10 Capturing and visualizing 3D dance data: Challenges and lessons learnt

Cláudia Ribeiro[1,2], *Rafael Kuffner*[1,3], *and Carla Fernandes*[1]

[1] *BlackBox Project, FCSH, Universidade Nova de Lisboa, Portugal*
[2] *LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal*
[3] *University College London, United Kingdom*

10.1 Introduction

Advances in 3D, motion sensing, and virtual reality (VR) technology, as well as easier access to affordable devices such as VR headsets, have triggered innovative initiatives in dance research around the world. These initiatives include new ways of documenting, archiving, and preserving dance to enhance approaches in creative and teaching practises. In the case of the most ephemeral and intangible art forms, such as contemporary performing arts and dance in particular, incorporating such advances is neither easy nor trivial. BlackBox as an Arts&Science project[1] has addressed new forms of mediation for collaborative practises and methodologies amongst the creative agents and different players involved. Not only have we crossed different scientific domains (from cognitive linguistics to performance studies, neurosciences, and computer vision), but we have also closely worked with artists and media designers. The main BlackBox platform hosts all outputs from the three case studies that were carried out: from animated infographic video clips and longer documentaries of rehearsal processes, to 360° footage, point clouds, 3D visualisations, and video annotation tools.

Various rehearsal sessions of three invited choreographers were taken as case studies to be documented, annotated, and published, according to the following theoretical principle: the formal cognitive structures (concept categorisation, work-oriented keywords, mental image schemata, conceptual metaphors, etc.) operating in the somatic bodies of the choreographers/performers should be reflected as much as possible into the architectural structures of the outputs to be developed. This task demanded designing and testing workflow processes, mediated by annotation procedures that functioned in accordance with the video capture in question and the envisaged online publication strategies.

In this chapter, we describe the main challenges and lessons learnt from capturing and visualising 3D dance data for four years in projects with three Portuguese contemporary choreographers and their dancers. The context in which these are described ranged from documenting creative processes and 3D

DOI: 10.4324/9781003106401-14

visualisation of dance movements to developing a VR annotator for contemporary dance.

10.2 Context

One of the aims of BlackBox was to create a new paradigm for the documentation of performance composition by augmenting a 2D video annotator with 3D visualisations of the resulting annotations. Therefore, our initial approach during the first case study with choreographer João Fiadeiro consisted in extending the existing functionality of the Creation Tool (2012) 2D video-annotator prototype to tackle this problem. The new system we developed places the 2D annotations on a 3D point cloud that is captured by depth sensors coupled with cameras around the stage. This enables the user to freely visualise the annotated performance in a 3D environment and from an arbitrary point of view. This system was then tested in a real environment, namely in two separate sessions using Fiadeiro's Real Time Composition (CTR) method (Fernandes et al., 2022, this volume), involving the choreographer and seven dancers from his company, who were recorded using video cameras and three Microsoft Kinect depth sensors. The data collected were used to generate the corresponding point clouds, which were analysed together with the videos. The main criteria we followed to choose the dance sketches for annotation was the richness of the data contained in the improvisation and their relevance level in relation to the choreographer's work. We produced point-cloud renderings for some of the main compositional concepts underlying Fiadeiro's CTR method (e.g., Position and Relationship) (Ribeiro et al., 2017).

In our second case study, we worked with choreographer Rui Lopes Graça, who comes from a neoclassical tradition of contemporary dance. We closely followed his creative process during the piece *15 Bailarinos e Tempo Incerto*, produced for the National Ballet of Portugal in 2016, and his respective interactions with the 15 dancers, namely the choreographic methods he followed to generate dance material, to guide and correct the dancers, and finally to compose the piece. We captured examples of small individual performances, and group and individual rehearsals, and used them to drive our research regarding the 3D reconstruction of point-cloud videos. Supported by that type of media, we developed techniques to visualise movement qualities of dance through rendering effects. Combining optical flow analysis with depth data information, we were able to estimate the 3D flow for each point in the cloud. This information was used to colour-code point-cloud videos, showing potential to be used for the creation of different rendering effects, which are also discussed in (Anjos et al., 2017, Anjos et al., 2018) (see Video 1, digital Support Material online).

