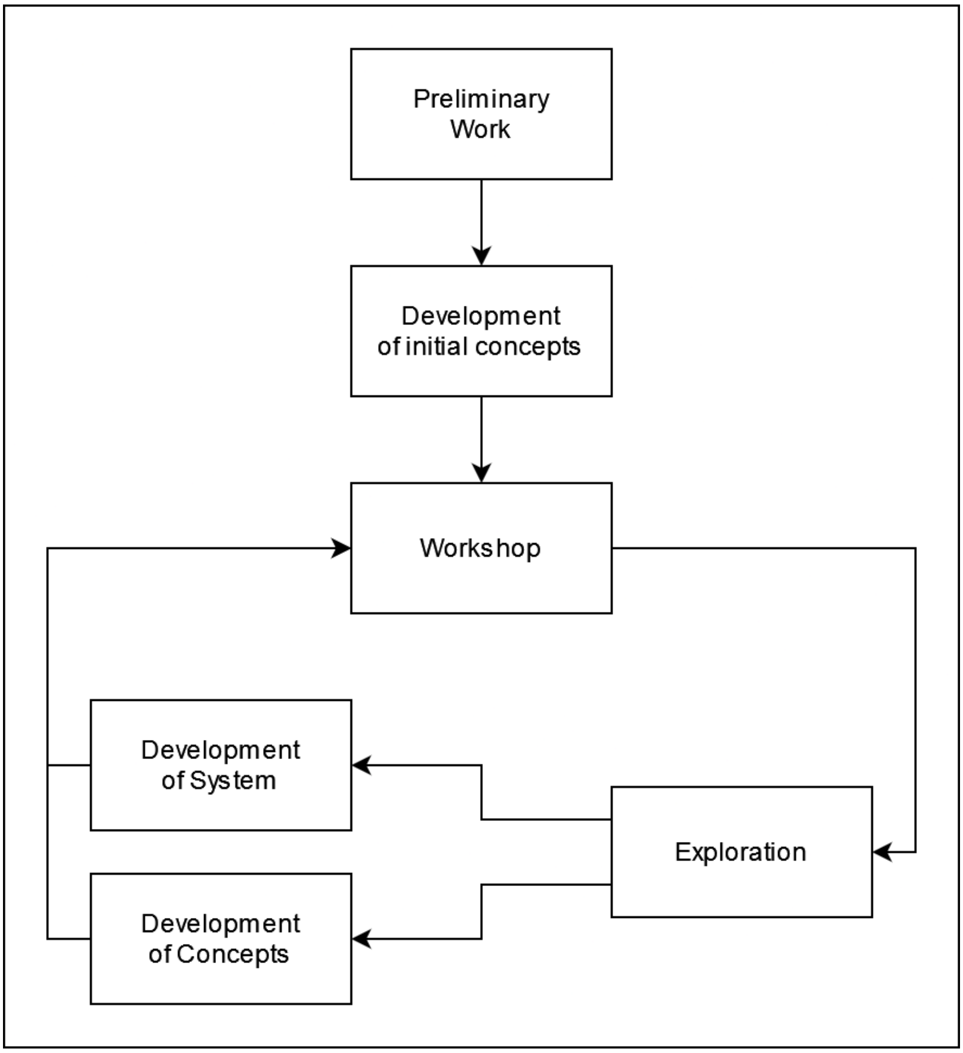


Figure 5.1: The conceptualisation of the iterative development process of *The Shapeshifter* performance, encompassing an iterative and participatory design of the interactive system.

and Jeon (2017). Masu et al.'s (2019) approach towards developing a set of guidelines for designers of interactive systems for dance is likewise framed in a relationship with participatory design, with the authors conducting workshops and interviews.

In view of the above, we structured our iterative process as depicted in figure 5.1. Based upon the the preliminary work and initial concepts detailed in chapters 2 and 4, we decided to structure the development process around a series of workshops and explorations. Meeting semi-regularly between January 2023 and May 2023 over a course of nine sessions ranging between two and six hours, we conceptualised each session as consisting of workshopping ideas that we had developed since the previous meeting and reflecting on the process as well as the emerging collaborative relationship. Due to the exploratory nature of developing the performance, the anti-solutionist stance with which we approached the system design, and the fact that the performance was being developed simultaneously with the technological components, instead of performing a formalised evaluation of iterations we framed the second half of the sessions as explorations, consisting of synthesising conceptual ideas that had been developed and system modules that had been designed since the previous meeting in view of the workshopping that had just taken place. As noted by Hansen et al., systematic and formal evaluation is rare in participatory design and, taking this within an anti-solutionist context, we therefore took a stance framing the evaluation as an exploration around which we could further orient development of the performance. The explorations continued after the session, where working separately we would develop concepts and system components in preparation for the next session. From the outset, we also decided that we would not develop the system for general use.[2] We viewed the system as an integral part of the performance, and therefore were not explicitly concerned with designing the system as a framework which other artists could incorporate into their work. We therefore did not preclude the incorporation of system components highly tailored to the collaborator's way of working, such as the use of machine learning techniques employing a personalised dataset (see section 6.3.3.3).

Data was collected at each session through a number of methods. These included observations and notes, written questions filled out by the collaborator, personal reflections, and video recordings of rehearsals of the performance. These methods were supplemented by informal discussions that took place throughout the process. Up until the sixth session, each session followed a similar pattern. I would come to the *Portal* an hour or two before the session was scheduled and set up the environment. This would include creating the physical performance space, as the laboratory is used for multiple purposes, and setting up and testing the system. When the collaborator arrived we would start with the workshop, involving a catch up on status, and share ideas and concepts that we had developed between sessions. Thereafter, we would run through a performance in its current state several times and explore possibilities for building the performance. At the end of each session, we would debrief and make a plan as to what needed to be done for the following session. Between sessions, I would work on the system in line with the decisions made in the previous debrief session, as well as test out new ideas that I could present to the collaborator in the next session. The collaborator would consider directions that the performance could take and bring these to the following session.

The results of the iterative process can be found in section 7.1.

[2] Although this doesn't preclude, and we welcome, the system's appropriation for other contexts. The codebase for the system is freely available at the link found in chapter 1.

5.2 Interview Methodology

The broad approach taken for the interview is a semi-structured, life-world interview following Brinkmann and Kvale (2018) and Kvale and Brinkmann (2009). Although this is grounded in a paradigmatic background of phenomenology, focusing on uncovering meanings related to the lived experience of the interview participant,[3] I do not position the interview as employing phenomenology as a method such as found in the broad phenomenological interview described by Creswell and Poth (2018), Giorgi et al.'s (2017) descriptive phenomenological psychological method, or the micro-phenomenological interview (Petitmengin, 2006) for reasons ranging from the drive towards generalisability in experience and thus requirement for a widely experienced phenomenon and multiple participants as required by Creswell and Poth, or the extensive time and training requirements for the performance of a micro-phenomenological interview.

However, I draw on several concepts from various phenomenological methods, as I believe these aid in keeping the interview focused on the experience of the performance. The first is the form of bracketing drawn from Giorgi et al. in which "we do bracket the co-positing of the physical or any other objectivities that are given thematically to consciousness. But the horizonal space-time world is not bracketed [...] the acts of consciousness are not bracketed and are considered to be real, but the thematic objects of consciousness are reduced even if the worldly horizon is not" (p. 181). I interpret this to imply a bracketing of thematic objects not directly related to the performance itself, in other words attempting to frame the interview in such a way that the performer does not draw comparisons to other phenomena or employ the use of allegories or metaphors, but that they still feel comfortable discussing the range of experiences that unfolded over the development of the performance. The second are the four principle phenomenological commitments that should inform the interview as proposed by Høffding and Martiny (2016), namely commitment to "the thing itself: Using the interview to acquire detailed first-person descriptions of an experience in question", "[i]nvariant structures: Using the interview to grasp the invariant structures of experience", "[s]ubjectivity cannot be reduced to objectivity: In the interview, the first-person perspective needs to be understood on its own terms", and "[e]naction, embodiment and embeddedness: Phenomenology construes subjectivity as embodied, enactive and embedded. The interview directly confronts us with these aspects of experience" (pp. 360-61). Høffding and Martiny build upon the concept that the interview itself should be viewed from a co-generative perspective involving the interviewer and participant, an approach well suited to the context of the interview in which the interviewer and participant are working in a collaborative relationship within the phenomenon that forms the topic of the interview.

[3] Or as Brinkmann and Kvale (2018) put it, aiming to "get as close as possible to precise descriptions of *what* people have experienced" (p. 21, emphasis in original) rather than focusing on the *how* found in approaches such as discourse analysis.

Finally, although I am not employing phenomenology as an explicit method for the interview, the paradigm upon which Brinkmann and Kvale's approach is based is still phenomenological in nature. I therefore take note of Gallagher's (2012) indication that an improper framing of questions can lead to obtaining opinions rather that descriptions of lived experience.

5.2.1 Ethical Considerations

A major ethical consideration required navigation in the design of the interview, namely the relationship between myself and the collaborator. As noted by Brinkmann and Kvale (2018), there is an inherent power imbalance between the researcher and the participant in an interview situation, requiring an awareness of several ethical issues. This was complicated in the context of this interview by the fact that the work done in this thesis takes place in the context of a research-creation project, which explicitly frames the participants in the project as collaborators rather than as researcher/study participant. Due to the shared aim of realising the artistic outcomes of the project, a risk emerges that the collaborator on the project feels that they are not able to, or are not allowed to, express thoughts that put the work that I have done in relation to the artistic outcomes in a negative light.

Moreover, my incentives for the project do not completely align with that of the collaborator. Although we are both incentivised by the artistic outcomes of the project, this represents the collaborators main incentive for the project. On the other hand, I am equally as incentivised by the research outcomes of the project, which is amplified by the fact that these outcomes will be used for the submission of a graded thesis upon which my graduation from a master's programme is dependent. As this is known to the collaborator, there is a substantial risk that the collaborator provides me with answers to questions and prompts that they feel that I want to hear, or that they feel will support my arguments, especially in light of the collaborative atmosphere that developed over the course of the project. Even more critically, I might also exert pressure upon the collaborator to answer in this way, even if this pressure may be exerted subconsciously.

Therefore, several measures were added to the interview guide in order to minimise the effects of these ethical issues. Firstly, the interview would begin with a prompt where I clarify my relationship to the project and collaborator and explicitly state that the answers should not be tailored to what the collaborator thinks that I might want to hear. In addition, several prompts were added asking the collaborator to reflect on their role within the project. Moreover, during the interview I would aim to keep a clear distinction between asking follow-up questions leading from the prompts and asking follow-up questions that were indications to the collaborator that they had not answered in a manner suited to my purposes. I would also be aware of body-language or tone of voice that could imply this.

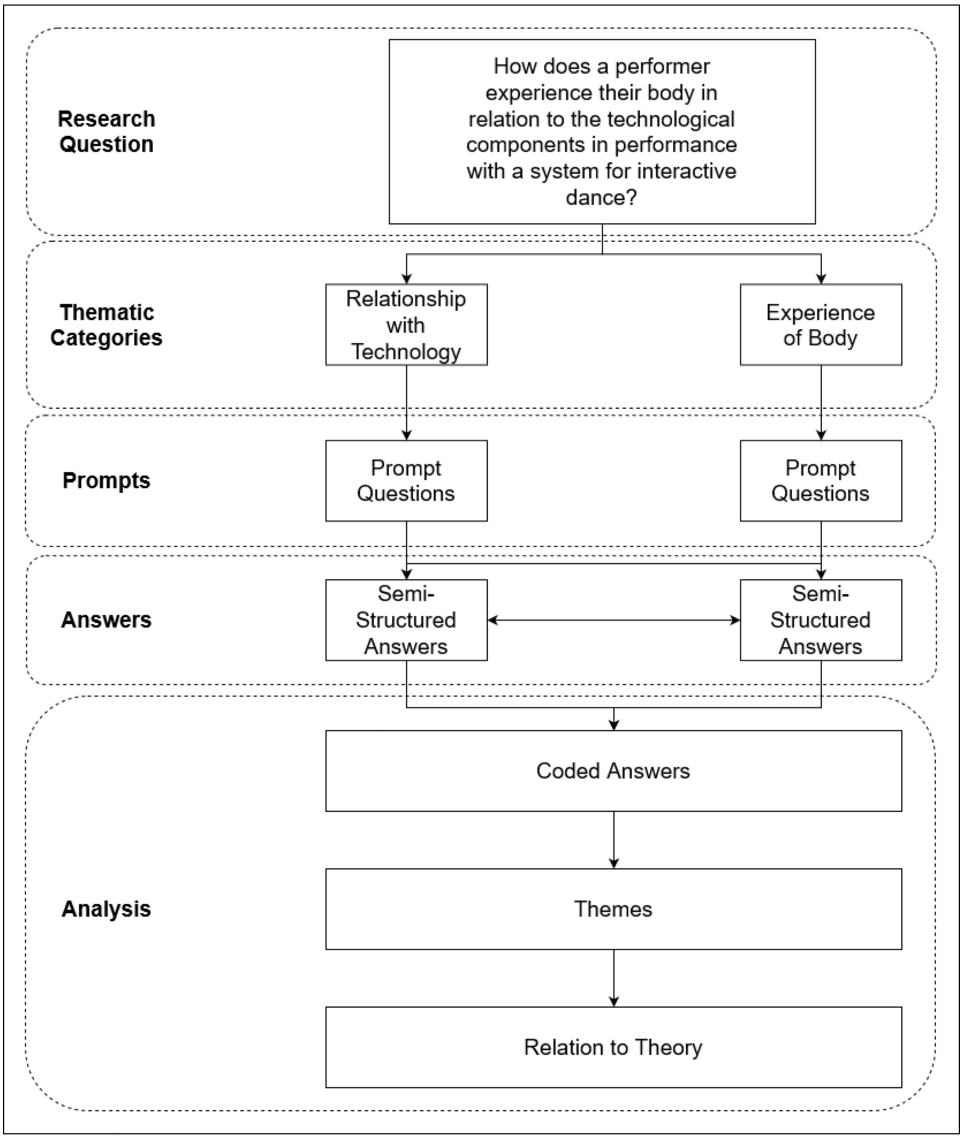


Figure 5.2: The conceptualisation of the interview and subsequent analysis.

5.2.2 Interview Guide and Structure

Figure 5.2 shows the conceptualisation of the interview process. From RQ3, two broad thematic categories were identified. From these categories, a number of prompt questions were drafted which could broadly be allocated to these categories. The prompt questions were not treated as a strict series of questions to be procedurally answered, but rather as a rough structure for topics to cover during the interview. The expectation was that there would be much thematic overlap in the answers provided to these prompts, and that the subsequent directions of the interview which derive from these prompts could transition across the boundaries of the categories. The prompt questions were organised into an interview guide that I could use during the interview, which is available in appendix

G. The interview guide was structured so that the interview would last for a duration of approximately 90 minutes. Audio of the interview was recorded using a portable audio recorder placed upon the table between myself and the collaborator.

5.2.3 Analysis Methodology

The first step of analysis was transcription of the audio recording which was carried out with the aid of AutoTekst, the automatic transcription service offered by the University of Oslo. This transcript can be found in appendix F. Thereafter, I reviewed the transcription while listening to the audio recording in order to correct errors, organise the transcription by speaker, and add timecodes. While performing these tasks, I started paying attention to recurring patterns, meanings, and ideas in the text that could form the basis for codes within the framework of an inductive coding approach. Thereafter, I reviewed the text multiple times and finalised the coding categories. I then coded the text in line with these categories. After coding the interview, I analysed the coded data to obtain emerging themes, contradictions, and juxtapositions in meaning. I then wrote descriptors of these, relating these to a wider embodied perspective as well as the co-construction model. This process is depicted in the lower portion of figure 5.2.

The interview results can be found in section 7.2.

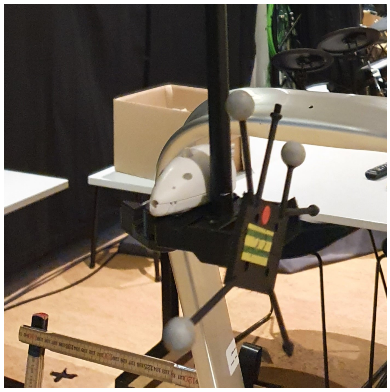
5.3 Evaluation of Motion to Visualisation System Latency

In this section, I provide a brief description of the method developed to measure the system's end-to-end motion to visualisation latency. A more comprehensive description of the method can be found in appendix E. Although a wide range of system metrics can be measured, I choose to focus on motion to visualisation latency, as the collaborator and I observed noticeable changes during rehearsals. These differences in latency contribute to the collaborator's experience of the system, and should therefore be characterised.

There is little in the literature on the measurement of the latency between the motion of a tracking object and a visual representation on a screen in interactive performance. I therefore draw and build upon the method of latency measurement in relation to head-mounted displays as proposed by Steed (2008) and Friston and Steed (2014). Steed's method is based around the use of a pendulum mounted with a light and the tracking object. A camera records the swing of the pendulum and a screen upon which a visualisation of the tracking object is displayed. From this, the horizontal displacement from the mean position of both the light and the visualisation is extracted, and the data are fitted with a sine function. The phase difference between the two sinusoids provides the latency.

(a) The rigid arm used as a pendulum.

(b) The laser pointer.



(c) The mounted tracking object.

Figure 5.3: The pendulum components.

I adapt Steed's method through the replacement of the light on top of the pendulum with a laser pointer which is aimed at the top of half of the screen. The visualisation of a tracking object also attached to the pendulum, shown in figure 5.3, is displayed in the bottom half, shown in figure 5.4. This allows the camera to be placed in front of the pendulum, minimising interference between the camera's view of the screen, bleed from the light placed on top of the pendulum, and the visualisation. In addition, this allows the camera to be placed closer to the screen meaning that the screen

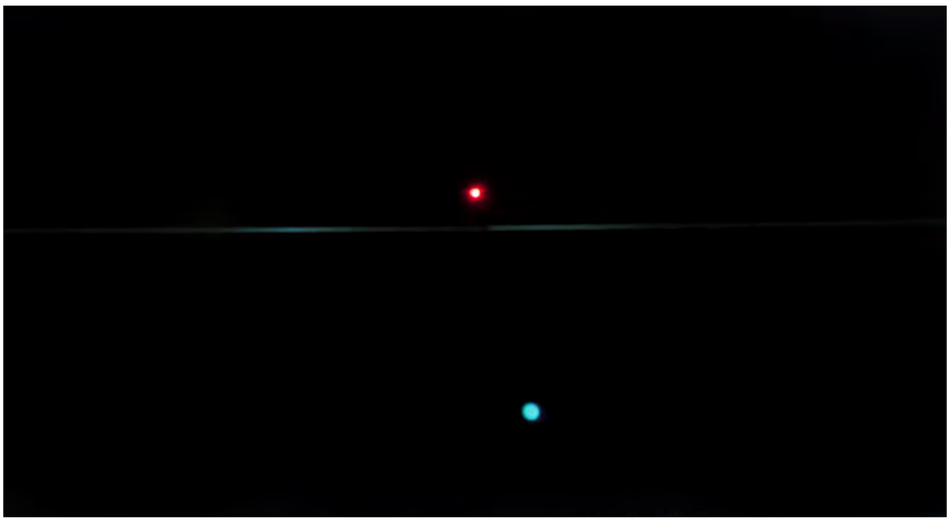


Figure 5.4: The screen as captured by the camera. Top is the laser point, bottom is the visualisation.

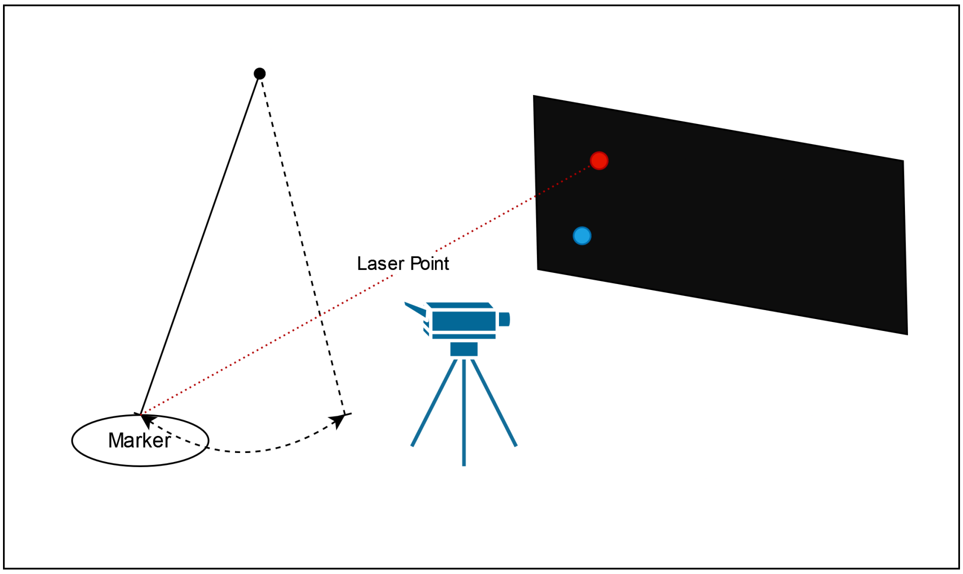
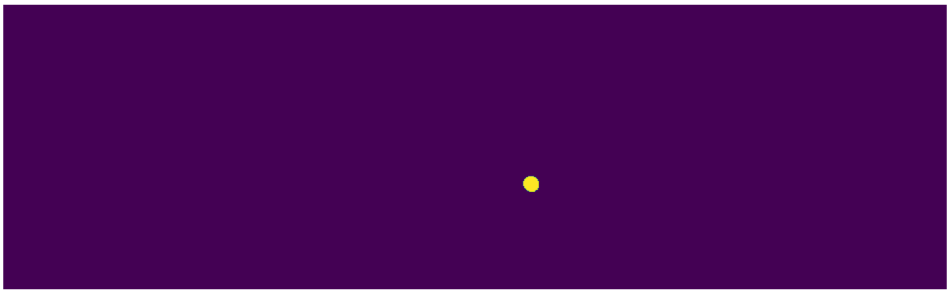


Figure 5.5: An overview of the fundamental setup for the method. A rigid pendulum is suspended with a position marker and laser pointer attached. The laser point is displayed on a screen above the visualisation of the position marker. A camera captures the screen. Not shown are various measurement tools to ensure that the swings are consistent in amplitude and position.

is captured within a greater area of the camera's resolution. This means that the position of the laser point and visualisation will be captured with a higher degree of accuracy. The evaluation set up is shown in figure 5.5.

(a) The laser point.



(b) The visualisation.

Figure 5.6: Processing applied to the video in order to isolate the laser point and visualisation within each frame.

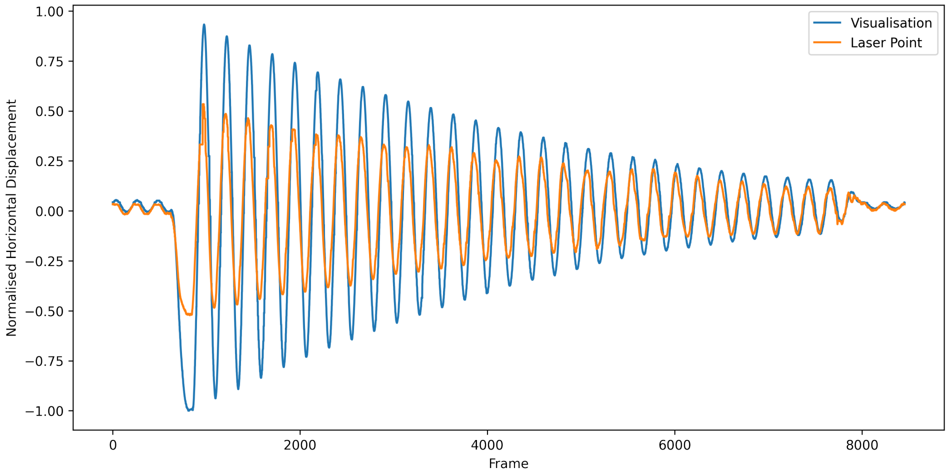


Figure 5.7: The horizontal displacement of the laser point and visualisation.

After performing several processing steps to isolate the laser point and visualisation, shown in figure 5.6, the horizontal displacement from the mean horizontal position is extracted. I omit the fitting of the data with a sine function, as the extracted data are already sinusoidal in form, shown in figure 5.7. This also, for example, enables analysis based upon parameters such as the smoothness of the slope of the sinusoid. Instead of calculating the phase difference between the two sinusoids, I draw on Di Luca (2010) and calculate the cross-correlation of the two signals.

The results of the latency evaluation can be found in section 7.3.

5.4 Summary

In this chapter I outlined the methods employed in this thesis relating to *The Shapeshifter*. These consist of the iterative and participatory methods for the design of the system and the development of the performance, the interview technique grounded in a phenomenological approach, and the quantitative methods to evaluate the system's motion to visual latency. In the following chapter, I present the thesis' first result, namely the design and implementation of *The Shapeshifter* system.

Chapter 6

The Shapeshifter Performance System

In this chapter, I provide an outline of *The Shapeshifter* system oriented around the flow of signals, describing each of the major modules in depth. Thereafter, I provide a description of how the system was implemented. *The Shapeshifter* system takes the motion input in the form of position data provided by the MoCap system and audio input and maps these to visual and audio outputs. These mappings shift across the performance, based upon the cumulative motion and audio input provided.

6.1 Terminology

The following chapter employs terminology to refer to concepts and components of the system that I will clarify to aid in understanding. The term *phase* refers to a single state of the system. Each repetition of the set of nine *phases*, with variations in the mappings and representations, is referred to as a *cycle*. The physical objects which are tracked by the MoCap system are referred to as *position markers*. The performance space in which the performer works is the *physical performance space*. Their counterparts within the system are *virtual markers* and the *virtual performance space*. To trigger the system to move to the next *phase*, all *position markers* must be positioned within corresponding *bounding boxes*, volumes demarcated within the *physical performance space* that track whether a *position marker* is located within them. The mechanism through which interpolation between mappings and representations takes place is through *phase weighting vectors*, which provide the amount of interpolation. *Phase weighting vectors* are determined through the *shifting algorithm*, which updates the *phase weighting vectors* at the end of each *phase*. An outline of the central functionality of the system in view of this terminology is provided in figure 6.4

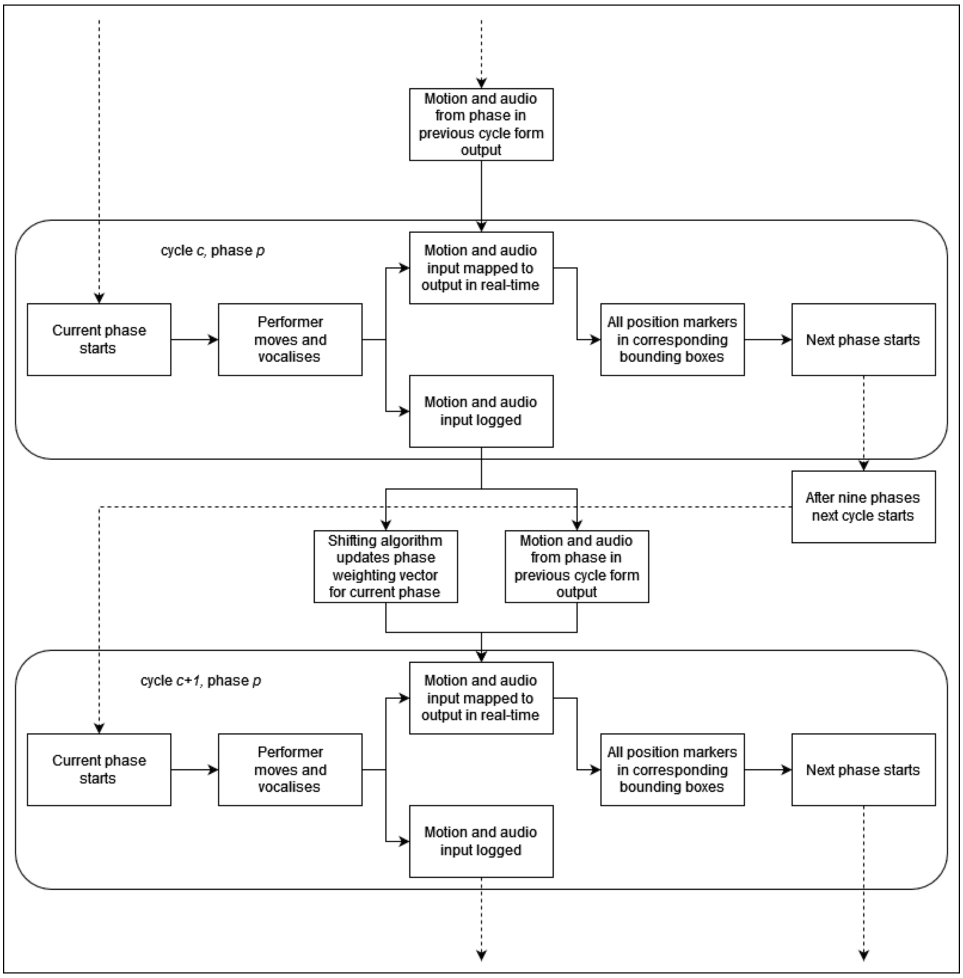


Figure 6.1: A functional overview of the flow of a performance.

6.2 System Overview

The system is designed with a modular approach, incorporating two distinct time scales of operation. Firstly, in real-time, the system maps input signals to output signals within a specific *phase* and *cycle*. Secondly, there are modules that operate offline, logging input signals from a given *phase* in a given *cycle* and processing them once the *phase* has ended. The outputs of the non-real-time modules determine system mappings and outputs in future *phases* and *cycles*. A performance synch module synchronises the system's modules, keeps track of the temporal position of the performance, and ensures that operations are executed and completed at the correct time. An overview of all system components can be seen in figure 6.2.

6.3 Modules

This section provides a description of each of the system's modules. Several modules consist of a number of individual components. If this is the case

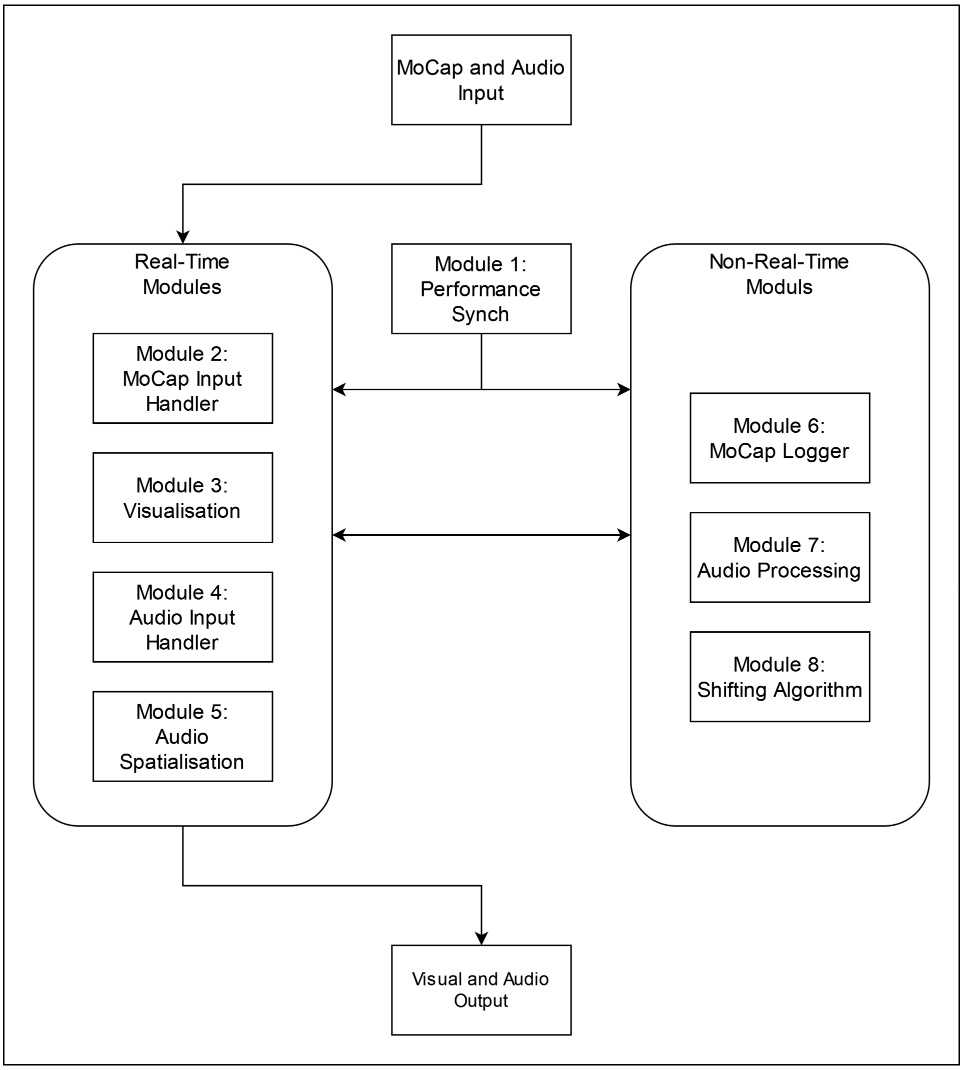


Figure 6.2: An overview of all system modules split between modules that operate under real-time and non-real-time constraints. The performance synch module synchronises the modules and tracks the state of the performance. Arrows signal any and all data-flows, regardless of data type.

for a given module, the components are described individually.

6.3.1 Module 1: Performance Synch

The performance synch module does not process any input signal nor output any processed signal. Rather, it operates purely in the domain of control signals, and acts as the centre of synchronisation for both the real-time and non-real-time components. Likewise, it also ensures that all data processing and output between the output modalities of the system remain synchronised. Upon receiving a control message to start the performance, the module begins to track the *phase* and *cycle*. When the performer manoeuvres all *virtual markers* within the *bounding boxes*, the module sends out a master control message to all other modules to proceed to the next phase of the performance. At the outset of each phase, it sends individual messages to each non-real-time components defining which logged data are to be processed. Likewise it sends individual messages to the real-time components defining which processed data are to be passed to the system's output. Moreover, it tracks which *phase weighting vectors* are to be employed within the real-time components of the system within a specific *phase* and *cycle*.

6.3.2 Modules 2-5: Real-Time Modules

There are four major real-time modules within the system. These are:

Module 2: MoCap Input Handler

Module 3: Visualisation

Module 4: Audio Input Handler

Module 5: Audio Spatialisation

The first module is the MoCap input handler, which relies on the MoCap system to capture the global position data of physical *position markers*. The position data is then post-processed and passed to the second module, the visualisation component. Here, the *virtual markers* and *virtual performance space* are rendered based on the position data, with exact specifications determined by the *phase weighting vectors* obtained from the non-real-time shifting algorithm.

The third module is the audio input handler, which pre-processes the input audio signal captured by the performer's microphone. The amplitude of the input audio signal is mapped to an interpolation between the updated *phase weighting vectors* and the initial vectors defined at the outset of a performance as well as the mix between the two audio outputs of the audio spatialisation module. The interpolated phase weights are then passed to the audio spatialisation module, the fourth and final real-time module. This module comprises two components: the *trajectory* component and the *diffusion* component, which correspond to separate spatialisation techniques.

The audio spatialisation module makes use of audio files and position data that are provided by the non-real-time audio processing and MoCap logger modules. The *phase weighting vectors*, as well as features extracted from the audio input, are also used to map parameters of the spatialisation. Additionally, the real-time component passes the input MoCap and audio data to the non-real-time audio processing and MoCap logger modules. An overview of the signal flows within the real-time modules is shown in figure 6.3.

