Motion Pointillism: The (Re/De)Construction of the Normative Body through Motion Capture

–– Hugh Alexander von Arnim
(UNIVERSITY OF OSLO)
ORCID: 0009-0005-7511-9019

–– Tejaswinee Kelkar
(UNIVERSITY OF OSLO)
ORCID: 0000-0002-2324-6850

–– Live Noven (INDEPENDENT ARTIST)

Marker-based, optical motion capture systems make use of reflective markers, interpreting them as clusters of dimensionless points in space. Before labeling and arranging these markers, potentially to fit a model of a kinematic chain, these markers possess little referentiality to objects in physical space. However, the construction of a kinematic model of the human body requires making several assumptions about the body and its affordances. In this article, we problematize the use of the kinematic model in dance performance that employs motion capture, placing focus on the referentiality of visual representations derived from markers and models while examining how motion capture contributes to the construction of the body through the embedding of assumptions and values about what a body is and can do within the technology. Through the design and conceptualization of two interactive dance performances titled *Reconfigurations* and *The Shapeshifter,* we develop an approach to working with motion capture that we term *motion pointillism*, which aims to resist the systemic assumptions embedded in the modeling process. This approach conceptualizes the emergence of the dimensionless points' referentiality to a human body as a collaborative component of system development and performance, which occurs both in the design of visual representations as well as in the viewers' perception.

Keywords: Motion Capture, Modelling, Body, Dance, Normative

Over recent decades, motion capture (hereafter also mocap) has become popular in contexts such as stage-setting, popular music concerts and film. The technology's use in dance and various types of staging has been relevant for both the artistic as well as the popular stage. In this article, we problematize several assumptions made by these technologies about bodies, and how potentials for dance exploration and creative possibilities are co-created by these technologies. We present two of our own dance works that employ optical, marker-based mocap. These works were developed through our research-creation practices, with each consisting of an interactive system used in a performance. Through these projects, we developed an approach towards the use of marker-based mocap that we term *motion pointillism*. We contextualize this approach in relation to works that employ the technology as a method to model the skeletal structure and kinematics of the body.

One of the most recent large-scale appearances of mocap in popular media is connected to the band ABBA. In 2022, ABBA began their first concert series in 42 years in support of their new album *Voyage*. The show, which is currently scheduled to last until 2025, consists of the members of ABBA performing their greatest hits along with accompanying dance routines in a dedicated arena in the Queen

Elizabeth Olympic Park. However, it is not the members of ABBA who are physically performing the show. Instead, animations of the band as they looked in the 1970s perform a routine captured from the band[1] using an optical, marker-based mocap system. Despite some issues relating to the use of the technology as well as several initial technical setbacks, the show was received positively by audiences and critics, and has sold over one million tickets to date (Matthews and Nairn 298).

The show is among the most high-profile productions at the intersection of marker-based mocap and choreography, using a form of mocap in live performance that had previously mostly been associated with film and video games. However, dance works involving this technology date back to the 1990s, while the technology itself was formed through its initial application for medical biomechanical analysis in the 1980s (Downie 306). In the following section, we introduce some fundamental characteristics of optical, marker-based mocap and describe how the technology functions.

Optical Marker-Based Motion Capture

The term motion capture covers a wide range of technologies and techniques, with Kristian Nymoen defining motion capture at a fundamental level as "the use of a sensing technology to track and store movement" (13).[2] The technique central to our work and this article is "optical marker-based motion capture", which 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 emitted 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 can be calculated through triangulation. The position of the marker takes the form of a three-dimensional point within a Cartesian coordinate system, expressed as distance relative to a user-defined origin representing a 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 or hidden from the cameras' view in a way that makes it visible to fewer than two cameras, the marker can no longer be registered by the system. This means that the physical properties of the capture volume must be considered to ensure that the cameras provide adequate coverage. 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 that is against the floor may not 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. Thirdly, any object that reflects or emits IR light will be registered as a marker by the system. In locations that contain many of either of these (for example spaces exposed to sunlight), this can add a lot of noise to a capture. This noise can be in the form of missing data or hallucinated data points (usually called ghost markers), or marker jumping, for example.

When passive markers are used, individual markers do not possess a distinct identifier. In the case of non-real-time use, once a recording has been completed markers can be labeled and any gaps in the capture can be filled. Depending upon the amount of noise in the capture, this can be quite a long and arduous process. If it is important that individual markers are consistently identifiable during the capture session (for example, in interactive performances that map a specific marker to a specific parameter), a method must be developed with which its identity can be preserved across any gaps. In a sense, all passive markers are calibrated and occasionally imagined by the system (Karreman).

A defining property of an optical, marker-based system is that the position of each marker is projected within a coordinate system.[3] This means that the position of each marker is not calculated 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 result, there is no inherent relationship recognized 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 by a method chosen by the system designer or 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 modeling of the human skeleton as a kinematic chain.

Rigid Bodies and Kinematic Modeling

Mocap has used an approximation of the human body in the skeletal form since its inception. These types of representation rely on bone-based approximations of the body, from a surface level of capture. Since markers are dimensionless points within the capture volume, a few steps need to be performed to obtain higher-order properties. To obtain the spatial dimensions and rotation 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.

To model more complex objects, a series of rigid bodies can be joined together to form a *kinematic chain*, which mathematically represents this series as connected to one another by joints that have predetermined degrees of rotational and transformational freedom. These are often organized hierarchically, with one rigid body serving as the root to which all other rigid bodies are chained. The modeling of kinematic chains can quickly become quite complex. A full discussion of this is beyond the scope of this article, however, and we refer to Müller for a more thorough description of this process in relation to the human body.[4]

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 modeled as consisting of either two or three connected rigid bodies in contrast to the 33 vertebrae commonly found in the human spine.

This representation of the body in its skeletal form is used by most markerless mocap technologies as well. Among others, pose estimation algorithms such as OpenPose use one approximation of the skeleton derived through training a machine learning model with a dataset of images of the human body with the labeled position of 135 keypoints (Martinez et al. 5), and the Microsoft Kinect skeleton tracker uses a skeletal model defined through the position and orientation of twenty joints extending from the hips in relation to the position of the device itself (Le et al. 341). It is hard to trace the exact origins of the skeletal structure that remains in a variety of these algorithms, but for example, the idea of approximating the human body through mocap in the form of a skeleton has been in the Vicon system since the inception of their software tools in 1979 (Vicon).