Finally, we prototyped a VR annotation system, where some of the most common documenting tasks during the creative process could be performed in an immersive environment.

In the third case study, we developed and provided alternative ways of seeing Sylvia Rijmer's choreographic process, both while exploring creative tools for deliberate choice-making within a contemporary dance-making process and in the form of a short choreographic excerpt. We used methodologies from discourse analysis and gesture studies at first, in order to subsequently offer alternative perspectives on a dance creation process using new media: 3D point clouds, VR, stereoscopic and high-speed photography, and animated infographics.

Based on her pre-choreographic concepts, Rijmer (Rijmer, 2022, this volume) has set up an analytical investigation to structure the empirical knowledge framed around cognitive attention and movement habitual patterns within an artistic dance practise poised to question and challenge the negotiation of 'choice' in dance. Specifically, she has generated tools for compositional (Dodging) and typological (Scanning) movement possibilities, and (partly) superimposed them on a re-enacted map of Cornelius Cardew's *Treatise* Score," which the dancers were invited to interpret.

Here, we have done a specific study with the title "Dance in VR – deconstructing choreographic objects through expanded media," which has led to the development of a VR installation gallery offering end users an opportunity to immersively experiment alternative visualisations of the creation process of a dance piece.[2]

10.3 Capturing and visualising 3D dance data

10.3.1 Challenges of capturing 3D dance data

With the goal of capturing the visually relevant information from a performance or a simple movement sequence, it is fundamental to address questions on the nature, context, and purpose of the data capture, in order to choose and install the most suitable setup. This will dictate what kind of hardware that will be used, what other equipment will be necessary, and possibly where the capture session should occur. Considering all these aspects beforehand is a crucial task to verify the viability of the proposed data capture session.

Video-based rendering (VBR), the process of generating novel views from the video capture of a dynamic scene, is a very broad research field, and it has very specific challenges and problems related to the technical process itself, independent of the context. In this section, we will refer specifically to the challenges of capturing dance data. For an overview of VBR and current challenges on a generic level, we recommend the survey paper by Anjos et al. (2018a).

10.3.1.1 Nature of the content

Different dance styles produce different types of movements and contents to be recorded. At a very basic level, one needs to consider speed. Very fast motion might not be well captured by burst photography or low frame-rate videos. Another very commonly overlooked aspect is the lighting conditions of a capture studio. High-speed motion may require longer exposure times in cameras, which usually generates blurred images.

On the other extreme, very slow-paced performances can run for extended periods of time. This may result in high amounts of data for dense data types, such as 3D reconstructions, which means that a normal 2D video capture, burst photography, or skeleton reconstruction may be better suited.

Independent of speed, one also needs to consider the subjects of the data capture. If one can make strong assumptions about the content that is going to be captured (e.g., people wearing certain costumes), it might be viable to use a simplified data representation for what is captured. If it is a solo piece where all that matters is the dancer, a tracked skeleton representation coupled with a 3D reconstruction of mesh is a viable alternative. However, for duets, trios, and so forth, markerless tracking might run into difficulties, thus requiring more precise tracking solutions (e.g., VICON, Optitrack). If no strong assumptions can be made about the content that is to be expected, or there is high uncertainty regarding its behaviour, making it hard to accurately track (e.g., usage of props, or the presence of complex objects, such as clothing and loose hair), one needs to resort to 3D video recording in order to generate sequences of point clouds. This is typically the case in not only contemporary dance and improvised performances, but also in a variety of dance styles where different props are an essential part of the performance (this was the case in our first study with Fiadeiro, for instance). Moreover, if costumes play a relevant part in the performance, the option of using specific bodysuits for motion tracking cannot be considered, once again restricting the possibilities for tracking technologies.

