Ghost in the Virtual Reality: Translating the human essence with motion captured dance

Eugenia S. Kim
School of Creative Media
City University of Hong Kong
eugenia.kim@my.cityu.edu.hk

In the world of dance and its related disciplines, the ability of a performer to successfully 'move' audience members enough to elicit an emotional reaction is vital. Certain factors such as the proximity of audience to performer and viewing a performance in real-time affect how this effect is achieved. When dance movement is digitised through moving image or animation, there is potential for loss in translation of the emotional feeling experienced during a live performance versus recorded live performance or animated performance. Conversely, a heightened sensation might occur through the use of cinematography, editing and special effects. This translation issue may be encountered when creating a virtual reality animation dance using motion capture since the technologies involved can both interfere or enhance the presentation of movement. Assuming that human essence needs to be captured along with physical motion in order to generate an emotional reaction, then the choreography and motion capture data become the 'ghost' that is transplanted from human into a new digital body. This separation then raises the question of how to maintain the subtleties required for communication that lead to generating empathy in viewers for a virtual performer's narrative. To address these issues, I engaged in a series of dance motion capture sessions for a virtual narrative about mental health as the basis for examining how a choreographer and motion capture dancer can work with the limitations of technology, rather than be limited, to produce useful data. Specific limitations included use of contemporary and somatic dance, a relatively low number of cameras and dots, no facial or hand data and the use of abstract humanoid figures. Although a universally applicable solution was not discovered, I was able to identify a set of strategies that would be useful to contemporary dance choreographers using motion capture technology for the first time. Furthermore, the strategies are intended for movement narratives rooted in portraying emotion rather than physical spectacle dependent on virtuosity and visual effects.

Dance. Motion capture. Virtual reality. Emotions. Somatic movement

1. INTRODUCTION

In the world of dance and its related disciplines, the ability of a performer to successfully 'move' audience members enough to elicit an emotional reaction is vital. This is a key factor regardless of whether one is watching live or recorded dance. The benefits of a live performance are that the creative team has more control over elements such as proximity to the audience, venue, perspective and viewing time. In other words, what the audience sees and when. When dance movement is digitised through pre-recorded moving image or pre-rendered animation, the viewer gains more control over their experience – which might mean losing any built-in emotional build-up or key

moments. Conversely, cinematography, editing and special effects can sometimes heighten a sense of emotion where little existed before.

One specific scenario where a loss in translation of emotion and human essence may be encountered is when using motion capture to create a virtual reality (VR) animated dance. Assuming that the human essence needs to be captured along with physical motion in order to generate an emotional reaction, then the choreography and motion capture data become the 'ghost' that is transplanted from human into a new digital body. This separation then raises the question of how to maintain the subtleties required for communication

that lead to generating empathy in viewers after viewing a virtual narrative.

To address these issues, I produced a set of VR experiences that use motion capture animations for a movement-based narrative about mental health. By taking on the role of both choreographer and performer I was able to directly explore how to work with the limitations of existing technology, rather than be limited by it, to generate effective movement. Ultimately, it became a question of how one choreographs and performs contemporary dance in a manner that suitably conveys 'human' emotion given technological and other constraints.

At the end of the process, it was clear that an affordable (less than \$1000USD) and easy-to-use (no technical or animation expertise required) hardware/software package would not be available in the immediate future. Nonetheless, I was still able to identify some strategies that might be useful for choreographers new to marker-based motion capture systems.

2. BACKGROUND

2.1 A brief overview of motion capture for dance

For the mainstream consumer, motion capture is usually encountered in major Hollywood films and video games. Embedded within those media is movement that is identifiable as either 'acting' or 'dance.' This includes games such as the 'Just Dance' series where players' movements are captured using a depth camera such as the Kinect if used with an Xbox 360 console (Ubisoft 2019).