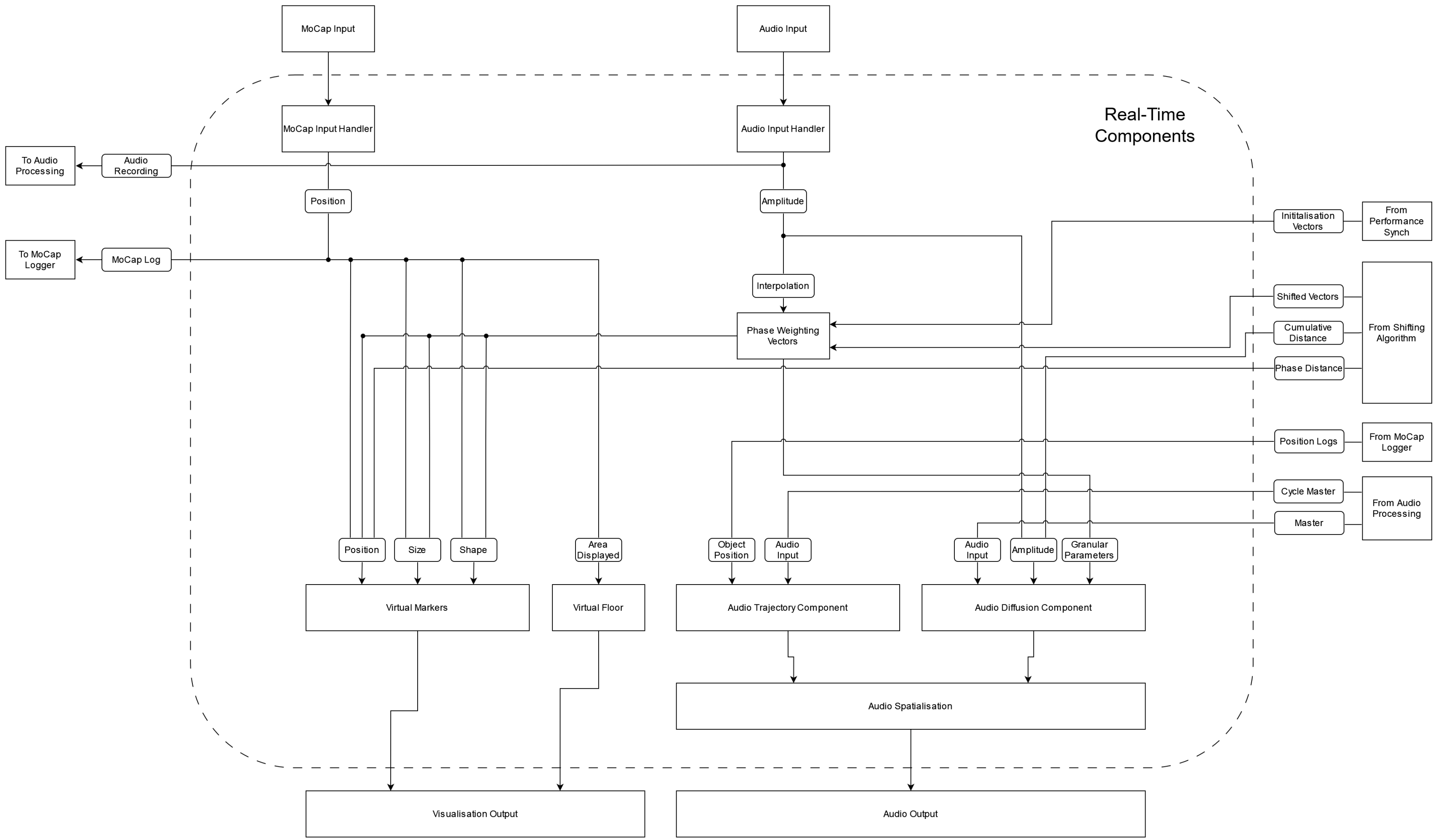


Figure 6.3: An overview of real-time system components. These are the mappings and signal flows that take place within the real-time system components within a given *phase* and *cycle* of the performance. Input signals flow from top to bottom. Signals passed to and from non-real-time modules and the performance synch module flow right to left. Large boxes represent modules or module components. Small boxes describe the signal type.

6.3.2.1 Module 2: MoCap Input Handler

The MoCap input handler module is centred around the MoCap system, with several components that pre-process and route the input. Within the MoCap software, *position markers* are defined and the position data for each communicated to a receiver within *The Shapeshifter* system. Within this receiver, the position data are scaled and have a rotational transform applied in order to match the position of the camera within the *virtual performance space* and are then routed to the *visualisation* and *MoCap logger* modules.

Position Markers The system has been designed to work with up to 30 *position markers*. The exact number of *position markers* to be used within a performance is designated within a calibration phase prior to performance.

Bounding Boxes A core component of the motion input module is the bounding boxes component. The bounding boxes act as the mechanism through which the performer can trigger the transition to the subsequent *phase* of the performance. This takes the form of a cubic volume defined in the *physical performance space* for each *position marker*. For bounding box b_i with centre point $c_i = x_{b_i}, y_{b_i}, z_{b_i}$ and edge length l, the volume in the *virtual performance space* which b_i encapsulates can be formulated as

$$V_{b_i} = x_{b_i} \pm \frac{l}{2}, y_{b_i} \pm \frac{l}{2}, z_{b_i} \pm \frac{l}{2} \tag{6.1}$$

When the *position marker* is positioned within the area of its assigned bounding box, the bounding box is considered active. Considering this as a function of the position $p_i = x_{p_i}, y_{p_i}, z_{p_i}$ for *position marker i* where

$$f(x_{p_i}, y_{p_i}, z_{p_i}) = \begin{cases} 1 & \text{if } x_{b_i} - \frac{l}{2} < x_{p_i} < x_{b_i} + \frac{l}{2} \\ & \text{and } y_{b_i} - \frac{l}{2} < y_{p_i} < y_{b_i} + \frac{l}{2} \\ & \text{and } z_{b_i} - \frac{l}{2} < z_{p_i} < z_{b_i} + \frac{l}{2} \\ 0 & \text{otherwise} \end{cases} \tag{6.2}$$

the transition to the next phase is triggered when

$$\sum_{i=1}^{I} f(x_{p_i}, y_{p_i}, z_{p_i}) = I \tag{6.3}$$

where I is the total number of *position markers*. This functionality is shown in figure 6.4.

The positions of the *bounding boxes* are set prior to performance in a calibration stage. The performer positions the *position markers* in the *physical performance space*, and a separate component logs their position upon the triggering of a command. These positions will be used as the centre point of the corresponding *bounding box* during performance. The edge length of the *bounding boxes* is also set in this calibration phase.

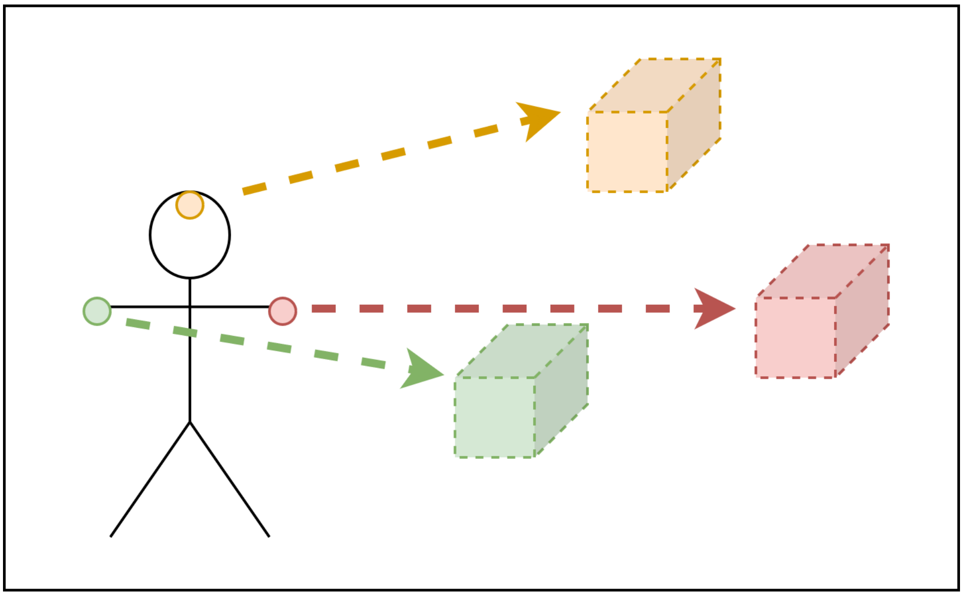


Figure 6.4: The functionality of the bounding boxes. When all *position markers* are simultaneously within their corresponding bounding box, the next *phase* is triggered. *Position markers* do not necessarily have to be positioned on the body.

6.3.2.2 Module 3: Visualisation

The visualisation module comprises one of the primary output modalities of the system. Within this module, both input modalities of motion and sound are mapped to various elements of the visualisation. The visualisation consists of a defined *virtual performance space*, shown in figure 6.5, that is intended to be spatially analogous to the *physical performance space*, mirrored along the axis of the video wall. In this way, the video wall can be seen as reflecting the *physical performance space*. Within the *virtual performance space*, each *position marker* is directly represented by a *virtual marker*, which is the weighted average of a series of geometric shapes, referred to as *phase shapes*.

Each *phase* is delineated by a number of *phase distinctions*. These consist of a colour that is applied to all elements of the *virtual performance space*, and a distinct geometric shape and motion pattern for each *phase shape*. The motion pattern consists of limitations placed upon the mapping of the position data provided by the *position marker*, for example in relation to the degrees of freedom of the *virtual marker*, or the mapping of the position of the *position marker* to the position of the *virtual marker's* vertices. As the *cycles* progress, the *phase distinctions* begin to shift and combine with one another. The amount of shift is determined through the *shifting algorithm*.

The Virtual Performance Space The *virtual performance space*, shown in figure 6.5, is designed to mirror the *physical performance space*. A minimalist design was aimed for in order to keep the focus on the MoCap

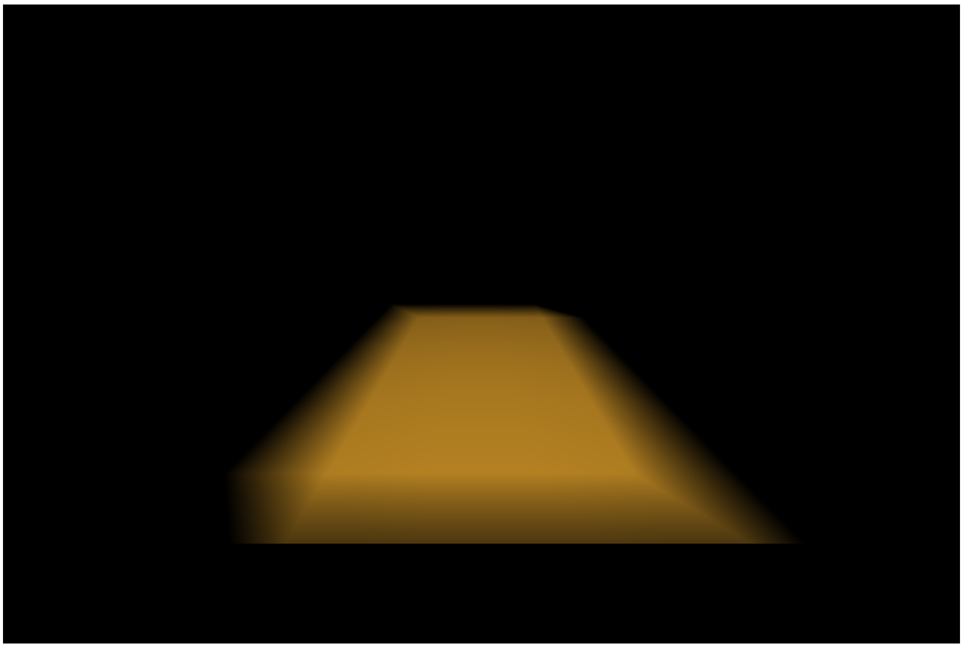


Figure 6.5: The virtual performance space in phase two of cycle one with a black backdrop and a virtual floor, which serves as a frame of reference. The virtual floor consists of a grid of 10×10 square-shaped polygons, and only those within a sphere defined by the centroid position of all virtual markers and a diameter corresponding to the range of the markers are visible.

input's visual representation. The space features a black backdrop and a *virtual floor*, which was added to the performance space in order to provide the performer and audience with a frame of reference. The *virtual floor* is composed of a grid of 10×10 square-shaped polygons, whose vertices' centroids are located at the MoCap system's origin. To reduce computational expense and offer an additional frame of reference, visible polygons are limited to within a sphere that encompasses the *virtual position markers*. This sphere has a center point defined as the centroid of the *virtual markers* and a diameter equal to the range of their positions.

To create a one-point perspective in the virtual performance space, a virtual camera is positioned 1 m above and 2 m away from the MoCap system origin. This effectively situates the MoCap system origin 2 m behind the video wall, which serves as the projection surface. Using a one-point perspective results in some perspective distortion when the performer moves close to the video wall. At the outset of the project, an alternative orthogonal projection was used. However, this projection was ultimately abandoned in favor of the one-point perspective, as it lacks the ability to visualise depth, in effect limiting the performer to working in two dimensions.

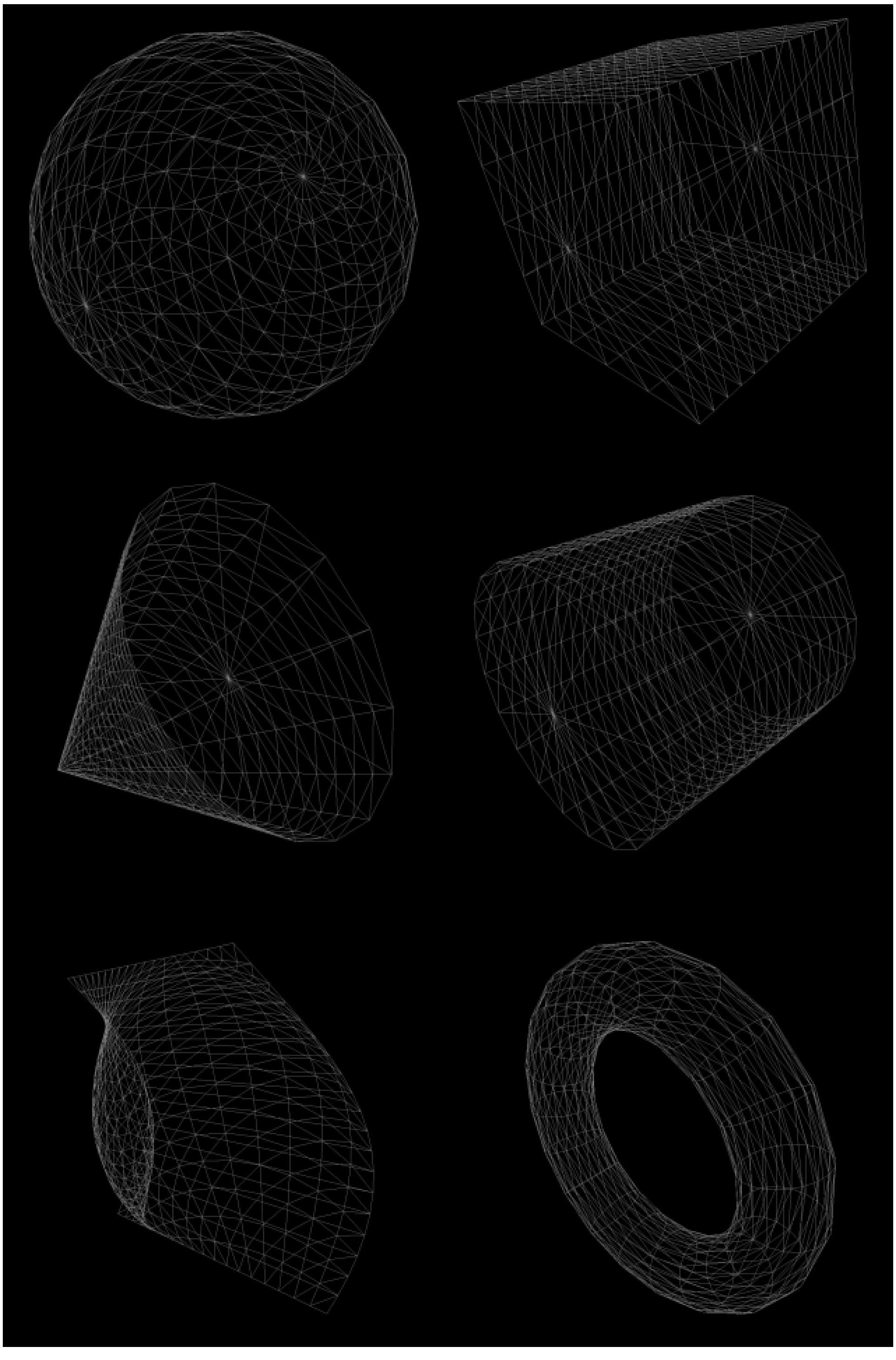


Figure 6.6: Several *phase shapes*, constructed with 400 vertices. The *phase shapes* for all nine *phases* are based upon these six geometric shapes. Several *phases* use the same *phase shape*, however the base form is altered through the mapping of other parameters to vertex position.

The Virtual Markers Each *virtual marker* is modelled using a polygonal rendering approach. This means that the *virtual marker* is defined as a polyhedron through a series of x,y,z vertices in the three dimensional *virtual performance space*, which are then joined together through a number of edges. The form that the *virtual marker* takes is determined through a mixture of nine *phase shapes*, shown in figure 6.6. A *phase shape* s_i is the geometric shape associated with each *phase* within the performance. This consists of N vertices comprising three coordinates x, y, z so that each individual *phase shape* can be represented by matrix

$$s_i = \begin{bmatrix} x_i & y_i & z_i \\ x_{i2} & y_{i2} & z_{i2} \\ \vdots & \vdots & \vdots \\ x_{in} & y_{in} & z_{in} \\ \vdots & \vdots & \vdots \\ x_{iN} & y_{iN} & z_{iN} \end{bmatrix} \tag{6.4}$$

for $i \in [1 : I]$, where I is the total number of *phases*.

Typically, the placement of the *phase shape* in the *virtual performance space* is established by mapping the position of the corresponding *position marker* to the centroid of the *phase shape*, although this approach can differ between *phases*.

In order to enable shifting of the shape of the *virtual markers* between *cycles*, the current state of the *virtual marker* is determined through a combination of each of the *phase shapes*. This is approached through setting the vertices of the *virtual marker* in phase i as a weighted average of the vertices of each *phase shape* so that for a given vertex v_n:

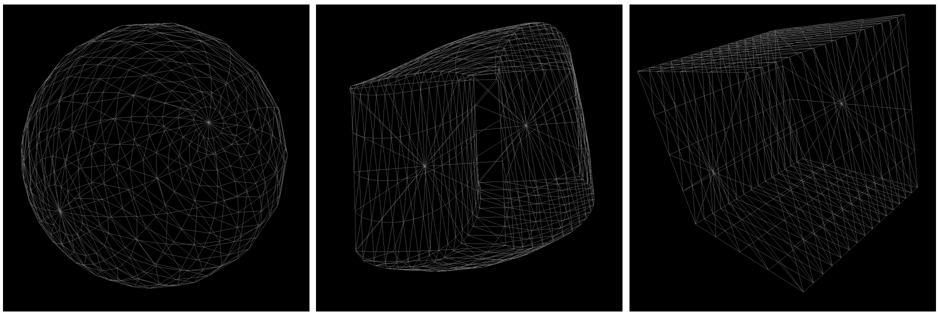
$$v_n = \begin{bmatrix} \frac{\sum_{i=0}^{I} V_{p_i} x_{s_{in}}}{\sum_{i=0}^{I} V_{p_i}} & \frac{\sum_{i=0}^{I} V_{p_i} y_{s_{in}}}{\sum_{i=0}^{I} V_{p_i}} & \frac{\sum_{i=0}^{I} V_{p_i} z_{s_{in}}}{\sum_{i=0}^{I} V_{p_i}} \end{bmatrix} \tag{6.5}$$

where V is the *phase weighting vector* for current *phase* p, for which a requirement is that $\sum_{i=0}^{I} V_{p_i} = 1$.

Therefore, within a given *phase* p, the state of *virtual marker* s_p can be defined as a matrix of vertices s_p, for which:

$$s_p = \begin{bmatrix} \frac{\sum_{i=0}^{I} V_{p_i} x_{s_{i_1}}}{\sum_{i=0}^{I} V_{p_i}} & \frac{\sum_{i=0}^{I} V_{p_i} y_{s_{i_1}}}{\sum_{i=0}^{I} V_{p_i}} & \frac{\sum_{i=0}^{I} V_{p_i} z_{s_{i_1}}}{\sum_{i=0}^{I} V_{p_i}} \\ \frac{\sum_{i=0}^{I} V_{p_i} x_{s_{i_2}}}{\sum_{i=0}^{I} V_{p_i}} & \frac{\sum_{i=0}^{I} V_{p_i} y_{s_{i_2}}}{\sum_{i=0}^{I} V_{p_i}} & \frac{\sum_{i=0}^{I} V_{p_i} z_{s_{i_2}}}{\sum_{i=0}^{I} V_{p_i}} \\ \vdots & \vdots & \vdots \\ \frac{\sum_{i=0}^{I} V_{p_i} x_{s_{i_n}}}{\sum_{i=0}^{I} V_{p_i}} & \frac{\sum_{i=0}^{I} V_{p_i} y_{s_{i_n}}}{\sum_{i=0}^{I} V_{p_i}} & \frac{\sum_{i=0}^{I} V_{p_i} z_{s_{i_n}}}{\sum_{i=0}^{I} V_{p_i}} \\ \vdots & \vdots & \vdots \\ \frac{\sum_{i=0}^{I} V_{p_i} x_{s_{i_N}}}{\sum_{i=0}^{I} V_{p_i}} & \frac{\sum_{i=0}^{I} V_{p_i} y_{s_{i_N}}}{\sum_{i=0}^{I} V_{p_i}} & \frac{\sum_{i=0}^{I} V_{p_i} z_{s_{i_N}}}{\sum_{i=0}^{I} V_{p_i}} \end{bmatrix} \tag{6.6}$$

given that N is identical for each *phase shape* s_i. An example of a *virtual marker* that is interpolated between two *phase shapes* with a weighting of $0.5, 0.5$ can be seen in figure 6.7.



(a) Phase shape one. (b) Interpolation. (c) Phase shape two.

Figure 6.7: Interpolation between phase shapes in order to create the virtual marker. This shows two phase shapes with a weighting of 1 towards themselves, and the interpolation which is weighted 0.5 towards each.

At initialisation of a performance, the *phase weighting vectors* are set so that $V_{p_i} = 1$ where $p = i$ and $V_{p_i} = 0$ where $p \neq i$. This means that for the first *cycle* of the performance $s_p = s_i$. Throughout a performance, updating of the *phase weighting vectors* is started at the completion of a *phase* based upon output from the shifting algorithm.

The matrices representing the *phase shapes* are further complicated by the mapping of various parameters to the positions of vertices in differing *phases*. For example, in both phase two and phase four, the *phase shape* is a cube defined with 400 vertices. However, in phase two, the horizontal position of the vertices is mapped to the parameters of the absolute position of the *position marker* in the relation to the origin of the MoCap coordinate system and the distance of the *position marker* from the barycentre of all *position markers*. This provides the *phase shape* of a cuboid, that extends and contracts as the *position markers* are moved closer together and further apart, and are moved further from the centre of the *physical performance space*.

In addition, several *phases* make use of connections extending from the *phase shapes*. For example, phase five utilises a spline the bisects all *virtual markers*, and phase seven draws a line outwards from the *phase shape* to the surface of an invisible sphere surrounding the *virtual performance space*. Both the additional parameters that modulate the *phase shape* and the connections between *phase shapes* were designed in view of the work of Schlemmer's methods. The exact principles are outlined in section 7.1.

6.3.2.3 Module 4: Audio Input Handler

The audio input captures the input signal from the performer's microphone and applies some basic pre-processing. The signal is passed to the audio processing module.

In addition, the signal is directly mapped to the amount of interpolation between *phase* representations and mappings through applying linear interpolation to the *phase weighting vector*. Upon breaking an amplitude threshold, the value of the current *phase weighting vector* begins a ramp towards the value of the *phase weighting vector* for the current *phase* in the first *cycle* over 2 s. Therefore, the current values of *phase weighting vector* V_p at time t after the amplitude threshold is broken can be calculated as

$$V_{p_t} = V_p + (V_{p_1} - V_p) \times \left(\frac{t}{2}\right) \tag{6.7}$$

where V_p is the current values of the *phase weighting vector*, V_{p_1} is the values of the *phase weighting vector* in the first *cycle*, and t is measured in seconds.

When the audio input falls below the threshold, the interpolation releases and a ramp towards the current *phase weighting vector* begins over 1 s, so that

$$V_{p_t} = V_{p_1} + (V_p - V_{p_1}) \times \left(\frac{t}{1}\right) \tag{6.8}$$

The amplitude of the audio input signal is likewise mapped to the output amplitude envelope of the two audio spatialisation components. Before the amplitude threshold hold is broken the output amplitude envelope of the *trajectory component* is set to 0 and the *diffusion component* to 1. In a similar manner to the ramps applied to the *phase weighting vectors*, when the threshold is broken a 2 s ramp is applied so that at time t after the threshold has been broken, the amplitude of the *trajectory component* envelope can be calculated as

$$A_{t_t} = \left(\frac{t}{2}\right) \tag{6.9}$$

and the *diffusion component* as

$$A_{d_t} = 1 - \left(\frac{t}{2}\right) \tag{6.10}$$

with the equations reversed and with a denominator of 1 when the input amplitude falls below the threshold.

6.3.2.4 Module 5: Audio Spatialisation

The system's audio output is designed to be reproduced spatially within the *physical performance space*. This is achieved through the audio spatialisation module, which processes the audio data received from the non-real-time audio processing module. For *The Shapeshifter*, the chosen method of spatialisation is *Ambisonics* as this is the format with which I have the most experience.

The spatialisation module comprises two main components. Although *The Shapeshifter* is not explicitly located within the field of electro-acoustic music and composition, the field provides a large body of literature on approaches to working with spatial audio which can inform the use of the spatial audio in the system. For example, Baalman (2010) provides an overview of spatial composition techniques, and each of the two components encourages the use of a different technique.

The first component is the playback of the individual recording made by the performer in each phase in each cycle. These are referred to as the *cycle master* tracks, for which there is an individual *cycle master* track for each *phase* in each *cycle*.[1] These are intended to be reproduced to leave a trace of the performer in the performance space, with the looper function representing not just the sound that was made but also the positions that the performer travelled through as it was being made. In Baalman's terminology, this can be seen as the composition technique of *creating trajectories*, for which "trajectories will introduce choreography of sounds into the piece" (p. 209).

The second component is an audio bed that envelops the performance space. This is intended represent the shifting elements of the system, with the parameters of the audio bed dependent upon the *phase weighting vectors*. The audio to be spatialised for this component is the sum of all of the recordings made in each phase across all cycles. This is referred to as a *master* track. There is therefore one *master* track for each phase. This is approached with the technique referred to by Baalman as diffusion, with the goal of "creating broad, or even enveloping, sound images" (p. 210).

The component passed to the audio output is determined in real-time depending upon the amplitude of the performer's vocalisations in the current *phase*.

The Trajectory Component Within a given *phase*, the trajectory component maps the MoCap position data vectors recorded during that *phase* in each previous *cycle* to the corresponding *cycle master* audio data. This comprises three central mappings, each focused on mapping the position of a *position marker* to an object in the Ambisonics environment. The first of these is a mapping of the mean vertical position of the *position markers* to the centre frequency of a band pass filter. This provide a mapping for the vertical position MoCap vectors, as only two dimensional Ambisonics is used for the system. The second mapping maps the position of the *position marker* on the horizontal plane to the azimuth of the Ambisonics object. The third maps the radial distance of the *position marker* from the MoCap origin to the Ambisonics object's gain and frequency domain processing. The exact method through which these mappings are achieved can be found in appendix D.

The Diffusion Component Within a given *phase*, the diffusion component applies a granular synthesis technique to the *master* audio track. This results in a diffuse field centred around the origin of the *Ambisonics* plane. For each *phase*, the parameters of the granular synthesis are set to specific values as part of the sonic design of the representation in that *phase*. These also shift throughout the *cycle*, interpolating between the parameters set for each *phase* on the basis of the values of the *phase weighting vectors*.

[1] With the exception of the 9th *cycle*, as any recordings made during this *cycle* will not be played back.

6.3.3 Modules 6-8: Non-Real-Time Modules

Several aspects of the system undergo changes during a performance, which affects both the mappings within the real-time component and the data processed and passed to the visualisation and audio spatialisation modules. These changes are based on the cumulative input to the system over the course of the performance. There are three major non-real-time components. These are:

Module 6: MoCap Logger

Module 7: Audio Processing

Module 8: Shifting Algorithm

The first module is the MoCap logger, which captures the real-time input of the MoCap system and creates logs of the position vectors for each *position marker*, *phase*, and *cycle*. The logs are then provided, as needed, to the audio spatialisation module of the real-time component. Similarly, the audio processing module logs the audio recording for the *phase* and *cycle*, processing and passing the audio files to the spatialisation component. The data logged by these modules serves as the basis for the shifting algorithm module. The shifting algorithm is comprised of two sub-modules: the Dynamic Time Warping (DTW) module and the Support Vector Machine Classifier (SVM) module, which operate on the logged motion and audio data respectively. The shifting algorithm outputs the updated *phase weighting vectors* and transmits them to the real-time component. An overview of the dataflows within the real-time modules is shown in figure 6.8.

6.3.3.1 Module 6: MoCap Logger

The MoCap logger module receives the continuous stream of MoCap data from the MoCap input handler module and stores it for use by the system's non-real-time modules. The data is formatted and stored in a way that is required for the processing by each individual module. When the performance synch module sends a control message, the MoCap logger passes the relevant logs to the appropriate module for further processing.

6.3.3.2 Module 7: Audio Processing

The audio processing module takes the input provided from the audio input module and logs the recorded audio for a *phase* and *cycle* in a corresponding buffer. From the buffers that accumulate over the course of a performance, there are two audio tracks that are created, the *cycle master* and *master* tracks, each of which are passed to a separate component of the audio spatialisaion module. In the creation of these tracks, an amplitude envelope is applied based upon the DTW path passed from the Dynamic Time Warping component of the shifting algorithm.

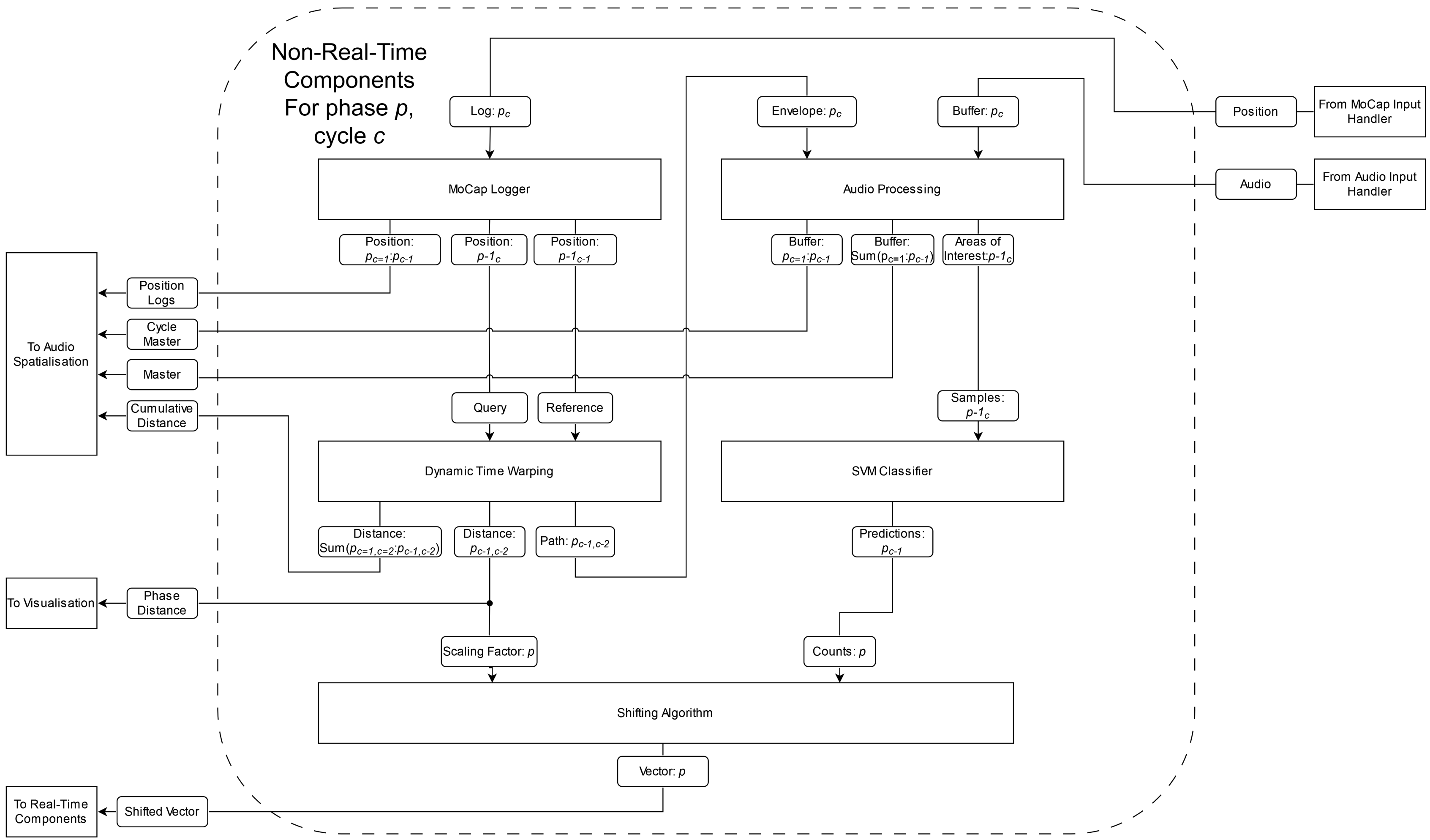


Figure 6.8: An overview of non-real-time system components. These are the mappings and data-flows that take place temporally for the non-real-time components within a given *phase* and *cycle* of the performance. Input signals flow from top to bottom. Signals passed to and from real-time modules and the performance synch module flow right to left. Large boxes represent modules or module components. Small boxes describe the signal type.

6.3.3.3 Module 8: Shifting Algorithm

The shifting algorithm is integral to the core concept of the performance, whereby the visual and sonic output of the system gradually transforms and takes on characteristics of the other phases throughout the duration of the performance. The performer defines a baseline for the parameters of the shifting algorithm, which are then compared against this baseline throughout the performance. For the motion input, the shifting algorithm compares the motion logs of a given *phase* of a *cycle* against the motion in the same *phase* in the previous *cycle*. For the sound input, the algorithm compares the recorded audio against examples of vocalisations defined for each *phase* prior to the performance.

The mechanism underlying the shifting algorithm is based upon the use of *phase weighting vectors*. For a given phase p, the *phase weighting vector* V_p of length P can be defined as

$$V_p = \begin{bmatrix} w_1 & \cdots & w_{p-1} & w_p & w_{p+1} & \cdots & w_P \end{bmatrix} \tag{6.11}$$

for $p \in [1 : P]$, where P is the total number of phases and $\sum_{p=1}^{P} w_p = 1$. At initialisation, the *phase weighting vectors* as set so as $w_p = 1$, in other words each phase is weighted fully towards itself. At the end of a phase, the *phase weighting vector* for that phase is updated so as to start providing weight towards other phases. Each *phase* has its own *phase weighting vector*.

There are two main components which form the shifting algorithm which determine the updates of the phase weights. These are the *Dynamic Time Warping* (DTW) component which is applied to the MoCap logs, and the *Support Vector Machine Classifier* (SVM) component, which is applied to the recorded audio. Figure 6.9 provides an overview of the functionality of the shifting algorithm.

The Dynamic Time Warping Component Dynamic Time Warping (DTW) is a method with which the similarity of two discrete time series of varying lengths can be evaluated. This is posed as an optimisation problem, for which the optimal solution aligns each data point in the first time series to the data point in the second which results in the lowest total cost value (the total cost is alternatively referred to as the DTW distance in the literature) (Müller, 2007a), as shown in figure 6.10.[2] This is calculated by creating a two-dimensional cost matrix, with each dimension corresponding in length to one of the time series. For each cell in the matrix, a cost value for each pair of data points can be calculated, with the cost value commonly presented as the absolute difference (Euclidean distance) of the values of the data points at that pair of indices. The solution is the path of indices through the cost matrix, show in figure 6.11, which results in the lowest sum of cost values and which satisfies the following conditions:

1. The first data point in the first time series must be at minimum matched to the first data point in the second time series and the final

[2] A full overview of the mathematics behind DTW is beyond the scope of this thesis. I refer to Müller (2007a) for a comprehensive description.

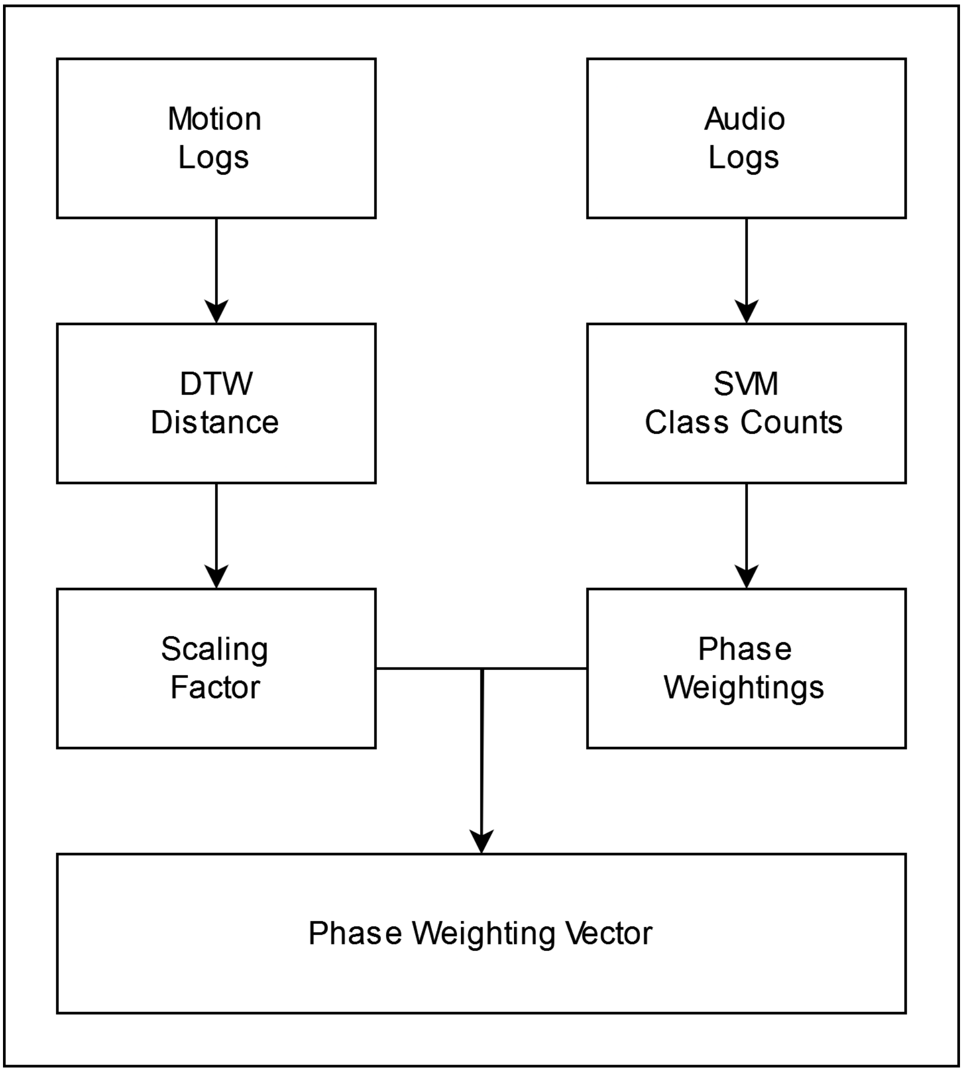


Figure 6.9: A block diagram of the shifting algorithm.

data point in the first time series must be at minimum matched to the last data point in the second time series.