Lastly, the possibility of occlusions often happening on stage needs to be taken into account. In group scenarios, one will preferably install the cameras spread out in a wider configuration to ensure that no part of one dancer's performance is occluded in a crucial moment. This is a key factor to consider if any sort of markerless tracking technology is being used, since losing sight of a dancer will typically cause tracking algorithms to fail. In solo scenarios, self-occlusions may need to be considered as well, and it is typically recommended to use more than a one-camera viewpoint to ensure the 3D reconstruction of the complete performer's body and surroundings.

10.3.1.2 Context of the capture

Similarly to *what* is being captured, *where* the data is captured has an impact as well, albeit lesser, on how the data capture process will look. When capturing with an anthropological mind-set, where one desires to directly observe what happens in the studio, the capture setup needs to be mobile and non-intrusive.

It is fair to assume that in general, dance studios offer very different conditions from those found in motion capture studios, as far as the infrastructures to connect computers and cameras are concerned. Also, mirrors and lighting can have a negative impact on the quality of the capture data. Moreover, as dance rehearsals are fast-paced environments, minimal preparation and calibration times are desired in order not to disturb the natural process one needs to observe.

On the other hand, when bringing dancers to a motion capture studio or lab, the main thing to consider is how to provide proper conditions for the dancers to perform in our space, trying to replicate the necessary space, the type of floor material and lighting conditions, as well as access to any necessary props used in the performance.

10.3.1.3 Purpose of the capture

Finally, before recording, it is extremely important to define the purpose of the data capture. Defining the user interaction paradigm is the usual starting point for deciding how a VBR process will occur (Anjos et al., 2017) regarding capture setup and data representation. In the context of dance, while simple 3D data such as skeletal information is adequate for interaction, since it is easy to manipulate, it might be too abstract of a representation for documenting and archiving a specific dance piece, in which case denser representations must be used.

On the other hand, interaction with point cloud data is more difficult to design, as it is an unstructured type of data. However, they are more adequate for archiving purposes, since they encode all the data in detail.

10.3.1.4 Use cases in the BlackBox project

The three different case studies we worked on during the BlackBox project were used as opportunities to implement and validate the available capture software and alternatives. From each study with the choreographers, we collected content of a different nature and with different purposes, but similar contexts.

Starting with context, the three case studies shared a similar problem: the captures were always performed in the dance studio and not in our own laboratory. This forced us to develop solutions that were quick to deploy and required minimal calibration time. The toolkit we developed in order to calibrate multiple 3D scanners has been described in Sousa et al. (2017).

In the first case study (with João Fiadeiro), the action was slow paced, and little could be assumed about the content that would be captured, given that improvisation with props was a core part of his method. For those reasons, the media used had to be image-based, given that it would be unfeasible to attach markers to every single object. However, the purpose of our capture was not only to visualise the captured data in two dimensions but to extend the visualisations to three dimensions in order to better illustrate his composition method (Ribeiro et al., 2017). Moreover, we wanted to be able to annotate the captured data in a similar way to what was done in 2D videos (as described in Ribeiro et al., 2016). Hence, we resorted to a combination of video and 3D