Motion Capture in Dance Performance: A Review of Works

With their focus on human body motion, mocap technologies may seem to be an inviting prospect to choreographers who work with multimedia performance. For this article, we are limiting our scope to live dance performances and artistic installations. Within this context, there are two main styles of working with mocap that emerge. The first involves mocap occurring prior to a performance, with the captured data then played back during the performance. The second is to use the mocap system live in a real-time interactive system. The latter, however, is relatively infrequent. Bevilacqua et al. emphasize the relative scarcity of the use of marker-based mocap in real-time, interactive dance systems, attributing this to the complexity of handling such systems and the requirement of the performer to wear markers during the performance. To this, we would also add that an optical, marker-based mocap system is quite expensive, with no companies making consumer-grade options, and has quite a high barrier of entry in terms of technical knowledge required for operating the system. Due to this, many of these works involve either a link to academic research or the involvement of a private company that specializes in animation.

The latter is the case for two seminal works of the late 1990s. The Riverbed group, formed by Paul Kaiser and Shelley Eshkar, was responsible for several of the earliest forays into the integration of mocap with dance (Dixon). With the group, Michael Girard and Susan Amkraut developed a kinematic modeling software named Biped, which formed the basis of Riverbed's collaborations with a number of prominent and influential dancers and choreographers. This software formed the basis of their first collaboration with Merce Cunningham, an animated installation named *Hand-drawn Spaces*. Based upon this successful collaboration, Cunningham proposed to use the mocap technology in a work that also involved live dance performers on stage. The result of this was a work titled *BIPED* in 1999, named after the modeling software that had been developed by Riverbed. The work featured dancers on stage, accompanied by projections of animated captures showcasing two or three dancers[5] executing a series of Cunningham's movement sequences onto a scrim. As reported by Abouaf, the process used to create the work involved a single afternoon of mocap recording with the dancers ("'Biped': A Dance with Virtual and Company Dancers 1" 1). After processing the captured data, kinematic models were created that 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" ("'Biped': A Dance with Virtual and Company Dancers 2" 5).

The second method was the creation of a 3D model by mapping a spline curve 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 modeling of kinematic effects, such as skin and tendon behavior, and even "foot to ground collision response" (188). This method of mapping a kinematic model to a 3D-animated model has proved influential to proceeding developments of dance work involving the use of optical, marker-based mocap, with Dixon, in reference to an image of dancers in front of one of the animated figures featured in *BIPED*, noting that "*BIPED* images such as these have been so admired and reproduced that they have become archetypical of the digital dance and performance movement" (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. 1). 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 modeling of the skin and muscle behavior (Baumgartner). After the premiere of the work as an installation at The Cooper Union in New York, the piece was later incorporated as a part Jones' *Breathing Show* tour. It sees multitudes of 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 including a re-envisioning of *Ghostcatching* in 2010 as *After Ghostcatching* (Barber); several works undertaken by Marc Downie including collaborations with Merce Cunningham and Trisha Brown for which he developed an agential approach towards kinematic modeling of the skeleton from marker positions; Vincs and McCormick's use of the model to drive representation outwards from the body of the dancer in a stereoscopic projection with a real-time system; Satore Studio's work on *HÁITA*, which incorporates mocap along within a wider ecosystem of sensing systems to capture a wide variety of dance styles; and Dan Strutt's 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.

There are, of course, several marker-based dance works that do not attempt to kinematically model the skeleton with the mocap system. For example, the work *Lucidity* (James et al.) uses a custom-built tracking engine to trace the position of a dancer as a cloud of points from which higher-level features such as dancer proximity and groupings are extracted, with limb motion modeled through statistical methods relating to the point cloud.

Our Explorations: *Reconfigurations* and *The Shapeshifter*

During our research-creation practice, focused on the development of interactive systems and performances that foreground human body motion, we developed two dance projects centered around the use of optical, marker-based mocap. Both projects originated from a desire to explore marker-based mocap as a technology and its relationship to the human body in a physical environment as well as in its virtual representation. The projects form a continuity, sharing fundamental conceptual ideas, as well as developed systems. The first, *Reconfigurations*, was developed over the first half of 2022 and served as an exploratory probe into ways of working with the technology, as well as a testing ground for our core systems. Following this, in the summer of 2022, we began to work on *The Shapeshifter*, with which we aimed to expand upon work done in *Reconfigurations*.

The technological components of both performances are built upon the top of an OptiTrack mocap system permanently installed at our institution. However, instead of making use of the modeling and processing techniques of the OptiTrack software, named Motive, we employed the software as a simple throughput to stream the position of the markers in the capture volume onward to our own software. To date, performances have only taken place at our institution and so the presentation of the works in this article refer only to these performances. However, we are also in possession of a portable OptiTrack system and are in the process of working to transport our work to other venues. In the following sections, we will first discuss our high-level design motivations, followed by a discussion of each work, detailing their origins and the ideas that we wished to explore and providing a brief overview of the performances and the systems that we developed for each.

Motivations for Design

Our motivations for these works originated from a desire to interrogate how the body was represented by the OptiTrack system. Having previously worked with this system and other optical systems for work on motion analysis projects, we started to think about how the mocap software visually presented the human form as a construct of the motion data that was captured. We likewise started to recognize several limitations that mocap systems imposed upon how the body could be represented, owing to both software design as well as hardware capabilities. In view of this, we considered the forms of data that can be acquired by the mocap system and how these are presented visually within the software.

The OptiTrack system that we use as a base mocap system fundamentally works with the position of markers captured by the system. Motive presents the position of these markers visually as small, colored spheres within the 3D capture volume. Beyond this, Motive allows the organization of a collection of markers as either a rigid body or a kinematically modeled skeleton. The skeleton can be represented as one selection of avatars such as a mannequin which encompasses the markers. Alternatively, the skeleton can be represented by a series of sticks, which join the spheres of the markers along the path of the skeleton's bones. A rigid body is similarly represented, with a series of sticks demarcating the boundary of the object.