2. The indices of the path must monotonically increase.
3. The indices of the path can only increase in steps of one, for either or both components of the index.

Due to the fact that a matrix must be evaluated with a dimensionality of the two time signals, classic DTW has both a quadratic time and space complexity. This makes it rather unsuited for use in a system with hard time constraints, especially within the context of audio signals with a sample rate of either 44.1 kHz or 48 kHz. For this reason, in *The Shapeshifter*, DTW is performed solely upon MoCap data sampled at a considerably lower 120 Hz.

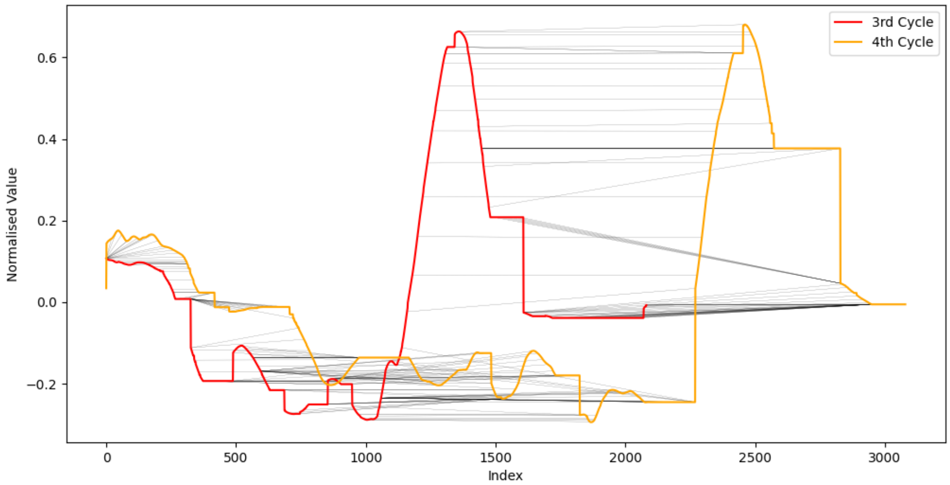


Figure 6.10: The alignment of the data points of two time series using DTW. Here the two time series represent the motion of a single *position marker* in relation to the lateral axis of the MoCap coordinate system, which was worn by the collaborator during the pilot performance. This is the motion in the 1st *phase* in the 3rd and 4th *cycle*. These present similar motion patterns, although differing in duration. The black lines represent the alignments between samples as calculated through the DTW algorithm as having the lowest cost value. The alignment of each 20th sample is plotted.

There are several methods through which the time and space complexity of the DTW algorithm are reduced. On the implementation side, a dynamic programming approach is employed (Müller, 2007a). Moreover, as described by Salvador and Chan (2007), there are several categories of approaches towards alterations of the algorithm that can reduce the time and space complexity. The first of these is constraining the number of cells that are evaluated in the cost matrix. This commonly take the form of various windowing functions (Hiyadi et al., 2016; Itakura, 1975; Ratanamahatana & Keogh, 2004b; Sakoe & Chiba, 1978). However, as noted by Salvador and Chan (2007), constraining techniques do not work so well for time series that possess great temporal variance. The second approach is abstracting the data upon which the DTW is performed, in effect reducing its resolution. Various approaches for abstraction include PDTW (Keogh & Pazzani, 2000), MsDTW (Müller et al., 2006), MrMsDTW (Prätzlich et al., 2016), FastDTW (Salvador & Chan, 2007), and CoarseDTW (Dupont & Marteau, 2016). Finally, indexing techniques such as SparseDTW (Al-Naymat et al., 2009), PrunedDTW (Silva & Batista, 2016), and EAPrunedDTW (Herrmann & Webb, 2021) can be used to reduce the number of times the algorithm must be run.

The system employs a fastDTW algorithm with a radius of one. The cost value is the square of the difference of the derivates of the two time series as proposed by Keogh and Pazzani (2001). Starting from cycle three, the derivatives of the min-max normalised MoCap vectors corresponding to each *position marker* and axis for the previous two *cycles* are calculated.

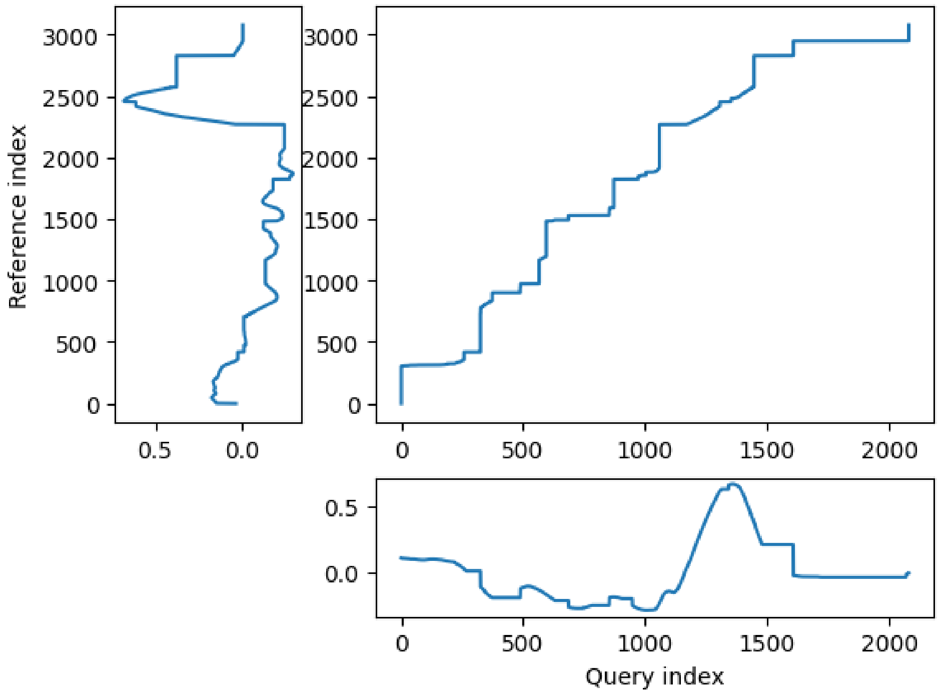


Figure 6.11: The path through the cost matrix with the lowest sum of cost values for the same two series as shown in figure 6.10. This alignment has a DTW distance of 0.0015, normalised for path length following Ratanamahatana and Keogh (2004a). Plot created with the Python version of the DTW package (Giorgino, 2009).

These vectors are then passed to the fastDTW algorithm. For each axis of each *position marker* there is an output of the DTW distance and the path through the cost matrix. The distances are normalised based upon path length as provided by Ratanamahatana and Keogh (2004a). As only a single distance value is required for each *position marker*, the mean across all axes is taken. The paths are passed to the audio processing module, and the distance is passed as part of the formula to update the *phase weighting vectors*.

The Support Vector Machine Component A Support Vector Machine (SVM) is a supervised machine learning algorithm that attempts to fit a hyperplane (or hyperplanes in the case of multiclass datasets) to a dataset of discrete classes that best separates these classes. A full description of the functionality of an SVM is beyond the scope of this thesis, however I refer to Pisner and Schnyer (2020) for an overview.

The aim of the SVM is to classify the vocalisations performed during each *phase* as belonging to one of the nine *phases* with which we worked in the system. Once the vocalisations have been classified, the counts of classes are passed as part of the formula to update the *phase weighting vectors*.

A personal dataset was recorded with the collaborator in the 6^{th} development session. The collaborator explored each of the *phases* with the *bounding boxes* disabled while performing vocalisations that they associated with each *phase*. These were recorded in a single audio file as they would be during a performance, and passed through an identical hardware and signal processing chain as used in rehearsal and performance. The dataset was small, comprising roughly 100 examples per class, however the idea is that the model can be updated or re-trained prior to each performance if desired.

The individual examples were extracted from the audio files through identifying regions of interest in the audio files' spectrograms. Using the timestamps provided for the regions of interest, the audio files were segmented, rms values and melspectrogram extracted, and the classifier was trained. The model was trained 1000 times, with a mean f1 score of 0.6.

Within the system, the processing pipeline for inference is identical to that of training.

Weight Updates At the end of *phase* p, once the DTW distance and SVM counts have been calculated, the *phase weighting vectors* are updated according to the following formula

$$V_p = \begin{bmatrix} \frac{w_0 + \alpha d\,(c_0/c_{\max})}{\sum_{p=1}^{P} w_p} & \cdots & \frac{w_p + \alpha d\,(c_p/c_{\max})}{\sum_{p=1}^{P} w_p} & \cdots & \frac{w_P + \alpha d\,(c_P/c_{\max})}{\sum_{p=1}^{P} w_p} \end{bmatrix} \tag{6.12}$$

for $p \in [1 : P]$, where P is the total number of phases, α is a scaling factor, d is the DTW distance, c_p is the SVM counts towards phase p, c_{max} is the maximum SVM count for the current *phase*, and w_p is the current weight towards phase p.

6.4 Implementation

6.4.1 Hardware

The hardware implementation of the system primarily consisted of the equipment available within the *Portal*, with the hardware schematic of the system shown in figure 6.12. The system component of *The Shapeshifter* is run from a *HP Spectre X360* laptop [3] positioned on a table set against the north wall of the MCT Portal. The laptop has the following specifications:

1. 11^{th} Gen Intel Core i7-1165G7 2.80 GHz CPU
2. 16 GB RAM
3. Intel Iris Xe integrated GPU

[3] https://www.hp.com/us-en/shop/slp/spectre-family/hp-spectre-x-360

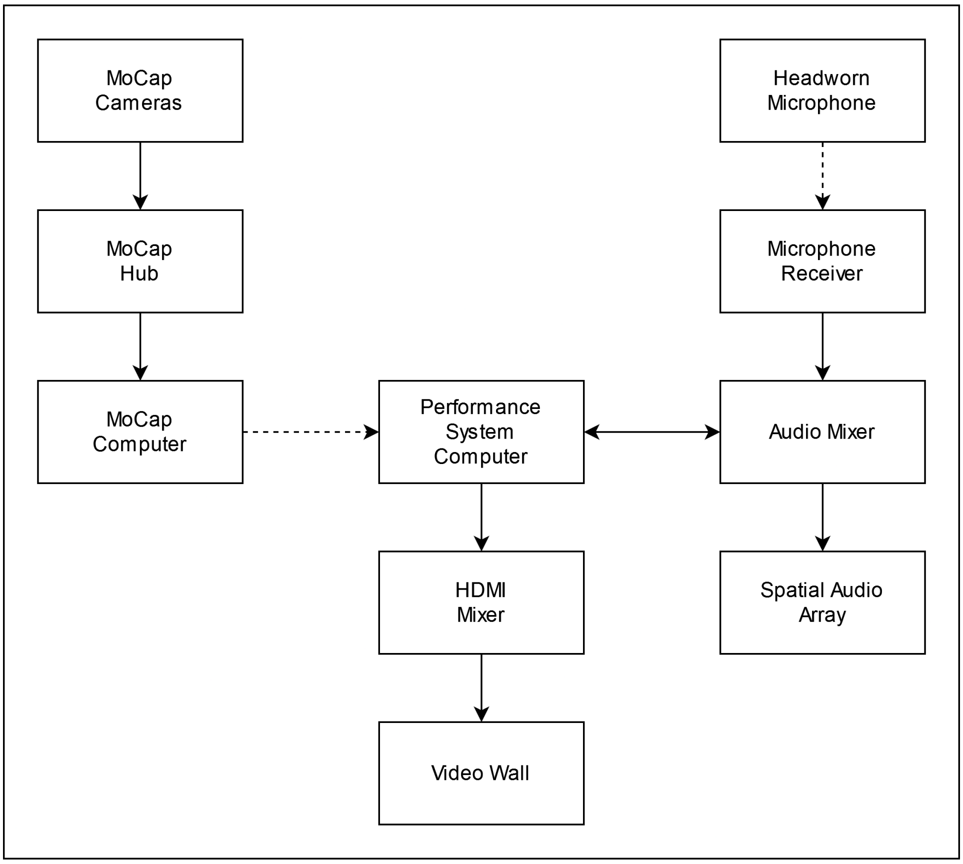


Figure 6.12: The hardware schematic of the system. Dotted lines represent wireless connections.

The visual output of the system was displayed on the video wall, connected to the laptop via HDMI and passing through a *Stoltzen SUHM44-H2* HDMI matrix.[4] The *Midas M32* mixer, which serves as the audio hub for the system, is connected directly to the laptop via USB, functioning as an audio interface. The mixer outputs the audio signal to the spatial array. A *Sennheiser SL Headmic 1*[5] wireless headset microphone is worn by the performer in order to ensure that the microphone has a consistent spatial relationship with the performer regardless of their position within the performance space. The *Sennheiser SL Rack Receiver DW-4-US*[6] for this is connected directly to the mixer.

The *OptiTrack* system runs on a desktop PC located in the control booth. This PC has the following specifications:

1. 6th Gen Intel Core i5-6500 3.2 GHz CPU
2. 8 GB Ram

[4]https://stoltzen.eu/?lang=en

[5]https://www.sennheiser.com/en-de/catalog/products/wireless-systems/sl-dw-headmic-set/sl-headmic-set-dw-3-eu-c-505880

[6]https://www.sennheiser.com/en-de/catalog/products/wireless-systems/sl-dw-rack-receiver-5362c/sl-rack-receiver-dw-4-us-505899

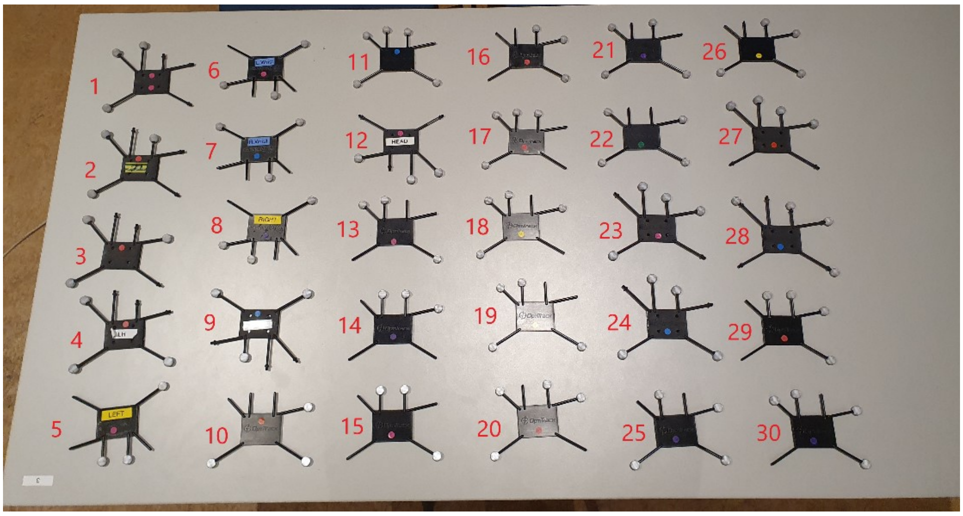


Figure 6.13: The 30 *Optitrack* Rigid Bodies that can be used in performance.

3. Intel HD Graphics 530 integrated GPU

The computer running the *Optitrack* system passes data to the computer running *The Shapeshifter* with the User Datagram Protocol (UDP) via a wireless connection over the *Portal's* local network. Although the UDP protocol is faster than the Transmission Control Protocol (TCP), which makes it well suited to real-time applications, it is a connectionless protocol, meaning that there is no acknowledgement or checking of errors by the receiver. This means that it is open to both data packets being lost, as well as packets arriving out of order.

Optitrack rigid bodies,[7] shown in figure 6.13, are used as *position markers*. These are configurable non-deformable objects for which up to five individual markers can be attached to the extending arms which can be defined as a rigid body within the *Motive* software. This provides several advantages over using individual markers. Firstly it enables each *position marker* to have an individual ID which allows for processing that would be more difficult if this were not the case, for example the drawing of a path through the *virtual markers* in a certain order which is found in the visualisation of *phases* five and nine. It also means that in the case of occlusions, when the rigid body is rediscovered by the system it assumes the identical ID. Additionally, the rigid bodies' marker configurations are easily reproducible, meaning that these can be reliably recreated for repeat performances.

A further aspect to consider is that the rigid bodies are a lot larger and more conspicuous than the individual markers. However, as we are not attempting to mask the use of MoCap system, this is not a problem for us. Moreover, their larger size, as well as the fact that the arms which do not have a marker attached can be used as a handle, makes them easier for the collaborator to hold and interact with during performance.

[7] https://optitrack.com/accessories/markers/#mcp1090

For performance, the collaborator wears an *Optitrack Motion Capture Suit Classic* jacket,[8] which is covered in a felt material which enables the attachment of the Velcro-backed rigid bodies. This is part of a full suit which also consists of a pair of trousers, however the collaborator decided against wearing these as they preferred the range of motion offered by their own trousers. If they decided to attach rigid bodies to parts of their body which were not covered by the jacket, a Velcro band was constructed out of small Velcro straps which could be wrapped around the body part.

6.4.2 Software

The software is primarily self-developed, although several off-the-shelf modules are employed. The real-time components are programmed within the *Max/MSP/Jitter* (MAX) visual programming environment,[9] with the visuals rendered with the *Jitter OpenGL* framework.[10] Several components are implemented with *Javascript* within the Max environment. The spatial audio components were implement with the use of two external MAX packages. The encoding of the trajectory component, as well as the decoding of the Ambisonics to the speaker array, was implemented with the *ICTS Ambisonics* package.[11] The diffusion component made use of the *hoa.syn.grain* encoder object from the *HoaLibrary* implementation for MAX.[12]

The non-real-time components are primarily implemented as a series of *Python* scripts, although certain functionalities are handled by *MAX*. The DTW component makes use of the *fastdtw* library adapted with a custom cost function.[13] Audio data handling is based upon the *librosa* library (McFee et al., 2015). The SVM component makes use of the implementation provided by the *scikit-learn* library (Pedregosa et al., 2011). Regions of interest in the audio recordings are identified with the use of the *scikit-maad* library (Ulloa et al., 2021).

The position data is broadcast on a loopback function from the *Motive*[14] software to a *Python* script which passes this to The Shapeshifter system via UDP with the *Open Sound Control* (OSC) protocol[15] which is built on top of the *NatNetSDK*,[16] *Optitrack's* software development kit. This script is a modified version of a script provided to me by Pedro Lucas Bravo, at that time a fellow student, in early 2022.[17] Communication between the scripts which are part of the *The Shapeshifter* system makes use of the *python-osc*[18]

[8] https://optitrack.com/accessories/wear/

[9] https://cycling74.com/products/max

[10] https://www.opengl.org/

[11] https://www.zhdk.ch/forschung/icst/software-downloads-5379/downloads-ambisonics-externals-for-maxmsp-5381

[12] https://github.com/CICM/HoaLibrary-Max

[13] https://github.com/slaypni/fastdtw

[14] https://optitrack.com/software/motive/

[15] https://ccrma.stanford.edu/groups/osc/index.html

[16] https://optitrack.com/software/natnet-sdk/

[17] https://pedrolucas.tech/

[18] https://github.com/attwad/python-osc

and *osc4py3*[19] libraries.

6.5 Summary

In this chapter, I presented a major result of this thesis, namely *The Shapeshifter* system. I described the design of the individual modules and components, and followed this with a description of how this design was implemented.

[19] https://sourcesup.renater.fr/scm/viewvc.php/osc4py3/

Chapter 7

Results

In this chapter, I present the results of the iterative process, the post performance interview, and the system latency evaluation.

7.1 The Iterative Process

In this section, I present an overview of the results of the iterative development process. This is structured around the key developments from the early, middle, and late sessions. These results, organised by session, are presented in table 7.1.

7.1.1 Early Sessions

In the early sessions, the workshops were focused upon finding a collaborative way of working, developing the concepts that we would like to explore in the work (as outlined in section 4.3), and familiarising the collaborator with the technologies involved. The explorations were centred around an early iteration of the performance system which was focused on the bounding boxes. Visual representations were limited to spheres with trailing angular lines, and the audio functionality of the system was limited to a simple looper which looped the vocalisations of the collaborator recorded during a phase, shown in figure 7.1. During these sessions, the collaborator worked with very few position markers, and these were treated as objects external to their body, in other words they mainly held the position markers in their hands. In these sessions, our focus was on the way in which the system's functionalities reflected limitations onto the collaborator's body in the physical space. These included the ways of working determined by the use of optical, marker-based MoCap, such as finding a way to work with marker occlusions, as well as the design of functionalities of the performance system, in particular the sizes and positions of the bounding boxes and the use of the video wall.

Figure 7.2 shows several of the collaborator's reflections after the second session. These already reflect several themes that would become apparent in the interview conducted after the pilot performance.

Session	Date	Major Outcomes
1	12.01.23	1. Definition of working relationship 2. Foundations of concepts underlying performance and system 3. Introduction to technologies, backgrounds, and ways of working
2	02.02.23	1. Perspective on performer's focus when treating markers as external to body 2. Validation of bounding box component
3	21.02.23	1. Integration of sound 2. Foundations of representational designs 3. Advances in workflow
4	16.03.23	1. Setting limits on developments 2. Procedures for validation 3. First visual and sonic designs
5	11.04.23	1. Integration of spatial audio 2. Integration of representation interpolations
6	18.04.23	1. Final visual and sonic designs 2. Recording of audio dataset for SVM training
7	26.04.23	1. Rehearsals with full system for pilot performance
8	04.05.23	1. Pilot performance with audience 2. Post-performance interview
9	09.05.23	1. Video recording of performance 2. Reflections on collaboration 3. Plans to move forward

Table 7.1: A summary of the sessions used for workshopping and exploration. The results as major outcomes are also presented.

7.1.2 Middle Sessions

In the workshops of the middle sessions we developed the visual and sonic designs that would function as the base representations. These stemmed from distinct moods developed by the collaborator in relation to ways of moving and vocalising originating in the scenic "worlds" that they perceived when relating to differing marker placements, ways of connecting markers together in the visual representations and colour schemes with which we worked in the exploration sessions, shown in figure 7.3. These included, for example, the development of a phase based around the earlier iterations, working with a deep blue colour and trailing lines, which the performer perceived as moving through a scene deep underwater, and another phase based around a dark yellow, with cuboid representations that would change in size and position relating to the proximity of physical markers to one another, which the collaborator perceived as a structure in a windswept desert. Viewing these in relation to the conceptual ideas inspired by Schlemmer (as described in section 4.3), in between sessions I would further develop these representations. For the visual representations, this included aspects such as applying restrictions to the visualised marker's degrees of freedom or working with ideas of

Figure 7.1: The collaborator during the exploration from the third session. Two *Optitrack* rigid bodies were in use at this time. The collaborator triggers the next phase by positioning these in bounding boxes in the physical space that we defined earlier. During this session, each phase was defined just through colour. Motion through space left trailing lines. A rudimentary version of vocal looping was being explored during this session.

balance through the visualised markers reacting to parameters such as the position of the centroid of all physical markers, with the relationship between these aspects and the results of Schlemmer's methods presented in table 7.2. Working with audio recordings of the collaborators vocalisations, I would then develop sonic designs to match these concepts and visual representations. The collaborator would provide feedback during the following explorations.

In addition, we developed the mechanism that would interpolate between the representations over the course of the performance in the shifting algorithm, and integrated the spatial audio components. This was part of a wider push to develop the performance as a whole, including considering how to stage a performance within the *Portal*.

Employing the framework offered by the co-construction model enabled the identification of several assumptions and values that were being embedded in *The Shapeshifter* system during the development process. To take one example, the marker position data received from the *Optitrack*

> **. . . [I focused a] lot on the screen on my body and its position, especially hands bc they had the sensors/reflectors. The spatial placement comes into it bc I wanted to know where I was standing, but really my bodily presence and position was the focus rather than the space . . .**
>
> **. . . big boxes I had a body/space focus, smaller I had a hands/screen focus . . .**
>
> **. . . I was more aware of how my body moved rather than taking movement from off the screen . . .**
>
> **. . . there was a huge gap [in intentionality] when I had no visual bc I was facing away from the screen. This I find v interesting bc it means that I must have used the screen as a reference point a lot after all, even tho it felt like I was not, as described above . . . The sight is such a taken-for-granted-sense, and even in the side-vision I guess I don't really notice I'm using it a lot . . .**

Figure 7.2: The collaborator's reflections with regard to their focus in relation to several questions I asked after the second session. From email correspondence with collaborator.

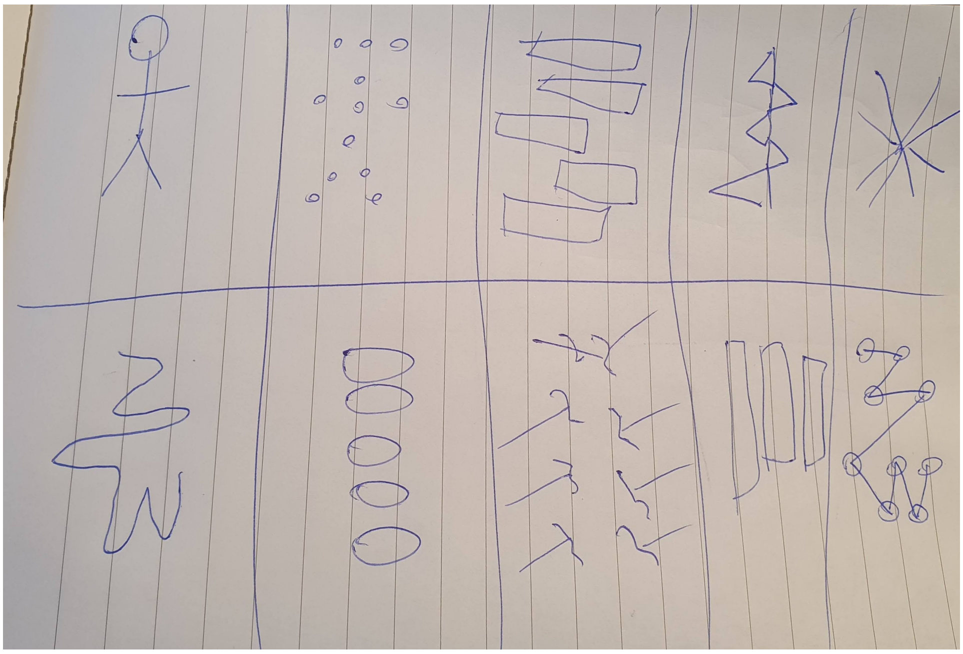


Figure 7.3: A rough sketch of some early ideas for visual designs.

system has several transforms applied when received by *The Shapeshifter*, primarily to shape the data to fit within the *virtual performance space* and have this mirror the *physical performance space*. However, I developed these components between sessions. During the following exploration, it quickly became clear that these transforms were rooted in my own sense of *my* body, the way that I experience my body in motion, and the way in which I interact with the *position markers*. The collaborator would frequently perform motions with a much larger range than I would, would move a lot faster, and move outwith the field of view of the virtual camera. Moreover,

Phase	Result	Method
1	1. The Technical Organism 2. Dematerialization	1. Spherical marker representations 2. Trailing splines representing the previous seconds of motion
2	1. Ambulant Architecture	1. Cuboid marker representations that extend and contract based on position
3	1. The Technical Organism	1. Conical marker representations that extend from and rotate around a central axis
4	1. Ambulant Architecture 2. Dematerialization	1. Horizontal cubic marker representations 2. Marker position connected to barycentre through extending lines
5	1. The Marionette 2. Dematerialization	1. Pointed marker representations that extend, contract, and change position in relation to marker barycentre 2. Connected through a spline passing through all marker positions
6	1. The Technical Organism	1. Disklike marker representation
7	1. Dematerialization	1. Flame shaped marker representation 2. Extending lines onto an invisible sphere surrounding performance space
8	1. Ambulant Architecture 2. Dematerialization	1. Vertical cubic marker representations 2. Expanding and contracting height based upon marker position
9	1. The Marionette 2. The Technical Organism	1. Torus shaped marker representations 2. Marker representations rotate based upon marker position and relation to marker barycentre 3. Linear path drawn between markers

Table 7.2: The relationship between the *phase* representations and the results of ways of the relationships between the body and space developed by Oskar Schlemmer.

the collaborator is noticeably shorter than I am, meaning that the transforms that I was applying when developing alone, using my own body as the model to test ranges of motion, were tailored to my own kinesphere. As a consequence, these developments were not well suited to the collaborator. The form of the body that we were co-constructing was rooted in my own, exposing an underlying power imbalance stemming from my role as system designer.

7.1.3 Late Sessions

The late sessions comprise the three final sessions. These consisted of a long rehearsal session in the lead up to the pilot performance, the pilot performance, interview, and audience survey,[1] and a wrap up session where we recorded a video of a performance, reflected on the process, and planned how to move forward with the project.

7.2 Interview Results

In this section, I present the results of the interview that took place after the pilot performance, organised by theme. Numbers in brackets after quotations refer to the transcript segments as presented in appendix F.

7.2.1 The Body Originating from the World

A recurring theme in the collaborator's depiction of their experience in relation to their conceptualisation of the representation of the body within the system is that this body is not isolated but rather emanates from and is closely linked to other aspects of the output of the system. The collaborator conceptualises this in scenic terms, thinking of elements such as the colour of the backdrop of the visual output of the system as contributing to the construction of a "world" or "universe" which guides their movements and vocalisations during improvisatory phases of motion, noting that "the movements in between [triggering the bounding boxes], because we built up these kind of worlds in each color or each section, a lot of it just comes from the little world that we built. And then I feel like that's enough of preparation. If I know which concept I'm working on, then I can just do whatever fits in my mind, what will fit in that concept" (16). The collaborator took the example of the fourth representation, which had concretised conceptually as an insect, "I think the movement is based on that but then that's also how we build the universe like the sound and the movement came together so I think they're very interconnected and I feel like they're a part of that same. I feel like immediately that I'm in that world and my whole body is this insect and that's including the voice as well" (41). Although framed in concrete terms ("my whole body *is* this insect"), the collaborator emphasises that this is not in the sense of a feeling of inhabiting the body, but rather the ways in which the construction of this conceptual body guides their performance, stating "I'm not so concerned about like how does it move like with the insect, I'm really concerned about like, how does all my different points move . . . I'm really like looking at the

[1]We originally planned for this thesis to include an evaluation of the audience's perception of the relationship between the performer's physical body and the virtual representation of the body constructed by the system over a series of performances for which this survey would have served to pilot several questions. As the collaborator had to leave the project in July 2023, this was unable to take place. The data from these surveys can be found in appendix H.

trying to, okay so if I turn it's going to look like that or if I go jump back and forth it's going to look like that" (58).

However, not every representation provides this sense in the same manner. For example, the second representation, that developed into a concept of a moving structure in a desert wind, was perceived by the collaborator entirely in scenic terms, acting more as backdrop for their physical body: "it's definitely some kind of person that is stuck in this storm and with that image I don't feel like I'm so much the screen. I feel like the screen is more just a backdrop even though I know I'm changing the lines and everything and I feel like that's just a backdrop and then I'm the person. That's just the desert in the background" (54). They highlight the importance of the representational layer to determine their embodied relationship with the representation, stating that "I don't feel such a strong connection no then I just feel like that's just some kind of you know like almost like a heartbeat thing at the hospital or like the radio signal or something. ... I'm doing this [their motions and vocalisations] because it's so far away from the human form anyway, because it's just these very long lines so if I'm moving my arm up the whole thing is moving up and it feels less like I'm moving. It feels more like a separate thing I think because of the way it's made visually" (56).

The collaborator noted the connection between the number of position markers that were used to construct the representation in the role of developing the conceptual worlds and bodies that were being constructed, stating in relation to working with greater or lesser numbers of position markers that "[t]he figure then doesn't feel so like a figure or character. It's more like just some kind of visual representation a bit like the yellow lines. They all feel a bit like that because it's not so apparent. Because when you have all the markers then or like more markers then you get like this character or this assembly of different points. So I think that definitely changes the way the representation feels" (80). With fewer position markers the representation becomes more scenic in nature, "more like an environment or just like a separate something" (82). With a greater number of position markers,[2] it becomes easier to conceptualise in terms of the human form: "I think like the representation would look more like human body. Which I think would feel different. I'm not sure if it would feel like a separate character then or if it will feel like it's more like a representation of me or like I'm moving it and getting affected by it. ... I could go either way I think" (84). However, the more human the representation is perceived as being, the more restrictive it becomes, as there is less opportunity for the collaborator in their role as an observer to contribute to the construction of the world and representation of the body: "I think it would feel like a lot like a character but I'm not sure if it would be a stronger relationship. Because it's this relationship with abstract object

[2] I asked them about their thoughts on using a set of 30 position markers in a performance, the maximum number that the system is capable of working with. We had not done this at the point of the interview, with the maximum number that we had used being 15 due to the issue of latency becoming too noticeable for the collaborator with higher numbers.

like in a laser or an insect or some kind of forest creature. This then it's the atmosphere is so strong and it's more abstract, so I'm creating this world around it. I feel like if it was more human humanoid figure it would be maybe more realistic and maybe not as like open as a world and as like an abstract image of the insect" (84). The feeling of loss of openness in their contributions was counteracted by their ability to determine position marker locations, and thus have an additional structure through which they can participate in the construction, stating that "you don't have to place them so that it looks like a body. ... you could make this very random thing going from the foot to the shoulder to the and then it would just be a shape. Instead of sometimes looking like a human body" (86).

When the system begins to interpolate between the representations, the collaborator's sense of the concepts of the worlds and bodies likewise begins to change, to the extent that they began to identify the desert concept less in scenic terms: "It actually then felt like it was a different person a little bit because I already had that detachment I think to those dots so it felt like that was just like a shape in the desert" (68). This also occurred in relation to other representations, for example "the red one is very different when it changes the shape because it doesn't have that pointy that I feel like are those like devil horns or some kind of like diabolic symbol. And then when it's rounding up it's just a bit more woolly and like not so clear anymore. ... I think that also affects how I move because also when the shapes that kind of goes some of them just goes bigger or more round and then I think I have the feeling that I want to collect all of them ... I'm not able to put them all so close together all the points but then when they're bigger it's a lot easier to just cram down and be and collect all the dots closer to each other" (68).

7.2.2 The Importance of the Visual Modality

The sense of the world is very much determined by the visual modality and the ability of the collaborator to see the representations, with the collaborator emphasising the split between their awareness of the source of the world depending upon their visual relationship with the video wall: "I feel like I am the whatever world it is and then it's not so much me. I feel like then what I'm seeing on the screen that is what I'm doing. That is my, well, that has been my focus a lot I think and then I just move according to that and then sometimes when I'm not looking at the screen I'm just moving. Then it's more like I'm in my own self world and that's just something apart from me" (43).

This is also reflected in extent to which they feel that they are in their body and in the space, noting that "I'm focusing a lot on the screen, that's what I've noticed. And when I'm not focusing on the screen, like I said before, then I feel kind of detached and them I'm just very aware of my body and my position and how I'm moving and that's in a way where I kind of zone out and feel more present in myself and in the room" (60) and "when I'm not looking at the screen I'm moving my arm I'm very aware of how I'm moving my arm but when I look at the screen I'm more aware of how the representation looks on screen than how I'm moving my arm" (62). Despite the importance of the incorporation of their voice into the construction of the world and the representation of the body that inhabits it, this relationship is not present with respect to the collaborator's relationship with the sound output of the system, with the collaborator stating that "I feel like a lot of the relationship is very strong with that screen and the visual. And maybe because the sound is so low[3] it's difficult like that that is somehow not so present for me. The sound which is obviously also part of the physical presence but I feel like the visual is it's stronger than the sound to me" (62).

7.2.3 Tensions of Control

The collaborator made multiple remarks characterising the relationship with the representations in manners such as "almost like that puppet master kind of relationship" (43), indicating that they perceived the relationship between themselves and the representation as one in which they are in control. However, there were several remarks that point towards this relationship being a little more complex.