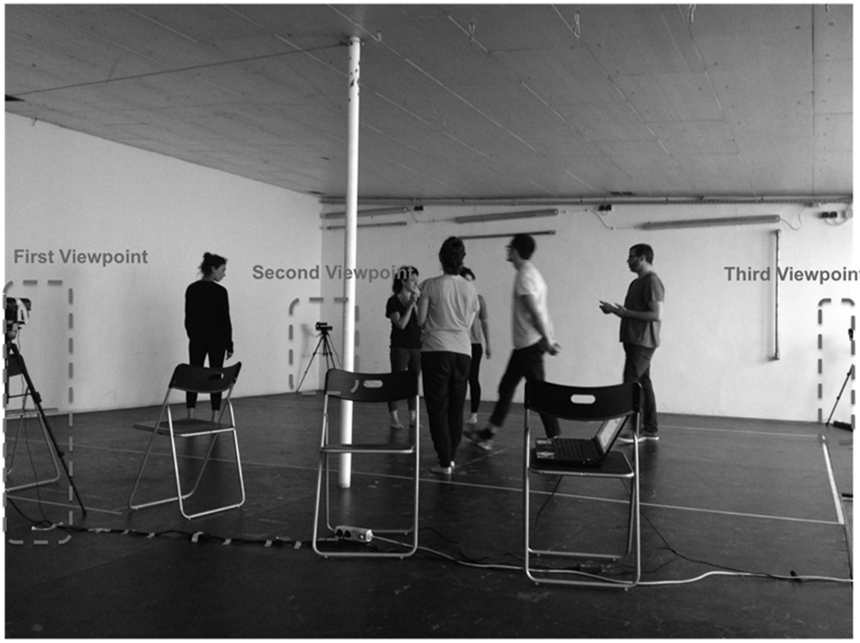

Figure 10.1 Setup used for our first case. Kinect V2 cameras to capture point cloud videos, paired with video cameras on top for standard 2D capture.

point clouds to be able to manually annotate the data using the video annotator, but also to visualise these annotations in 3D later on.

The setup used is represented in Figure 10.1. The main challenge regarding the capture we dealt with in this project was the length of the 3D captures, which would occupy close to 3GB/minute. This resulted in the development of a compression technique for this type of data (Anjos et al., 2017), which then allowed visualising the full sequences on a common computer and interacting with the data in real time.

The second case study (Rui Lopes Graça) had shorter sequences with faster dancer movements covering larger areas of the stage. The performances we accompanied were solos and synchronised group dance sequences. Our two main goals were: (a) to generate a visualisation of movement qualities in the group performances and solos (Anjos et al., 2017); and (b) to expand our annotation approach to a VR context (Ribeiro et al., 2018). We used the same capture setup from the previous use case for the 3D capture, but we did not use a paired video. Examples of the generated data can be seen in Figure 10.2. In this case, we started to explore capturing the dancer's joint information using markerless tracking from the Kinect sensors. However, due to the high-paced sequences we recorded in this case study, we found markerless tracking to be lacking in precision.

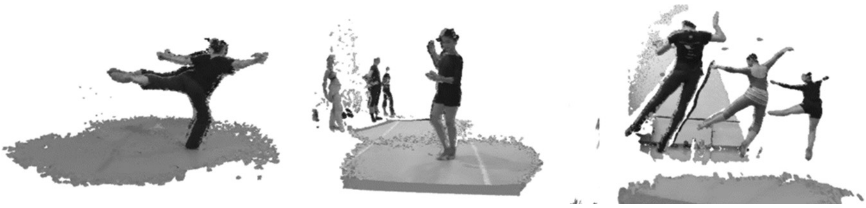

Figure 10.2 Three camera setups. Left to right: Isolated dancer with wide camera setup; group setting with a narrow camera setup; single camera for a group setting.

Finally, with our third case study with Silvia Rijmer, the nature of the content was similar to the previous one, although the dance movements were generally slower. This allowed us to better synchronise the 3D and skeleton data for VR annotation, visualisation, and interaction.

10.3.2 Challenges of visualising 3D dance data

The continued advances in 3D technology has allowed the development of low-budget motion capture systems, which are able to capture a dance performance or individual movements from several points of view. By using this data in computer vision and computer graphic algorithms, it is possible to reconstruct an almost realistic virtual representation of what was captured, and therefore preserve those moments in time.

As a result, in recent years several datasets have emerged, consisting of different types of dance data, including video data, motion capture data, and burst photographs. The relevance of this data is particularly important when considering the preservation of the intangible nature of dance through archiving processes, documenting creative processes, or innovative ways of teaching/learning dance. This raises issues related to how these data should be visualised, specifically which approaches should be considered to prevent losing the important and subtle qualities inherent in dance movements.

10.3.2.1 Visualisation of point-cloud data

Point clouds are unstructured collections of sparse 3D points located in space. The relevance of point-cloud data has increased since the advent of cheaper and ubiquitous depth-sensing solutions (e.g., Microsoft Kinect, Asus Xtion, and PMD CamBoard), which allow the creation of 3D virtual representations of the real world in an ever more accessible way.