The use of motion capture specifically for dance performance works and cinema, however, has its own separate history. Distinguished artists such as Merce Cunningham and Bill T. Jones have engaged in ground-breaking 'virtual dance' projects utilising animation and optical motion capture since the 1990s; the creation process was described by Birringer (2002) as "The manipulated data become the ghost of the dance." More remarkable was the translation of Cunningham's 1971 'Loops' into a motion capture animation from 2001-2011 (OpenEnded Group 2019). Because the work was never taught to any other dancer and was only hand and finger data, it stands as both preservation and performance object in one for a body part that generates a lot of data.

Since that time, motion capture for dance at the academic research level has continued to grow. The OpenEnded Group continued their work through collaborations with institutions like Arizona State University for live interactive dance

performances incorporating motion capture (James et al. 2006). Chan et al. (2011) took motion capture from the stage and into the classroom by proposing a dance training system. More recently, the WhoLoDancE (Whole-Body Interaction Learning for Dance Education) team in Europe uses motion capture and virtual reality (VR) to create an immersive database of dance for accessibility, education and preservation purposes (Cisneros et al. 2019). These are but a few of the innovations taking place.

Table 1: Summary of select motion capture technologies and services accessible from HK during 2018–2019.

Company/System Type	Costs	Characteristics
<i>Rokoko – Wireless suit</i>	\$2295 USD (student price) \$2495 USD (standard price) Extra fee for software	A full wireless body suit, no facial or hand data option included.
<i>Notch – Moveable sensors</i>	\$379 – \$1137 USD depending on configuration	Ranges from 6-18 points on elastic bands. Uses smartphone or tablet to capture motion. No facial or hand data option.
<i>Synertial – Suit with HTC VIVE trackers</i>	Pricing requires quote. Requires custom software.	A multi-piece suit; hand data option available for additional price.
<i>Deakin University (Australia) – Optitrack camera-based system; early 2018 quote</i>	\$4000 AUD for day rate \$10 AUD/second for clean up	Marker-based system using a suit and optical cameras.
<i>Freespace/Tracklab (Australia) – Optitrack camera-based system; early 2018 quote</i>	\$2000 AUD for day rate \$10 AUD/second for data clean-up	Same system as above; data clean-up requires using Motivebody software. Extra cost for hand data.
<i>MetaObjects (HK) – HTC Vive tracker, goes on-site to whichever facility is available</i>	\$500 HKD/hour special rate for artists	Provides technical services using either an HTC Vive setup for on-site shooting or would learn the setup for an existing facility.

To give context to the level of expertise and technology required, it is helpful to look at the array of options available. For an independent artist with limited funding, the reality of the logistics could potentially prevent a small-scale project from taking place. In the following table, I present the motion capture options that were available to me in Hong Kong over a 12-month timespan.

One reason why a depth camera option was not presented here was simply that it did not appear in the literature as a preferred technology for capturing dance nor as an immediately available service. An example of its effective use for a dance project would be Maria Takeuchi's *Asphyxia*, which made use of several Kinect cameras (2015). The post-production process requires stitching together the data from multiple camera sources.

2.2 Empathy, virtual embodiment and movement

For this project, empathy with the animated figures was the main emotional response that I was seeking from viewers. Edward Warburton (2011) proposed that there were three ways to experience empathy through dance: somatic, kinaesthetic and mimetic. For viewers, it is kinaesthetic empathy – the feeling of experiencing the movement even though they are not actively moving – that matters the most. For virtual reality experiences, embodiment is a desirable way to influence a viewer's perception of a situation and therefore elicit a specific feeling (Baily, Bailenson & Casasanto 2016). This quality can be achieved through both the first and third person perspectives with evidence suggesting that the first-person perspective yields stronger results (Debarba et al. 2017). This leads to a dilemma then as to which perspective is best for viewing a motion capture animation in an immersive environment and if kinaesthetic empathy can lead to emotional empathy.