For example, the collaborator noted that the major points of focus on their body during performance were their feet, eyes, and head, the former because it "changes so much of the other of the rest of your body, the way you put your feet" (71), and the latter due to the desire to look at the visualisation, noting "usually they [the eyes] will be in the body somehow but now they're kind of either in the periphery or directly looking at the screen. So they are kind of a focus in that way because it's affecting the whole way I'm thinking and if I'm not looking at it then it's also affecting how I'm moving" (50). However, at no point in the performance did the collaborator have position markers placed on the head or on the feet. When questioned about how this fact effects the shaping of the representation, the collaborator emphasises the reciprocal relationship between their body and the technological elements in shaping the representation, stating that "now I'm just really using my head to look at the screen or not look at the screen. So if I had something on my head I would have to place my head in a certain way as well. And because I'm focusing so much on the relationship between me and the screen then the head kind of becomes detached from that" (77).

However, the relationship with the (lack of) position markers on the feet adds a complicating element. The collaborator described that they perceive their feet as playing a determining role in their full-body motion, that they "think the hands and the upper body is usually just a lot of time just following what you're doing with your feet" (75) and that "the arms are so naturally moving when you're moving the feet" (75). As the feet are not directly providing motion to the position markers themselves, the perceived source of much of the collaborator's motion that contributes to the construction of the representation is only indirectly represented.

[3]The volume of the sound output was relatively quite in order to control acoustic feedback from the headset worn microphone.

In addition, the shifted representations can provide an unexpected element for the collaborator beyond their control, with the collaborator noting that they "remember last time it [the shifted representations in the second phase] was really like covering the whole screen with those lines at some point so I really expected that to happen. So that kind of surprised me that that didn't happen" (50).

7.2.4 Limitations and Frustrations

The collaborator noted several moments within the performance that were a cause of frustration within the performance. Several of these relate to the ability to move the position markers into the bounding boxes. They relate this specifically to smaller sizes of bounding boxes, emphasising the way in which this brings them out of the moment and sense of their body: "when they're too small it definitely feels a lot of this moving into my head and the intellectual kind of thinking getting frustrated instead of trying to be in the task and being in my body in the room in the space. Then the frustration of not getting it then gets too much" (91). They contrast this with earlier iterations of the systems in which they held the position markers in their hands and there were fewer phases, even if the bounding boxes were equally as small, stating that "I remember that was also kind of a fun challenge if it was very small because then it almost became just like 'get the position' and then try to remember exactly how you move so that you get in the exact right position for the next one" (93). They frame their relationship with these bounding boxes explicitly in terms of a choreography, and that they perceive this as an interference with contribution to the construction of the representation: "that feels more like a task I think and less improvised. Yeah, that feels more like a choreography or like a set score that you have to do. Instead of focusing on the emotion and the feeling and the exploration of the world or the thing that you're creating" (95).

This association of the bounding boxes forcing a form of choreography is also expressed by the collaborator in a further context. Reflecting on the need to define the bounding boxes and their sizes prior to beginning to explore the construction of the world and the representation of the body in the improvasitory phases, the collaborator considers that "I think I was bending down to lower something so I developed my way of doing that phase into a direction that didn't fit into that box anymore. So I think if I could redefine the box I would have done it differently now because that felt more natural. And then, but I didn't realize that that was what was happening so I thought that I was doing it right but I was actually, I had actually developed the thing to something else" (23). Their reflection frames their need to adapt their motion to reach the boxes in terms of right and wrong, in view of which they draw an explicit link to the performance of a choreography, stating that "it's a bit of a frustration because you can't change the limits that you've set which is very often the case in performance anyway, I find, because you have to set some boundaries because if it's too free then yeah. But then these are kind of limited so you don't feel like okay but it was this and now it's turning into this but

Position Markers	Min	Mean	Median	Max	SD	SE
1	0.055	0.111	0.112	0.142	0.02	0.003
5	0.067	0.121	0.112	0.205	0.028	0.004
10	0.087	0.189	0.183	0.348	0.053	0.007

Table 7.3: Summary statistics of motion to visualisation latency with respect to number of *position markers*. Values are in seconds.

it still has to go back to that one shape. So it's almost like a choreography where you have to hit different spots, like you have to hit those spots and then you can be very free within those boxes but you have to come back to that one spot" (31).

A further source of frustration was the provided by the latency of the system. The collaborator implied that it was not the absolute value itself, but rather the inconsistency of the latency between phases, stating "[t]here was some point when it started to be delayed again when I thought, oh no now it's going to be. But then it somehow it went away again. I think it was just like one of the phases or something that was a bit delayed" (50).

7.3 Evaluation of Motion to Visualisation System Latency

In this section, I present the result of the latency evaluation of the performance system. As the mapping and number of computational processes vary across the course of a performance, the end-to-end latency of the system could vary greatly depending upon the point in time during a performance the measurement is taken. Under ideal circumstances, measurements would be taken for each *phase* in each *cycle*, as well as during transitions between *phases*. However, doing this while simultaneously assessing the latency across a range of *position markers*, along with taking enough samples that result in significance for statistical tests between *phases*, quickly results in an absurd number of measurements to take. For example, taking 30 samples for each condition would result in $7 \times 9 \times 30 \times 30 = 56{,}700$ samples. As the time taken to measure one sample is approximately 1 min, completing this would take approximately 945 hours, not to mention the space required for data storage. Instead I collected 5 samples for each *cycle* and *phase* for 1, 5, and 10 *position markers*, as this was the range across which we primarily worked in rehearsals and development sessions. This results in a total of 945 samples.

The system component of *The Shapeshifter* was running as if in a performance, with a few minor variations to simplify data processing. All objects that were not the visualisation of the *position marker* were altered so as not be included in the scene lighting and had their colour set to black. This meant that they were being rendered, but were not visible. For conditions with higher numbers of active *position markers* than one, all except the marker affixed to the pendulum were provided dummy

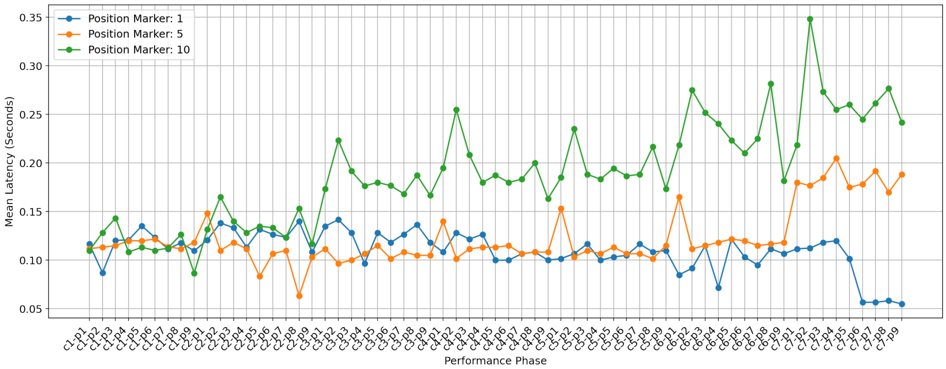


Figure 7.4: The mean latency of each *phase* across a performance.

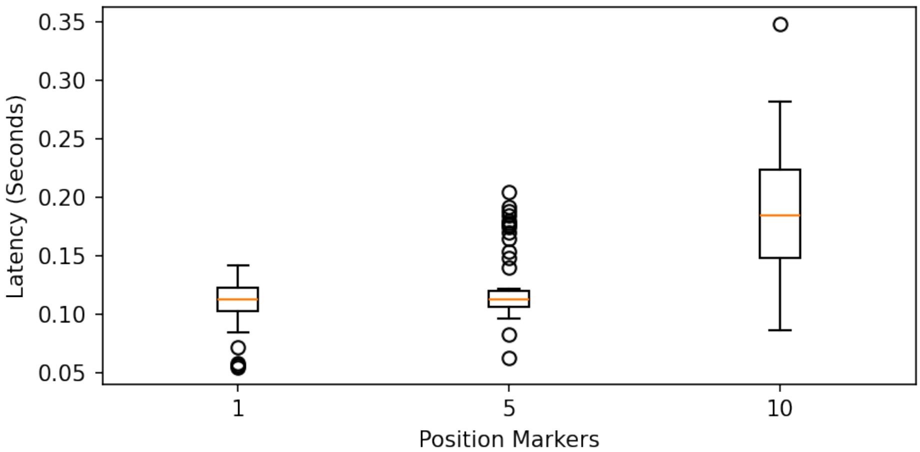


Figure 7.5: Box plot of mean motion to visualisation latency across *position markers*.

data that moved them to a corner of the display. The camera used was a Panasonic HC-X2000.[4] Video was captured at a frame rate of 120 Hz, resulting in a possible error of approximately 8 ms, and a resolution of 1920×1072.

The results of the measurements can be viewed in terms of the relationship between the three independent variables, namely the number of *position markers*, the *phase*, and the *cycle*. An overview of mean *phase* latency across a performance is shown in figure 7.4. Table 7.3 presents summary statistics of the latency with respect to the number of *position markers*. Summary statistic relating to *position marker*, *cycle*, and *phase* respectively are likewise presented in figures 7.5, 7.6, 7.7.

A Levene's test relating to the variances of the *position markers* groups returns a statistic of 84.117 with $p < 0.05$, supporting a rejection of the null hypothesis that the variances of the groups are equal. As homoscedasticity

[4]https://www.panasonic.com/no/consumer/kameraer-og-videokameraer/videokameraer/4k-videokameraer/hc-x2000.html

Position Markers	Statistic	P-Value
1	9.037	< 0.05
5	4.576	< 0.05
10	11.815	< 0.05

Table 7.4: The results of a Levene's test in relation to *cycle* regardless of *phase*.

Position Markers	Statistic	P-Value
1	16.392	< 0.05
5	151.949	< 0.05
10	105.159	< 0.05

Table 7.5: The results of a Welch's ANOVA in relation to *cycle* regardless of *phase*.

Position Markers	Statistic	P-Value
1	6.538	< 0.05
5	0.432	0.901
10	1.022	0.419

Table 7.6: The results of a Levene's test in relation to *phase* regardless of *cycle*.

Position Markers	Statistic	P-Value
1	6.288	< 0.05
5	3.227	< 0.05
10	4.559	< 0.05

Table 7.7: The results of a Welch's ANOVA for one *position marker* and a one-way ANOVA for five and 10 *position markers* in relation to *phase* regardless of *cycle*.

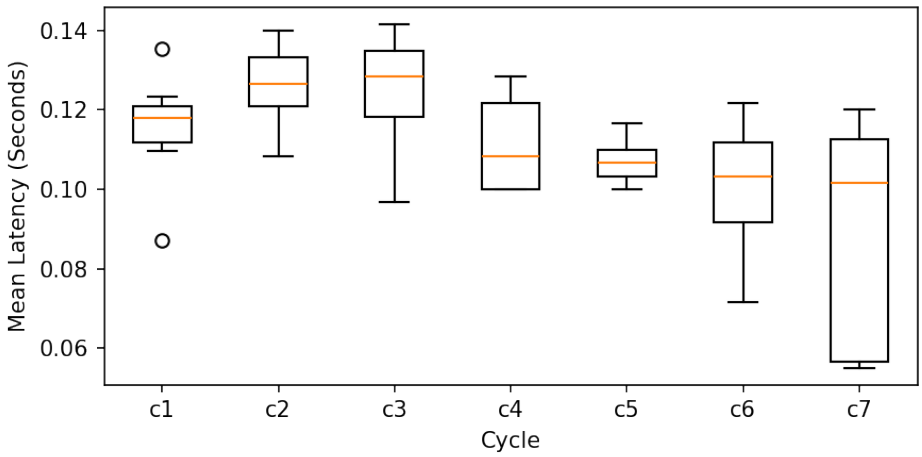
cannot be assumed for the groups, a Welch's ANOVA, which does not assume homoscedasticity, returns a statistic of 226.164 with $p < 0.05$, supporting a rejection of the null hypothesis that the means of all groups are equal. This supports a conclusion that the number of *position markers* has a significant effect on motion to visualisation latency.

For each value of *position marker*, the results of a Levene's test relating to the *cycle* of the performance regardless of *phase* are shown in table 7.4. These results support a rejection of the null hypothesis that the variances of the latency values between *cycles* are equal. The results of a Welch's ANOVA for each *position marker* are shown in table 7.5. As $p < 0.05$ for all *position markers*, these results support a conclusion that the *cycle* of the performance has a significant effect on motion to visualisation latency.

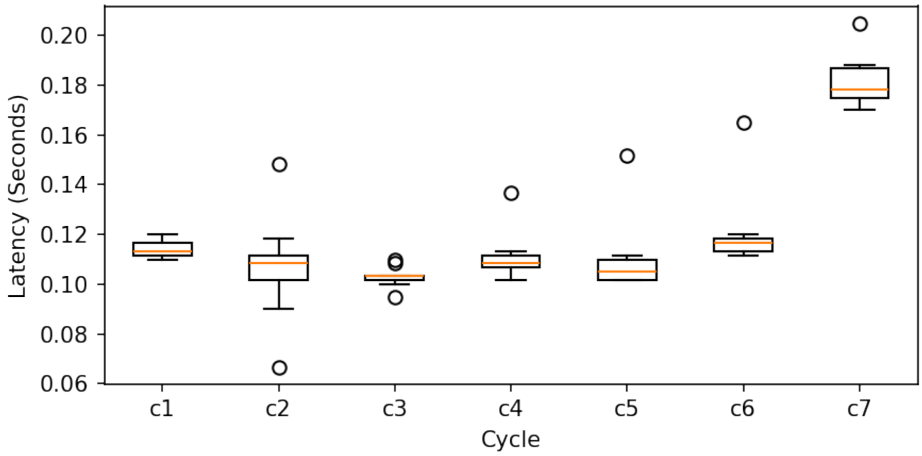
In addition, for each value of *position marker*, the results of a Levene's test relating to the *cycle* of the performance regardless of *phase* are shown in table 7.6. For the variances between *phases* for one *position marker*, $p < 0.05$, supporting a rejection of the null hypothesis that the variances of the latency values between *phases* are equal. However, for the variance between *phases* for five and 10 *position markers*, $p > 0.05$, supporting a rejection of the alternative hypothesis that the variances of the latency values between *phases* are not equal. In view of this, the results of a Welch's ANOVA for one *position marker* and a one-way ANOVA for five and 10 *position makers* are presented in table 7.7. These results support a conclusion that the *phase* of the performance has a significant effect on motion to visualisation latency.

7.4 Summary

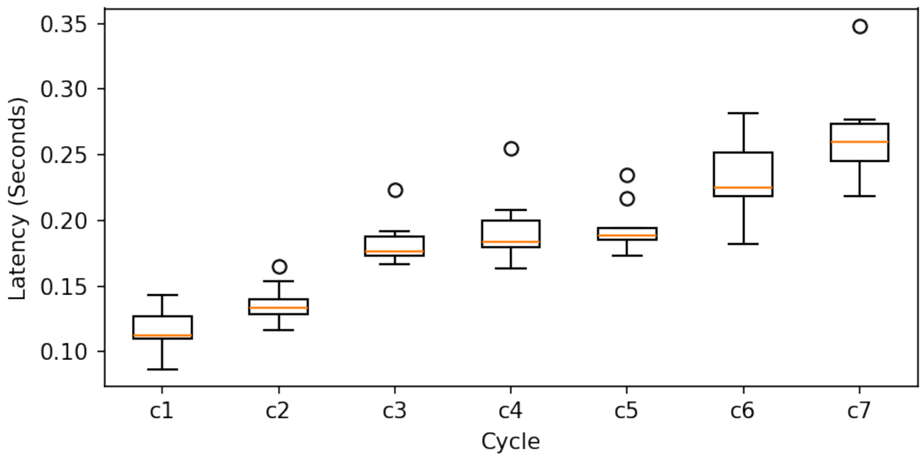
In this chapter, I presented the results of the iterative and participatory process of system design and performance development, the thematic categories induced from the coded interview with the collaborator after the pilot performance, and the results of an evaluation of the system's motion to visualisation latency. In the following chapter, I discuss these results and relate them back to the thesis research questions.



(a) Box plot of mean motion to visualisation across *cycles* - 1 *position marker*.

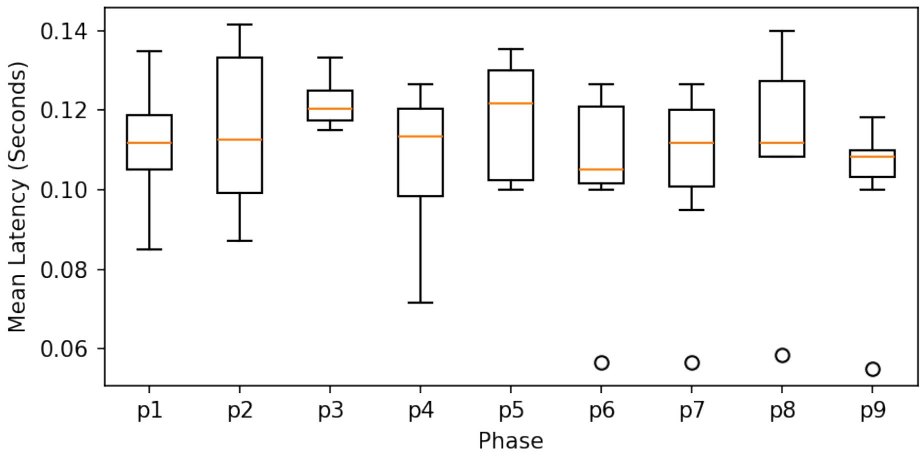


(b) Box plot of mean motion to visualisation across *cycles* - 5 *position marker*.

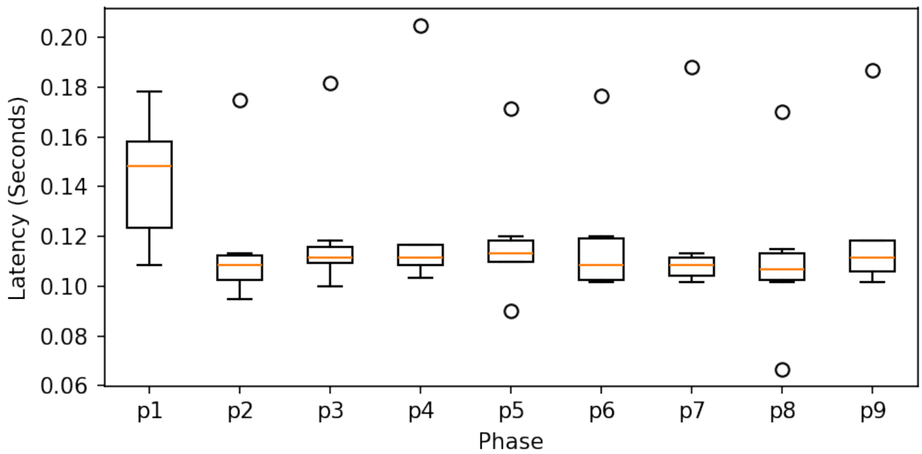


(c) Box plot of mean motion to visualisation across *cycles* - 10 *position markers*.

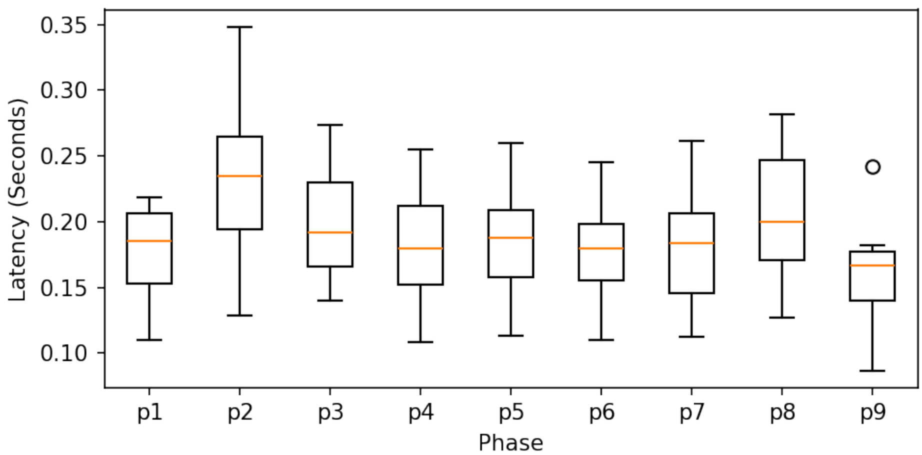
Figure 7.6: Box plot of mean motion to visualisation across *cycles* regardless of *phase*.



(a) Box plot of mean motion to visualisation across *phases* - 1 *position marker*.



(b) Box plot of mean motion to visualisation across *phases* - 5 *position marker*.



(c) Box plot of mean motion to visualisation across *phases* - 10 *position markers*.

Figure 7.7: Box plot of mean motion to visualisation across *phases* regardless of *cycle*.

Chapter 8

Discussion and Conclusions

In this chapter, I discuss and synthesise the results of the iterative and participatory development, the post-performance interview, and the system latency evaluations. Thereafter, I draw several conclusions in relation to the research questions around which I centre this thesis. Finally, I discuss several limitations of this thesis and possible paths of future work.

8.1 Discussion

Framing the development process specifically around the construction of bodily representation rather than around functionalities and mappings led to a system which on the surface runs counter to several of the guidelines developed by Masu et al. (2019). Firstly, there are only three major real-time mappings within the system: motion-to-visualisation, sound-to-sound, and sound-to-visualisation. Moreover, the motion-to-motion mapping is distinctly contrary to Masu et al.'s guideline that mappings do not reproduce information visible in the physical performance, with a direct mapping of position in the *physical performance space* to position in the *virtual performance space*. However, complexity is added through the integration of the cumulative effect of motion and audio captures occurring throughout the performance which slowly and subtly shift these mappings and their representations at the system output, which, moreover, can be completely novel between performances. This creates room for the performer to explore these new spaces and creates the opportunity for them to find new meanings in the representations of the body through these mappings.[1] From a perspective grounded in Schlemmer, these mappings place restrictions upon the ways in which the representations interact with the virtual space, in effect applying Schlemmer's laws regarding costuming and the *Tänzermensch* in subtly different and unpredictable ways in each performance which propagate outwards into the performer's performance.

The collaborator's focus on their sense of embodiment stemming from

[1] The results from the piloted audience surveys indicate that the audience members additionally interpreted the work in a variety of ways. As a full audience evaluation was unable to take place, these results have been excluded from this thesis. However, the raw data are available in appendix H.

concepts originating in the world of the representation can be linked to their background in physical theatre. As noted by Sanchez-Colberg (1996), physical theatre's roots in German expressionist dance places much emphasis on movement extending outwards from the body in space, in the sense that "dance is about the body, in space, through movement" (p. 45). In view of this, it is important to not just consider the body that is being constructed through the use of MoCap, but how this construction influences the perception of the space which it inhabits. Through eschewing techniques on the construction layer that require precise numbers of markers and static positioning, we moved a part of the process of this layer into the performance itself. This highlighted the performer's role in reciprocally constructing the representation of the body and the role that this plays in terms of their embodied relationship with the constructed body, stemming from a holistic view of the system's output modalities, their motion, and their vocalisation.

The collaborator's focus on the screen and feelings of frustration when not able to trigger the bounding boxes as planned can be related to Purser's (2017) finding that when analysing interviews conducted with contemporary dancers from an approach grounded in Merleau-Pontian phenomenology, many of the dancers employed the terminology of some variation of "being in your body" (p. 45) in order to describe their subjective, embodied experience. The interview participants contrast the sense of "being in your body" with "making shapes" for observers external to the body (including viewing oneself in the mirror or being under the gaze of a camera), the latter of which instantiates a sense of bodily awareness more instrumental in nature.[2] Specifically, the collaborator employed exactly this terminology when discussing the missing bounding boxes, framing this as a negative.

However, the collaborator's statement that they were more in their body when they looking away from the screen provides a contrast to the feeling of frustration that develops when they were brought out of their body by the bounding boxes. Although they instrumentalise their body to an extent based upon feedback from the screen, "making shapes" in order to explore the representation, they made clear that they felt a much stronger sense of the world and representation of the body that they are co-constructing when focusing on the visuals, and to a lesser extent the sound. This sense structured their movement and vocalisations, returning them to a bodily focus.

Thinking about the bounding boxes in terms of "making shapes" can also reveal their relationship with the kinematic model. The bounding boxes required specific position markers to be in specific positions simultaneously in order to function and trigger the subsequent phase. Framed in these terms, the functionality resembles that of the construction of a kinematic model when using an optical, marker-based system. However, in-

[2]Purser also notes that a sense of "being in the moment" was likewise important for the dancers, and comes to the conclusion that the space between these two senses of being results in access to a further mode of being she terms "inhabited transcendence" (p. 50).

stead of these requirements being located in relation with the body, they are located in relation to the space. This is closely related to one of the major affordances of this type of motion capture, namely that it provides a coordinate system independent of the object being tracked. In this sense, it functions as a form of "inverse" of the kinematic model, relating the construction of the representation of the body to the space instead of the physical body present in that space. However, in contrast to the static and calcified positioning of markers required by kinematic modelling, defined during development of the MoCap system by it's designers, the positioning required through the use of the bounding boxes is malleable, determined by the collaborator before each performance.

The collaborator's frustration with the bounding boxes ties this relation to the kinematic model to the conceptualisation of the sensing device. During the pilot performance the collaborator did not interact to a major extent with the position markers, instead primarily leaving them in static positions on their body. This led to them conceptualising the bounding boxes as forcing a sort of choreography and placing limitations on the improvisations between the boxes, indicating that the collaborator was conceptualising the position markers as an extension of their own body. This was not the case, however, in the early exploration sessions in which the collaborator held the position markers in their hands and interacted with them as objects external to the body, which they found more like a "game" or "challenge", creating an environment in which the collaborator could be in their body. They only became perceived as limiting when the way of thinking about constructing the representation of the body became heavily linked with the physical body. When the sensors stand in for what is being sensed, the collaborator found themselves making shapes.

The collaborator's frustration with the system latencies reveals a similar pattern. There is little in the literature on motion to visual latency from a performer's perspective in a live, interactive performance, possibly linked to the fact that in most multimedia performance involving visuals these are projected behind the performer (Birringer, 2015). However, for example, in relation to motion tracking involving head-mounted displays such as for alternate reality, extended reality, and virtual reality applications, Adelstein et al. (2003) propose a maximum threshold of 17 ms above which latency becomes noticeable. In reference to pointing tasks Pavlovych and Stuerzlinger (2009) note that motion based video game controllers such as the Wii remote display latencies of 106.3 ms. The mean system latencies for one and five position markers are roughly in line with the latency of a Wii remote in a pointing task. However, for 10 position markers, within the range of markers with which we would commonly work, the mean latency value was approximately 1.8 times this value, reaching a maximum of 348 ms for one phase. However, the baseline values in relation to head-mounted displays and pointing tasks in video games should be a lot lower than would be acceptable for interactive dance, as both require a much higher degree of singular focus, and in the case of head-mounted displays can also lead to motion sickness. Moreover, the perception of this latency is complicated through other factors which alter the motion of the virtual marker, such as the interpolation mapped to the audio input. This follows the collaborator's experience, where latency only became frustrating when there was a sudden change. Instead, the stable values of latency seemed to be incorporated into their performance, conceptualised as a part of the functioning of the system and therefore feeding into the process of co-constructing the representation of the body.

8.2 Conclusions

In this thesis, I aimed to critically explore the process of the construction of a virtual representation of the body through the use of optical, marker-based motion capture in live dance. I formalised this through three research questions which were investigated in the context of a research-creation project in which I employed a number of theoretical, qualitative, and quantitative methods as well as the development of a theoretical model and a performance including the design of an interactive system. The work done in this thesis provides the following answers to these research questions:

RQ. 1 How does the use of a motion capture system construct a virtual representation of the body in live performance, and which assumptions about the body does it make?

From a perspective grounded in a view of the human body informed by disability studies, I outlined how the wider and social environment shapes assumptions about what the body is, how it can move, and how it can be represented. MoCap technology employed in a context of dance represents a cultural reproduction of the body which is imbued with the ethical and aesthetic values of the network of participants in its production. This does not just occur at the representation of the output of the systems employed to construct the body, but rather is embedded in the affordances with which the technologies are designed and the ways of working that they encourage. In addition, both the performer and observer in a performance contribute to this construction through their engagement and understanding of the relationship between the performer and the virtual representation of the body. Based upon this, I framed the use of MoCap in live dance performance as a co-constructive process, and provided a model of how the values and assumptions of participants in this process interact with each other and with the technology. This model is intended to provide a framework as a supplement to commonly found systems-oriented models of MoCap in domains such as animation, interactive music, and interactive dance, in order to provide the developers of performance works, interactive systems, and underlying technologies with a lens not just to view the flows of information through systems but also the ways of thinking about the body that are embedded in their design and functionalities.

RQ. 2 How can a multi-modal interactive system be iteratively designed from a perspective which foregrounds motion capture as co-constructing a representation of the body?

Together with the collaborator in the research-creation project, we developed a dance and physical theatre performance which incorporated an interactive system built upon an optical, marker-based MoCap system. We grounded our development using the framework provided by the co-construction model, to identify assumptions and values which we were embedding in the system. We conceptualised the markers to have a fluidity in their role as either standing in for a body part or as objects external to the body, viewing them as objects which constrain the body's movement. We also drew upon previous work done in the development of the guidelines for *motion pointillism*, and the work of Oskar Schlemmer when considering how the the markers and representations at the system's output can restrict the performer's motion and thus contribute to the construction of the representations in relation to the performer's physical body. Although this resulted in a system in which there are only three rather direct one-to-one real-time mappings, a complexity was achieved through considering how the performer is contributing to the construction of the representation temporally across the performance, and working with motion and audio captures taken throughout the performance to influence how these mappings are actualised. We developed a system in which bounding boxes in the physical performance space served as triggers within the performance. I relate this to the practicalities of the kinematic model, contextualised, however, not just in relation to the performer's body but to the spatial properties of the physical performance space and a maleability of form. The performer's conceptualisation of the position markers likewise plays a role in the ways in which the bounding box system contributes to the construction of the body. An evaluation of the system's motion to visualisation latency showed that the system latency changes across the performance in relation to the underlying computational operations occurring.

RQ. 3 How does a performer experience their body in relation to the technological components in performance with a system for interactive dance?

The collaborator experienced an alternating sense of being in their body and instrumentalising their body. They displayed a tendency towards conceptualising the technological components with which they were interacting as elided with their body, removing them from their conceptual chain and treating them as stand-ins for the body parts to which they were attached. This became more prominent as we moved through the development process, using increasing numbers of position markers. This also resulted in feelings of frustration as the unique affordances, properties, and functionalities of the technological systems became perceived as limitations placed upon their body.

The collaborator framed their relationship with the representations of the body within the performance layer as tightly connected with ideas of space. Their perception of the body originated in the "worlds" that these representations created, conceptualising the representation of the body and the space as interwoven phenomena that contribute to structuring their motion and vocalisations. This relationship between their body, the representation, and the virtual and physical space was perceived by the collaborator in a holistic manner, until this sense was interrupted by perceived limitations in the system, which brought them out of the feeling of being in their body. The system's latency was not perceived as such as frustration, unless there were sudden changes in the latency values.

8.3 Limitations and Future Work

The conclusions reached in this thesis must be taken in the context of the following limitations. First and foremost, the work done in this thesis represents a single project undertaken with a single collaborator. This means that the findings, especially relating to the performance development and collaborator's experiences, are not generalisable.

The co-construction model developed in chapter 3 was developed along a single structure of a perspective on the body taken from disability studies. Further theoretical perspectives can be explored which can further support this model. Moreover, this thesis was solely concerned with optical, marker-based MoCap, in particular the use of kinematic modelling, so the applicability of the model to additional forms of MoCap presents a path of future work. This could be, for example, in relation to the conceptualisation of the IMU in inertial systems and the models and methods employed to construct the body with this technique. Moreover, the work done in this thesis is primarily concerned with motion tracking involved in interactive performance. However, the co-construction model also relates to the use of motion capture. Although captures were employed in this thesis, they were integrated as a component of the interactive system, so further work can be undertaken to investigate the applicability of the model to performance in which the physical body of the performer is solely related to a representation constructed from motion captures.

The project was also limited by several factors which impacted the time frame available for performances, and by extension any forms of data collection reliant upon performance. Firstly, the *Portal* underwent a significant renovation in July 2023. As the installed technological infrastructure was significantly altered, this limited the ability to rehearse performances and explore the system further without significant alterations to the system code. In addition, the collaborator had to leave the project from the end of July 2023 for at least one year. Although we had made provisions to continue developing the performance and take it to external, local venues, these plans instead present future work. This also limited any evaluation of audience perspectives on the performance and co-construction model. As the observer plays a significant role within the model and the *motion poin-*

tillism guidelines that were employed, full evaluation of the perceptions of the audience in view of the relationship between the performer's body and the virtual representations presents a possibility for future work.

The machine learning component of the system remains relatively unexplored. Firstly, the model performs sufficiently, although not exceedingly well. As the time period to collect the dataset and train the model was limited, there is a lot of room to optimise the features extracted from the data and the model in order to obtain a better performance. There is also the question of how to properly evaluate the model's use in the context of a performance. The model is employed in order to provide the shifting algorithm with a mechanism through which to weight the shift in representations towards other representations, based upon the performer's vocalisations resembling those in other representation phases. The use of machine learning is to provide this with a functionality somewhere between a random number generator and a fully pre-determined rules-based approach where the vocalisation parameters relating to each phase must be consolidated into the system design and therefore provide little flexibility. A hypothetical perfectly trained model would classify the vocalisations correctly according to the features with which it has been presented. However, the performer's perception of a vocalisation might not align with these classifications, as the performer's experience of the holistic performance environment might provide them with an entirely different understanding of the link between their vocalisations and the phases of the performance. Moreover, these relationships will probably not be identical across or even within performances. From this view, metrics relating to the accuracy and precision of the model become less important, and instead evaluation should focus on whether the model's "vibe", as put by Grietzer (2017), matches the performer's vibe.

The effects of the spatialisation of the audio output of the system are likewise unexplored in this thesis. For *The Shapeshifter*, we decided to spatialise the audio for aesthetic reasons, placing the audience within the soundfield to mirror the way in which they are positioned within the performance. However, future work could comprise investigating how the performer relates to the sonic representation of the co-constructed body, specifically in view of the spatialisation.

There are several additional technical metrics for the system that can also be evaluated. For example, the non-real-time processing must be completed within a hard time limit, namely that all processing for a given *phase* in a given *cycle* must be complete before the *phase* is reached in the next cycle. At present, based upon print outs during performance, it appears to be the case that this is achieved. However, a more systematic evaluation of this metric relating to the number of position markers and the length of the *phase* is necessary, as the time complexity of several of the non-real-time processing modules is not linear. Moreover, the evaluation of the system latencies presents opportunity for future work. Firstly, the methods have not been validated, so validating these methods is a priority. Secondly, the analysis of the motion to visualisation is quite limited in this thesis, based upon taking the correlation between the two sinusoidal signals. However, other parameters of the sinusoids could be evaluated, for example the smoothness of the slope in order to evaluate the smoothness of the visualisation and relate this to the performer's perception of the representation and how this influences and effects their performance. Moreover, I have not been able to find any baseline values for the perception of motion to visual latency in interactive performance. Investigating these values and how various latency values are perceived, not only by performers but also by observers and audiences, and the ways in which the system latency influences the performance present an opportunity for future work.

Finally, for *The Shapeshifter* we grounded the development of the performance and the design of the system in the works of Oskar Schlemmer. As noted by Norman (2015), Schlemmer's work has achieved little influence within the field of motion computing, especially in comparison to his contemporary Rudolf Laban. Further exploration into the integration of Schlemmer's ideas, concepts, and methods relating the body to space in motion analysis and interactive system design offers a wide range of possibilities for future work.

Works Cited

ABBA Voyage. (2021). Abba voyage: How abba used motion capture to create their avatars. Retrieved September 19, 2023, from https://www.facebook.com/ABBAVoyage/videos/abba-voyage-how-abba-used-motion-capture-to-create-their-avatars/399264848445591/

ABBA XXI. (2022). Abba voyage interview with the choreographer of the show wayne mcgregor. Retrieved September 19, 2023, from https://www.youtube.com/watch?v=hI6GuKxVCkI

Abouaf, J. (1999a). "biped": A dance with virtual and company dancers. 1. *IEEE MultiMedia*, *6*(3), 4–7. https://doi.org/10.1109/93.790605

Abouaf, J. (1999b). "biped": A dance with virtual and company dancers. 2. *IEEE MultiMedia*, *6*(4), 5–7. https://doi.org/10.1109/93.809227

Adelstein, B. D., Lee, T. G., & Ellis, S. R. (2003). Head tracking latency in virtual environments: Psychophysics and a model. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *47*(20), 2083–2087. https://doi.org/10.1177/154193120304702001

Al-Naymat, G., Chawla, S., & Taheri, J. (2009). Sparsedtw: A novel approach to speed up dynamic time warping. *Proceedings of the Eighth Australasian Data Mining Conference - Volume 101*, 117–127.

Arlander, A. (2009). Artistic research — from apartness to the umbrella concept at the theatre academy, finland. In S. R. Riley & L. Hunter (Eds.), *Mapping landscapes for performance as research: Scholarly acts and creative cartographies* (pp. 77–83). Palgrave Macmillan UK. https://doi.org/10.1057/9780230244481_9

Auslander, P. (2023). *Liveness: Performance in a mediatized culture* (Third edition). Routledge.

Baalman, M. A. (2010). Spatial composition techniques and sound spatialisation technologies. *Organised Sound*, *15*(03), 209–218. https://doi.org/10.1017/S1355771810000245

Balsamo, A. (1995). Forms of technological embodiment: Reading the body in contemporary culture. *Body society*, *1*(3–4), 215–237.