Nevertheless, the results achieved display significantly lower resolutions compared to high-end commercial solutions that produce extremely high-resolution scans. This lack of resolution raises several challenges regarding point-cloud visualisation. The greatest shortcoming is the loss of definition in close-ups, meaning that, when zooming in a certain region of the point cloud, shapes are distorted and do not accurately represent what was originally captured. Another important challenge is solving foreground/background interference and depth perception due to occlusions across the rendered points.

When considering capturing dance data, it is also important to be able to create relations between different elements of the captured scene, for example, to relate multiple dancers to the props they used. Since this type of data is simple Cartesian coordinates and their respective colour-coded information, there is no additional contextual information about what is actually represented in the point cloud, and therefore there is no straightforward approach to separate individual elements of the point clouds from each other.

10.3.2.2 Visualisation of skeleton data

The Kinect sensor supports skeleton data represented by joint information of 20 joints (head, arms, hands, legs, hips, and body mass centre, amongst others). With this data, it is possible to use 3D primitives, such as spheres and cylinders, to construct a representation of the skeleton data. Nevertheless, when there are occlusion and self-occlusion problems, such as when the Kinect sensor cannot track a joint because it is occluded by another body part or it is out of range, the tracking information results in figurative movements that may not correspond to the human body's. In the specific case of dance data, this represents a considerable disadvantage.

Another challenge is related to the synchronisation of point-cloud data and skeleton data. Depth information is captured at a higher rate than skeleton information, since the latter needs to be generated through complex AI algorithms. When collecting very fast movements, this becomes an even greater challenge because it is virtually impossible to synchronise those data with tracking information alone.

10.3.2.3 Use cases in the BlackBox project

Within the framework of the BlackBox project, we devised new visualisation techniques that would allow alternatives and improvements on how choreographers visualise and document their respective creative processes. With this in mind, the approach used for each of the case studies attempted to address different challenges, specifically documenting creative processes, annotating dance data, and visualising dance data in a VR environment.

Throughout the past 30 years, choreographer João Fiadeiro has developed a method to generate and compose dance performances in real time that comprise a set of underlying fundamental concepts (Jürgens et al., 2016). When working with his dancers, Fiadeiro resorts to the use of different props, such as office material and other commonly used objects (e.g., chairs, tables, duct tape), which are intertwined and create relationships between the dancers' movements and with the unfolding of the performance itself. It was therefore important to be able to computationally isolate each dancer from the objects and from the other dancers, in order to accomplish a valid visualisation of his concepts.

As explained previously, amongst the challenges of visualising point clouds is the ability to overcome foreground/background confusion. To address this challenge, we developed a stroke-based rendering technique (Anjos et al., 2017) (see Video 2, digital Support Material online) for visualising point-cloud data. This technique allows for high perception of objects and shapes at different resolutions in real time. It was inspired by the aesthetics of the Impressionist movement and uses brush strokes instead of simple points as a graphical primitive to render point clouds with a painterly appearance (for a more detailed description, see Anjos et al., 2017). To isolate the dancers and the objects, we developed a cluster detection algorithm, which basically separates unorganised point clouds into individual clusters (sets of points) based on their respective distances and the distance to the centre of the AABB box between two frames (current and previous) (Ribeiro et al., 2017) (see Video 3, digital Support Material online). Two examples of how we used those techniques to visualise Fiadeiro's concepts are shown in Figure 10.3 and Figure 10.4 and Video 4, Video 5, Video 6, Video 7, digital Support Material online.

Figure 10.3 Composition in Real Time concepts: on the top, two shots of the *Position* example: each participant represents a consecutive position. On the bottom, two shots of the *Relation* example: the second participant slowly assimilates the colour of the first as they establish a relation.

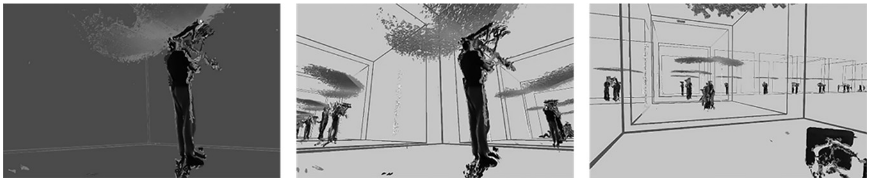

Figure 10.4 Possible futures and pasts: Possibilities are unveiled through the sides of the cube, where possibilities are organised in a 3D space.