While this representation of the kinematic model is recognizable 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 by joining these points with a series of predetermined connecting lines. Importantly, these connecting lines are also fixed in terms of properties relative to the body that they are portraying, such as their length and the points which they connect, freezing it in this uncanny form. We began to consider how we could unfreeze this form and remove the constraints imposed by the modeling process. What if the markers were not assumed to be in fixed positions on the body so that the modeling 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 conceptualize a single human body as the boundary of the modeling process, enabling the encompassing of inanimate objects, parts of the environment, and even a second body as part of the construction of the form?

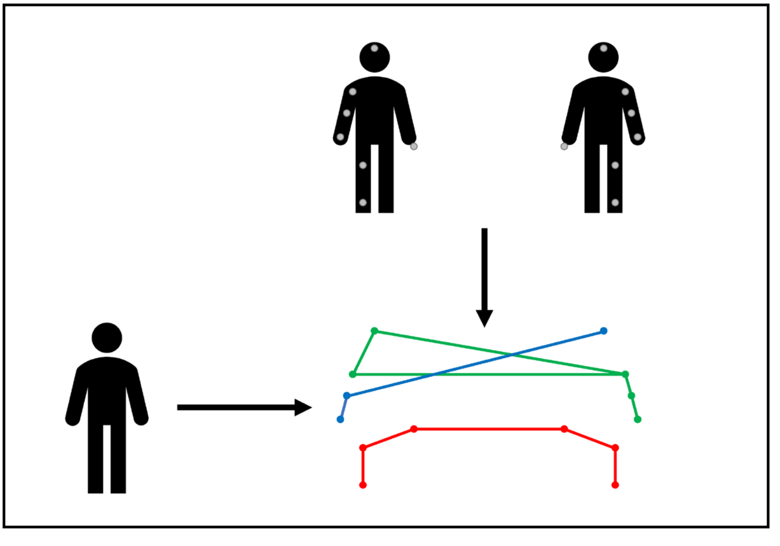
Unfortunately, the Motive software is quite obstinate with its modeling process, being especially inflexible when it comes to reconfiguring these in real-time. There is no manual way to create connections between markers. Rigid bodies, once defined, stop being tracked if a marker moves outside of the 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 which consequently fixes the parts of the body to be modeled. 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. These thoughts motivated the design of *Reconfigurations*, the first of our explorations.

Reconfigurations

In *Reconfigurations*, we employed Motive to stream marker positions to our own software with the aim of enabling malleability in the formation of marker constellations. Using the programming language Python, we in effect recreated a simplified version of the Motive display. Each marker is represented in a pseudo-3D virtual capture volume. Depth is simulated by altering the size of each circle. However, we added an element of interactivity, which functions as a reconfigurable "modeling" process. Markers can be assigned to a "body" on the fly, with each body represented by a different color. Markers belonging to the same body can be joined together with a line representing a bone by clicking on one marker and drawing it 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 recognized 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 (von Arnim, “Reconfigurations”). Two mocap performers improvise dance phrases within a performance area, employing up to thirty markers[6] which they are free to 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 incorporate inanimate objects to which they attach a marker. Instead of being pre-defined at the outset of a performance, the form 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

Figure 1: The two dancers position the markers in the performance space, either by 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 (2023), © Hugh Alexander von Arnim



created a sound synthesis engine to which motion parameters are mapped to generate musical material across a performance.

A performance was held in May 2022 as part of a concert of telematic music, that is, musical performance over a network connection. For this, the performers were in the laboratory where the OptiTrack system is installed and the virtual representation, a camera feed of the performers in the physical space, and the audio of the sound generated by the system were streamed over a network to a second location. There were audience members physically present in both locations.

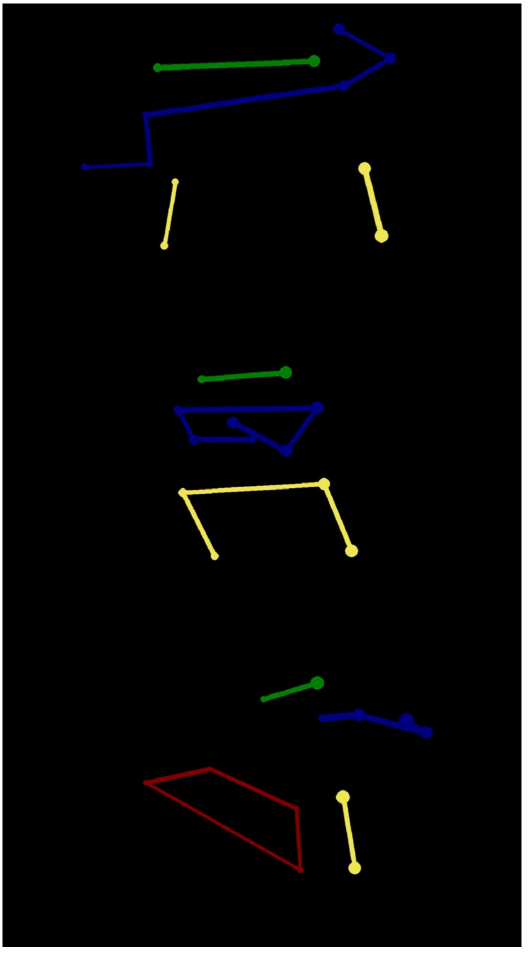


Figure 2: Three stills of the visualization display from a performance of Reconfigurations. Each dancer was wearing markers along one side of their body divided across the lateral plane, resembling the positioning of markers required for full-body kinematic modeling on one half of the body. Initially, this was used to create a combined body for the dancer, with the dancers coordinating their motions 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 (2023), © Hugh Alexander von Arnim

The Shapeshifter

Aiming to build on the work done for *Reconfigurations*, in the summer of 2022 we began work on a follow-up project: *The Shapeshifter*. Our first goal was to create a system that presented a full 3D rendering of the virtual capture volume instead of the pseudo-3D rendering used for *Reconfigurations*. This would also allow us to have greater control over parameters such as camera placement and the projection of the data representing three-dimensional coordinates onto a two-dimensional screen. To this end, along with the Python components that we developed for *Reconfigurations*, we integrated an additional system component built in the Max/MSP/Jitter programming environment, as this enables a simplified rendering process based upon the Open Graphics Library (OpenGL). Moreover, as the programming language is primarily used for music and sound applications, we could easily integrate the mocap data with audio processing to expand the sonic components of the work (von Arnim, "The Shapeshifter").