Barber, T. E. (2015). Ghostcatching and after ghostcatching , dances in the dark. *47*, 44–67. https://doi.org/10.1017/S0149767715000030

Bardzell, S. (2010). Feminist hci: Taking stock and outlining an agenda for design. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1301–1310. https://doi.org/10.1145/1753326.1753521

Barton, B. (2017). Introduction i: Wherefore par?: Discussions on "a line of flight". In A. Arlander, B. Barton, M. Dreyer-Lude & B. Spatz (Eds.), *Performance as research* (pp. 1–19). Routledge.

Baumgartner, H. (1999). How to catch a ghost. *Mechanical Engineering-CIME*, *121*(4), 108.

Bevilacqua, F., Schnell, N., & Fdili Alaoui, S. (2011). Gesture capture: Paradigms in interactive music/ dance systems. In G. Klein & S. Noeth (Eds.), *Emerging bodies: The performance of worldmaking in dance and choreography* (pp. 183–194). transcript Verlag. https://doi.org/doi:10.1515/transcript.9783839415962.183

The bigger picture: Abba voyage. (2022). *Engineering Technology*, *17*(6), 14–15. https://doi.org/10.1049/et.2022.0622

Birnbaum, D. M. (2007). *Musical vibrotactile feedback* [Master Thesis]. McGill University [tex.entrytype: other]. https://escholarship.mcgill.ca/concern/theses/xd07gx894

Birringer, J. (2004). Interactive dance, the body and the internet. *Journal of Visual Art Practice*, *3*(3), 165–178. https://doi.org/10.1386/jvap.3.3.165/0

Birringer, J. (2015). Gestural materialities and the worn dispositif. In N. S. Sutil & S. Popat (Eds.), *Digital movement: Essays in motion technology and performance* (pp. 162–185). Palgrave Macmillan.

Bisig, D. (2022). Generative dance - a taxonomy and survey. *Proceedings of the 8th International Conference on Movement and Computing*, 1–10. https://doi.org/10.1145/3537972.3537978

Bjerknes, G., Ehn, P., & Kyng, M. (Eds.). (1987). *Computers and democracy - a scandinavian challenge* [Citation Key: bca59c80d1ff11df8cb9000ea68e967b]. Gower Publishing.

Blythe, M. (2017). Research fiction: Storytelling, plot and design. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 5400–5411. https://doi.org/10.1145/3025453.3026023

Bodenheimer, B., Rose, C., Rosenthal, S., & Pella, J. (1997). The process of motion capture: Dealing with the data. In D. Thalmann & M. Van De Panne (Eds.), *Computer animation and simulation '97* (pp. 3–18). Springer Vienna. https://doi.org/10.1007/978-3-7091-6874-5_1

Borgdorff, H. (2006). The debate on research in the arts. *Sensuous Knowledge 2*.

Brenton, H., Kleinsmith, A., & Gillies, M. (2014). Embodied design of dance visualisations. *Proceedings of the 2014 International Workshop on Movement and Computing*, 124–129. https://doi.org/10.1145/2617995.2618017

Brinkmann, S., & Kvale, S. (2018). *Doing interviews*. SAGE Publications Ltd. https://doi.org/10.4135/9781529716665

Camurri, A., Volpe, G., Piana, S., Mancini, M., Niewiadomski, R., Ferrari, N., & Canepa, C. (2016). The dancer in the eye: Towards a multi-layered computational framework of qualities in movement. *Proceedings of the 3rd International Symposium on Movement and Computing*, 1–7. https://doi.org/10.1145/2948910.2948927

Candy, L. (2006). Practice based research: A guide. *CCS report*, *1*(2), 1–19.

Copeland, R. (2004). *Merce cunningham: The modernizing of modern dance*. Routledge.

Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry research design: Choosing among five approaches* (Fourth edition). SAGE.

Daniels, G. S. (1952). *The average man?* (Tech. rep.). Air Force Aerospace Medical Research Lab Wright-Patterson AFB OH. https://apps.dtic.mil/sti/citations/tr/AD0010203

Davis, L. J. (1995). *Enforcing normalcy: Disability, deafness, and the body*. Verso.

Delbridge, M. (2014). The costume of mocap: A spatial collision of velcro, avatar and oskar schlemmer. *Scene*, *2*(1), 221–232. https://doi.org/10.1386/scene.2.1-2.221_1

Di Luca, M. (2010). New method to measure end-to-end delay of virtual reality. *Presence*, *19*(6), 569–584. https://doi.org/10.1162/pres_a_00023

Dils, A. (2002). The ghost in the machine: Merce cunningham and bill t. jones. *PAJ: A Journal of Performance and Art*, *24*(1), 94–104.

Dixon, S. (2007). *Digital performance: A history of new media in theater, dance, performance art, and installation*. The MIT Press. https://doi.org/10.7551/mitpress/2429.001.0001

Doutreligne, S. (2020). From critical manifesto to transformative genderplay: Feminist futurist and dadaist strategies countering the misogyny within the historical avant-gardes. *Documenta*, *38*(2). https://doi.org/10.21825/documenta.81885

Downie, M. N. (2005). *Choreographing the extended agent: Performance graphics for dance theater* [Ph.D Thesis]. Massachusetts Institute of Technology [publisher: Massachusetts Institute of Technology, School of Architecture and Planning ...]. https://www.media.mit.edu/publications/choreographing-the-digital-agent-live-performance-graphics-for-dance-theater/

Dupont, M., & Marteau, P.-F. (2016). Coarse-dtw for sparse time series alignment. In A. Douzal-Chouakria, J. A. Vilar & P.-F. Marteau (Eds.), *Advanced analysis and learning on temporal data* (pp. 157–172). Springer International Publishing.

Eldridge, A., Kiefer, C., Overholt, D., & Ulfarsson, H. (2021). Self-resonating vibrotactile feedback instruments | |: Making, playing, conceptualising :| |. *NIME 2021*. https://doi.org/10.21428/92fbeb44.1f29a09e

Erdem, C., Wallace, B., & Jensenius, A. R. (2022). Cavi: A coadaptive audiovisual instrument–composition. *NIME 2022*. https://doi.org/10.21428/92fbeb44.803c24dd

Fasciani, S. (2014). *Voice controlled interface for digital musical instrument* [Ph.D Thesis] [Citation Key: $10635_1 18256$*tex.entrytype : other*]. https://scholarbank.nus.edu.sg/handle/10635/118256

Fdili Alaoui, S. (2019). Making an interactive dance piece: Tensions in integrating technology in art. *Proceedings of the 2019 on Designing Interactive Systems Conference*, 1195–1208. https://doi.org/10.1145/3322276.3322289

Fdili Alaoui, S., Carlson, K., & Schiphorst, T. (2014). Choreography as mediated through compositional tools for movement: Constructing a historical perspective. *Proceedings of the 2014 International Workshop on Movement and Computing*, 1–6. https://doi.org/10.1145/2617995.2617996

Fdili Alaoui, S., Jacquemin, C., & Bevilacqua, F. (2013). Chiseling bodies: An augmented dance performance. *CHI '13 Extended Abstracts on Human Factors in Computing Systems*, 2915–2918. https://doi.org/10.1145/2468356.2479573

Fels, S. (2000). Intimacy and embodiment: Implications for art and technology. *Proceedings of the 2000 ACM workshops on Multimedia*, 13–16. https://doi.org/10.1145/357744.357749

Fels, S., Gadd, A., & Mulder, A. (2002). Mapping transparency through metaphor: Towards more expressive musical instruments (2003/01/17). *Organised Sound*, *7*(2), 109–126. https://doi.org/10.1017/S1355771802002042

Finley, S. (2007). Arts-based research. In J. G. Knowles & A. L. Cole (Eds.), *Handbook of the arts in qualitative research: Perspectives, methodologies, examples, and issues* (pp. 71–81). Sage Publications.

Frayling, C. (1993). Research in art and design. *Royal College of Art research papers*, *1*, 1–5.

Friston, S., & Steed, A. (2014). Measuring latency in virtual environments. *IEEE Transactions on Visualization and Computer Graphics*, *20*(4), 616–625. https://doi.org/10.1109/TVCG.2014.30

Gallagher, S. (2012). Taking stock of phenomenology futures. *The Southern Journal of Philosophy*, *50*(2), 304–318. https://doi.org/10.1111/j.2041-6962.2012.00108.x

Garland-Thomson, R. (1997). *Extraordinary bodies: Figuring physical disability in american culture and literature*. Columbia University Press.

Giorgi, B., Giorgi, A., & Morley, J. (2017). The descriptive phenomenological psychological method. In C. Willig & W. Stainton Rogers (Eds.), *The sage handbook of qualitative research in psychology* (2nd ed., pp. 176–192). SAGE Publications Ltd.

Giorgino, T. (2009). Computing and visualizing dynamic time warping alignments in r: The dtw package. *Journal of Statistical Software*, *31*(7). https://doi.org/10.18637/jss.v031.i07

Gleicher, M., & Ferrier, N. (2002). Evaluating video-based motion capture [Citation Key: 1017510]. *Proceedings of computer animation 2002 (CA 2002)*, 75–80. https://doi.org/10.1109/CA.2002.1017510

Gleicher, M. (1999). Animation from observation: Motion capture and motion editing. *SIGGRAPH Comput. Graph.*, *33*(4), 51–54. https://doi.org/10.1145/345370.345409

Goldman, D. (2003). Ghostcatching: An intersection of technology, labor, and race. *Dance Research Journal*, *35/36*, 68–87.

Gould, J. D., & Lewis, C. (1985). Designing for usability: Key principles and what designers think. *Communications of The Acm*, *28*(3), 300–311. https://doi.org/10.1145/3166.3170

Grietzer, P. (2017). A theory of vibe. *Glass Bead*, *1*. https://www.glass-bead.org/article/a-theory-of-vibe/?lang=enview

Gunkel, D. J. (2018). The relational turn: Third wave hci and phenomenology. In M. Filimowicz & V. Tzankova (Eds.), *New directions in third wave human-computer interaction: Volume 1 - technologies* (pp. 11–24). Springer International Publishing. https://doi.org/10.1007/978-3-319-73356-2_2

Hall, M. C. (2018). *The bioethics of enhancement: Transhumanism, disability, and biopolitics*. Lexington Books. http://www.vlebooks.com/vleweb/product/openreader?id=none&isbn=9781498533492

Hansen, N. B., Dindler, C., Halskov, K., Iversen, O. S., Bossen, C., Basballe, D. A., & Schouten, B. (2019). How participatory design works: Mechanisms and effects. *Proceedings of the 31st Australian Conference on Human-Computer-Interaction*, 30–41. https://doi.org/10.1145/3369457.3369460

Hansen, P. (2017). Research-based practice: Facilitating transfer across artistic, scholarly, and scientific inquiries. In A. Arlander, B. Barton, M. Dreyer-Lude & B. Spatz (Eds.), *Performance as research* (pp. 32–49). Routledge.

Hayles, N. K. (1999). *How we became posthuman: Virtual bodies in cybernetics, literature, and informatics*. University of Chicago Press.

Herrmann, M., & Webb, G. I. (2021). Early abandoning and pruning for elastic distances including dynamic time warping. *Data Mining and Knowledge Discovery*, *35*(6), 2577–2601. https://doi.org/10.1007/s10618-021-00782-4

Hiyadi, H., Ababsa, F., Montagne, C., Bouyakhf, E. H., & Regragui, F. (2016). Adaptive dynamic time warping for recognition of natural gestures [journalAbbreviation: 2016 Sixth International Conference on Image Processing Theory, Tools and Applications (IPTA)]. *2016 Sixth International Conference on Image Processing Theory, Tools and Applications (IPTA)*, 1–6. https://doi.org/10.1109/IPTA.2016.7820971

Høffding, S., & Martiny, K. (2016). Framing a phenomenological interview: What, why and how. *Phenomenology and the Cognitive Sciences*, *15*(4), 539–564. https://doi.org/10.1007/s11097-015-9433-z

Hunt, A., & Kirk, R. (2000). Mapping strategies for musical performance. In *Trends in gestural control of music* (pp. 231–258, Vol. 21). Ircam.

Itakura, F. (1975). Minimum prediction residual principle applied to speech recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, *23*(1), 67–72. https://doi.org/10.1109/TASSP.1975.1162641

James, J., Ingalls, T., Qian, G., Olsen, L., Whiteley, D., Wong, S., & Rikakis, T. (2006). Movement-based interactive dance performance. *Proceedings of the 14th ACM international conference on Multimedia*, 470–480. https://doi.org/10.1145/1180639.1180733

Jones, B. T., Kaiser, P., & Eshkar, S. (1999). Ars electrónica application: Ghostcatching a-74. https://archive.aec.at/prix/showmode/33648/

Karpashevich, P., Hornecker, E., Honauer, M., & Sanches, P. (2018). Reinterpreting schlemmer's triadic ballet: Interactive costume for unthinkable movements. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–13. https://doi.org/10.1145/3173574.3173635

Kaye, J. (2009). *The epistemology and evaluation of experience-focused hci* [Ph.D Thesis]. Cornell University. https://hdl.handle.net/1813/11657

Keevallik, L. (2014). Having a ball: Immaterial objects in dance instruction. In M. Nevile, P. Haddington, T. Heinemann & M. Rauniomaa (Eds.), *Interacting with objects* (pp. 249–268). John Benjamins Publishing Company. https://doi.org/10.1075/z.186.11kee

Keogh, E. J., & Pazzani, M. J. (2000). Scaling up dynamic time warping for datamining applications. *Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining*, 285–289. https://doi.org/10.1145/347090.347153

Keogh, E. J., & Pazzani, M. J. (2001). Derivative dynamic time warping, 1–11. https://doi.org/10.1137/1.9781611972719.1

Kitagawa, M., & Windsor, B. (2008). *Mocap for artists: Workflow and techniques for motion capture*. Elsevier/Focal Press.

Klein, J. (2017). What is artistic research? *Journal for Artistic Research*. https://doi.org/10.22501/jarnet.0004

Kvale, S., & Brinkmann, S. (2009). *Interviews: Learning the craft of qualitative research interviewing* (2. ed). Sage.

Ladd, G. W. (1979). Artistic research tools for scientific minds. *61*, 1–11. https://doi.org/10.2307/1239494

Lahusen, S. (1986). Oskar schlemmer: Mechanical ballets? *Dance Research*, *4*(2), 65–77. https://doi.org/10.2307/1290727

Landry, S., & Jeon, M. (2017). Participatory design research methodologies: A case study in dancer sonification, 182–187. https://doi.org/10.21785/icad2017.069

Louppe, L. (2010). *Poetics of contemporary dance*. Dance Books.

Louppe, L. (2015). *Poetik des zeitgenossischen tanzes*. Transcript-Verlag.

Lucas Bravo, P. P. (2022). *A human-machine music performance system based on autonomous agents* [Master's Thesis]. University of Oslo. https://www.duo.uio.no/handle/10852/96115

Magnusson, T., Kiefer, C., & Ulfarsson, H. (2022). Reflexions upon feedback. *NIME 2022*. https://doi.org/10.21428/92fbeb44.aa7de712

Marshall, M. (2009). *Physical interface design for digital musical instruments* [Ph.D Thesis]. McGill University [tex.entrytype: other]. https://escholarship.mcgill.ca/concern/theses/h128ng18h

Masu, R., Correia, N. N., Jurgens, S., Druzetic, I., & Primett, W. (2019). How do dancers want to use interactive technology?: Appropriation and layers of meaning beyond traditional movement mapping. *Proceedings of the 9th International Conference on Digital and Interactive Arts*, 1–9. https://doi.org/10.1145/3359852.3359869

McCormick, J., Vincs, K., Nahavandi, S., Creighton, D., & Hutchison, S. (2014). Teaching a digital performing agent: Artificial neural network and hidden markov model for recognising and performing dance movement. *Proceedings of the 2014 International Workshop on*

Movement and Computing, 70–75. https://doi.org/10.1145/2617995.2618008

McFee, B., Raffel, C., Liang, D., Ellis, D. P., McVicar, M., Battenberg, E., & Nieto, O. (2015). Librosa: Audio and music signal analysis in python [Citation Key: mcfee2015librosa]. *Proceedings of the 14th python in science conference*, *8*, 18–25.

Meador, W. S., Rogers, T. J., O'Neal, K., Kurt, E., & Cunningham, C. (2004). Mixing dance realities: Collaborative development of live-motion capture in a performing arts environment. *Computers in Entertainment*, *2*(2), 12–12. https://doi.org/10.1145/1008213.1008233

Mine, M. R. (1993). *Characterization of end-to-end delays in head-mounted display systems*.

Miranda, E. R., & Wanderley, M. M. (2006). *New digital musical instruments: Control and interaction beyond the keyboard*. A-R Editions.

Moore, F. R. (1988). The dysfunctions of midi. *Computer Music Journal*, *12*(1), 19. https://doi.org/10.2307/3679834

Mulder, A. G. (2000). Towards a choice of gestural constraints for instrumental performers (M. Wanderley & M. Battier, Eds.). *Trends in Gestural Control of Music*, 315–335.

Müller, M. (2007a). Dynamic time warping. In M. Müller (Ed.), *Information retrieval for music and motion* (pp. 69–84). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-74048-3_4

Müller, M. (2007b). Fundamentals on motion capture data. In *Information retrieval for music and motion* (pp. 187–209). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-74048-3_9

Müller, M., Mattes, H., & Kurth, F. (2006). An efficient multiscale approach to audio synchronization. *546*, 192–197.

Naccarato, T. J., & MacCallum, J. (2017). Critical appropriations of biosensors in artistic practice. *Proceedings of the 4th International Conference on Movement Computing*, 1–7. https://doi.org/10.1145/3077981.3078053

Ng, J., & Bax, N. (2023). Spooker trouper: Abba voyage, virtual humans and the rise of the digital apparition. *Paragraph*, *46*(2), 160–175. https://doi.org/10.3366/para.2023.0427

Nielsen, J. (1993). Iterative user-interface design. *Computer*, *26*(11), 32–41. https://doi.org/10.1109/2.241424

Norman, S. J. (2015). Oskar schlemmer's programmatic gesture research. In N. S. Sutil & S. Popat (Eds.), *Digital movement: Essays in motion technology and performance* (pp. 21–34). Palgrave Macmillan.

Nymoen, K. (2013). *Methods and technologies for analysing links between musical sound and body motion* [Ph.D Thesis]. Department of Informatics, Faculty of Mathematics and Natural Sciences, University of Oslo. https://www.duo.uio.no/handle/10852/34354

Nymoen, K., Skogstad, S. A., & Jensenius, A. R. (2011). Soundsaber – a motion capture instrument. *Proceedings of the international conference on new interfaces for musical expression*, 312–315. https://doi.org/10.5281/zenodo.1178125

OptiTrack. (2022, November). Skeleton tracking [Accessed on 10.07.2023]. https://docs.optitrack.com/motive/skeleton-tracking

Pavlovych, A., & Stuerzlinger, W. (2009). The tradeoff between spatial jitter and latency in pointing tasks. *Proceedings of the 1st ACM SIGCHI symposium on Engineering interactive computing systems*, 187–196. https://doi.org/10.1145/1570433.1570469

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, *12*(85), 2825–2830.

Petitmengin, C. (2006). Describing one's subjective experience in the second person: An interview method for the science of consciousness. *Phenomenology and the Cognitive Sciences*, *5*(3–4), 229–269. https://doi.org/10.1007/s11097-006-9022-2

Pisner, D. A., & Schnyer, D. M. (2020). Support vector machine. In *Machine learning* (pp. 101–121). Elsevier. https://doi.org/10.1016/B978-0-12-815739-8.00006-7

Plaete, J., Bradley, D., Warner, P., & Zwartouw, A. (2022). Abba voyage: High volume facial likeness and performance pipeline. *ACM SIGGRAPH 2022 talks*. https://doi.org/10.1145/3532836.3536260

Prätzlich, T., Driedger, J., & Müller, M. (2016). Memory-restricted multiscale dynamic time warping [journalAbbreviation: 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)]. *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 569–573. https://doi.org/10.1109/ICASSP.2016.7471739

Purser, A. C. E. (2017). 'being in your body' and 'being in the moment': The dancing body-subject and inhabited transcendence. *Journal of the Philosophy of Sport*, *45*(1), 37–52. https://doi.org/10.1080/00948705.2017.1408018

Ratanamahatana, C. A., & Keogh, E. (2004a). Everything you know about dynamic time warping is wrong. *32*.

Ratanamahatana, C. A., & Keogh, E. (2004b). Making time-series classification more accurate using learned constraints. *Proceedings of the 2004 SIAM International Conference on Data Mining (SDM)*, 11–22. https://doi.org/10.1137/1.9781611972740.2

Reynolds, J. M. (2019, November). The normate. In G. Weiss, A. V. Murphy & G. Salamon (Eds.), *50 concepts for a critical phenomenology* (pp. 243–248). Northwestern University Press. https://doi.org/10.2307/j.ctvmx3j22

Risoud, M., Hanson, J.-N., Gauvrit, F., Renard, C., Lemesre, P.-E., Bonne, N.-X., & Vincent, C. (2018). Sound source localization. *European Annals of Otorhinolaryngology, Head and Neck Diseases*, *135*(4), 259–264. https://doi.org/10.1016/j.anorl.2018.04.009

Sakoe, H., & Chiba, S. (1978). Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics,* *Speech, and Signal Processing*, *26*(1), 43–49. https://doi.org/10.1109/TASSP.1978.1163055

Salvador, S., & Chan, P. (2007). Toward accurate dynamic time warping in linear time and space. *Intelligent Data Analysis*, *11*(5), 561–580. https://doi.org/10.3233/IDA-2007-11508

Sanchez-Colberg, A. (1996). Altered states and subliminal spaces: Charting the road towards a physical theatre. *Performance Research*, *1*(2), 40–56. https://doi.org/10.1080/13528165.1996.10871489

Schacher, J. C., & Kocher, P. (2006). Ambisonics spatialization tools for max/msp. *Omni*, *500*(1).

Schlemmer, O. (1987). Man and art figure [Original work published 1961]. In W. Gropius & A. S. Wensinger (Eds.), *The theater of the bauhaus* (pp. 19–46). Wesleyan university press.

Schmid, S., Studer, D., Hasler, C.-C., Romkes, J., Taylor, W. R., Lorenzetti, S., & Brunner, R. (2016). Quantifying spinal gait kinematics using an enhanced optical motion capture approach in adolescent idiopathic scoliosis. *Gait Posture*, *44*, 231–237. https://doi.org/10.1016/j.gaitpost.2015.12.036

Scully, J. L. (2008). *Disability bioethics: Moral bodies, moral difference*. Rowman Littlefield.

Silva, D. F., & Batista, G. E. A. P. A. (2016). Speeding up all-pairwise dynamic time warping matrix calculation, 837–845. https://doi.org/10.1137/1.9781611974348.94

Simonsen, J., & Robertson, T. (Eds.). (2013). *Routledge international handbook of participatory design*. Routledge.

Steed, A. (2008). A simple method for estimating the latency of interactive, real-time graphics simulations. *Proceedings of the 2008 ACM symposium on virtual reality software and technology*, 123–129. https://doi.org/10.1145/1450579.1450606

Sterne, J. (2001). Mediate auscultation, the stethoscope, and the "autopsy of the living": Medicine's acoustic culture. *Journal of Medical Humanities*, *22*(2), 115–136. https://doi.org/10.1023/A:1009067628620

Sterne, J. (2021). *Diminished faculties: A political phenomenology of impairment*. Duke University Press.

Stévance, S., & Lacasse, S. (2018). *Research-creation in music: Towards a collaborative interdiscipline*. Routledge.

Straus, J. N. (2011). *Extraordinary measures: Disability in music*. Oxford University Press.

Strutt, D. (2022). A simple tool for remote real-time dance interaction in virtual spaces, or "dancing in the metaverse". *Critical Stages/Scènes critiques*, *25*. https://www.critical-stages.org/25/a-simple-tool-for-remote-real-time-dance-interaction-in-virtual-spaces-or-dancing-in-the-metaverse/

Sutil, N. S. (2014). Mathematics in motion: A comparative analysis of the stage works of schlemmer and kandinsky at the bauhaus. *Dance Research*, *32*(1), 23–42. https://doi.org/10.3366/drs.2014.0085

Ting, Y.-W., Lin, P.-H., & Lin, R. (2021). A study of applying bauhaus design idea into the reproduction of the triadic ballet. In P.-L. P. Rau (Ed.), *Cross-cultural design. applications in arts, learning, well-being, and social development* (pp. 65–83, Vol. 12772). Springer International Publishing. https://doi.org/10.1007/978-3-030-77077-8_6

Ting, Y.-W., Lin, P.-H., Lin, R., & Shi, M.-H. (2022). A study of 3d stereoscopic image production of "triadic ballet" of the theater of the bauhaus. In P.-L. P. Rau (Ed.), *Cross-cultural design. applications in learning, arts, cultural heritage, creative industries, and virtual reality* (pp. 283–293, Vol. 13312). Springer International Publishing. https://doi.org/10.1007/978-3-031-06047-2_20

Tragtenberg, J. N., Calegario, F., Cabral, G., & Ramalho, G. L. (2019). Towards the Concept of Digital Dance and Music Instruments. *Proceedings of the International Conference on New Interfaces for Musical Expression*, 89–94. https://doi.org/10.5281/zenodo.3672878

Ulloa, J. S., Haupert, S., Latorre, J. F., Aubin, T., & Sueur, J. (2021). Scikit-maad: An open-source and modular toolbox for quantitative soundscape analysis in python. *Methods in Ecology and Evolution*, *12*(12), 2334–2340. https://doi.org/10.1111/2041-210X.13711

Vincs, K., & McCormick, J. (2010). Touching space: Using motion capture and stereo projection to create a "virtual haptics" of dance. *Leonardo*, *43*(4), 359–366. https://doi.org/10.1162/LEON_a_00009

Zemp, R., List, R., Gülay, T., Elsig, J. P., Naxera, J., Taylor, W. R., & Lorenzetti, S. (2014). Soft tissue artefacts of the human back: Comparison of the sagittal curvature of the spine measured using skin markers and an open upright mri. *PLOS ONE*, *9*(4), e95426. https://doi.org/10.1371/journal.pone.0095426

Appendix A

Modelling MoCap

At a fundamental level, a MoCap system can be modelled as a linear pipeline, in which input provided from the system's sensors is processed and stored. Such a model can be found, for example, in Nymoen's (2013) model of the motion capture process as a method in musicological research, shown in figure A.1. Under this model, motion tracking takes place when sensor data are acquired and then undergo a processing step to transform them into motion data. If these data are stored, the process becomes motion capture. Crucially, however, there are two steps that are not found in this model that are important to the use of MoCap in artistic contexts. Firstly, there is no mention of the motion that should be acquired by the sensing component of the system, and secondly, the processing required to reproduce the stored or tracked data in an output modality is omitted.[1] In the following sections I examine models of MoCap from the fields of computer graphics and animation and interactive music and dance, all of which encompass these wider steps.

A.1 Computer Graphics and Animation

The computer graphics and animation literature takes a prevailing perspective similar in nature to Nymoen's, conceptualising the process as linear and goal-oriented. This process is usually framed as comprising two broad stages; first the extraction of motion data from a recording of a performer made via a sensing apparatus in order to construct a model of the performance and second the application of this model to an animated figure (Gleicher, 1999).[2] This predominantly takes the form of performing processing on the raw sensor data in order to obtain a model of the performer's skeleton and then using this skeleton to rig the animation as a reproduction of the capture.

[1]Reproduction used here in the broadest sense of the term, in that the data tracked or stored are then used in some manner.

[2]Gleicher also notes that historically this has involved a degree of tension between the motion capture technicians performing the first part of this process and the animators carrying out the second, with both sides unclear on the extent of the other's remit as well as the results that they can expect from each other.

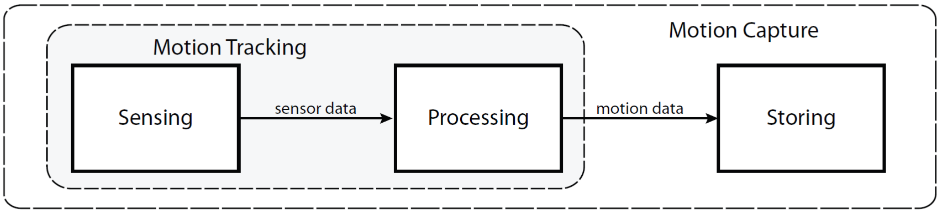


Figure A.1: A model of the MoCap process, showing a linear process of acquiring sensor data, processing the data to obtain motion data, and then optionally storing it. From Nymoen (2013, p. 13).

This perspective slides into viewing the achievement of these goals as problems to be solved. Gleicher and Ferrier (2002) succinctly summarise this perspective, stating that the "goal of motion capture is to record the movement of a performer (typically, but not always, human) in a compact and usable manner" (p. 1), framing this in terms of extracting the "essence" of the performer through properties such as "mood, expression, and personality" from the extraneous "myriad of details" added to this essence by the system within which it is embedded (p. 1). Through over-enthusiastic processing techniques, animators can obscure this essence, and thus the process of MoCap becomes a problem to solve: which procedural techniques can be developed that best extract and preserve this essence?

This view of motion capture likewise implies a strong sense of linearity to the process. The optimal motion capture system extracts the essence of the performer's motion, abstracting the motion from its appearance, and directly applies it to the animation. This is a view quite explicitly reflected in Gleicher's (1999) summary of the motion capture process as the employment of "[o]ptical, mechanical or magnetic sensors [to] record the movements that can then be transferred to animated characters" (p. 1), encapsulated in a five step process of "1. Plan the motion capture shoot and subsequent production. [...] 2. Capture the motion. 3. Clean the data. 4. Edit the motions. 5. Map the motions to the animated characters" (p. 2).

The reduction of the process to a series of procedural steps can be found across the literature related to motion capture for animation, ranging from encompassing the entire process to more granular views of sub-processes. For example, Bodenheimer et al. (1997), detailing their process to create a skeleton that can be used to rig an animation, describe this series of steps as placing the sensors, measuring and building a skeleton, optimising the skeleton and applying inverse kinematics. This procedural approach to the process of motion capture is encapsulated in the importance placed upon developing a workflow conceptualised as a pipeline, implying a linear flow of data from one end of a pipe to another, with textbooks on motion capture for animation such as Kitagawa and Windsor (2008) dedicating whole chapters to good practice in constructing pipelines such as the one shown in figure A.2.

Although these perspectives encourage a procedural, linear viewing of the MoCap process, sites of reflexivity and feedback are noted. In

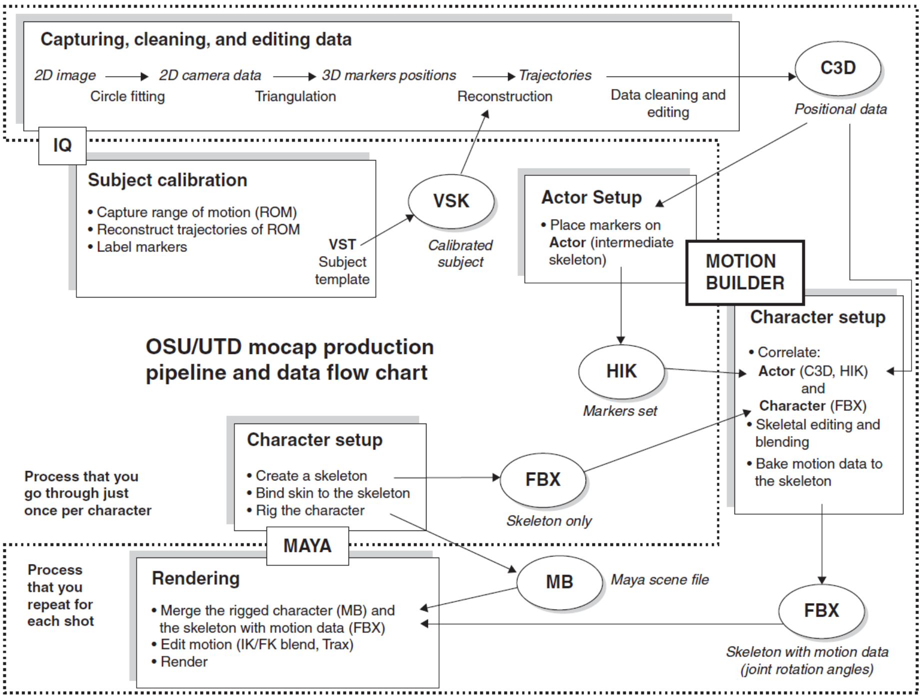


Figure A.2: A model of the motion capture pipeline developed for the University of Texas at Dallas and Ohio State University, showing the procedural steps required for the linear flow of data from sensors to animation with an optical, marker-based system. From Kitagawa and Windsor (2008, p. 199).

reference to his five step process, for example, Gleicher (1999) notes that steps four and five (edit the motions and map the motions to the animated characters) are sometimes performed iteratively. However, key here is that such sites of reflexivity are not seen as integral to the process. As a result, the performer is shut out from the process of motion capture, existing as a provider of abstracted essence of motion that is captured through the sensor array. This is visible in Gleicher's (1999) description of a further point of reflexivity in the process, namely that the motion capture process necessitates a step of motion tracking and that the tracked data is often visualised while the capturing of performance takes place. Instead of taking the perspective that under such observation in which the performer becomes acutely aware of their own motion, representations of this motion feed back into the way in which the performer moves, Gleicher remains focused on the goal of mapping essence to animation while minimising the influence of such observation on performance and emphasising utility to the process, stating that "[e]ven if the final result will require adjustment and production, instant feedback to the performer is useful" (p. 1).

In light of the above, the perspective taken in motion capture for animation indicates a high degree of mediated auscultation. From this view, the performer instrumentalises their body to provide the desired essence to the observational sensor array. At this point, the physicality of the body providing the motion is discarded from the process, with the essence embedded in the data instead becoming representative of the body itself. The process progresses until the body actualises in the finished animation. This positions the process as inherently collaborative, with various contributors extracting the essence and shaping the body throughout the process.

Viewing this perspective in relation to the co-construction model, it is the designers and operators of the systems who play a major role in the construction of the body, with great preference given to operations occurring in the sensing, construction, and representation layers. Consideration of the performance layer is minimised.

A.2 Interactive Music Systems

The Digital Musical Instrument (DMI) is a key concept in the field of interactive music. DMIs are musical instruments which consist of a "control surface or gestural controller, which drives the musical parameters of a sound synthesizer in real time" (Miranda & Wanderley, 2006, p. 1). Here, the term gestural controller is used in reference to musical gestures, that is "any human action used to generate sounds" (Miranda & Wanderley, 2006, p. 5). Although this can cover a broad range of possible inputs to the system, this definition implies a requirement to track the performer's motion. Likewise, although the term gestural controller can cover a wide array of input devices, most of which do not fall under the umbrella of technologies which are delineated specifically as MoCap technologies as defined in Nymoen's (2013) taxonomy, several DMIs, such as *SoundSaber* (Nymoen et al., 2011), have leveraged these technologies either in the construction of the controller or as the controller itself. In cases such as *SoundSaber*, the MoCap system is used to track the motion of an object external to the performer's body. However, these technologies are particularly well-suited to be implemented within a framework of what Mulder (2000) terms immersive controllers, which "impose few or no restrictions to movement" (p. 319), which can be further categorised based upon the visualisation strategy employed by the performer to conceptualise their interaction with the system; internal controllers for which the control interface should be conceptualised as the body itself, external controllers for which the control interface should be conceptualised as external to the body (as in air instruments), and symbolic controllers in which the interface is too complex to visually conceptualise and which instead relies upon the extraction of formalised gestures from the performer's motion.

Miranda and Wanderley's (2006) model of the DMI, shown in figure A.3, provides the most condensed overview of the motion tracking process as viewed from a DMI perspective. This model views the process of tracking human motion as a process of sensing a performer's gestural input, and then mapping the acquired data to parameters of a sound

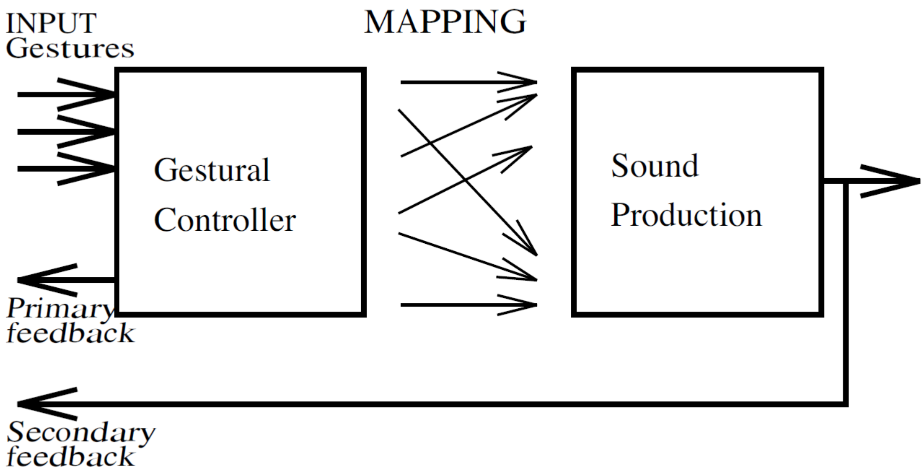


Figure A.3: Miranda and Wanderley's (2006) model of the digital musical instrument. From Miranda and Wanderley (2006, p. 3).

synthesis engine. This is not modelled as an entirely linear process, as feedback provided to the performer by the system to influence their continuing production of input gestures, both in the form of primary feedback provided by the gestural controller such as haptics, as well as the secondary feedback of the sound produced by the DMI, are included in the model. Moreover, due to the nature of these systems as feedback systems, the performer is by necessity pulled into the system schematic and takes equal standing to the technological components and work done by the system designer. However, as noted by Mulder, immersive controllers are not best suited for providing primary feedback, meaning that in these cases it is the secondary feedback of sound that predominantly draws the performer into the feedback loop of the DMI. This model has gained much traction within the field of interactive music systems, and many related models have been constructed that build upon, expand, and produce this model at various levels of granularity by authors such as Birnbaum (2007), Fasciani (2014) and Marshall (2009).