While working with Rui Lopes Graça, we focused on how to transpose the functionalities of video annotation systems into a VR environment. Specifically, we explored how stroke and voice annotations could be enhanced in 3D environments. Regarding stroke annotation in a video scenario, a user is only able to annotate what is focused and foregrounded in the video, therefore losing the information of spatial organisation and the ability to relate different dancers or objects that are in the same line of depth.

To augment this type of annotation, we took advantage of the 3D environment, specifically implementing the possibility to annotate directly on the skeleton data (specific joints) and the point-cloud representation at specific frames, supporting static and dynamic annotations (Figure 10.5) (see Video 8, digital Support Material online).

With Sylvia Rijmer, we explored how different dance data could be visualised and organised in a VR environment in a manner that would also express the choreographers' ideas and concepts. We created a virtual environment of a gallery or exhibit hall, composed of three 3D architectural structures (Figure 10.6 left) corresponding to each individual dancer, arranged in four different "floors" or levels, and one 3D structure composed of three levels where the user can experience the trio's compositional material (see Video 9, digital Support Material online).

On the ground level, users can experience and experiment with different visual representations of the graphic elements of *Treatise*, a piece composed by Cornelius Cardew,[3] which were selected and "danced" by the dancers during

Figure 10.5 From left to right: an example of contextualized drawing; interaction with the main menu; and 3D drawing.

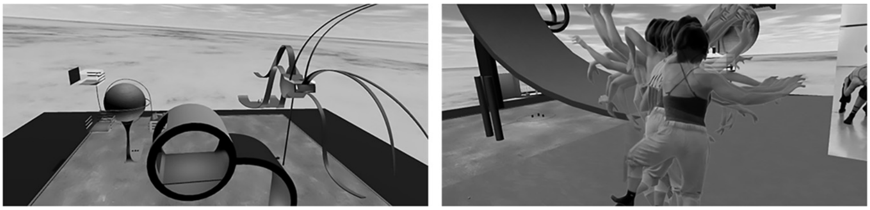

Figure 10.6 From left to right: a general visualisation of the 3D environment; a representation of a dancers' movement using high-speed photo animation.

the rehearsals with Rijmer (see Rijmer, 2022). Based on that experience, each of the three dancers created a solo composition and a trio piece, which were captured using different media, namely 2D video, high-speed photography, stereoscopic photography, and motion/depth sensors. In each of the upper "floors" or levels of the dancers' 3D structure, users can explore interactively different aspects of their solos, either by visualising individual movements represented as high-speed photo animations (Figure 10.6 right) or annotated videos (which users are invited to follow or to create their own original material) and finally, visualise a 3D representation of the complete solo. The trio material is similar, with an added possibility to interact with 3D representations of the different graphic elements and compose new material by manipulating point-cloud data. For a more detailed description, please refer to Jürgens et al. (2020).

10.4 Conclusions

Using real-world data to enhance virtual experiences is becoming more and more commonplace in our increasingly tech-infused society. During the BlackBox project, we were able to introduce this technology to the fields of dance and present concept prototypes and applications of how future researchers and developers can create new experiences based on these novel paradigms. We believe that as our research shows, this technology is mature enough to start being used in other contexts of the performing arts.

As technology advances in the future, we expect 3D scanners to be as ubiquitous as common cameras are. This will allow more users to capture data in more varied scenarios that are currently not supported (e.g., street performance). Moreover, as the quality of head-mounted displays for virtual and augmented reality experiences increases, we expect these experiences to be accessible outside research labs, and with considerably higher visual quality. Finally, machine learning progresses on body tracking will diminish the dependence on high-end mo-cap studios and devices to perform structured tracking. We believe these will allow the next generations of choreographers and dancers to work more closely with immersive technology.