With this follow-up project, we aimed to answer several questions that arose over the course of our work on *Reconfigurations*. Although we had moved away from using a pre-defined kinematic model of the skeleton, our work on *Reconfigurations* was still influenced by the visual grammar of the modeling process. Markers were represented by spheres and were connected by lines representing "bones". But what if we moved away from this component of that visual grammar? We would still work within a virtual recreation of the capture volume, but what if we displayed the markers and connections in a variety of manners? How would these influence where the dancer positioned the markers, and the bodies that they attempted to configure? It was at this point that the third author, who is a practicing dancer and physical theater performer, joined the process to collaborate on the work and create continuity through the development process.

The Shapeshifter is an improvisatory dance work for a single dancer. Prior to the performance, the dancer positions up to thirty markers[7] 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 dancer improvises a motion pattern and accompanying vocalizations. To trigger the end of a phase, each of the markers must be located within a corresponding space in the physical performance area. Each phase presents a different visualization style both for the virtual representation of the marker and any connections drawn between markers. At the end of the nine phases, the cycle begins again. During the second cycle of the phases, the performer's vocalizations for each phase from the previous cycle are looped within the corresponding phase in the current cycle. Starting in the third cycle of the motion phases, the representations of the markers and connections begin to shift, interpolating between the representations of all nine phases. The interpolation is based upon several factors, relating to the similarity of the performer's motion and vocalizations to the motion patterns and vocalizations performed in the previous cycles, with both the similarity measures influencing the amount and direction of the interpolation. The looped vocalizations begin to twist and distort away from the original recordings. As the number of cycles increases, it becomes difficult for the dancer to purposefully control the representations, building to a climax in the seventh and final cycle of the nine motion patterns.

A performance takes place with a similar setup to *Reconfigurations*, with the performer facing a video wall that mirrors the physical capture volume with a virtual capture volume. The audience is also positioned within the performance space. We took a similar approach to marker occlusions as we did in *Reconfigurations*, building these "errors" into the functionality of the system so that having the performer move around the audience becomes a part of the performance. Aiming to envelop the audience within the performance space, as well as provide audible "traces" of the performer's motion through the space, the looped vocalizations 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 vocalization at the point of the centroid of all markers at the time that the vocalization was recorded. Both have different processing applied and develop in different manners as the performance progresses.

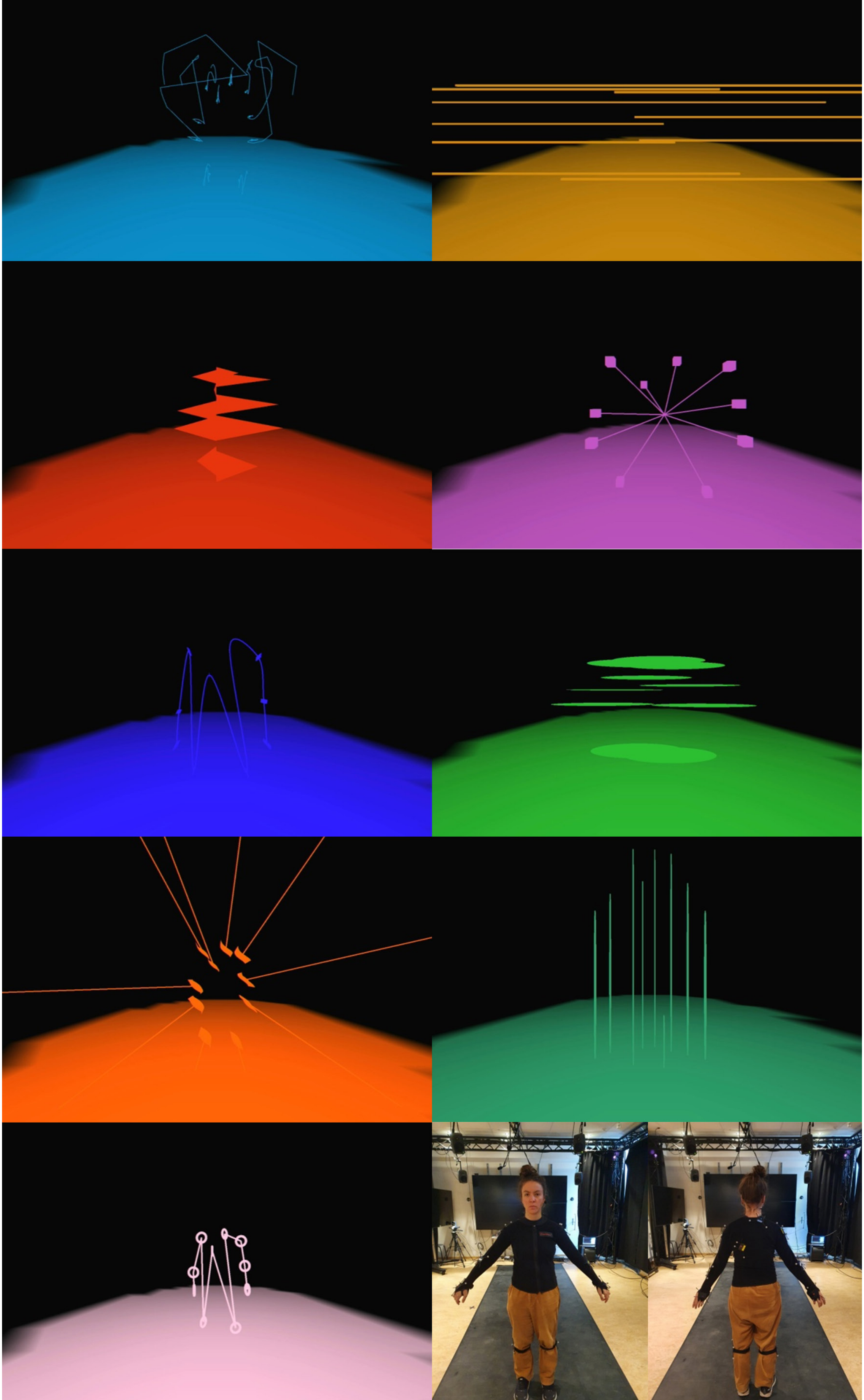
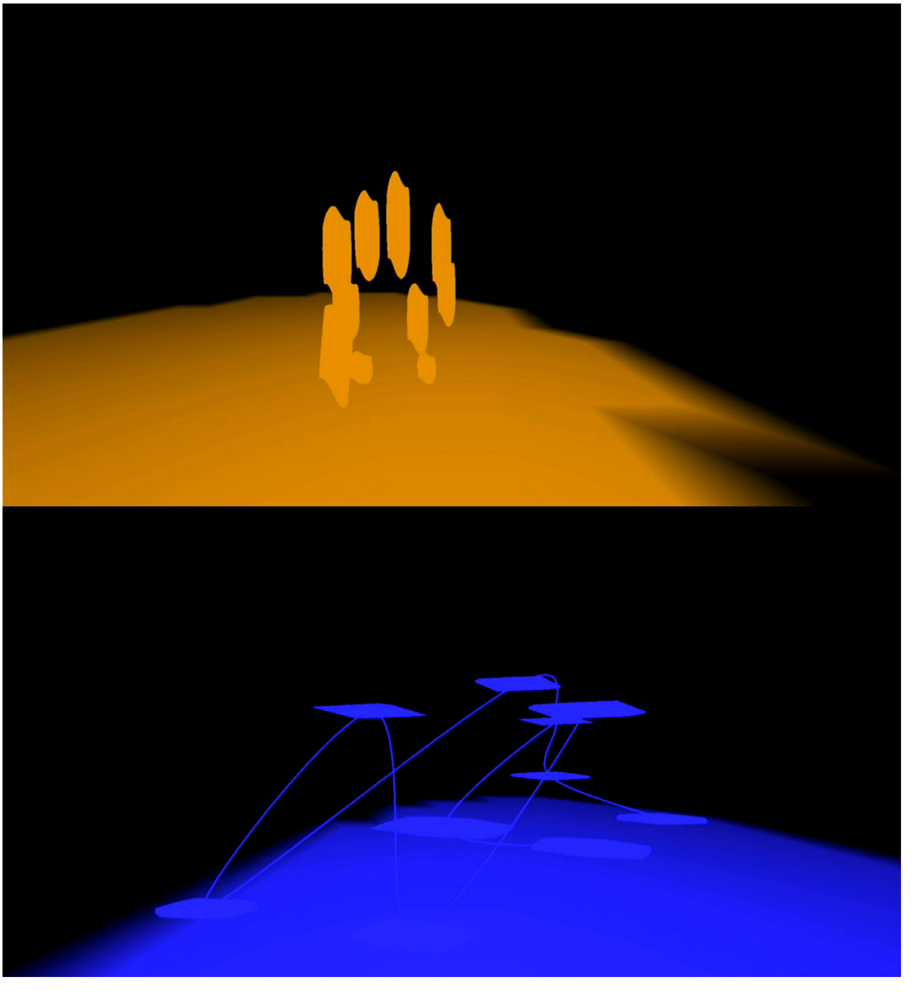