These models of the DMI offer several divergences in their conceptualisation of the participants in the MoCap process in contrast to the process as viewed from an animation perspective. Notable is the conceptualisation of the role of the performer in the process. Instead of serving as a source of "essence" acquired through the MoCap technology, where they are subsequently excluded from the process, the DMI model positions the performer as assuming a more significant role, establishing a direct feedback loop with the technological components involved in the process. Rather, it is other participants in the process who step back once the process of MoCap begins, with, for example, the designer of the DMI serving as determining a framework within which the MoCap process can take place before removing themselves from the process.

The relationship between the input and output of the system is much more complex than is seen in the animation perspective. From that perspective, the relationship between the motion data captured from the performer and the animated figure should map directly, with data capture from one part of the body being used to rig one part of the animation. However, from an interactive music perspective, this relationship should be more complex, with a tendency to aim for more complex mappings that map input to many parameter of the sound output (Hunt & Kirk, 2000).

The idea of a complex mapping is tied to another key concept in the design of DMIs. As noted by Fels (2000), in reference to Moore (1988), a key desideratum for the designer of a DMI is the construction of an instrument that enables a high degree of control intimacy. A method through which this control intimacy can be achieved is through the use of a transparent mapping, namely the employment of a mapping strategy in which the conceptual link between the input gesture and the output sound is directly understandable, critically both by the performer *and* any observers (Fels et al., 2002). If this is achieved, it is possible that the performer can obtain a sense of embodiment described by Fels (2000) as "the person embodies the object. That is, they have integrated the object and its behaviour into their own sense of self. The object becomes part of them. The object is an extension of their own bodies and mind" (p. 14). This phrasing points towards a high degree of mediated auscultation when viewing the DMI as a sensing system, a literal situation where the sensing apparatus is erased from the consciousness of the performer. A key difference to the conceptualisation of the MoCap process from an animation perspective, however, is that this is not only aimed for as a fundament of DMI design, but its achievement is also the mark of a successful instrument.

A.3 Interactive Dance

In comparison to the more well-defined field of interactive music, interactive dance is much more dispersed. As noted by Tragtenberg et al. (2019), the terminology relating to the field is diverse, even with regard to its name. However, Tragtenberg et al. provide an umbrella definition as "digital systems used in interactive dance performances developed to enhance the expressive possibilities of dancers with sensors to capture their movements and produce sounds, visuals or movement through robotic actuators" (p. 89). In a similar manner to the conceptualisation of the system in interactive music, this definition is not explicitly concerned with MoCap, however MoCap technologies are encompassed within the definition, for example in the interactive works outlined in section 2.1.3.

As Tragtenberg et al. note, this diversity also applies to conceptual frameworks, with no singular system's model or conceptual framework having gained prevalence in quite the way that the model of the DMI has in interactive music. Such frameworks include Bisig's (2022) inclusion of performance within a taxonomy of generative dance, relating the category to other categories as well as detailing approaches that can be employed within this category, and Fdili Alaoui et al.'s (2014) inclusion of interactivity

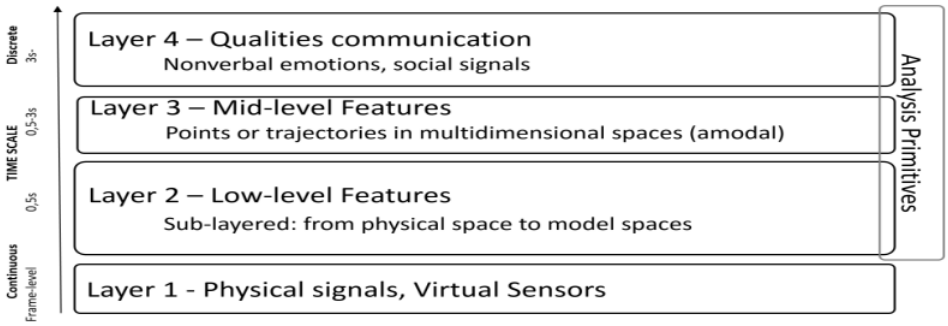


Figure A.4: Camurri et al.'s (2016) model of their conceptual framework relating to the expressive qualities of movement. From Camurri et al. (2016, p. 2).

as a class of feature in a wider framework to categorise choreographic tools with regard to Laban Movement Analysis components.

Camurri et al. (2016) offer a comprehensive model of a conceptual framework for interactive system designers, shown in figure A.4. This model is based around the extraction of expressive qualities from motion data, operating across hierarchical layers representing multiple timescales. Camurri et al. note that this model should be understood from an observer's perspective (that descriptors of qualities refer to how they are perceived by the observer and not by the performer), and moreover state that an intended aim for this model is to aid in cross-modal mappings (for example, motion data can be acquired on layer one, processed so that the designer is working on layer three, and then mapped from layer three back down to an audio signal in layer one).

A key aspect of this model is that there is an explicit recognition of the fact that the features and qualities constructed on the higher levels of the model originate in the signals acquired from the sensors on layer one. This contrasts with the models of the DMI presented above, which place focus on the gesture as system input. Even in the variants of the model that recognise the sensing components, this enables a slide into removing the sensing device from the conceptualisation of the interaction, encouraged by the drive towards expressivity. However, under this model, mapping

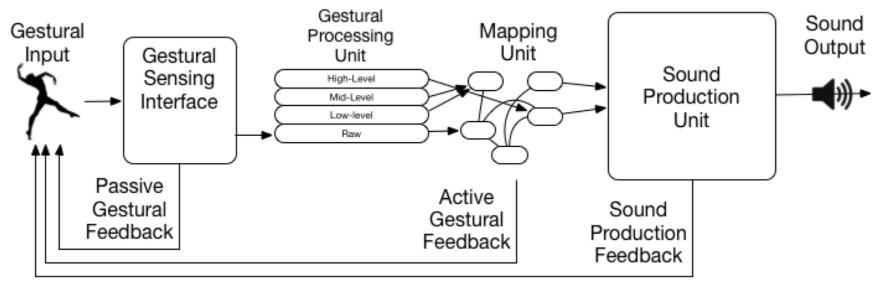


Figure A.5: Tragtenberg et al.'s (2019) model of the digital dance and musical instrument. From Tragtenberg et al. (2019, p. 91).

is conceptualised as moving between temporal layers. This obscures the way in which the mapping process is itself a choice that can affect how the expressive qualities are perceived.

Tragtenberg et al. (2019) combine Camurri et al.'s model with the model of the DMI to develop the dance and digital musical instrument, shown in figure A.5. Although focusing on the mapping of gestural input specifically to audible modalities, this model acknowledges the multiple layers of data that are extracted from the sensing apparatus and that these may be individually mapped to parameters of the output modality. When thinking in these terms, it becomes difficult to remove the sensing device from the conceptual chain.

Appendix B

The Marker as Constraining Object

In a passage concerning the dance solo, Laurence Louppe (2010) contends that the solo is the closest that dance comes to performance art. As a result, framing the dance solo as comparable to the art object, she draws a link between the idea of constraints in both choreographic and performance practice, stressing the importance of defining constraints to avoid falling back into known schemata. She makes the point that "[i]n 'performance' the notion of an obstacle or difficulty is used for its power to directly affect an action, or to change the body through taking a risk, or to make and impact on the performer's behaviour" (p. 207). To provide examples to support her point, Louppe chooses to make reference to several works in which the constraint is provided through an object; Dominique Dupuy with a human bone, Sophie Taeuber-Arp and the mask designed by Jean Arp[1] (which is described by Louppe as "more like an art object attached to the body than an innocent and undifferentiated costume accessory" (p. 208) and shown in figure B.1), and two works by Reinhild Hoffmann. It is upon the mentioning Hoffman's works, *Solo mit Sofa*, shown in figure B.2, which has Hoffman wrapped in fabric which is attached to a sofa so that her movement is framed through her relation to the object, and a further solo work,[2] in which she is encumbered with heavy rocks, that Louppe saliently points towards the materiality of the objects themselves:

[1]There is some uncertainty on who exactly designed and made the mask. Doutreligne (2020), for example, attributes its construction to either Jean Arp or Marcel Janco.

[2]The English translation of Louppe's work appears to have several errors. Among others such as mis-gendering Hoffmann and spelling her name incorrectly, the second work that I refer to as a further solo work is named by the translator as *Soloabend*. This seems to be a misunderstanding by the translator. A *Soloabend* in German refers to a performance of individual solo works. For example, a performance of *Solo mit Sofa* is referred to as a *Soloabend* by the Steirischer Herbst festival archive in reference to the type of performance that it was. The other work that Louppe refers to appears to be *Steine (Der Weg)*. As I have no knowledge of French, I am unable to check Louppe's original writing. However, the German translation (Louppe, 2015) does not have the errors present in the English translation, and refers to what the English translator refers to as *Soloabend* as "einem anderen Solo aus derselben Zeit" [another solo from the same period] (p. 252).

> In both cases the performance has a second, sculptural dimension due to the qualities of the material from which the object is made - which makes these solos (visual) 'art' works on several levels (in their visual arrangement and their process).
>
> (Louppe, 2010, p. 208)

This use of objects to influence, guide, and determine their movements is present in additional dance works. Works such as Oskar Schlemmer's *Pole Dance*, shown in figure B.3 in which the dancer is affixed with long white poles, which not only extend the body but also create imbalance and restrict its motion (Lahusen, 1986), and Merce Cunningham's *Antic Meet*, shown in figure B.4, which involves a section in which Cunningham dances with a chair attached to his back, and in which even "performers seemed to acquire the emotional reticence and palpable physicality of objects" (Copeland, 2004, p. 29), place focus on the relationship between the dancer and the object. The use of objects to guide and constrain movement is also present in dance teaching, and here the objects do not have to have a physical presence, with, for example, immaterial objects metaphorically employed by dance teachers in order to aid in the comprehension of technique by their students (Keevallik, 2014).

Birringer (2015), building on Louppe, draws a link between the sensor apparatus worn by dancers in works involving an interactive system and the concept of the *dispotif*, "a conceptual category for examining environments (material, technological, medial) of regulating, strategic frameworks that are configured in certain ways making it possible for certain types of phenomena to occur" (pp. 170-71). Connecting this to the Schlemmer's costuming and the "technical apparatuses" of post-revolutionary Russian constructivists, he notes that

> tools and appendages are amplificatory choreographic and scenographic elements that may become fundamental for the plasticity of articulation of movement and movement relations to space.
>
> (Birringer, 2015, pp. 164-65)

Additionally, he notes that "tangible matter always mattered in performance, not just in relation to instruments and hardware but also through material processes and techniques of enhancement" (p. 172). Moreover, he argues against ideas that interactive systems that work with visuals enhance dancers' bodies, noting that this enhancement is often not visible to dancers during performance and therefore cannot be experienced as enhancing. In this sense, it is the materiality of the sensor itself which primarily experienced by the dancer.

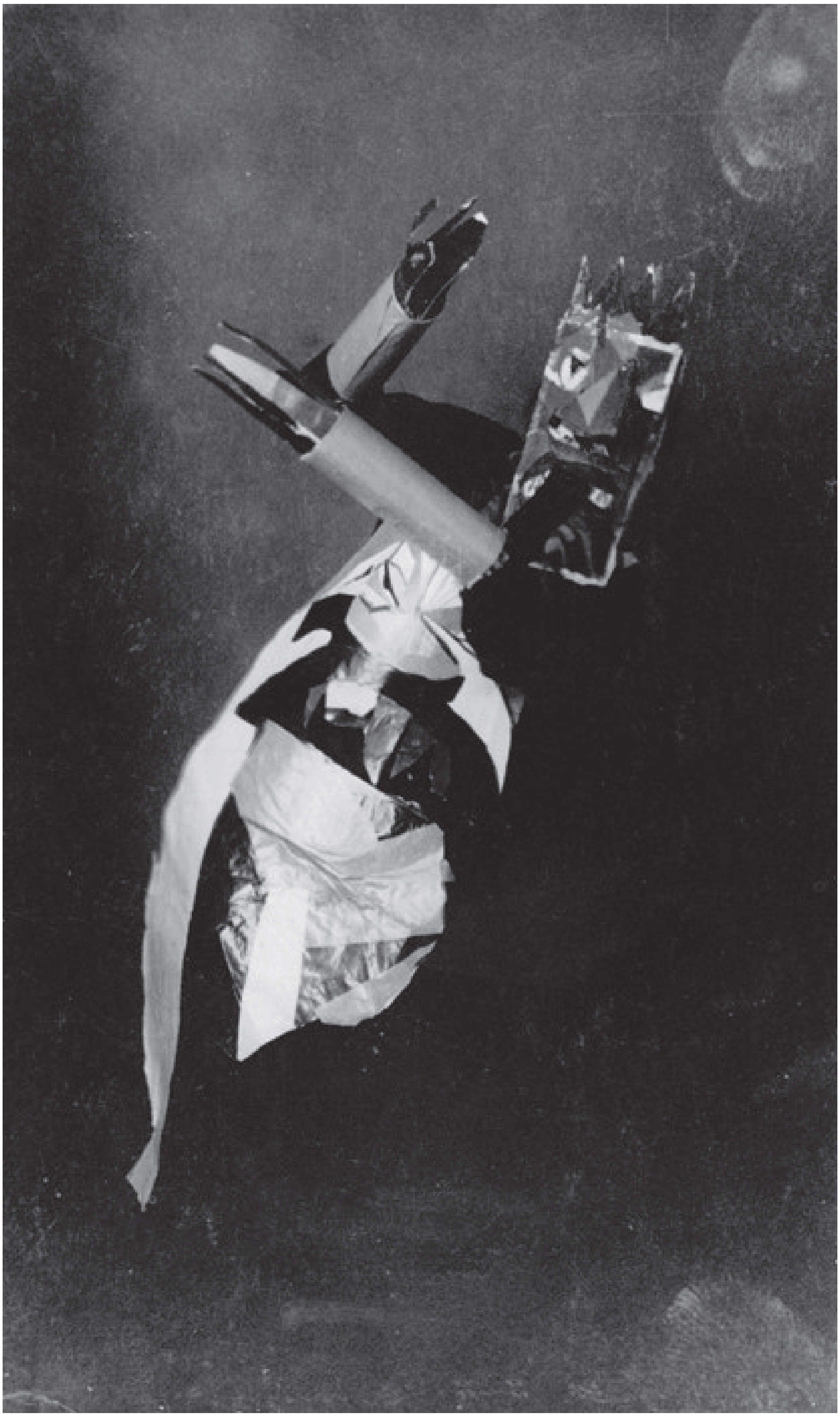


Figure B.1: Sophie Taebuer wearing the mask designed by either Jean Arp or Marcel Janco. From Doutreligne (2020).

Figure B.2: Reinhild Hoffmann's *Solo mit Sofa*. From reinhild-hoffmann.de.

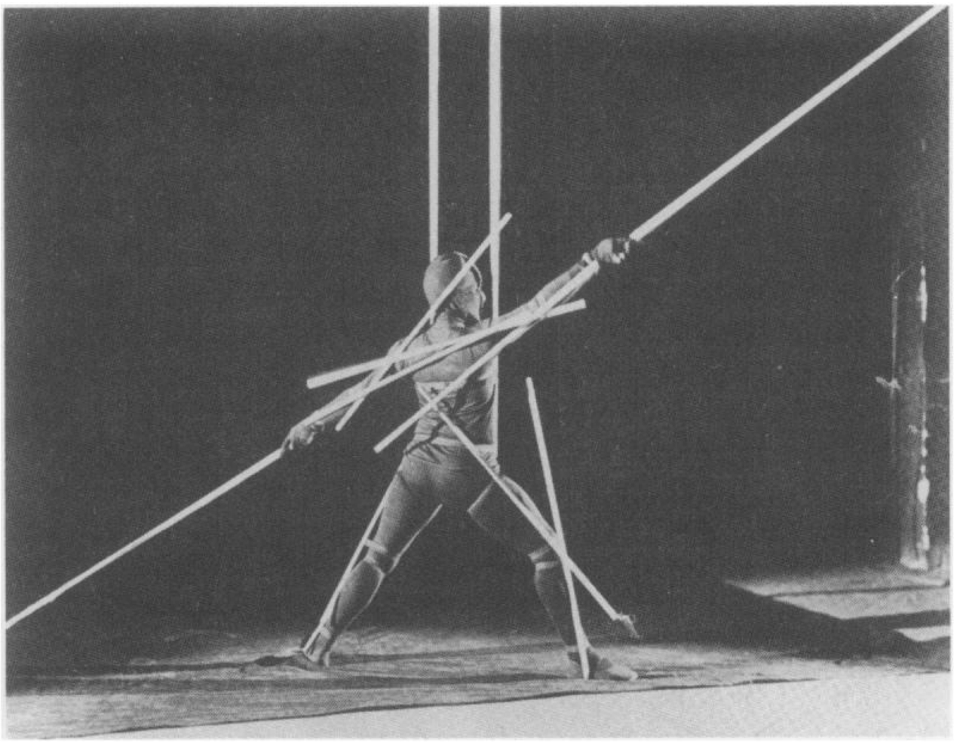


Figure B.3: The costuming from Oskar Schlemmer's *Pole Dance*. From Lahusen (1986).

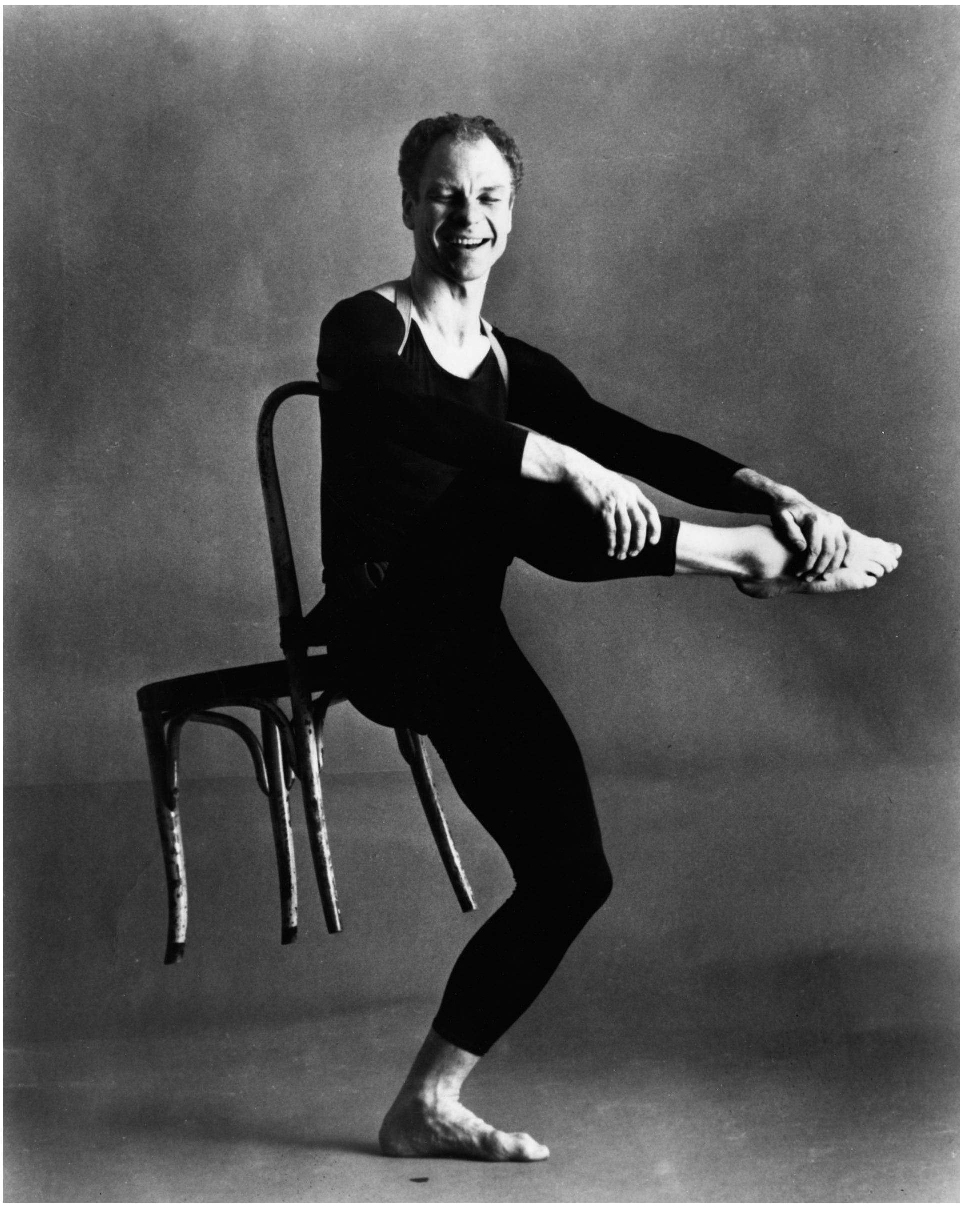


Figure B.4: The costuming from Merce Cunningham's *Antic Meet*. From rauschenbergfoundation.org.

Appendix C

Oskar Schlemmer, Tänzermensch, and Technology

Oskar Schlemmer, a major figure in the Bauhaus movement of the early twentieth century, worked across a variety of media such as dance, sculpture, and visual arts. His work had as one of its main concerns an exploration of space and its interaction with the human body in motion. His dance work was formalist, abstracting the human form and motion patterns and relating this to the cubic geometry of the space in which the dance took place. Central to Schlemmer's thoughts was the *Tänzermensch* (*Man as Dancer*) who embodies these abstractions and relationships:

> *He obeys the law of the body as well as the law of space; he follows his sense of himself as well as his sense of embracing space.* As the one who gives birth to an almost endless range of expression, whether in free abstract movement or in symbolic pantomime, whether he is on the bare stage or in a scenic environment constructed for him, whether he speaks or sings, whether he is naked or costumed, the Tänzermensch is the medium of transition into the great world of the theater (*das grosse theatralische Geschehen*).
>
> (Schlemmer, 1987, p. 25, emphasis in original)

Schlemmer posits that costuming is integral to the transformation of the human body into the *Tänzermensch*. He differentiates between *native* costuming, as "produced by the conventions of religion, state and society" (Schlemmer, 1987, p. 25), and *stage* costuming. For Schlemmer, it is only stage costuming that effects a transition to *Tänzermensch*. He notes that the two forms of costuming are commonly confused and that the use of genuine stage costumes has been extremely rare practice, limited to the standardised costumes of the *commedia dell'arte*. Schlemmer outlines four methods fundamental to the use of costuming in the transformation of human body to *Tänzermensch*:

1. The laws of the surrounding cubical space

Figure C.1: The costuming from Oskar Schlemmer's *Triadic Ballet*. From Lahusen (1986).

2. The functional laws of the human body in their relationship to space
3. The laws of motion of the human body in space
4. The metaphysical forms of expression

which are shown along with their explications and results in figure 4.4.

These methods were applied by Schlemmer through a number of costume designs in a variety of manners across several works, for example, in the *Pole Dance*, in which a dancer is affixed with long white poles, which not only extend the body but also create imbalance and restrict its motion (Lahusen, 1986), as well as in the *Space Dance* in which three dancers are costumed in three separate colours and each must perform with distinctive floor patterns and movement characteristics. The *Triadic Ballet*, one of Schlemmer's major works, places the dancers in heavy, solid costumes which limit the dancer's movements in a variety of ways, shown in figure C.1.

As described by Sutil (2014), Schlemmer's work has a heavy mathematical element to it. Moreover, Norman (2015) describes Schlemmer's approach as programmatic, arguing that it can be seen as a forerunner to dance works involving motion computing. Additionally, she argues that it can serve as a source of inspiration for future research, although this remains relatively unexplored in relation to Schlemmer's contemporary Rudolf Laban. In view of this there is comparatively little in the way of technologically integrated dance work that draws on Schlemmer's work. Several works have examined the spatial aspects of Schlemmer's work. For example, for a reconstruction of the *Triadic Ballet* at the Theatre of the Bauhaus, dancers were superimposed onto a computer generated background (Ting et al., 2021) and projected during performance using 3D stereoscopy (Ting et al., 2022). More theoretically, Delbridge (2014) explores Schlemmer's theories on costuming to explore the relationship between a performer wearing a motion capture suit and their relationship to physical and virtual space.

Placing focus on the body of the performer, Karpashevich et al. (2018), created a wearable technology version of one of the costumes from the Triadic Ballet. They emphasise the restrictions that such costuming places on the dancer's motion, and centre the dancer's experience of these restrictions in their evaluation of the system.

Appendix D

Mapping MoCap Position to Ambisonics Object Position

The first step to mapping MoCap position data to sound source position within the Ambisonics plane is to define the coordinate system of each component and a method of aligning these. In contrast to the MoCap position data, which is defined as three Cartesian axes in relation to an origin, the position of sound sources within Ambisonics encoding is usually defined within a spherical coordinate system. Within this system, defining the position of a point in three dimensional space consists of defining an azimuth (ϕ), which measures the anti-clockwise angular position of a point on the horizontal-plane in relation to the origin, the elevation (θ), which is analogous to the azimuth except upon the vertical plane, and the radius (r), which is the Euclidean distance from the origin to the point. As *The Shapeshifter* works with Ambisonics in two dimensions, the relevant spherical coordinate components for the system are ϕ and r.

To directly simplify alignment, the origin of the MoCap coordinate system and the origin of the Ambisonics coordinate system can be set to the same location in physical space, so that

$$(0,0)_{x,y} = (0,0)_{\phi,r} \tag{D.1}$$

which in this case requires calibrating the MoCap system so that the central point of the speaker system is defined as the origin. Additionally, the axes are aligned so that ϕ_0 is aligned with the positive x axis.

A further requirement is to ensure that distance from the origin is measured on an identical scale within the Ambisonics and MoCap environments. For this, assuming regular and equidistant loudspeaker placement, each loudspeaker can be assumed to be located $r = 1$ unit from the origin in the Ambisonics environment. Therefore a scaling factor s to be applied to each Cartesian position can be defined as

$$s = \frac{r}{\sqrt{\sum_{i=1}^{I}(p_i - o_i)^2}} \tag{D.2}$$

where $i \in [1 : I]$ is the Cartesian axis, I is the total number of axes, p

is loudspeaker position in the MoCap environment for any loudspeaker included in the array and o is the origin. As $r = 1$, $o = (0, 0)$, and $I = 2$, this can be simplified to

$$s = \frac{1}{\sqrt{\sum_{i=1}^{2} p_i^2}} \tag{D.3}$$

The position data vectors must be applied to the Ambisonics encoding in real-time. Therefore, a playback component is configured to read through the MoCap vectors at a rate that synchronises the MoCap sample rate to the audio sample rate. In addition, both the audio playback and MoCap vector playback receive a synchronized trigger from the performance synch module in order to ensure that $t_{MoCap} = t_{audio}$.

As there are a potential 30 *position markers* included in a MoCap vector and only one audio track in a *cycle master*, a method is required through which to map these 30 vectors to a single Ambisonics object. In the current iteration this takes the form of a single Ambisonics object for each *cycle master*, which is passed through a bandpass filter to provide a vertical mapping. The center frequency of the filter is determined through

$$f_c = \alpha \times \frac{\sum_{i=1}^{I} z_i}{I} + b \tag{D.4}$$

where z is vertical position, I is the number of *position markers*, α is a scaling factor, and b is an offset. This also provides an opportunity to create a mapping for the vertical position MoCap vectors, as the three-dimensional Cartesian MoCap vectors are mapped onto the two-dimensional Ambisonics environment.

The MoCap angular position in relation to the origin for each position marker on the horizontal plane is mapped to the azimuth of the corresponding Ambisonics object. The angular rotation (θ) of the pair of Cartesian coordinates (x, y) can be obtained through

$$\theta = \text{atan2}(x, y) \tag{D.5}$$

However, as this provides $-\pi \leq \theta \leq \pi$ rad and the range of φ in the Ambisonics environment is $0 \leq \varphi \leq 2\pi$ rad, and in order to obtain the mean θ, this must be developed as

$$\theta = (\text{atan2}(\frac{\sum_{i=1}^{I} x_i}{I}, \frac{\sum_{i=1}^{I} y_i}{I}) + 2\pi) \mod \pi \tag{D.6}$$

The radial distance (r) of a MoCap *position marker* on the horizontal plane within the Ambisonics environment is calculated as

$$r = s\sqrt{\sum_{i=1}^{2} {p_i}^2} \tag{D.7}$$

where s is the scaling factor defined above, p is the position of the *position marker*, and i is the Cartesian axis.

Component	Distance Gain Coefficient
Omni, $d \leq 1$	$1 - \left(1 - \frac{\sqrt{2}}{2}\right) d^2$
Omni, $d > 1$	$\frac{\sqrt{2}}{2} \cdot 10^{\frac{(d-1)u}{20.0}}$
Higher order, $d \leq 1$	d
Higher order, $d > 1$	$10^{\frac{(d-1)u}{20.0}}$

Table D.1: Gain coefficients for simulation of distance for objects within the *Ambisonics* field. d represents distance, u represents dB per unit. From Schacher and Kocher (2006, p. 2).

The aim with the mapping of the radial distance is to simulate the distance of the sound source represented in the Ambisonics object corresponding to the distance of the *position marker* from the origin. It is important to note that this is intended to be a relative distance. This means that the aim is not to simulate a sound source at location of the *position marker* as it would be perceived by the listener at the origin, but rather to create an illusion that a sound source at a location closer to the listener sounds closer than another.

There are several acoustic properties that relate to the perception of the distance of a sound source. As summarised by Risoud et al. (2018), this includes the ratio of direct sound to reverberant sound, the initial delay time gap between the direct sound and the first reverberant sound, the sound intensity level, and the spectral qualities of the sound among others. Approaches towards achieving an illusion of distance for stereo reproductions of sound sources are commonly based around modelling these properties. For the reproduction of sound sources within an Ambisonics soundfield, approaches include applying gain compensation with varying coefficients for the zeroeth and higher orders within and without the unit circle (Schacher & Kocher, 2006).

For *The Shapeshifter* the radial distance is mapped to a gain compensation as found in Schacher and Kocher (2006), with the coefficients for various orders and the location of the sound source within or without the unit circle shown in table D.1.

Appendix E

Full Description of Latency Measurement Methodology

In the literature, the measurement of latency between an input gesture provided through MoCap and an output visualisation is primarily discussed in a context of Virtual Reality (VR), Augmented Reality (AR), or Extended Reality (XR). Specifically, I was unable to find any literature on motion to visual latency in a context of interactive performance involving MoCap. Therefore, the baselines and methods discussed here will mainly draw on resources from these fields.

E.1 Overview

Within a real-time graphics system which maps motion input to a visualised output there are several sources that can determine the latency of an input gesture to an output visualisation. In reference to head-mounted display systems, Mine (1993) provides an overview of these as:

1. Tracking system delay
2. Application host delay
3. Image generation delay
4. Display system delay

These are expanded by Steed (2008) as:

1. Sensor reading and computation
2. Sensor data communication
3. Application computation
4. Rendering computation
5. Display refresh

Within the visualisation component of *The Shapeshifter*, the motion to display chain can be split into several subsystems, each corresponding to one or several of Steed's latency sources. These subsystems are summarised as follows:

1. **The Motion Capture Subsystem**
   This subsystem comprises:
   (a) Camera reading and internal pre-processing
   (b) Communication to the MoCap software
   (c) MoCap software processing
   (d) Processing for transmission with the OSC protocol

   The data from this subsystem is communicated to the subsequent subsystem over the wireless network. This subsystem is viewed as a single component corresponding to the latency sources of *sensor reading and computation* and *sensor data communication*.

2. ***The Shapeshifter* System**
   This subsystem comprises:
   (a) Signal pre-processing
   (b) Mapping
   (c) Rendering

   This subsystem is viewed as corresponding to the latency sources of *application computation* and *rendering computation*.

3. **The Display Subsystem** This subsystem receives data from the preceding subsystem over a wired, HDMI connection. It comprises:
   (a) Processing through a HDMI matrix
   (b) Display

   This subsystem is viewed as corresponding to the latency source of *display refresh*.

For the real-time mapping of the performer's input gesture to the output visualisation, there is a requirement that the end-to-end latency of this chain is as low as possible.

Although measuring the end-to-end latency of such a system represents a complex task (Friston & Steed, 2014), several authors have contributed methods to achieve this. Many of these involve the use of a video camera to simultaneously capture an object being tracked and its visualisation on a screen. Lucas Bravo (2022), measuring the latency of a tracking object to a visualisation of the spatial placement of a sound source in a spatial audio environment, offers a method which involves placing the tracking object at the top on an incline in front of a screen and releasing it in view of the camera. Steed (2008), offers a methodology termed the *sine fitting method*.

This involves attaching a light source to the tracking object, which is then suspended to form a pendulum. A video camera is placed that can record both the tracking object and the visualisation on the screen. The recorded video is then processed to find the centroid of the pendulums swinging motion for both the light source emitted from the tracking object and the on-screen visualisation. For each frame, the horizontal displacements of each of these from their respective centroids are calculated. A sine function is then fitted to the data and the phase difference is calculated to provide the latency.

E.2 Data Collection

The method employed by the author derives from Steed's method. A pendulum is built from a rigid arm that has only one degree of rotational freedom, shown in figure 5.3a. Mounted upon the end of the arm is a laser pointer, shown in figure 5.3b that has been attached to a rigid body, shown in figure 5.3c. This is mounted upon a ceiling truss and pointed towards the video wall used in the performance. A measure is placed below the pendulum, shown in E.2. This offers both a way in which to align the rest position of the pendulum, as well as consistent release of the the pendulum from an identical position.

The screen is split horizontally. The laser pointer is calibrated so that the entire arc of its swing is contained within the top segment. On the bottom segment, the visualisation is displayed, shown in figure 5.4. A laser level is employed to ensure that the laser pointer does not have an angular offset. This is achieved through aligning the centre point of the measure with the laser pointer mounted upon the pendulum and the laser point on the screen, shown in figure E.3. A camera is placed in front of the screen and framed so that only the screen is visible. The laser level is again employed to ensure that the camera does not have an angular offset. The camera is placed 2 m from the screen, and the laser pointer 3 m from the screen, hanging 1 m below the pivot point of the pendulum. After the collection of each data sample, the pendulum is re-calibrated, shown in figure E.4.

This system offers several advantages in comparison to that offered by Steed (2008). Firstly, the use of the laser pointer means that the camera does not have to be placed behind the pendulum. This means that there is no opportunity for the visualisation to be occluded by the pendulum. Moreover, as the camera can be framed so that solely the screen is in view, the data that is collected for the analysis step requires less cleaning, especially if the screen background is set to a colour that contrasts heavily with both the laser point and the visualisation.

It is important to note that there are several limitations which are imposed by the camera settings. Firstly the camera frame rate must be set to at least the capture rate of the lowest sampling or refresh rate within the MoCap chain. For *The Shapeshifter* this is 60 Hz. Moreover, the camera's resolution determines the minimum distance that can be captured, as motion which is contained within the area captured by a single pixel will

Figure E.1: The laser pointer is aimed towards the top half of the video wall.

not be discernible. However, the time taken to process large video files is long, and therefore the camera's resolution should be kept as low as possible. This is another advantage of the use of the laser pointer over Steed (2008). The camera can be positioned much closer than the screen and therefore a lower resolution can be used.

E.3 Data Processing

Similarly to the data collection, the data processing heavily draws upon Steed (2008). Thereafter, the video is duplicated and cropped to isolate the laser point and the visualisation in separate files. After this the two video files are converted to grayscale and a thresholding function is applied

$$f_t(p_{(x,y)}) = \begin{cases} p_{(x,y)}, \text{if } p_{x,y} \geq t \\ 0, \text{if } p_{(x,y)} < t \end{cases} \tag{E.1}$$

where $0 \leq p_{x,y} \leq 255$ is the pixel value at position x, y and t is a threshold value. This isolates the the area of interest, as shown in figure 5.6.

For each frame of each video, a heuristic of

$$c_{(x,y)} = \lfloor \frac{x_{\max} - x_{\min}}{2} + x_{\min} \rfloor, \lfloor \frac{y_{\max} - y_{\min}}{2} + y_{\min} \rfloor \tag{E.2}$$

is used to calculate the centre point of the area of interest.

Figure E.2: The measure below the pendulum. The black bars signify the centre of the measure and +/- 10 cm

.

As the data points collected from a swinging pendulum with one degree of rotational freedom should represent a multivariate normal distribution, the probability density function of the data is calculated and any data points below a threshold are discarded as outliers. Any gaps in the data are filled by repeating the last existing sample.

The centroid of the pendulum's motion in the horizontal plane for each video is calculated as the mean of the x coordinate of centre points across all frames. The displacement of each data point from the centroid for each video provides two sinusoids. These are normalised and zero-meaned. The maximum sample value of the cross-correlation of the two sinusoids is the calculated, and converted from samples to seconds in order to provide the latency.