2.3 Artistic and narrative case studies

Several case studies were used as a reference point for designing both the content and interface. In terms of movement, the works developed by Katherine Boydell, Alex Mermikides, Genevieve Smith-Nunes, and Kevin Turner demonstrated how a wide range of approaches were used to generate empathy for various illnesses. Mermikides (Weitkamp & Mermikides 2016) and Smith-Nunes (Smith-Nunes & Neale 2017) both take an interdisciplinary approach to describing elements of having a physical illnesses, with the latter taking place in a VR environment. Boydell (2011) and Turner (Company Chameleon 2016) use mental health as their topics, with Turner specifically

focusing on his own bipolar disorder which may have started as ABPD.

Although there is a growing number of medical related VR experiences, it is still uncommon to see one that makes use of motion capture and dance movement. Of the case studies listed above, Smith-Nunes' pain[Byte] is especially relevant in that it uses a Kinect-based setup rather than a marker-based system and uses ballet, which is immediately identifiable as a dance technique (Smith-Nunes, Shaw & Neale 2018). The question that remained unanswered was what happens when the main creator is performing movement that describes their personal experience with an illness and how the emotions of that come through in the virtual realm.

3. MOTION CAPTURE PROCESS

For this scenario, 'embodiment' was not presented through a first-person perspective of being the patient in the real world. Instead, the viewer is trapped inside the patient's mind and observing movement representations of symptoms, feelings, thoughts, and actions. This echoes the feeling that I had of being trapped with myself as a young patient, as though I was observing myself both externally and internally at the same time. It also has its roots in my 2005 dance work, '(A)Typical Day', which has a similar staging. To place a viewer in a first-person perspective would have required using a haptic suit and surrounding them with virtual 'mirrors.' The technology required to reproduce the sensation of breathing, spinal contractions, and gestures of any kind, however, would render the experience impractical for general viewing. Thus, for the sake of accessibility, I chose to keep the viewer in third-party perspective but still 'inside' the mind and able to interact with a life-sized figure. Any 'empathy' that might be derived is not from having "walked in someone else's shoes" but rather closer to 'sympathy' in that there are gestures, poses and movements that are common to the average human experience.

The prototype has two parts: 'Symptoms' and 'Coping.' In 'Symptoms', there are six short movement phrases on loop. The narratives were derived from my personal memories and other patient accounts from medical reports of specific ABPD symptoms. Wooden mannequins represent each symptom with the hard edges meant to echo the clinical nature of the reports. For 'Coping', a more traditional linear dance narrative is used to show my memories of coping with ABPD. Some of these experiences were ones I had in common with other patients I had met over time. A particle cloud effect is used to give the feeling of constantly coming undone and pulling back together. The two

parts were designed to contrast with each other conceptually, choreographically and visually.

This contrast in styles was originally intended for the purpose of user testing to determine which style was more effective. Over time, however, I realised that the narrative nature behind each part demanded these stylistic differences. 'Symptoms' is rooted in medical science – that is, the jargon and observations that the medical community agrees is descriptive of ABPD. 'Coping', however, is completely rooted in the daily experience and emotions of a patient. These differences informed the nature of the movements performed and by extension, the motion capture process.

Figure 1: From top to bottom – v0 prototype, v1 'Symptoms' and v1 'Coping'.

In version 0 of the prototype, I worked with MetaObjects to use an HTC Vive system that was modified to do motion capture. It required use of four sensors and two hand controllers. Although the data was clean enough to be used as is, the visual result was stilted because it was difficult to dance with full range of movement while large sensors were strapped to the feet and hands occupied by controllers. I had also taken it upon myself to do the modelling, animation, and programming on top of the choreography, performance, and sound design

within a one-month span, which resulted in loss of quality for that prototype.

For v1 of the prototype, I worked again with MetaObjects to use the motion capture lab in the Centre for Applied Computing and Interactive Media (ACiM) at the City University of Hong Kong. This lab uses a marker-based 24-camera Raptor system, which at full capability can:

- Use 99 markers for each actor
- Record hand/finger data (but not facial data)
- Record multiple sessions a day
- Dedicated recording and clean-up technician. Previous projects have had multiple animators assigned to it.