Figure 3: The initial visualizations of each of the nine motion phases for The Shapeshifter. These comprise a distinct style for representing each marker, sometimes including connections to other markers or points in space. These connections are not just calculated spatially, for instance the first phase 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 third author with the locations of the markers, all of which were positioned on her body (2023), © Hugh Alexander von Arnim

Figure 4: Two examples of the representations shifting to interpolate between the initial visualization styles. This is based upon a combination of the motion patterns and vocalizations performed by the dancer 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 (2023), © Hugh Alexander von Arnim



Marker Referentiality and Co-Construction of the Body

These two performances illustrate an attempt to move beyond the technology's origins as intended for biomechanical analysis with respect to the ways of working encouraged by the system and the methods through which a visual representation of the body is constructed from motion data, especially in view of the focus that is placed upon the modeling process by most producers of the technology. And although suitable for many applications, kinematic modeling as the primary method through which to construct the body results in a representation for which the form is already to an extent determined. There is a strain of thought that is repeatedly found in the discourse around mocap centered upon kinematic modeling for the performing arts that elides the motion of the dancer with the motion of the animated figure. Such discourse is present, for example, in the artists' statement for *Ghostcatching*, where Jones, Kaiser and Eshkar write that "the body of Jones is multiplied into many dancers" (1), or in comments made by ABBA's *Voyage* show producer, Ludvig Andersson, that "it is not a version of, or a copy of, or four people pretending to be ABBA. It is actually them" (ABBA Voyage).

Statements such as these point towards a situation where the collection of markers with which the dancer is outfitted are viewed as *invisible mediators* between the dancer's motion and the virtual representation. They move towards standing in for the body parts to which they are attached, referring the data that they collect not to their own motion but to the body of the dancer.

To underline this point, we can draw an analogy to another technology used in artistic performance: the microphone and loudspeaker. In her history of these technologies as musical instruments, Cathy van Eck outlines several approaches towards their usage. Building upon work on mediation technologies by Jonathan Sterne (*The Audible Past*), van Eck outlines that the thought underlying two of her approaches, reproducing and supporting, is that in their idealized form the technologies underlying the reproduction or support of sound are transparent. She details the train of thought underlying this concept, writing that.

> When I speak about this transparency in the reproducing and supporting approach, I must underline that I am not referring to technical possibilities, but about how the technology is perceived or even the cultural consensus of how it should be perceived, which means, in this case, that the technology should not be perceived at all. The music should sound as if produced by a human body interacting with a musical instrument, not with technology. … hearing a singer amplified through microphones, amplifier and loudspeaker rarely results in the audience perceiving a musical instrument consisting of singer, microphone, amplifier and loudspeaker. The main perception will remain that of somebody singing, whatever other technology is added to the voice. … The sound produced is affected by a combination of all of these [technological] elements, but the semantic acts of sound creation are associated with only the singer. (41–42)

Framed in these terms, we can view the mocap system as a sensing system which aims to either reproduce (in the case that the capture takes place prior to performance) or support (in the case of a real-time interactive system) the semantic acts of motion that are associated with the dancer. The markers and camera system, following the same logic, should remain transparent, leaving as little mark on the production as possible.