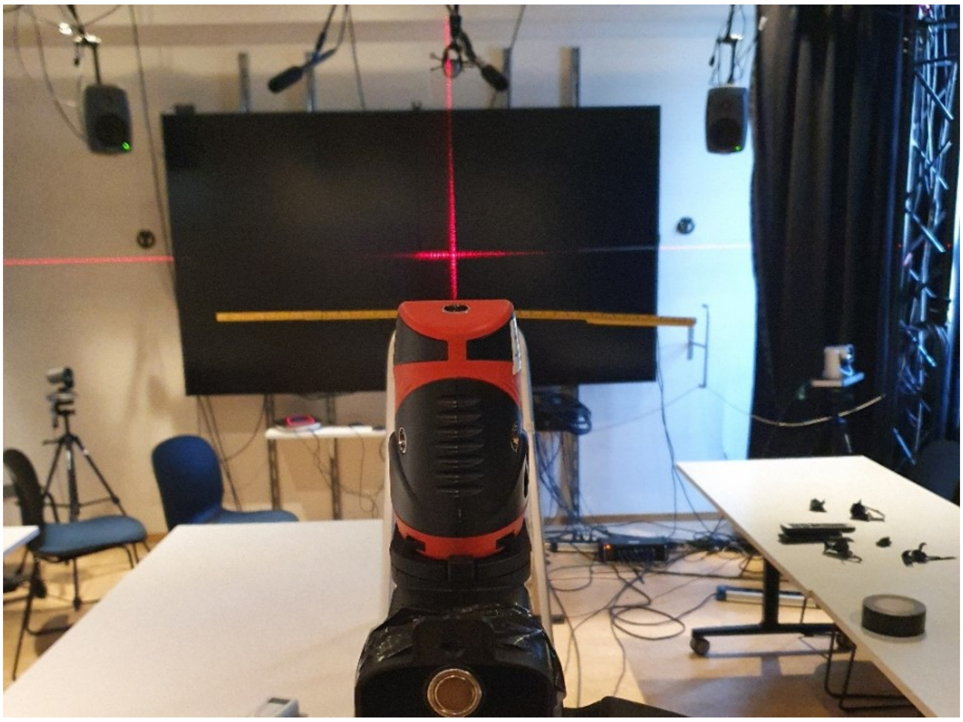


Figure E.3: Alignment of the centre of the measure, the laser pointer attached to the pendulum, and the laser point on the screen.

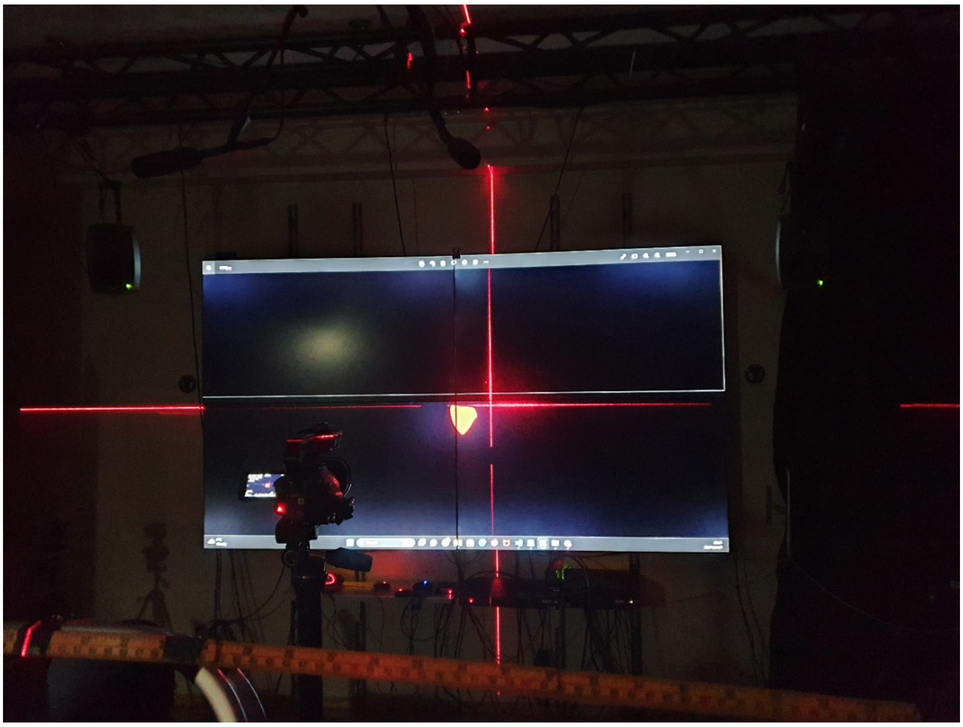


Figure E.4: The recalibration of the pendulum after the collection of each data sample. This involves realigning the laser point with the laser-pointer and the centre of the measure in order to account for any change in position of the laser pointer during the pendulum's motion.

Appendix F

Interview Transcript

1	**Interviewer:** *00:00:02 – 00:00:08*	So, thank you for doing the performance and for doing this and for taking part in the project in general.
2	**Performer:** *00:00:08 – 00:00:08*	You’re welcome.
3	**Interviewer:** *00:00:10 – 00:00:28*	So this interview is divided into two main parts. In the first part we’re going to talk about the performance that just took place, and then in the second part we’re going to talk about the whole process a little bit. And it should last roughly around an hour, maybe a little bit more.
4	**Interviewer:** *00:00:30 – 00:00:55*	And I also then was participating in the project, so I also want you to, I don’t want you to feel that you have to answer correctly, as long as you answer however you want, because we do already have a relationship. So, answer honestly and don’t feel that you have to give me sort of an answer that I want.
5	**Performer:** *00:00:56 – 00:00:58*	Oh yeah, yeah. Okay, nice.

6 **Interviewer:** *00:01:00 – 00:01:22*

So yeah, the first part is, and it's quite an open interview, I'm going to ask some questions, but we can go off and build sort of whatever direction we want to from that. So we have these prompts, but they're not the be all and end all. If you feel that you're finished as well saying something, you can just say I'm done.

7 **Interviewer:** *00:01:25 – 00:01:33*

So, the first part is then dealing with the performance that we just did. And firstly, how did you prepare mentally for the performance today?

8 **Performer:** *00:01:39 – 00:02:20*

I think a part of it was, well, it was kind of mentally, physically when we met before to check if everything was functioning with the technique, which is because that has been a lot on my mind, because sometimes it just breaks or it doesn't work or something is wrong. So it's kind of taking up some of the mental space, I think, to think about what if nothing works and what if it's completely wrong. So that was like some part of the mental preparation, I think, was that part as well, so that we knew that, well, at least it's working now. If it doesn't work in one hour, that can also happen, but at least it's working now and we know some of the things that can happen. So that was a part of it, I think.

9 Performer:
00:02:20 – 00:03:18

And then having lunch was really nice, just sitting outside, I think. I think nature is a really good way to relax and you don't have any distractions. I went to the park, had some food and then just looking at the view, the trees and all the seagulls and stuff. So I think that was the main part. It was good to have that break, actually, so that we knew the technical things and then to have a break where you can just relax. That's what I usually like before a performance as well, that I have some time to just sit or lie down where I don't have any other focus, so that I'm not on my phone or I'm not talking to anybody. It's actually quite nice. So I think that's good. And not really thinking so much about the performance either, just trying to think about as little as possible.

10 Performer:
00:03:21 – 00:03:57

And then when the time is very close to the performance, when you start putting on the clothes and you have a lot of preparation, that's when I go through what I need to remember, like about breathing and about finding how to stand. And that's when I'm thinking through that part. So I think it's good for me at least to have this period where I'm not thinking about it. And then I can have a short period where I just quickly go through everything so that I know it and then I'm ready to start.

11 Interviewer:
00:03:59 – 00:04:04

And you mentioned that you lay down on the grass as well, so there's also a physical element.

12 **Performer:** *00:04:04 – 00:05:05*

Yeah, I guess. Yeah, so I think to either sit comfortably or lie down so that you're not using so much tension. This time I didn't like warm up so much. I just did some stretches and like for the physical warm up. But I think also just relaxing because it's not really, it's not like a physical challenge, this performance. So I think it's more important with the mental kind of and with the body relaxed and breathing so that I'm not, because I can feel in the performance now even that I sometimes I get too stressed and my breath is too short. So then I try to use that in the sound as well. But then, yeah, so that's more important, I think, for this. That was more important for me in this than the physical aspect. But yeah, it's the physical aspect with the mental because it's if it's a calm environment or if you can lie down or lean back or something, it's more the physical helps you to get into the mental kind of relaxation and preparation. So, yeah.

13 **Interviewer:** *00:05:07 – 00:05:30*

Thank you. So the performance that we just did, or that you just did, it mostly consists of improvisatory phases. So you improvise between hitting these predetermined triggers. How much thought did you give to what you would do in each phase before you started the performance?

14 **Performer:** *00:05:31 – 00:05:32*

Before in each phase?

15 **Interviewer:** *00:05:33 – 00:05:34*

So before the performance?

16	**Performer:** *00:05:21 – 00:06:18*	Yeah. I didn't really think a lot about it. It's just about when we did the technical run an hour before the performance. Then I had some idea, then I, from doing it then, I had some ideas of how I could do that going into this one. Okay, the next one, I'm going to go into it like that. But at the movements in between, because we built up these kind of worlds in each color or each section, a lot of it just comes from the little world that we built. And then I feel like that's enough of preparation. If I know which concept I'm working on, then I can just do whatever fits in my mind, what will fit in that concept.
17	**Performer:** *00:06:19 – 00:07:04*	So I didn't think about, yeah, so I thought a little bit about this, like, how can I get into it? And then that, of course, doesn't always work because you're trying to, but then it doesn't work. So you have to work on the, yeah, trying to go out of it and try to get back in again. But yeah, so I think, so I didn't think, I'm not thinking so much about the different, yeah, what I'm doing between. It's just, it's really just from the preparation that we did that, okay, so I know I'm going to go from here until like going from low until raising my hands. So then how will that fit in that world that we created in this space?
18	**Interviewer:** *00:07:05 – 00:07:21*	You mentioned sort of the going, sometimes it doesn't work and you have to go back in and go back out because it hasn't triggered the boxes. How does that then affect what you do in the next phase?

19	**Performer:** *00:07:23 – 00:08:03*	Oh, interesting. Maybe triggers the, how I'm going into the next phase a lot because then I'm kind of stressed and I'm, you know, I think I'm going to get it right the first time and then it doesn't work and you have to do it again and then you don't know how it's going to work. So then it's almost like a hesitation and a bit of a stress when the next one starts because you kind of feel like it's never going to happen but then obviously, of course, at some point it does happen and then you kind of caught off guard because now you're kind of into this trying and failing period.
20	**Performer:** *00:08:03 – 00:08:41*	So then the first part is either a little bit stressful or it's a bit, yeah, or it's just, what was the other word I used? I used some other word but I forgot it already. Yeah, maybe a little bit stressful but not like, I feel like a bit like you're thrown into it then more because you can't control it like now I'm going into the next phase. You have to be like, oh that's, you know, now it happened and then the next phase is a little bit slow to take off or it feels like, yeah.
21	**Interviewer:** *00:08:43 – 00:08:57*	And do you feel that you then have to prepare for that transition the next time around that you have to feel, the fact that it didn't work last time, does that then have an effect on the next time through when you're there?

22 **Performer:** *00:08:57 – 00:09:34*

There was a few things that, there was this one phase that almost every time was a problem and then I worked out some kind of vocabulary that is like, okay, I'm going to use this movement then until I get it right because that worked the last time. So I'm definitely thinking about it when I get to that phase the next time. I don't think I think about it so much because then I'm focused on the next, on whichever phase I'm in but then when I get there it's like, okay, I can remember that I did this to get it the last time so I did that movement or that motion to try to see if that works.

23 **Performer:** *00:09:36 – 00:10:28*

And also there was another one that was not a problem at the beginning but that for some reason I did it wrong. I think I was bending down to lower something so I developed my way of doing that phase into a direction that didn't fit into that box anymore. So I think if I could redefine the box I would have done it differently now because that felt more natural. And then, but I didn't realize that that was what was happening so I thought that I was doing it right but I was actually, I had actually developed the thing to something else. So then when I realized that it was easier again to get it but then yeah, it kind of, no I'm not sure. So your question is like if it's affecting, no, how was it? Did I answer it? Maybe? Yeah, okay, yeah, yeah. Maybe.

24 **Interviewer:** *00:10:29 – 00:10:34*

But kind of building on that a little bit, you said that then sometimes it has to go off in a new development.

25 **Performer:** *00:10:35 – 00:10:35*

Yeah.

26	**Interviewer:** *00:10:35 – 00:10:39*	There was the phase where you said you would redefine the bounding box.
27	**Performer:** *00:10:39 – 00:10:39*	Yeah, yeah.
28	**Interviewer:** *00:10:40 – 00:10:44*	Then you had to start working with that during the performance.
29	**Performer:** *00:10:44 – 00:10:44*	Yeah, yeah.
30	**Interviewer:** *00:10:45 – 00:10:48*	Do you see that then as a positive or a negative thing? Or?
31	**Performer:** *00:10:49 – 00:11:25*	No, like you said, it's a bit of a frustration because you can't change the limits that you've set which is very often the case in performance anyway, I find, because you have to set some boundaries because if it's too free then yeah. But then these are kind of limited so you don't feel like okay but it was this and now it's turning into this but it still has to go back to that one shape. So it's almost like a choreography where you have to hit different spots, like you have to hit those spots and then you can be very free within those boxes but you have to come back to that one spot.

32 Performer:
00:11:25 – 00:12:37

So it's difficult to define it at the very beginning because you don't know if you feel like that's going to actually fit into the phase or whatever you're making. So it feels a bit frustrating that you're like, yeah, I just wish I could move on to the next one instead of catching it. So that's when your brain starts working which I don't like because it feels more like it should be a bodily experience and it's better when I'm just in that and kind of accepting that it's not working or like this is, I wish it would be more like this or I wish I could do it more like that but now I have to do it like this. But then when I start to intellectualise it and think about it then I don't like that so I'm trying to get away from it but it's difficult sometimes because you feel like it should just be working and you just want to work with the emotions and the feelings of each experience of each phase instead of having to be in that box. So it's definitely a challenge to not switch on the brain and start looking at it from that point of view. Yeah.

33 Interviewer:
00:12:40 – 00:12:51

It's on a slightly different track, then in the first cycle through the performance, there's no sound that's made by the system at all. It's just you who's producing the sound.

34 Performer:
00:12:51 – 00:12:51

Yeah, yeah.

35 Interviewer:
00:12:53 – 00:12:58

And how do you get a feeling for the sound that is produced in each phase?

36 Performer:
00:12:58 – 00:13:21

So now I think it's because we have built it up during several times, especially some of the sounds we have been working on a few times. Now it just feels very natural to do those sounds but then sometimes it's kind of surprising as well that I'm doing this sound now. Like this feels maybe different but it's still connected to that universe or that phase or whatever.

37 Performer:
00:13:24 – 00:14:34

But the first time around I think I'm focusing a lot on what we have been working on. So I can kind of place the sounds where we have placed them before and I can kind of just go with those kind of safer sounds. And then after that I feel like that's when more I can experiment with what are the boundaries of these sounds or what can I do instead or how can I complement or go against. Yeah, like make a clash with these existing sounds already. So from the beginning I think it's more free in a way because you don't have the recorded sound yet but it's also more like I feel like I'm putting down what we have already worked on so I think I'm founding it from what we worked on the other times. And then that I think was based a lot on also what sounded good on the feedback. Like what sounded good and also what we felt matched the visuals.

38 Interviewer:
00:14:35 – 00:14:57

And do you consider then, because when in a later cycle of the performance when you make a sound it reduces the volume down to silence of the sound of the speaker, right? Do you consider what you're making in the first sound as well in terms of how you're going to answer it later on?

39 Performer:
00:15:00 – 00:16:01

A little bit like the very first time, because like as an example the first phase is very much like underwater whale-y sounds but I feel like if the bass has deeper tones it's nicer to contrast with light sounds but if the bass has light and if the bass contains all of it then there's not really anything. It's more difficult to work on it later so it's more like that I'm thinking about maybe the first round should be a little bit more simple and a little bit more like defined and then the other ones can just add or do something to that. So it's not like I'm thinking if I'm making this sound then that will be nice to work on this but I'm definitely just trying to do something a little bit simple at the beginning and not go too much of which I think is more successful maybe sometimes but not all the time.

40 Interviewer:
00:16:02 – 00:16:08

And how do you then also get a feeling for how you should move in that first cycle?

41 Performer:
00:16:09 – 00:17:14

So it's very connected I think to the sounds and the universe that we created so I think just those sounds and how the visuals pan out. Like the visuals. When it looks like an insect it feels like you can do this rapid movement but when it looks like a kind of laser thing then I think the movement is very based on the visuals because it's what looks nice with the some of them look really nice if you turn it around or some of them look really nice if you put all of the markers close together or far apart so I think the movement is based on that but then that's also how we build the universe like the sound and the movement came together so I think they're very interconnected and I feel like they're a part of that same. I feel like immediately that I'm in that world and my whole body is this insect and that's including the voice as well.

42	**Interviewer:** *00:17:17 – 00:17:41*	So do you feel then that the, you said you feel like you are that insect for example in the fourth phase and in the past we've been talking about it that the visual on the screen is the insect. Yeah. Do you feel then any relation to what is on the screen then?
43	**Performer:** *00:17:41 – 00:18:38*	Yeah I feel like I am the whatever world it is and then it's not so much me. I feel like then what I'm seeing on the screen that is what I'm doing. That is my, well, that has been my focus a lot I think and then I just move according to that and then sometimes when I'm not looking at the screen I'm just moving. Then it's more like I'm in my own self world and that's just something apart from me but every time I look at the screen I feel like that's a mirror of me or that's like just this is me up there but it's yeah I have to move the puppet strings or like yeah it's almost like that puppet master kind of relationship I feel.
44	**Interviewer:** *00:18:40 – 00:18:44*	And do the sounds then coming from the system feel the same or do they feel ?
45	**Performer:** *00:18:44 – 00:18:47*	I think, yeah, I think when they're played back you mean.
46	**Interviewer:** *00:18:47 – 00:18:47*	Yes.

47	**Performer:** *00:18:48 – 00:19:39*	So they obviously some of them sound more distorted than some others so some of them sound more like what I did I feel and they yeah I feel like they fit and especially today the last time number nine was so much better when it was transformed than it was when I was doing it live so that was very interesting I thought so that was very like the distortion worked really well with that image so then I could didn't in the end I didn't have to say like do any sounds on that part because it was yeah it just it just sounded a lot better with just the played back sound. Yeah. Yeah.
48	**Interviewer:** *00:19:39 – 00:19:41*	Yeah. Thank you.
49	**Interviewer:** *00:19:42 – 00:19:53*	And were there any moments in the performance in which the way the system behaved surprised you?

50 Performer:
00:19:55 – 00:21:20

Well when it jumped over I thought that was the system but that was actually you just clicking it next to the next place so that was that, that surprised me. There was some point when it started to be delayed again when I thought, oh no now it's going to be. But then it somehow it went away again. I think it was just like one of the phases or something that was a bit delayed, had a bit of a delay. The yellowy part I remember last time it was really like covering the whole screen with those lines at some point so I really expected that to happen. So that kind of surprised me that that didn't happen and I don't know if that was a change in the code or the way we had the different positions or what made that happen but it surprised me a little bit and then but then yeah no I don't think it was any other surprise. In a way the fact the sound is quite low I think it's even better if you stand in the middle and it's like with the feedback you can't bring it that much up but it's not a surprise but it's just you know sometimes I'm struggling to actually hear like what's being played back because I'm also moving and breathing and then I can't actually hear like what sounds is going on so that's not like a surprise because it happened it always happens like that but it's also something that I definitely thought of today that it's sometimes difficult to hear.

51 Interviewer:
00:21:21 – 00:21:26

How do you deal with the fact that it's sometimes difficult to hear? Do you have any strategies?

52 **Performer:**
00:21:27 – 00:22:06

I feel like then I'm just either just I feel like it's one or two so either just making some kind of sound anyway and like don't like pay attention to how it's going to relate or whatever just which is nice because then you don't think too much you just like do whatever you feel like this is a moment instead of always relating it to what's been before so it's either that or just stay quiet and then just okay well whatever it is it can just play out and I can just do my own thing. And there's already making a lot of sound just moving around and breathing so I think even with that it's live sound and the other sound as well.

53 **Interviewer:**
00:22:09 – 00:22:13

Were you attempting to create a narrative at all with the performance?

54 **Performer:**
00:22:18 – 00:23:14

Maybe a little bit with the yellow part I feel like it's very much like a desert storm. So it's not so much a narrative but it's definitely some kind of person that is stuck in this storm and with that image I don't feel like I'm so much the screen. I feel like the screen is more just a backdrop even though I know I'm changing the lines and everything and I feel like that's just a backdrop and then I'm the person. That's just the desert in the background and some of the other one has like some kind of character like the insect or the blue or the line going through them that feels more like a character but not really like a narrative like it's not really developing into something but it's more like a character.

55 **Interviewer:**
00:23:18 – 00:23:35

When you said that for the yellow one you feel that it's like a desert and you're the person stuck in the desert storm what do you feel that your motion is doing to that background as opposed to the ones which are more like a figure?

56 **Performer:** *00:23:36 – 00:24:22*

It's an interesting question because I think that's I'm just quickly trying to think if it happens in more of them I don't think so I think that's the only one what I don't feel such a strong connection no then I just feel like that's just some kind of you know like almost like a heartbeat thing at the hospital or like the radio signal or something that is just even though I'm doing this because it's so far away from the human form anyway, because it's just these very long lines so if I'm moving my arm up the whole thing is moving up and it feels less like I'm moving. It feels more like a separate thing I think because of the way it's made visually. And yeah then it just yeah.

57 **Interviewer:** *00:24:24 – 00:24:31*

And how do you think that that then affects the way that you move differently to the ones which are more human?

58 **Performer:** *00:24:33 – 00:25:45*

No then maybe it's well I'm not really sure like the image itself gives me also the I'm slowing down a lot because I'm following this resist the wind the you know the sand so it's like a resistance so then it naturally slows down. But I think also because I'm not so concerned about like how does it move like with the insect, I'm really concerned about like, how does all my different points move so that or like the laser thing like how does those moves and I'm really like looking at the trying to, okay so if I turn it's going to look like that or if I go jump back and forth it's going to look like that. But with this one then I'm less concerned about how it will just be a still image which I think is in a way nicer with this particular visual with the yellow lines and instead of trying to figure out what could be an interesting way of moving this these lines I'm really just focusing on they can just stay still and I feel like it's just the sound and the backdrop is like the main parts. Yeah.

59 **Interviewer:** *00:25:49 – 00:25:54*

How physically present in the room there do you feel during the performance?

60 **Performer:** *00:26:00 – 00:26:27*

How physically present? So I'm focusing a lot on the screen, that's what I've noticed. And when I'm not focusing on the screen, like I said before, then I feel kind of detached and them I'm just very aware of my body and my position and how I'm moving and that's in a way where I kind of zone out and feel more present in myself and in the room. And like what that's when I focus a lot on the physical.

61 **Performer:** *00:26:28 – 00:26:42*

And then I thought I was going to focus a lot more on the people standing there but they were just really just standing there I didn't think about them at all except for like trying to not walk into them or like yeah.

62 **Performer:** *00:26:44 – 00:27:43*

But I feel very present with the screen I feel like a lot of the relationship is very strong with that screen and the visual. And maybe because the sound is so low it's difficult like that that is somehow not so present for me. The sound which is obviously also part of the physical presence but I feel like the visual is it's stronger than the sound to me like the playback sound. So I feel present in this creating this space between me and the screen in a way. And when I'm not looking at the screen I'm moving my arm I'm very aware of how I'm moving my arm but when I look at the screen I'm more aware of how the representation looks on screen than how I'm moving my arm. Yeah. Maybe that's it.

63 **Interviewer:** *00:27:44 – 00:27:56*

So then is it when you're looking at the screen that the screen is almost guiding your movements or do you still feel that you're guiding the screen s movements?

64 **Performer:** *00:27:57 – 00:28:36*

I feel like the vibe and the world is affecting a lot how I'm moving and also the figure because if that looks if this kind of movement looks nice then I will try to explore that more and be like okay I'm going to focus on that movement because that looks the visual then looks very interesting if I'm doing this or like the visual looks more interesting. So yeah. So I think it's definitely me who's moving it but I'm really affected by how it looks.

65 **Interviewer:** *00:28:38 – 00:29:00*

And does that change when it starts to shift into different shapes? For example in what you were saying about the yellow one beforehand that it is the one that feels more like a you're a figure in a backdrop and all the rest feel that you're with, that you're almost, that it's a person being represented on the screen or something.

66 **Performer:** *00:29:01 – 00:29:01*

Yeah.

67 **Interviewer:** *00:29:02 – 00:29:10*

Does the yellow one ever start shifting so that you feel that you're that it's more person like and do other ones start to shift more like a backdrop?

68 **Performer:**
00:29:09 – 00:31:10

Yeah. Today it was a little bit because last time it just filled up the screen a lot and then this time it was more round and smaller shapes. It actually then felt like it was a different person a little bit because I already had that detachment I think to those dots so it felt like that was just like a shape in the desert or something. I didn't feel so I didn't feel like I affected it so much still. And with some of the other ones like the red one is very different when it changes the shape because it doesn't have that pointy that I feel like are those like devil horns or some kind of like diabolic symbol. And then when it's rounding up it's just a bit more woolly and like not so clear anymore. So that definitely changes the atmosphere a little bit but then I know I can just be the person to that affects it to come back to what it feels like. But I think that also affects how I move because also when the shapes that kind of goes some of them just goes bigger or more round and then I think I have the feeling that I want to collect all of them all of the points because that you have the possibility to do that when they're just like tiny dots you can't really collect them all together because you can't. I'm not able to put them all so close together all the points but then when they're bigger it's a lot easier to just cram down and be and collect all the dots closer to each other. So that's affecting then the way of moving in that scenario or like that face change it as well.

69 **Interviewer:**
00:31:13 – 00:31:18

And are there particular parts of your body that you focus on during the performance?

70 Performer:
00:31:19 – 00:31:37

So the breath definitely because that's. Yeah so today I felt like a little stressed and a little bit like I had to produce a lot of sound at the beginning and became short of breath. And so I have to you know that's something you always have to come back to that's like the main focus I think.

71 Performer:
00:31:37 – 00:32:29

And then I'm trying the feet as well I think I'm always thinking about because it changes so much of the other of the rest of your body, the way you put your feet. So I think it's also something I'm very conscious of. And then my eyes and my head it's not really so much my eyes are really fixed on the thing instead of. Yeah, usually they will be in the body somehow but now they're kind of either in the periphery or directly looking at the screen. So they are kind of a focus in that way because it's affecting the whole way I'm thinking and if I'm not looking at it then it's also affecting how I'm moving.

72 Interviewer:
00:32:31 – 00:32:36

So the two parts you've mentioned that you focus on are parts which don't have the markers on.

73 Performer:
00:32:37 – 00:32:37

Yeah.

74 Interviewer:
00:32:39 – 00:32:51

How do you think that your focus being placed on parts of your body where there aren't the markers affects how the visuals appear?

75 Performer: *00:32:50 – 00:33:40*

Well with the feet it's a little bit on the markers because if they're further apart you know there's one marker on each foot. And I think because the arms are so naturally moving when you're moving the feet. So then I think it's almost better and then there's some parts when I'm actually moving my arms and like focusing on how they are moving in the like especially in the robot like the ninth phase and in the blue line which I can't remember the number of. But yeah I think the hands and the upper body is usually just a lot of time just following what you're doing with your feet. So yeah I don't know how to answer that question anymore but yeah.

76 Interviewer: *00:33:41 – 00:33:45*

Do you think it would change if you had markers on your head or on your feet?

77 Performer: *00:33:45 – 00:34:47*

Definitely, I think so maybe especially on the head because it's now I'm just really using my head to look at the screen or not look at the screen. So if I had something on my head I would have to place my head in a certain way as well. And because I'm focusing so much on the relationship between me and the screen then the head kind of becomes detached from that. It's just used to observe the screen instead of... Because I'm also focusing on my hands to get them in into the right position that's especially with this one when I'm trying to get into this blue line from the insect from the purple to the blue. And then especially that one it's a lot with the hands because I'm not getting it right. And then yeah so then the focus comes on like where the issues are also in some parts of the performance the focus is on what.

78 Performer: *00:34:51 – 00:35:17*

And then you know because the arms you can use as an extreme like an extension as well. So if I'm putting my hands up then it's going to change the visual a lot. So I guess I'm focusing on it as well. It's like yeah one of the focuses as well because you have to... Yeah because then how does the visual look and then I'm trying to do something with my arm to see how it's looking. But I think yeah.

79 Interviewer: *00:35:21 – 00:35:53*

And also related to the markers throughout the process of developing and working with the system you've used quite a large number. You've used from one or two up to eleven or twelve I think at some points. And I mean the system has the possibility of going up to 30. If the computer can handle it. Are there any differences in your experiences of performing with different numbers of markers?

80 Performer: *00:35:55 – 00:36:35*

Yeah I think when there's few when there's for instance only two you have... The figure then doesn't feel so like a figure or character. It's more like just some kind of visual representation a bit like the yellow lines. They all feel a bit like that because it's not so apparent. Because when you have all the markers then or like more markers then you get like this character or this assembly of different points. So I think that definitely changes the way the representation feels. Yeah different.

81 Interviewer: *00:36:36 – 00:36:46*

Do you feel as well if it's then a few, if it's just one or two then, do you still feel like you're doing the yellow one that is an environment?

82 Performer: *00:36:46 – 00:36:56*

Yeah then it's more like an environment or just like a separate something. And it's yeah doesn't really become its own character. Yeah.

83 **Interviewer:**
00:36:59 – 00:37:05

And can you imagine what do you think would be different performing with say 30?

84 **Performer:**
00:37:07 – 00:38:27

Yeah I think then it would be then some of the shapes would be so much closer to a human body. I think like the representation would look more like human body. Which I think would feel different. I'm not sure if it would feel like a separate character then or if it will feel like it's more like a representation of me or like I'm moving it and getting affected by it. I'm not really I could go either way I think. But I would I think it would feel like a lot like a character but I'm not sure if it would be a stronger relationship. Because it's this relationship with abstract object like in a laser or an insect or some kind of forest creature. This then it's the atmosphere is so strong and it's more abstract, so I'm creating this world around it. I feel like if it was more human humanoid figure it would be maybe more realistic and maybe not as like open as a world and as like an abstract image of the insect or whatever.

85 **Interviewer:**
00:38:28 – 00:38:34

And do you think that that would change depending on where you placed the markers?

86 **Performer:**
00:38:34 – 00:39:08

It could be because if you place them very like for instance the one that's creating a line between all of them. That could obviously go from one hand to the next hand and you don't have to place them so that it looks like a body. I mean I think that could be interesting to try anyway with the 10 or 11 or 12 because then you could also have like instead of it looking like a body you could make this very random thing going from the foot to the shoulder to the and then it would just be a shape. Instead of sometimes looking like a human body.

87	**Interviewer:** *00:39:17 – 00:39:39*	So the boxes that are used then to trigger the next phase. We've also used quite a number of sizes of them ranging from about 0.2 meters to a meter or a meter and a half. What are the differences in experiences of performing with the system?
88	**Performer:** *00:31:13 – 00:31:18*	With the different sizes?
89	**Interviewer:** *00:39:41 – 00:39:41*	Or are there any?
90	**Performer:** *00:39:42 – 00:40:09*	Yeah I think there's a difference and I think we found a good midway point because if it's too big you're kind of just walking into it even though you didn t meant to at all and then you can't have any control of it. And then it feels very much like it's always just jumping to the next thing and you kind of want to be in that...
91	**Performer:** *00:40:12 – 00:40:41*	It feels like I should have some control over when I want to move into the next one but if it's just happening then... And then when they're too small it definitely feels a lot of this moving into my head and the intellectual kind of thinking getting frustrated instead of trying to be in the task and being in my body in the room in the space. Then the frustration of not getting it then gets too much.

92 Performer:
00:40:42 – 00:40:23

And I think when we tried the small boxes it was... At first we didn't try so many places so if it's just going from one to two and then back to one and two or like one two three one two three it's easier but I think remembering nine different very specific positions and getting those that would be... Maybe not as interesting because I think I would focus too much on getting the position right and it will never get on to the next one and the way between the boxes wouldn't be interesting to work on because it would always be like a huge barrier to get through the next time you get to.

93 Performer:
00:40:24 – 00:42:38

But now it's more you can also focus on the different world and the transition into the next box instead of just focusing on getting the next point. But then when we just did for instance three different phases then it was easier to remember exactly how to be placed. I remember that was also kind of a fun challenge if it was very small because then it almost became just like 'get the position' and then try to remember exactly how you move so that you get in the exact right position for the next one and it would be quite close because you can put the different points on the floor so close to them because the boxes are so tiny anyway. So you have to be exactly there so there's no chance of catching them on the way to the next. So then it's a different challenge because you can... But then I think that would be too frustrating if you're moving a long way if the boxes are... if the new phase is a long way away and also if you have too many phases. I will go into my head kind of thing.

94 Interviewer:
00:42:41 – 00:42:50

Does it then feel more or less improvisatory? What you were describing with the three phases together.

95 **Performer:** *00:42:52 – 00:43:20*

So that feels more like a task I think and less improvised. Yeah, that feels more like a choreography or like a set score that you have to do. Instead of focusing on the emotion and the feeling and the exploration of the world or the thing that you're creating. Yeah.

96 **Interviewer:** *00:43:24 – 00:43:25*

Okay, thank you.

97 **Performer:** *00:43:25 – 00:43:25*

That's it?

98 **Interviewer:** *00:43:26 – 00:43:28*

The first part of the interview.

99 **Performer:** *00:43:27 – 00:43:29*

Oh yeah, that's on the performance, yeah.

100 **Interviewer:** *00:43:29 – 00:43:34*

There's just a few questions, not as many, on the whole process.

101 **Performer:** *00:43:35 – 00:43:35*

Okay, yeah.

102 **Interviewer:** *00:43:36 – 00:43:37*

And sort of the development.

103 **Performer:** *00:43:37 – 00:43:37*

Yeah, nice.

104 **Interviewer:** *00:43:39 – 00:43:44*

So the first is, just can you tell me about your performance background?

105 Performer:
00:43:45 – 00:44:50

Okay, so I studied for three years drama and theatre in Wales, in Aberystwyth University. So I did drama in high school as well actually, so that was when I started doing some kind of drama thing. But then I decided to do it in university and my initial thought was to go into scenography. But then I decided to do drama and theatre because you can choose a lot of subjects within scenography but also within other things. So then I could do, I did mainly just scenography and physical theatre. So I think that like, and that kind of stayed with me as well, the space and the body kind of focus. So you have the scenography part where you focus on the space, what's in it, what it sounds like and what it feels like, and the light and the set and everything. And then you have the body in that space and the physical presence, which is the physical theatre. So I think that's how it started and I just naturally just picked those subjects in school.

106 Performer:
00:44:52 – 00:45:54

And then I knew I wanted to move back to Norway. So when I came back I just contacted people that I would like to work with. And then through that and through other contacts that I got from that, I started just working with some different people. And I also continued working with my friend from university as well, who's Swedish. So we made some different things after uni as well, which was nice. So I kind of kept in touch with that environment, but it's difficult because everyone moved to different places. So at the beginning it was a little bit like that, like staying with some of the uni contacts and then trying to establish a little bit in Oslo. Or the area around here.

107 Performer: *00:44:55 – 00:46:27*

And then I worked with... And from the beginning I started working with [a theatre] in Oslo, which is where I still work. And at the beginning I was just an assistant and this kind of work. And then I started performing with them a bit later on. But they do physical theatre, Grotowski and Artaud theatre work and then it developed from there.

108 Performer: *00:46:28 – 00:46:50*

And then with the improvisation part I think has come through different environments. Or like different people I've met and been in more dance related improvisation or more voice related improvisation. Through different people I've worked with or different people I've met in different forms.

109 Interviewer: *00:46:54 – 00:47:02*

And what term would you use then to describe yourself as a performer? In what art forms?

110 Performer: *00:47:03 – 00:47:42*

Yeah, so I usually just say that I'm an actor. Which I think is the closest to what I can get, because it's a quite open term anyway. And you can use it for so many things. But some of... I would also say like performer or performance artist sometimes. Or maybe sometimes I say I work with dance and theatre or physical theatre. So I think it's different terms depending on who I'm talking to and who I'm describing it to. And what are their backgrounds.

111 Interviewer: *00:47:49 – 00:48:01*

You have these many terms that you use. Is there a difference in the way you feel about yourself and how you describe it?

112 Performer:
00:48:02 – 00:48:52

No, I feel like all of them are true in different situations. Or it's just a way of describing it to someone. Say they don't know too much about physical theatre, they've never heard that term before. I feel like it's a better conversation if you say you're an actor and a dancer. Or I work with theatre and dance and the mix between those. Just because those are things they can immediately understand or have some kind of preconception about. And then you can actually have a conversation instead of just saying something that feels maybe more true to what I'm actually doing. But then that's not really interesting to be... So I'm not really so set.