When operating at the full level, the captured data can require about a month for clean-up of each session. Actors did not have to change any of their planned choreography, but they did modify their movements because the capture area was limited. In our sessions, our team used approximately 40 markers and a single technician did the data clean-up, animation, and programming. We had two 4-hour sessions with the first for calibration and testing and the second for actual capture. It took at least 20 hours to clean up the 4-hour session, but less time to do the mapping and programming.

The reason why these technical issues were considered for the actual choreography and performing was largely to minimise costs. Before shooting, I decided to take an improvisational approach with a basic score that I sent to the technician ahead of time. Having a tightly choreographed set of sequences like I did for the v0 prototype would not only have meant rehearsing 20+ minutes of choreography, but also investing time and money into studio space and work time. I also recalled having to modify movement based on technical glitches that were only discovered while recording. It would be easier to make quick choreographic changes if the movement was not set. Finally, using a complex dance technique like contemporary ballet or jazz was likely to result in a large amount of data clean-up. By using mainly pedestrian and somatic movement while keeping dance technique to a minimum it would be possible to reduce the amount of clean-up required.

Despite anticipating these factors, we still encountered the very real fact that computers will choose to malfunction unexpectedly and that occlusion will occur regardless of careful planning. An interesting phenomenon was that although no facial or hand data was being recorded, I still had a high amount of facial expression and detailed hand gestures during recording. It emerged that in order to fully express an emotion with the body, it meant

that the entire body literally had to be engaged regardless of whether it was being recorded. One unexpected outcome was the impact that my technician had on the choreography. I found that I had to rely on him quite a bit to know what movement was recording easily and what movement did not get picked up that well. His recommendations during the animation mapping and programming phases also helped me explore new options since he was able to quickly show changes in Unity that would have taken me longer to discover.

4. CHALLENGES AND STRATEGIES

After completing the prototype, I was forced to accept the inevitable: the reality is that “more cameras + more dots + more technicians + more polygons = better data.” This is of course simplifying the situation and is only a summary of why affordable high-quality motion capture solutions for individual artists is still elusive. When enough resources are available, more solutions are available. For example, WETA Digital's work with high-definition cameras demonstrate that it is possible to capture realistic facial expression data (Weta Digital n.d.). The high level of expertise and facilities required would make it difficult for an independent artist to access or reproduce their system. Likewise, while depth cameras like Intel's RealSense line are priced for the average consumer (\$149-199USD per camera at the time of writing; see Intel 2019), one still needs to have a certain level of technical ability to do anything with the resulting data.

At this time, as a dance practitioner, one must be strategic in balancing artistic, financial and technical requirements. The following is a list of suggestions for first-time dance practitioners and technicians working together on a low-budget project to help make the process a bit more efficient:

- Before the choreography session even begins, try a session with the motion capture system being considered; if there are signature steps or phrases, those should be included. Change to another system if necessary.
- Articulation of body parts as seen in South East Asian dance forms (i.e. eyes, fingers, feet) and other dance styles may not be possible for the system being used. Movement style and choices should be made in conjunction with the system selection.
- Choreograph for the space constraints if there are any as well as the suit if one is being worn. Structured improvisation might

work better for certain cases such as if there is not a lot of time available in the lab or for data clean-up.

- If a realistic human feel is desired, then it may be necessary to exaggerate the breath and other subtle gestures so that the final animated figure is not artificial or doll-like. More specifically, when touching another body part, that may come across in the animation or it may look like it is floating and require data manipulation.
- Rely on the technician to help ‘edit’ the choreography. This applies more for personal narratives rather than dance preservation. If a specific dance movement is essential, then plan for additional time to be spent on troubleshooting it.
- Expect performers to need breaks and lose performance energy over time just like a regular dance cinema shoot. It may be inevitable to schedule all takes for a single day to reduce costs so prepare performers for extra endurance and stamina if necessary.
- Unless you are willing to invest in face and hand data, it is inevitable that a lot of human expressiveness will be lost. Therefore, every other body part should be utilised to its maximum capability. For example, how would sadness be shown using the shoulders, chest, and rib cage rather than the face?