As noted by Naccarato and MacCallum, treating a sensing technology as transparent can lead to troubling implications. Framed through a discussion of the appropriation of medical sensing technologies for artistic purposes, they enter into dialogue with Sterne's discussion of the stethoscope and his claim that during mediated sensing the sensor apparatus must be "erased from consciousness" ("Mediate Auscultation" 123) and that as a consequence "the tool stands in for a whole process from which it erases itself" ("Mediate Auscultation" 123). 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 the 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" (6) 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. (6)

Naccarato and MacCallum's discussion of the propagation of ethical and aesthetic values points towards a broader discussion of what a human body is, and how discourse forms around what it can do and be. Here, a pertinent concept is the normate. In her book *Extraordinary Bodies*, Rosemarie Garland-Thomson examines social and cultural representations of bodies marked as disabled, contending that disability is attributed through socially layered exclusionary discourse, "not so much a property of bodies as a product of cultural rules about what bodies should be or do" (6). To aid her analysis, 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" (8). Joel Michael Reynolds further explicates the normate in terms of ability, describing the figure as "the tain of the mirror of ableism … the invisible mechanism that allows slippage from being to being-able" (244). He also notes that the specter of the normate haunts not only those designated abnormal 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" (244). Reynolds folds these social forces back into the proprioceptive and kinaesthetic sense of the body, 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" (255).

Returning to the kinematic model, we can then begin to see the mechanisms that are already in place to begin imbuing the model of the dancer with ethical and aesthetic values long before the point of its use to drive the motion of an animation. To design a kinematic model, two difficult questions must be answered: 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 generalizable, that is that they can be used by more than one person, and that they represent a simplified model of human kinematics.

For kinematic modeling 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 generalizability, 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 modeled 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 modeling process will not function correctly, or in some cases, is incapable of functioning at all.[8] 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 the figure of the normate. To take one example, the spine is often modeled as two to four rigid bodies connected by joints with either one or two degrees of rotational freedom. These rigid bodies are often modeled 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.). Although models have been developed that can reproduce the spine in more detail,[9] 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.

As a cumulative effect of these factors, assumptions start to form about the types of bodies for which kinematic modeling is intended. This is namely a body that possesses all of the body parts required by the model, can reach poses and perform motion patterns that are recognizable to the model, and have a body whose nuances of form and motion can be represented through the simplifications required. This is not to say that bodies that do not meet these requirements cannot be modeled, 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 that does fit the mold for the generalized 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.

Over the course of the development of *Reconfigurations* and *The Shapeshifter*, we began to work towards an alternative conceptual approach towards representing the body in dance work involving marker-based mocap. We conceptualize this approach as an alternative to kinematic modeling when representing the human body with creative application of marker-based mocap. We view this approach as an attempt to shift the emergence of a visual representation that refers to the physical body of the performer away from a concrete, bounded, and normatively-inclined process embedded within the mocap system's software10 and towards a process wherein that emergence occurs during the development and performance of an artistic work. Instead of fitting the motion data to a model of the body, we allow the motion data to play a role in shaping the body that emerges.

Motion Pointillism

We can start by returning to consider what we are capturing when working with a marker-based system, namely dimensionless points in space. Within capturing software, these are commonly represented with small, unconnected spheres of identical color. On captures of people wearing full-body marker sets that are required for kinematic modeling, it is possible to recognize a human form from these representations before any modeling has taken place. This recognition can even take place at quite a personal level. For example, Jeannie Steele, one of the dancers who performed the mocap 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 data before any modeling had occurred (Abouaf, "'Biped': A Dance with Virtual and Company Dancers" 2). However, although a human form can be recognizable, if markers are removed from the representation one by one, slowly the form that was visible dissolves. The work done in shaping the motion capture data into a representation of the human form is being performed in part by the viewer.

This thought stands at the center of our approach. Instead of conceptualizing points as referring to the motion captured from a human body to which we can apply a pre-defined model to drive an animation, with all the assumptions that brings, we instead envision these points as referring to the point within the physical capture volume at which they 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 colors 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 system designer, as well as the mocap system itself.

This idea of playing with the referentiality of the visual representation of the markers can be linked to work done by researchers across several domains. Rebecca Solnit posits paths of walking as traces of motion in which abstraction "dematerializes" bodies and motion (29). Laura Karreman approaches a similar idea with the "motion capture imaginary" (245), an investigation of the discourse on the transmission of dance knowledge mediated through technologies which "take the performer's body as their main point of reference" (8) and yet which "present the absence of the dancing body" (253).

We intend our approach, however, to function on a practical level, and have therefore formulated *motion pointillism* as five guidelines that can be applied to future artistic work employing a mocap system:

1. **The mocap system should be acknowledged and the data it provides taken literally**

Everyone involved in a performance should be aware of the fact that any individual visual representation of a marker refers to a specific point in space. For example, the visualization of a marker placed upon a performer's hand refers to the motion of the marker and should not be considered to refer to the hand itself directly, neither by system designers, performers, nor audiences. To encourage and emphasize this, the mocap system should be as opaque as possible. This means that no attempts should be made to conceal aspects of the mocap system, neither markers nor cameras. This is connected to:

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 modeling 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 hold it in a fixed position.

3. **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 translational transforms applied to them, and be represented by any object or series of objects. What is important is that any work done with the motion data does not assume that this motion originates from or refers to any specific source within the physical capture volume apart from the markers themselves.

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.

5. The performer must be able to see the forms

In many multimedia works involving mocap in real-time, the visual representation created from the motion data is 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 the performer is also able to take part in this process and view the configurations that extend from their participation.