113 Interviewer:
00:48:53 – 00:50:07

I feel like... I remember this one time I was at this voice laboratory. And I just mentioned that I play the flute. I play the... whatever it's called. In the region it's called a tidefl yte[?]. Because it's going across. And then this is called something else, flute. And then I said, yeah, so I kind of play the... I'm kind of sometimes a flautist, or I'm playing the flute sometimes. And then this other person who knew me said, but you are a flautist, you do play the flute. So clearly she had an idea about that I am, because I do play the flute, that I am a flautist. But I don't feel like that myself. But every term I'm using to describe myself I think it feels true. But I think then other people could see me as something else. If they have seen me play the flute, maybe they think I am. Or maybe their opinion is... which is fine too. And it's interesting to see what other people would describe me as, or when they do that.

114 Interviewer:
00:50:09 – 00:50:20

Did you then come to this process, or this project, sorry, from a certain perspective? Or was it very...

115 Performer:
00:50:26 – 00:51:12

So I think what I just heard is that this kind of... With the markers and how it's used in popular blockbuster films, it was the only relation I had to this using these kind of markers and this kind of system before. But I knew it was going to obviously not be like that. But I think my idea was a bit that it was going to be more like this, like a representation of the whole body, or focus on just how it was made or how this... But it's a whole different concept obviously, this one, than I thought. Because that was the only thing I had heard about using these markers before. And then... And then I knew because of the person that contacted me, [name of supervisor], your supervisor, and she said it would fit with what I'm doing. And I met her through this improvisation group, [name of group]. And so then I thought, okay, so it's going to have something with movement, connected with the voice and the voice as movement. That's how it's going to be, because that's how she knows me as a performer anyway. Because that's when we have been working in that environment, we have... I feel like that was the impression she must have had of me. That I'm like the body sound and the sounding body. Kind of... Thing to work from. So that was the only things I've... Before going to, but other than that, I didn't think so much about what it was going to be or what the project was.

116 Interviewer:
00:52:15 – 00:52:22

Have you ever worked with technological systems or interactive systems in performance?

117 Performer:
00:52:23 – 00:53:49

Not really, I just... I don't think so. I think I've never... I've only tried these... I've mentioned these rings before, [name of performer who works with the rings], she does this with the sounding rings. I've tried those, which I thought was very difficult. They're a very specific way of using them. So it's a little bit like having the smallest possible box here, and you have to do it in that way. Just moving your hand like that, wearing that ring, it's going to change the sound a lot. I think it's a different direction, it's up and down, and it's tilting. It's kind of a different direction, but a small movement has a lot to say. And you have to remember which one feels really restrictive. So this is more open than that, and also because every movement doesn't affect the imagery, or every sound doesn't affect... It's just like, as long as you go into the next phase right through these boxes, then everything else feels more open. Whereas with these things, it's like, as soon as you make a small sound, and I can't remember, I'm having a lot of trouble remembering, is it this now, or is it that way, or is it up, or is it that way? But I haven't worked with any other technical features, I don't think. I'm trying to think, but I can't. No, I can't remember.

118 Interviewer:
00:53:52 – 00:54:04

So then how did developing the work here with the technological system differ from your normal way of developing?

119 Performer:
00:54:11 – 00:55:06

So I think it's, what I mentioned before as well, usually it's like, if the movement material, or the sounding, or whatever, the text, whatever, is changing and taking a different direction, usually you can just build on that and see how it leads, or you can go back and then you can take a new direction, and you can always come back, or you can take a little bit of this direction, and you can always put it together. But here you also have to come into that next phase again. You can't go just starting doing something completely different. So it's almost like the choreography is pulling you back the whole time, like, okay, now you have to start that, and then you start the next one, and it's this kind of going back and forth. So that's a bit different, because usually it's a bit more that you can follow something for a longer time.

120 Performer:
00:55:08 – 00:55:40

But then, obviously, that's the process, I think. And the process has also been to test out how big should the boxes be, how many boxes can we have, how many points, how many markers can we use. So that's been a part of the research as well. But usually there's other kind of rules when you research something, like your voice can't be too much in this direction, or that doesn't fit.

121 Performer:
00:55:44 – 00:56:47

So maybe the boxes, they're so fixed, so then you have to come back to them. But usually you kind of don't work with such fixed boxes. It's more like the listening. And then obviously when you do come to having an actual performance, then everything is quite fixed anyway, and you have to do this for four minutes. And then you do stand there for one minute, and then you sit down. So then it's fixed, but the process is quite different then, because then to get to that point, like how long you should stand there, you tried to stand there for ten minutes, and you tried to stand there for ten seconds, and then you found in the end that one minute was the best option. So I think then the process can go. But here you always consider the boxes in the next phase. And you know that's going to be in the result as well. But usually when I work with some kind of other creative process, you don't know if that's going to be in the result, or if it's going to just be abandoned altogether.

122 Performer:
00:56:49 – 00:57:23

Although I have had some performances where you also, you know this text is definitely going to be at the end. So you have to find a solution of how to do that text, so that it would fit with, or like how to present it. Or you know, it's in this, you decided that this table is going to be there, and then, okay, so then you use that table in the process, and then you try to. So there's sometimes things that are going to be there, which is a bit like the boxes, you know they're going to be there in the end.

123 Interviewer:
00:57:25 – 00:57:38

So would you say then that it's, that the process here was trying to fit the boxes into the performance, the same way that you learned to fit the table into the performance?

124 Performer:
00:57:38 – 00:58:27

Yeah, yeah, in a way it's like you have these boxes, and then each box is a transition into a new world. So it's kind of the entrance. You have to get through this entrance, so then in that way it's like, yeah, compared to this text or something like, you have to say this text and then you can move on to, or you know, you're sitting on the table, and then the next thing has to happen. So it's some kind of, but usually it's not so fixed, because you can sit so many ways on the table, and you can say a text in so many ways. But the boxes, even though they're in different heights, or different, they are kind of, I don't know, it just feels a bit more like they're even more solid and certain.

125 Interviewer:
00:58:29 – 00:58:36

Which The boxes can be defined at the beginning of each performance.

126 Performer:
00:58:37 – 00:58:37

Yeah.

127 Interviewer:
00:58:38 – 00:58:52

So is it then that when you commit to making that decision at the beginning of the performance, then it feels concrete?

128 Performer:
00:58:53 – 00:59:38

Yeah, I think so. Yeah, I think, but also you know that it's going to be there. But we could also decide, I mean, you could also do it like, whenever you want, you can just click through five phases, or whatever, I give some certain sign, you skip three phases and you go directly into the next one. So, you know, you don't have to do it like that anyway. So, like, if you try it a certain amount of times, then just continue and never hit that box, and just be in that one phase for the rest of the performance. And then have some other sign that shows that the performance, so it doesn't have to be like that, but like the way we did it now, the process has been to go through the phase until the phases are done.

129 Performer:
00:59:40 – 01:00:06

So, you know, in a way it doesn't have to be like that. So that's an interesting question, I should think about that, like how, like, okay, so then what if you just want to stay in that one world, and then that's going to take too much time. If I stay there for five minutes, that means that you have to skip two rounds, because you know that's going to be how long that would take, and then go on from there, or like that we could have some kind of dynamic.

130 Performer:
01:00:07 – 01:00:52

But, yeah, so it's defined at the beginning, but then during the performance it's, but that's usually the case with any performance I do, that, you know, when the performance starts, you can't just change everything. There are certain things that you have to do, or there are certain things that are going to happen, if it's not entirely like improvisation performance, which I have done as well, but usually there is improvisation, but like within these four minutes you do this, or within these eight minutes you do this, which is kind of the same thing, like within this space you're doing that, and then you have to go through this gate, and you find a new world, and then you have to do this, or like these are your boundaries within this.

131 Performer:
01:00:54 – 01:01:28

So I'm not really sure, like it's difficult to define like how, why it feels like, because yeah, like you say, it's also defined at the beginning, so you could just make some boxes that you know it's going to flow into next, but then you don't always know how it's going to, because it's so improvised, sorry, I can't say that word, but because of all the improvising, then you have to, you can't just follow what it develops into, you kind of also have to find the next, yeah, next box in a way, but yeah.

132 Interviewer:
01:01:32 – 01:01:38

So on a bit of a different track, can you describe to me your role in the project?

133 Performer:
01:01:40 – 01:01:40

My rule?

134 Interviewer:
01:01:40 – 01:01:41

Your role.

135 Performer:
01:01:41 – 01:03:23

Oh, role. So I think what I have, what I can bring is that, I'm very interested in the relationship between my voice and my body, and how it's so connected, and how to not actually in fact separate those two, that it's, yeah, the vocal cords obviously just the body part too, like the fingers or whatever, so it's just a movement, that's just movement anyway, and the breath obviously is just movement, which is then making the sound. So I think that's what I can bring, because you need that kind of presence in that technology I think, because or else the technology is just some kind of, yeah, not really like interesting in living, it's just, you can just program it to be whatever, and then program some interesting sounds, but then when you have a living body that's focusing on that relationship, and also making a relationship with the screen, then you have a new kind of, I don't know, interaction, and another kind of input, and a different way of just doing things, instead of thinking about good shape or graphics, sound. So you have that, yeah, the voice and the body moving the image, so I think that that would be different if you didn't have, so I think that's the main part.

136 Performer:
01:03:26 – 01:05:05

And then definitely during the conversations, and you also had some questions for me at the beginning that I answered, then I think just our conversation has developed a lot, like, okay, maybe we should try it like this, like try the small boxes, bigger boxes, different positions, or different placement of the markers, how to use the space, what the different worlds feel like, and you definitely had a lot of, I think all your visual representations, we also talked about some of them, like this one is cool, and maybe we should have something more like that, and you described, so I think it's just our conversation, we have developed a lot of it, either like conversation, and then trying it, or just trying it and then discussing it afterwards, like how can we do it, and then trying it again, or waiting until next time. And just like a lot of just practical adjustments as well, like it would be really good if we could do it like, like one of the very first thing you did I think was to just catch the spot, so we didn't have to stand still and wait for, and wait for the, to like define the boxes for each phase, and that was just like, yeah, so some practical things as well, that comes with having another human in the space, instead of just as a concept, or, yeah. Thinking, yeah.

137 Interviewer:
01:05:09 – 01:05:21

And we've just got a few more questions, just a few more. So this project's lasted approximately half a year, and what influenced you then to commit to a project of that length?

138 Performer:
01:05:22 – 01:06:10

I think it's, I have this, if I commit to something I have a hard time not going through with it, so I think that's like my main, my main reason, because usually when I start something I always finish it, so that's like a main, that's an important value for me anyway, because I feel like it's important to see things through sometimes, or like most of the time anyway, because it's, yeah, or else you don't, you're not giving it a chance, and obviously you have to give something up if it's not for you, or if it's destroying you or whatever, but yeah. I think that's, yeah, that's like a main thing for me.

139 Performer:
01:06:11 – 01:06:58

And initially I was very intrigued by the project, so I think also, and then at the first time I can remember that it was very, like, felt very technological then, because I'm not working, I've never worked with any of these, like, technological kind of things, I felt very foreign, but I also thought that, you know, this is going to be a new thing for me, and new things are always good to find new ways, and to also discover more about who you are and what you're not, and what you, so I think that was, like, intriguing to begin with, and it also, like, after the first meeting it was something that really, was so different for me, and I always like to do something that's different.

140 Performer:
01:07:02 – 01:07:57

And also it was frustrating to work with this technology, because it's, yeah, like we discussed a lot now as well, with the process, it's quite different, and it's also creating this program, and it's kind of slow, because you have to make some, and then, you know, you have to build it up over a long period, which is different for me, so then, yeah, I think, yeah, just experiencing something so different, I think, was very, yeah, helpful as well, because it's nice to feel like, like you're feeling this frustration, and then trying to sit with that, and what is it, and why is it like that, and how can it still bring something new, or like bring something exciting, or how can it develop, even though it feels, like, restrictive, or it feels, like, different, and, yeah, yeah.

141 Interviewer:
01:07:59 – 01:08:06

Do you think you would have developed things differently if you weren't to see the screen in the performance?

142 Performer:
01:05:22 – 01:09:25

I think so, yeah, I think so, definitely so. But then, because it's so interesting to see how the different visuals, what kind of movement looks good on the different visuals, like some of them look really good when you're turning, and some of them look really good when you're stretching, some of them look really good when you're slow or fast, so, I don't think, also building the worlds would have been so different, the worlds would have been so different, I don't know how we would have built the worlds then, because then you wouldn't have the colours and the feeling of the thing, so, then I would have had to work on a different way of creating different worlds, because then what would it go, like what would it be based on, maybe just based on the position of the boxes, and then into the next phase, but then because the position would be different every time, then maybe the different phases would be different every time, so, yeah, I think it would be very different, but I'm not sure how it would, if it would just be the same movements, but going through the phases, or if it would just, yeah.

143 Interviewer:
01:09:27 – 01:09:30

Do you think your relationship to the markers would be different as well?

144 Performer:
01:09:34 – 01:10:21

Yeah, I don't know, maybe, because then, because you can't see what it actually looks like when you're doing that, maybe you would try to do so many different things, because now it was kind of, ok, this looks cool, or this representation looks really interesting, then I want to do that, like many times just turning over and over and over and over, or like moving faster and faster, but if you can't see that, then you don't know what representation is more interesting, so maybe you will just skip through a lot of different movements, yeah. Kind of like the same with the world, so you would just, you don't know what world will fit, so you just have to try a lot of things, yeah.

145 Interviewer:
01:10:25 – 01:10:36

And would you then pay as much attention to where you were placing the markers, or would you just be a lot freer with that?

146 Performer:
01:10:40 – 01:11:44

No, I think it would be I'm not sure that would be so different, the markers actually. I was thinking about the sound actually, because if I could still hear the sound, I think that would define a lot of the world, because it's, that all we tried earlier, in the early, early in the process, we tried just different sounds without any kind of world or any kind of thing, and then, when, then the next time you come back to the same face, you just have some kind of random sound that is just maybe different sounds, and then the next, so then when you're trying to lay something on top of that, there's nothing really to build on, because it's so, everything is just very eclectic and just put together weird sounds, but now that we have the sound, it's so much easier to build on that, to contrast it or to work with it, but the markers, I think I'm not so sure, because if you can see it visually anyway, I think I would just be like, okay, so my arm is moving and my markers are there, but I, yeah.

147 Interviewer:
01:11:47 – 01:12:02

One more thing that's just, that I'd like to ask as well, is that the sound is all from you, in the performance too, then to what extent do you feel like you're hearing yourself?

148 Performer:
01:12:04 – 01:12:40

When it's so distorted, I don't feel like it a lot, I feel like that is, but sometimes I can remember, like, it sounds a lot like the sound I was making, and I can remember the sound I was making when I'm hearing it back as well. So the sound, yeah, so it's very different to how distorted it is actually, I think, because some of it just sounds like white noise or some kind of, yeah, kind of what I was making, but different. So I think the sound definitely sounds, yeah, very different.

149 Interviewer:
01:12:41 – 01:12:47

Do you think that changes the way that you feel about the visual representation?

150 Performer:
01:12:51 – 01:13:24

A little bit, because the sound that's already then recorded and played back already makes some kind of image with the screen without me doing anything, I can just be standing still and that's already something going on. And then sometimes I feel like this distortion, it really fits with this image, and sometimes I feel like it doesn't fit with the image, because the way it's distorted just changed it from something I wanted to do, so then it's the letting go of the control, I think, to try to think about it.

151 Interviewer:
01:13:25 – 01:13:27

Does it feel more like you then or less like you?

152 Performer:
01:13:32 – 01:14:37

When if it's... I'm not sure. I think just sometimes it's just distorted and it feels still like what I was trying to make, and sometimes it doesn't sound distorted at all and it sounds the right way, and then that can go both ways as well, like the distorted sound is strange, usually the distorted sound is really good, it fits really well, I think. I don't think I ever heard a sound where I felt like I was completely wrong, except for some part of the process where we didn't stop the recording, when the phase started we also swapped the sound recording started straight away, which is better now, I think, because sometimes you can't stop the sound straight away and then that carries into the next phase. So that's the only time I felt like it was something wrong, or like something out of that, that didn't fit that world, because it was meant for the previous world, and then it was just dragging because I couldn't stop it.

153 Interviewer:
01:14:39 – 01:14:58

So final question, which is, to what extent did that performance space that we were... which was also the space we were developing in between, was also the space we performed in? To what extent did that influence the way that you developed the work?

154 Performer:
01:15:00 – 01:15:37

So it was a bit restrictive as well, because we knew that there's only certain places you can't stand, because the cameras won't catch you, and there's also certain positions that you can't do, and it's not really any point going so low to the floor, you're not going to catch it anyway, so it's better to have something that you know is going to be picked up. So that definitely restricted it a little bit, but again, restriction is not really necessarily a bad word, because that just means that you have more defined and more boundaries to work with.

155 Performer:
01:15:42 – 01:16:12

So I think that definitely did influence it a lot, and then the restriction of how you were seen by the camera, and also when we started using the microphones, which we didn't in the beginning, but it was going... or like the speakers, that's all the way around, because then you also wanted to stay in there to hear it yourself, and then the boxes, well, it was also within the reach of the cameras as well.

156 Performer:
01:16:14 – 01:17:15

So I think it would be very different in different spaces as well, just because if it was more open, if you were more free to go, I think we would have a lot more... let's say the circle was bigger, then you could go bigger distances, or run, or do... It doesn't feel like a space for such big sounds as well, because the movement and the voice is so connected, then if there's a big movement, then you can have a big sound maybe, or you can always contrast it, and it can still be connected, but I think if it was a big room, it would feel more like a run or a shout, it could be more appropriate. So I think that would be different, and again, if you have a smaller space, I think it would be a lot more restricted, and then for the boxes not to hit each other, we would have to find some way around that, so I think it's... especially the size affected it a lot.

157 Interviewer:
01:17:17 – 01:17:18

Okay, thank you.

158 Performer:
01:17:18 – 01:17:19

Okay, yeah, thank you.

159 Interviewer:
01:17:19 – 01:17:21

That's the end of the interview. Thank you very much.

Appendix G

Interview Guide

This appendix presents the prompt questions that were employed for the interview. Part one consists of prompts relating to the performance. Part two consists of prompts relating to reflections on the collaborator's role in the project.

Part One

1. State relationships to each other and project
2. State should not fit answers to me
3. How did you prepare mentally for the performance?
4. How did you prepare physically for the performance?
5. The performance mostly consists of improvisatory movement phases, however you do have to hit the predetermined positions to trigger the next phase. How much consideration did you give to what you would do in the improvisatory phases before the performance?
    (a) If there was any consideration:
        i. How did these considerations manifest in the performance?
6. In the first cycle of the performance there is no sound being produced by the system, it is just you producing sound.
    (a) In the first cycle of the performance there is no sound being produced by the system, it is just you producing sound.
        i. How did you get a feeling for the sound that should be produced at a given moment in the first cycle?
        ii. How does this differ from later cycles in the performance?
        iii. How did you get a feeling for how you should move?
7. Can you tell me about a moment in the performance in which the system surprised you?
    (a) If yes:

    i. How did you react to this?
    ii. Elaborate on further points of surprise
  (b) If none:
    i. Were there any moments where you attempted to get it to surprise you?

8. Were you attempting to creative a narrative throughout the performance?
  (a) If yes:
    i. Can you tell me a version of this narrative?
    ii. How successfully do you feel that you performed this narrative?
    iii. If positive answer elaborate on what facilitated this
    iv. If negative answer elaborate on what impeded this
  (b) If no:
    i. What was your guideline through the performance?

9. How physically present in the room did you feel during the performance?

10. What was your relationship to the screen during the performance?

11. Was there a particular part of your body that you were focusing on?
  (a) If yes:
    i. How consistent was this focus across the performance?

12. Were there any moments that frustrated you during the performance?

13. How do you think the performance would change if you placed the markers in different positions?

14. How do you think the performance would change if the bounding boxes were different sizes?

Part Two

1. Can you tell me about your performance background?

2. Did you come to the project with a certain perspective?

3. Have you ever worked with technological systems in performance?

4. Can you describe your role in the project?

5. This project has lasted approximately half a year so far. What made you commit to a project of at least that length?

6. To what extent did the performance space influence the way that you developed the work?

Appendix H

Audience Survey Pilot Data

The following appendix contains the raw data from an audience survey carried out after the pilot performance of *The Shapeshifter* in May 2023. This thesis was originally intended to include an evaluation of the audience's perception of the relationship between the physical body of the performer, the medial representation of the body, and the technologies involved in a performance. However, due to several limitations outlined in section 8.3, as well as a desire to reduce thesis scope, these plans were postponed. Nonetheless, the audience for the pilot performance were asked to fill out a short survey consisting of pilot questions. As these responses could be of use for future work, they are included in the following pages.

Survey - The Shapeshifter

Thank
you for attending this performance of The Shapeshifter. We have some questions
for you about your experience of the performance. The survey should take
approximately 15 to 20 minutes to complete. You are free to skip any question. You can withdraw
from answering this survey at any point. All answers are anonymous.

Background

What is your age?

25

What is your gender?

Male

Do you have any formal training in a form of performance involving movement (e.g. dance, theatre)?

- (•) Yes
- ( ) No

If yes, what form(s) of performance?

Theathre

If yes, for how many years did you receive formal training?

2 years at high-school

Do you have any formal training in musical performance?

- (•) Yes
- ( ) No

If yes, for how many years did you receive formal training?

4 years

How often do you attend live dance performance?

- ○ Weekly
- ○ Monthly
- ○ Yearly
- ◉ Less often than yearly

How often do you attend live theatre performance?

- ○ Weekly
- ○ Monthly
- ◉ Yearly
- ○ Less often than yearly

How often do you attend live music performance?

- ○ Weekly
- ◉ Monthly
- ○ Yearly
- ○ Less often than yearly

Do you have any prior experience with motion capture?

- (•) Yes
- ( ) No

If yes, can you briefly describe these prior experiences?

I have completed a course at university, and used it for two projects. Mainly to capture data, not to create artisic performances.

Do you have any prior experience with spatial audio?

- (•) Yes
- ( ) No

If yes, can you briefly descbe these prior experiences?

I have completed a course at university, and used it to create music. Not at a performance level.

The Performance

Can you briefly recap the performance?

The performer walked arround the room with a microphone, and used virtual "boxes" to shift between different figures representing her body on a screen. The spatial audio output changed on the basis of what figures were represented on the monitor.

What do you think was the central idea/theme that the performance was conveying?

Representing movement and sound in a spatialized 2D and ambisonics environment.

How many performers took part in the performance?

There were only one, but she was represented by figures on the screen. The number of figures, I guess were about 5?

Can you list the performers?

Line, sircles, triangle, tubes, and squares

What were you primarily focussing on during the performance?

The figures on the monitor

How would you describe the relationship between the performer and the visuals?

1 to 1. The figures changed on the basis of the performer.

Did this relationship change during the performance?

- (●) Yes
- ( ) No

If yes, can you describe how this relationship changed?

The figures change on the basis of where the performer was located. If she stepped into one of the virual boxes, the figure changed.

How would you describe the relationship between the performer and the sound?

Not as 1 to 1 as the visuals. The sound changed when the figures changed. But I did not experience that the sound had a direct correlation to the sound generated by the performer. It must have been some sort of trigger, and not a sampler.

Did this relationship change during the performance?

- (x) Yes
- ( ) No

If yes, can you describe how this relationship changed?

When the figures changed, the sound changed. Also the position of the performer made the sound change.

What parts of the performer's body were used to interact with the system?

All movements of a basic shape.

How clear was the relationship between the performer's motions and what was shown on screen?

Yes. 1 to 1.

How clear was the relationship between the performer's motions and the sound made by the system?

Some positions triggered the sound output.

How clear was the relationship between the performer's vocalisations and what was shown on screen?

No relation

How clear was the relationship between the performer's vocalisations and the sound made by the system?

Not as clear. Percived as a trigger.

Did you perceive a relationship between the performer's motion and the spatial placement of the sound?

- ○ Yes
- ◉ No

If yes, can you describe this relationship?

Which statement do you most agree with:

- ( ) The performer was in control of the performance.
- ( ) The system was in control of the performance.
- (•) The performer and the system were cooperatively in control of the performance.
- ( ) The performer and the system were competing for control of the performance.
- ( ) Neither the performer nor the system was in control of the performance.

In view of your answer to the previous question, can you provide an example from the performance that demonstrated this control relationship?

Both the system and the controller affected the output of visuals and sound. Whenever a new figure was presented, the performer could either use the figure, or choose not to, by moving into a new virtual box.

Survey - The Shapeshifter

Thank
you for attending this performance of The Shapeshifter. We have some questions
for you about your experience of the performance. The survey should take
approximately 15 to 20 minutes to complete. You are free to skip any question. You can withdraw
from answering this survey at any point. All answers are anonymous.

Background

What is your age?

42

What is your gender?

Male

Do you have any formal training in a form of performance involving movement (e.g. dance, theatre)?

- ○ Yes
- ◉ No

If yes, what form(s) of performance?

If yes, for how many years did you receive formal training?

Do you have any formal training in musical performance?

- ◉ Yes
- ○ No

If yes, for how many years did you receive formal training?

10

How often do you attend live dance performance?

- ○ Weekly
- ○ Monthly
- ○ Yearly
- ◉ Less often than yearly

How often do you attend live theatre performance?

- ○ Weekly
- ○ Monthly
- ◉ Yearly
- ○ Less often than yearly

How often do you attend live music performance?

- ○ Weekly
- ◉ Monthly
- ○ Yearly
- ○ Less often than yearly

Do you have any prior experience with motion capture?

- (●) Yes
- ( ) No

If yes, can you briefly describe these prior experiences?

Attending demonstrations of various student projects involving motion capture.

Do you have any prior experience with spatial audio?

- (●) Yes
- ( ) No

If yes, can you briefly descbe these prior experiences?

Student projects and professional sound design project for films

The Performance

Can you briefly recap the performance?

A dance, movement and vocal sounds witch got represented some way on a screen, in various stages. Seemed like a loop that was repeated 5 times, and each time the output from the system and speakers got more intense. First iteration contained no audio from the system, just from the performer, after which sound from the prior loops was output through the system in interplay with the performer's live actions and utterances.

What do you think was the central idea/theme that the performance was conveying?

The name of this form said shapeshifter, so I guess that will give me some bias. I was considering the output on screen to sometimes be very organic and other times more machine-like, as if it was some sort of play/conflict between the two. Also, how individual shapes often was broken up into many small shapes led me to think of nature, cells and swarms.

How many performers took part in the performance?

Hard to answer, but I would say 2.

Can you list the performers?

One human and a system of some kind, consisting of mo-cap, graphics and spatial audio.

What were you primarily focussing on during the performance?

Mostly the screen and the sound, but I looked more at the performer as the performance developed.

How would you describe the relationship between the performer and the visuals?

The screen represented the performers movement in various way, but the major scene changes seemed to be decided by the visuals and not the performer. That way the performer seemed to be led by the visuals as well.

Did this relationship change during the performance?

- ○ Yes
- ◉ No

If yes, can you describe how this relationship changed?

How would you describe the relationship between the performer and the sound?

The speaker output started in the second loop, and seemed to echo the performers sounds from the last loop. This way the performer seemed to build a little choir of voices and sounds. The sounds were however distorted and warped, and I was not able to get an understanding of whether the performers movement controlled or was influenced by the spatial positioning of the sound from the speakers.

Did this relationship change during the performance?

- (selected) Yes
- No

If yes, can you describe how this relationship changed?

It felt as if the sound grew as more audio from various loops were stacked on top of each other.

What parts of the performer's body were used to interact with the system?

The whole body and voice

How clear was the relationship between the performer's motions and what was shown on screen?

Quite clear

How clear was the relationship between the performer’s motions and the sound made by the system?

Not very clear

How clear was the relationship between the performer’s vocalisations and what was shown on screen?

Unclear however as it was repeated it got more familiar

How clear was the relationship between the performer’s vocalisations and the sound made by the system?

Clear

Did you perceive a relationship between the performer’s motion and the spatial placement of the sound?

- ○ Yes
- ◉ No

If yes, can you describe this relationship?

Which statement do you most agree with:

- ◯ The performer was in control of the performance.
- ◯ The system was in control of the performance.
- ◯ The performer and the system were cooperatively in control of the performance.
- ◉ The performer and the system were competing for control of the performance.
- ◯ Neither the performer nor the system was in control of the performance.

In view of your answer to the previous question, can you provide an example from the performance that demonstrated this control relationship?

The performer seemed in control of how their body movement controlled the visuals on screen, and of the various vocal sounds produced. However the major changes in the graphical representation seemed controlled by the system, and also some irratic less logical graphical behaviour seemed to happen out of the performers control, which led to some feeling of conflict, which however was interesting and felt like a part of the performance. Sometimes it felt like the performer had to push harder, like they were struggling with the system in order for the graphical output to change.

Survey - The Shapeshifter

Thank
you for attending this performance of The Shapeshifter. We have some questions
for you about your experience of the performance. The survey should take
approximately 15 to 20 minutes to complete. You are free to skip any question. You can withdraw
from answering this survey at any point. All answers are anonymous.

Background

What is your age?

30

What is your gender?

Male

Do you have any formal training in a form of performance involving movement (e.g. dance, theatre)?

- ○ Yes
- ◉ No

If yes, what form(s) of performance?

If yes, for how many years did you receive formal training?

Do you have any formal training in musical performance?

- ◉ Yes
- ○ No

If yes, for how many years did you receive formal training?

7

How often do you attend live dance performance?

- ○ Weekly
- ○ Monthly
- ○ Yearly
- ◉ Less often than yearly

How often do you attend live theatre performance?

- ○ Weekly
- ○ Monthly
- ○ Yearly
- ◉ Less often than yearly

How often do you attend live music performance?

- ○ Weekly
- ○ Monthly
- ◉ Yearly
- ○ Less often than yearly

Do you have any prior experience with motion capture?

- [x] Yes
- [ ] No

If yes, can you briefly describe these prior experiences?

Creation of animations for virtual characters, and related clean up.

Do you have any prior experience with spatial audio?

- [x] Yes
- [ ] No

If yes, can you briefly descbe these prior experiences?

Creation of spatialized audio stories using spatial plugins with Reaper.

The Performance

Can you briefly recap the performance?

A dance performer was creating sounds and moving around an audience. The sounds made went through about 8 themes or textures, each with their own characteristic. The sounds were recorded and processed to be replayed in the spot the performer was standing when making them, to gradually build a soundscape, enveloping the audience.

What do you think was the central idea/theme that the performance was conveying?

Evolution/growth/change

How many performers took part in the performance?

1, but if each recording counts as a performer I'd say 5 because I think there were 5 tracks running in the end.

Can you list the performers?

What were you primarily focussing on during the performance?

How sounds affected the visuals on the screen and how the looped sounds moved around the speakers.

How would you describe the relationship between the performer and the visuals?

The sounds made by the performer affected the size and shape of figures in the visuals, which moved around the screen to reflect where the performer was in the room.

Did this relationship change during the performance?

- ○ Yes
- ◉ No

If yes, can you describe how this relationship changed?

How would you describe the relationship between the performer and the sound?

The sound made by the performer was recorded and looped, and the performer made new sounds between the gaps of previous loops to gradually make a continuous soundscape.

Did this relationship change during the performance?

- ○ Yes
- ◉ No

If yes, can you describe how this relationship changed?

What parts of the performer's body were used to interact with the system?

Everything, but outer limbs (hands, feet) were mostly in control.

How clear was the relationship between the performer's motions and what was shown on screen?

Very clear.

How clear was the relationship between the performer's motions and the sound made by the system?

Very clear.

How clear was the relationship between the performer's vocalisations and what was shown on screen?

The piece had to loop 3-4 times before I started to see the relationship.

How clear was the relationship between the performer's vocalisations and the sound made by the system?

Very clear once it looped.

Did you perceive a relationship between the performer's motion and the spatial placement of the sound?

- (●) Yes
- ( ) No

If yes, can you describe this relationship?

I believe the sound was played back in the spot which it was recorded.

Which statement do you most agree with:

- (●) The performer was in control of the performance.
- ( ) The system was in control of the performance.
- ( ) The performer and the system were cooperatively in control of the performance.
- ( ) The performer and the system were competing for control of the performance.
- ( ) Neither the performer nor the system was in control of the performance.

In view of your answer to the previous question, can you provide an example from the performance that demonstrated this control relationship?

From the very beginning, the performer laid the foundations for the rest of the piece, since their movements and sounds were the building blocks for both visuals and the sound. While the system was a part of the performance, the dancer was in control of what the system produced.

Survey - The Shapeshifter

Thank
you for attending this performance of The Shapeshifter. We have some questions
for you about your experience of the performance. The survey should take
approximately 15 to 20 minutes to complete. You are free to skip any question. You can withdraw
from answering this survey at any point. All answers are anonymous.

Background

What is your age?

31

What is your gender?

Female

Do you have any formal training in a form of performance involving movement (e.g. dance, theatre)?

- ○ Yes
- ◉ No

If yes, what form(s) of performance?

If yes, for how many years did you receive formal training?

Do you have any formal training in musical performance?

- ○ Yes
- ◉ No

If yes, for how many years did you receive formal training?

How often do you attend live dance performance?

- ○ Weekly
- ○ Monthly
- ◉ Yearly
- ○ Less often than yearly

How often do you attend live theatre performance?

- ○ Weekly
- ○ Monthly
- ◉ Yearly
- ○ Less often than yearly

How often do you attend live music performance?

- ○ Weekly
- ○ Monthly
- ○ Yearly
- ◉ Less often than yearly

Do you have any prior experience with motion capture?

- ○ Yes
- ◉ No

If yes, can you briefly describe these prior experiences?

Do you have any prior experience with spatial audio?

- ○ Yes
- ◉ No

If yes, can you briefly descbe these prior experiences?

The Performance

Can you briefly recap the performance?

The performance was carried out through a technological interaction system where image, movement and sound influenced the development of the performance. A dancer carried out the performance. There were various visual "moving" scenarios that were shown on a screen om alternately several times and the dancer both influenced how these scenarios moved through her movements and adapted her movements to the various scenarios that appeared on the screen along the way. In addition, the dancer made customized sounds based on each scenario and after a few rounds the sounds the dancer made were integrated into the system and loudspeakers replicated bits of these sounds when the same scenario appeared again. This led to a gradually an increasing amount of sound during the performance.

What do you think was the central idea/theme that the performance was conveying?

The different scenarios created different environments in the room which the dancer felt, adapted to and made adapted sounds for. I think that what is presented in this performance is the way digital stimulation both limits but also inspires a dancer to bodily express themselves in specific directions.

How many performers took part in the performance?

One

Can you list the performers?

There was just one dancer, even though the audience were placed in the middle of the room where the dancer moved around.

What were you primarily focussing on during the performance?

Mostly the screen, but also the way the dancer moved and looked at the screen and adapted her expressions to the digital stimulation.

How would you describe the relationship between the performer and the visuals?

As described above, it was experienced as
a mutual adaptation between both elements. During the performance I got to think about the experience of dancing alone vs. dancing with a dance partner as an analogy to understand the mutual adaptation happening.

Did this relationship change during the performance?

- ○ Yes
- ◉ No

If yes, can you describe how this relationship changed?

How would you describe the relationship between the performer and the sound?

The dancer made inspired sounds based on the scenarios shown. These were the same each time the same scenario was shown, but the digital system recorded and duplicated the sounds for each time the same scenario was shown. This led to an eco-like sound experience that developed the further into the performance one was.

Did this relationship change during the performance?

- (•) Yes
- ( ) No

If yes, can you describe how this relationship changed?

More eco-like sound experience the further into the performance. More layers to the same type of sounds. However, the system didn't seem to duplicate the exact same sound, if felt slightly twisted.

What parts of the performer's body were used to interact with the system?

All of her body seamed to get used during the performance. The performer used large movements during the performance, which alternated between hard and fast and slower and softer movements.

How clear was the relationship between the performer’s motions and what was shown on screen?

There was a clear relationship. Moreover, not all scenarios showed an equally clear view of the body parts of the performer. While in some scenarios it was possible to see a correlation between the visuals and the body parts, in other scenarios this relationship didn’t exist. Here these visuals responded to the dancer's larger or smaller movements but not specific body parts.

How clear was the relationship between the performer’s motions and the sound made by the system?

I think the performer got slightly influenced by the sound made by the system. I observed at some points that when a new sound appeared, she took a small break in her movements before she started to move again.

How clear was the relationship between the performer’s vocalisations and what was shown on screen?

I could see how she chose to use the specific sounds to the specific visuals, but I didn't notice any relationship between the visuals in the various scenarios and the sound she made.

How clear was the relationship between the performer’s vocalisations and the sound made by the system?

As mentioned above, it was experienced as a duplication with a small twist. The twist was perceived as the length of the audio being recorded was cut and therefore became shorter when the speakers produced the sound that got recorded.

Did you perceive a relationship between the performer’s motion and the spatial placement of the sound?

- (•) Yes
- ( ) No

If yes, can you describe this relationship?

I experienced the sound to come from more than one place. However, I didn't notice any relation between the performers placement and where the sound came from.

Which statement do you most agree with:

- ○ The performer was in control of the performance.
- ○ The system was in control of the performance.
- ◉ The performer and the system were cooperatively in control of the performance.
- ○ The performer and the system were competing for control of the performance.
- ○ Neither the performer nor the system was in control of the performance.

In view of your answer to the previous question, can you provide an example from the performance that demonstrated this control relationship?

They influenced each other. However I would say that the system "framed" the performance to a larger degree, but that the dancer worked for being in control of the performance "within" these frames.