5. CONCLUSION

Realistically speaking, motion capture and related technologies have simply not yet reached a point where sophisticated results can be achieved without expensive equipment, highly skilled specialists and vast amounts of time. For this simple project, the total cost was approximately \$3000USD even with the in-kind support of ACiM and a minimal team of two. If all three types of resources are freely available, then choreographers and performers are also freed from most movement constraints. The reality is that very few individuals can currently produce a motion capture project on their own. That is not to say that in the future that a reasonable high-quality solution will not be feasible; what is unknown is the precise timing of when that future will arrive.

Technology, however, is not the only factor for helping to induce sensations of empathy and embodiment in a virtual environment. To that end, choreographers and performers must still be mindful about how even a simple gesture will be perceived once it is translated as either video or animation. Ethics and legalities must still also be adhered to when presenting sensitive content such

as medical narratives. These are elements that can be developed now while technology continues to evolve. The acting discipline has been able to develop its own methodology for motion capture; it should be feasible for dance to do the same.

6. REFERENCES

- Bailey, J. O., Bailenson, J. N., and Casasanto, D. (2016) When Does Virtual Embodiment Change Our Minds? *Presence*, 25(3), pp.222–233.
- Birringer, J. (2002) Dance and Media Technologies. *PAJ: A Journal of Performance and Art*, 24(1), pp.84–93.
- Boydell, K. M. (2011) Using performative art to communicate research: dancing experiences of psychosis. *Canadian Theatre Review*, 146, pp.12–17.
- Chan, J. C. P., Leung, H., Tang, J. K. T. and Komura, T. (2011). A Virtual Reality Dance Training System Using Motion Capture Technology. *IEEE Transactions on Learning Technologies*, 4(2), pp.187–195.
- Cisneros, R. E., Wood, K., Whatley, S., Buccoli, M., Zandoni, M. and Sarti, A. (2019). Virtual Reality and Choreographic Practice: The Potential for New Creative Methods. *Body, Space & Technology*, 18(1), 1–32.
- Company Chameleon (2016) Witness: Behind the scenes. <https://vimeo.com/185556787> (retrieved 19 February 2019).
- James, J., Ingalls, T., Qian, G., Olsen, L., Whiteley, D., Wong, S., and Rikakis, T. (2006) Movement-based Interactive Dance Performance. *MM '06 Proceedings of the 14th ACM International Conference on Multimedia*, pp.470–480.
- OpenEnded Group (2019) Loops (2001-11 DOWNIE ESHKAR KAISER). <http://openendedgroup.com/artworks/loops.html> (retrieved 16 March 2019).
- Smith-Nunes, G. and Neale, C. (2017). [pain]Byte VR Storytelling & Classical Ballet. *Proceedings of EVA London*, pp. 415–419.
- Smith-Nunes, G., Shaw, A., and Neale, C. (2018). PainByte: Chronic Pain and BioMedical Engineering Through the Lens of Classical Ballet & Virtual Reality. *TEI '18 Proceedings of the Twelfth International Conference on Tangible, Embedded, and Embodied Interaction*, pp.493–497
- Takeuchi, M. and Phillips, F. (2015) as·phyx·i·a. <http://www.asphyxia-project.com/film> (retrieved 16 March 2019).
- Ubisoft. (2019). Supported controllers for Just Dance 2019. <https://support.ubi.com/en-GB/Faqs/000038593/Supported-controllers-for-Just-Dance-2019-JD2019> (retrieved 16 March 2019).
- Weitkamp, E. and Mermikides, A. (2016). Medical performance and the 'inaccessible' experience of illness: an exploratory study. *Medical Humanities*, 42(3), pp.186–193.
- Weta Digital (n.d.) Facets | Weta Digital. <https://www.wetafx.co.nz/research-and-tech/technology/facets/> (retrieved 16 March 2019).