Conclusion

With *Reconfigurations* and *The Shapeshifter* we aim to raise critical questions that relate to the appropriation of technology developed for non-artistic purposes as an artistic method. These two works present an attempt to design a built-in way of reshaping, or remolding these systems to explore and demonstrate the boundaries created by their intended use and as a matter of artistic material we hope that these works help to spotlight the constraints by playing with the system's limits. Specifically, with these works we intend to highlight the complex interplay between the referentiality of the motion of markers, their visual representation, and the human body when employing marker-based mocap for artistic purposes. To this end, the work that we present here aligns with broader research on the phenomenological nature of the body, as well as mocap as a "legitimate source of knowledge" about the body (Karreman 99). With the formalization of our theoretical contribution of the five guidelines of *motion pointillism*, we hope to provide a framework that can be employed by other researchers and artists who work within this space.

We have framed this article from our perspective as system designers and performers, extending *motion pointillism* from an examination of the functionality of optical, marker-based mocap systems and our attempts to work around the assumptions, values, and limitations embedded within the technology. However, the work that we present in this article is situated within a wider discourse, in which visual representations referring to the body in performance, and the systems employed for deriving these representations, both shape and are shaped by wider societal views on what constitutes a body. Moving forward, we intend to turn our perspective to this wider discourse and examine how the material and perceptual representations of the human body that arise through our work relate to emerging theoretical bodies.

Finally, it is also important to note that we are presenting *motion pointillism* as one of a range of possible approaches towards modeling the body within creative applications of marker-based mocap. We do not mean to imply that modeling the skeleton as a kinematic chain cannot result in fruitful results or exemplary works of art. However, it is by far the most common approach to working with mocap, and the approach towards which most mocap software is nowadays oriented. Through the requirement that markers explicitly refer to the body to which they are attached, it is also an approach heavy-laden with assumptions about the bodies that can use the technology, consequently also contributing to the construction of those bodies that it claims to transparently represent.

Open-Source Contributions

Both *Reconfigurations* and *The Shapeshifter* can be accessed and used as open-source code repositories (von Arnim, “The Shapeshifter”; von Arnim, “Reconfigurations”). We welcome anyone interested in working with motion capture and dance to engage with the works and use them in their own artistic and research practice.

Works Cited

ABBA Voyage. "ABBA Voyage: How ABBA Used Motion Capture to Create Their Avatars." *Facebook*, 13 Oct. 2021, www.facebook.com/ABBAVoyage/videos/abba-voyage-how-abba-used-motion-capture-to-create-their-avatars/399264848445591/. Accessed 7 July 2023.

Abouaf, Jeffrey. "'Biped': A Dance with Virtual and Company Dancers. 1." *IEEE MultiMedia*, vol. 6, no. 3, 1999, pp. 4–7, doi.org/10.1109/93.790605.

---. "'Biped': A Dance with Virtual and Company Dancers. 2." *IEEE MultiMedia*, vol. 6, no. 4, 1999, pp. 5–7, doi.org/10.1109/93.809227.

Barber, Tiffany E. "Ghostcatching and After Ghostcatching, Dances in the Dark." *Dance Research Journal*, vol. 47, no. 1, 2015, pp. 45–67, doi.org/10.1017/S0149767715000030.

Baumgartner, Henry. "How to Catch a Ghost." *Mechanical Engineering-CIME*, vol. 121, no. 4, 1999, p. 108. *Gale Academic OneFile*, link.gale.com/apps/doc/A54482570/AONE?u=googlescholar&sid=googleScholar&xid=24012d69. Accessed 3 July 2023.

Bevilacqua, Frédéric, et al. "Gesture Capture: Paradigms in Interactive Music/ Dance Systems." *Emerging Bodies: The Performance of Worldmaking in Dance and Choreography*, edited by Gabriele Klein and Sandra Noeth, transcript Verlag, 2011, pp. 183–94.

Dixon, Steve. *Digital Performance: A History of New Media in Theater, Dance, Performance Art, and Installation*. The MIT Press, 2007, doi.org/10.7551/mitpress/2429.001.0001.

Downie, Marc Norman. *Choreographing the Extended Agent: Performance Graphics for Dance Theater*. 2005. Massachusetts Institute of Technology, Ph.D Thesis.

James, Jodi, et al. "Movement-Based Interactive Dance Performance." *Proceedings of the 14th ACM International Conference on Multimedia*, ACM, 2006, pp. 470–80, doi.org/10.1145/1180639.1180733.

Jones, Bill T., et al. *Ars Electrónica Application: Ghostcatching A-74*. Ars Electronica Archive, 1999, archive.aec.at/prix/showmode/33648/. Accessed 10 July 2023.

Karreman, Laura. *The Motion Capture Imaginary: Digital Renderings of Dance Knowledge*. 2017. Ghent University. Faculty of Arts and Philosophy, Ph.D Thesis.

Le, Thi-Lan, et al. "Human Posture Recognition Using Human Skeleton Provided by Kinect." *2013 International Conference on Computing, Management and Telecommunications (ComManTel)*, IEEE, 2013, pp. 340–45, doi.org/10.1109/ComManTel.2013.6482417.

Martinez, Gines Hidalgo, et al. "Single-Network Whole-Body Pose Estimation." *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, IEEE, 2019, pp. 6981–90, doi.org/10.1109/ICCV.2019.00708.

Matthews, Justin, and Angelique Nairn. "Holographic ABBA: Examining Fan Responses to ABBA's Virtual 'Live' Concert." *Popular Music and Society*, vol. 46, no. 3, 2023, pp. 282–303, doi.org/10.1080/03007766.2023.2208048.

Müller, Meinard. "Fundamentals on Motion Capture Data." *Information Retrieval for Music and Motion*, Springer Berlin Heidelberg, 2007, pp. 187–209, doi.org/10.1007/978-3-540-74048-3_9.

Naccarato, Teoma J., and John MacCallum. "Critical Appropriations of Biosensors in Artistic Practice." *Proceedings of the 4th International Conference on Movement Computing*, ACM, 2017, pp. 1–7, doi.org/10.1145/3077981.3078053.

Nymoen, Kristian. *Methods and Technologies for Analysing Links between Musical Sound and Body Motion*. 2013. University of Oslo, Ph.D Thesis.