10.5 Rights and permissions

Funding

This work was supported by the European Research Council under the project "BlackBox: A collaborative platform to document performance composition: From conceptual structures in the backstage to customisable visualisations in the front-end" (ERC Grant Agreement No. 336200). This publication was in part supported by FCT through project AWESOME, ref. PTDC/CCI-INF/29234/2017, and the LASIGE Research Unit, ref. UIDB/00408/2020 and ref. UIDP/00408/2020.

Notes

1 http://blackbox.fcsh.unl.pt/home.html
2 For more detailed information on the VR study with the title "Dance in Virtual Reality – deconstructing choreographic objects through expanded media": http://blackbox.fcsh.unl.pt/a-vr-dance-study-with-sylvia-rijmer-body-logic-in-virtual-reality.html
3 https://en.wikipedia.org/wiki/Treatise_(music)

References

Anjos, R. K., Pereira, J., & Gaspar, J. (2018, December). A navigation paradigm driven classification for video-based rendering techniques. *Computers & Graphics*, 77, 205–216. https://doi.org/10.1016/j.cag.2018.10.017

Anjos, R. K., Pereira, J. M., Gaspar, J. A., & Fernandes, C. (2017). Multiview layered depth image. *Journal of WSCG*, *25*(2), 115–123.

Anjos, R. K., Ribeiro, C., & Fernandes, C. (2018, June). Three-dimensional visualization of movement qualities in contemporary dance. *Proceedings of the 5th international conference on movement and computing*. https://doi.org/10.1145/3212721.3212812

Anjos, R. K., Ribeiro, C., Lopes, D. S., & Pereira, J. M. (2017). Stroke-based splatting: An efficient multi-resolution point cloud visualization technique. *The Visual Computer*, *34*(10), 1383–1397. https://doi.org/10.1007/s00371-017-1420-7

Fernandes, C., Evola, V., & Skubisz, J. (2022). What makes dancers extraordinary? Insights from a cognitive science perspective. In C. Fernandes, V. Evola, & C. Ribeiro (Eds.), *Dance data, cognition, and multimodal Communication* (pp. 244–273). Routledge. https://doi.org/10.4324/9781003106401-21

Jürgens, S., Fernandes, C., & Anjos, R. K. (2020). Experiencing choreographed 4D visuals in a VR dance installation. *Performance Research*, *25*(4), 94–97. https://doi.org/10.1080/13528165.2020.1842603

Jürgens, S., Fernandes, C. M. D. J. M., & Henriques, F. (2016). Re-constructing the choreographic studio of João Fiadeiro through animated infographic films. *PARtake: The Journal of Performance as Research*, *1*(1), 1–16.

Ribeiro, C., Anjos, R. K., & Fernandes, C. (2017). Capturing and documenting creative processes in contemporary dance. *Proceedings of the 4th international conference on movement computing - MOCO '17*, 1–7. https://doi.org/10.1145/3077981.3078041

Ribeiro, C., Anjos, R. K., & Fernandes, C. (2018). Virtual reality annotator: A tool to annotate dancers in a virtual environment. *Digital Cultural Heritage*, 257–266. https://doi.org/10.1007/978-3-319-75826-8_21

Ribeiro, C., Anjos, R. K., Fernandes, C., & Pereira, J. (2016). 3D annotation in contemporary dance. *Proceedings of the 3rd international symposium on movement and computing – MOCO '16*, 1–4. https://doi.org/10.1145/2948910.2948961

Rijmer, S. (2022). Negotiating deliberate choice- making: Insights from an interdisciplinary and multimodal encounter during the making of a New Contemporary Dance. In C. Fernandes, V. Evola, & C. Ribeiro (Eds.), *Dance data, cognition, and multimodal communication* (pp. 15–37). Routledge. https://doi.org/10.4324/9781003106401-3

Sousa, M., Mendes, D., Anjos, R. K., Medeiros, D., Ferreira, A., Raposo, A., Pereira, J. M., & Jorge, J. (2017). Creepy tracker toolkit for context-aware interfaces. *Proceedings of the international conference on interactive surfaces and spaces – ISS '17*. https://doi.org/10.1145/3132272.3134113