OptiTrack. "Skeleton Tracking." *OptiTrack Documentation*, 18 Nov. 2022, docs.optitrack.com/motive/skeleton-tracking. Accessed 10 July 2023.

Plaete, Jo, et al. "ABBA Voyage: High Volume Facial Likeness and

Performance Pipeline." *Special Interest Group on Computer Graphics and Interactive Techniques Conference Talks*, ACM, 2022, pp. 1–2, doi.org/10.1145/3532836.3536260.

Reynolds, Joel Michael. "The Normate." *50 Concepts for a Critical Phenomenology*, edited by Gail Weiss et al., Northwestern University Press, 2019, pp. 243–48, doi.org/10.2307/j.ctvmx3j22.

Satore Studio. "Háita - Satore Studio." *Satore Studio*, 2022. satorestudio.com/portfolio_page/haita-at-sonar-lisboa/. Accessed 03 July 2023.

Schmid, Stefan, et al. "Quantifying Spinal Gait Kinematics Using an Enhanced Optical Motion Capture Approach in Adolescent Idiopathic Scoliosis." *Gait & Posture*, vol. 44, 2016, pp. 231–37, doi.org/10.1016/j.gaitpost.2015.12.036.

Solnit, Rebecca. *Wanderlust: A History of Walking*. Penguin Books, 2001.

Sterne, Jonathan. "Mediate Auscultation, the Stethoscope, and the 'Autopsy of the Living': Medicine's Acoustic Culture." *Journal of Medical Humanities*, vol. 22, no. 2, 2001, pp. 115–36, doi.org/10.1023/A:1009067628620.

---. *The Audible Past: Cultural Origins of Sound Reproduction*. Duke University Press, 2003.

Strutt, Dan. "A Simple Tool for Remote Real-Time Dance Interaction in Virtual Spaces, Or 'Dancing in the Metaverse'." *Critical Stages/Scènes Critiques*, vol. 25, 2022, critical-stages.org/25/a-simple-tool-for-remote-real-time-dance-interaction-in-virtual-spaces-or-dancing-in-the-metaverse/. Accessed 10 July 2023.

van Eck, Cathy. *Between Air and Electricity: Microphones and Loudspeakers as Musical Instruments*. Bloomsbury Academic, 2017.

Vicon. "About Us." *Vicon*, vicon.com/about-us/. Accessed 10 July 2023.

Vincs, Kim, and John McCormick. "Touching Space: Using Motion Capture and Stereo Projection to Create a 'Virtual Haptics' of Dance." *Leonardo*, vol. 43, no. 4, 2010, pp. 359–66, doi.org/10.1162/LEON_a_00009.

von Arnim, Hugh Alexander. "Reconfigurations." *Github*, 19 Feb. 2024, github.com/Hughav92/Reconfigurations. Accessed 23 Apr. 2024.

---. "The Shapeshifter." *Github*, 20 Feb. 2024, github.com/Hughav92/The-Shapeshifter. Accessed 23 Apr. 2024.

Zemp, Roland, et al. "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*, vol. 9, no. 4, 2014, p. e95426, doi.org/10.1371/journal.pone.0095426.

Notes

1 As noted by Plaete et al., a number of younger stand-ins also provided additional motion capture for the animations.

2 Nymoen makes the distinction between motion capture and motion tracking, noting that the latter refers solely to the sensing and processing of motion without the storage of the data. However, he notes that these terms are often used interchangeably. For the sake of simplicity, in this article we will use the term motion capture (mocap) to refer both to the real-time tracking and processing of motion data as well as its storage and any processing that does not take place in real-time.

3 Key here is that this is a global coordinate system, which tracks the position of the marker in relation to a defined origin. The alternative is a local coordinate system, which tracks position against another position within the global coordinate system, usually another point in a kinematic chain. If two chains do not share a global coordinate system, there is no way of knowing their position relative to another, and this is a disadvantage of many other techniques used for

motion capture such as the use of Inertial Measurement Units (a combination of an accelerometer, gyroscope, and magnetometer).

4 To briefly summarize Müller, the kinematic chain used to model the human body is *open*, meaning that the chain does not loop back and connect with itself. This allows the organization of multiple rigid bodies into a hierarchical tree, connected by joints with determined degrees of freedom, therefore enabling the parametrization of the model. These parameters are divided between skeletal parameters (relating to the chain's topology and length of the bones) and free parameters (relating to the position and orientation of the chain and the relative orientation of individual bones). Importantly, as the chain is organized hierarchically, the position and orientation of child rigid bodies can be defined in a local coordinate system relative to their parent, all the way back up to the root of the chain. Therefore, through the observation and manipulation of free parameters of a single rigid body, the motion of rigid bodies either further down or further up the chain can be calculated through *forward* and *inverse kinematics* respectively. Müller provides a mathematical description of a simple kinematic model, with joints modeled as possessing reduced degrees of freedom. He notes that more complex models exist to account for joints with a greater number of degrees of freedom, as well as phenomena such as skin deformation and muscle force. Müller frames these more complex models in relation to providing "enhanced realism in computer animation" (198).

5 Abouaf reports that the captures took place with two dancers ("'Biped': A Dance with Virtual and Company Dancers. 2"), whereas Dixon reports three.

6 To be able to preserve marker identity across gaps, we used the position of a pre-defined rigid body consisting of three markers in a fixed position as a stand in for a marker. These are pre-produced by *OptiTrack*, only a little larger than a single marker, and likewise have adhesive tape affixed to their rear.

7 Here, as we did for *Reconfigurations*, we again employed the 30 pre-defined rigid bodies to stand in for the markers.

8 This is the case, for example, in the modeling function of the *Motive* software, which is used for both animation and biomechanical research (OptiTrack).

9 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 (Zemp et al.) and can be used to model the spine of a person with scoliosis.

10 A pertinent question to raise here is *why* the bodies modeled by mocap software take the normatively inclined form that they do, or more precisely, why mocap software designers narrowly limit the range of bodies for which they implement a kinematic model. It is possible here to begin to draw links to various values placed upon certain body types, such as the commercial value placed upon non-normative bodies within a capitalist ecosystem, as well as the broader social and cultural implications of these value judgments, however a full examination of this is beyond the scope of this article.