A Webcam Interface for Somatic-Technological Dance Experiences

Joana Martins
Instituto Superior Tecnico, Universidade de Lisboa, Portugal
joanabmartins@gmail.com

Todd Cochrane
Digital Technologies, NMIT, Nelson, New Zealand
Todd.Cochrane@nmit.ac.nz

Isabel Valverde
Research Centre of Arts and Communication, Lisbon Open University, Portugal
icvalverde@gmail.com

Ana Moura Santos
Instituto Superior Tecnico, Universidade de Lisboa, Portugal
ana.moura.santos@tecnico.ulisboa.pt

Abstract

This paper discusses a recent mixed reality interface developed within the collaborative art-technology project Senses Places, a participatory performance environment. The present webcam interface works by motion detection algorithms created through the frame difference method and responds to a critical need from the somatic-technological dance approach. Captured by a webcam, the participant's movements are exported into a 3D Virtual World where they trigger pre-set animations in an avatar, according to choreographic improvisation principles. Senses Places fosters mixed reality performance events that gather performers and visiting physical and online participants from all over the world in a virtual place, where they engage kinaesthetically in new shared embodied mediated experiences towards expanding awareness to cross geographic, cultural, disciplinary, artistic and human boundaries.

Introduction

Senses Places (SP) is a long term dance-technology project, producing new and transgressive artistic concepts, designs, and experiences aiming to bring the focus to *posthuman corporealities*. [1], [2], [3] The project's somatic-technological dance approach is geared to expand, instead of shrinking, the subjects' physical body communication capacities through its virtualizations. It departs from dance, performance and new media theory-practice studies, Eastern-Western contemporary movement and dance practices, such as Contact Improvisation and Butoh, interweaving somatic practices, like Body-Mind Centring, Release, and Alexander Techniques, and ancient Tai Chi and Yoga. [4], [5]

Through scheduled events, welcoming anyone who wants to join, participants from all over the world, with different cultures and areas of expertise meet virtually and perform together. They focus on the perceived sensations from theirs and others' physical, video recorded and avatar embodiments and the overall hybrid environment. The main collaborators come from Portugal, New Zealand, Japan, Newfoundland, the Netherlands and Brazil, intensifying and diversifying the cross-cultural and disciplinary somatic-technological performance experimentation.

Background

SP engages with multiple areas of study, both artistic and technological. Within the technological field, we approach subjects related to computer vision and more recent with biometric statistics. For example, through a biometric device, we are able to connect the participant's cardiac beat with the lights and sounds in the physical and in virtual worlds.

Interface device creative experimentation focuses on choreographing/designing the interactivity of inter-subjective experiences. [1] The SP project wants to bring attention back to the participants' and avatars' embodied interaction with and through their environment, producing kinaesthetic exchanges. [6] The quality of the experience is raw, unfinished, unpolished and emergent, due to the multiple components involved and its participatory nature, conducive to artistic creation. [7]

SP started in 2007 as Real Virtual Games, focusing on breaking video game interface stereotypes through cooperative embodiment interfacing. [1] With the integration of 3D Virtual World (3DVW) Second Life® (SL), a Multi-User Virtual Environment (MUVE), it got more ambitious, developing avatar animations, through motion capture and spreadsheet methods and creating new interfaces and gadgets that involve connections via Webcam, Wii Remote, Kinect and a biometric device.

First performance experiences experimented with physical world video streams fed into the virtual world, while the virtual world was projected back into the physical world, converging embodiment and realities (see Figure 1).

Although initially the avatar's movements were predetermined, from the very beginning of the project the concept of perceiving the avatar as a dance partner with whom it was possible to interact helped the participants to move in an open improvised manner. That was the motivation to create new ways to merge both worlds via a webcam, among others interfaces.

Fig 1. *A dancer testing a biometric device with responsive colour LEDs and interacting with an avatar in an SL projection,* 2010, Isabel Valverde, original snapshot.

Choreographic Approach

The project's inclusive cross-cultural intertwining of technological and somatic-based contemporary dance aims to aesthetically intervene in the creation of new channels for kinaesthetic empathy and communication through/between our senses and places, opposing the generalized physical body instrumentalization and replacement. [1] The overall dynamic, rooted in kinesthesis, is conducive to the participant's increasingly shifting and encompassing focus within the complexity of the intermodal embodied experience.

By engaging in this atmosphere where improvisation is crucial, performers are stimulated to move in ways that might be unfamiliar. It encourages the participants to go out of their comfort zones and explore other embodied engagements. We approach a new way of choreography as an interactive experience based on the unpredictability, resulted by the human-avatar and the avatar-avatar connections. The movement flows from following a multiplicity of internal and external bodies' stimuli, amplifying the awareness of each moment of the immersive experience and intensifying the unexpected, intersubjective and collective engagements.

Webcam Interface

The webcam interface within SP is used to expand the subject's movements by triggering different animations on an avatar, and was first presented as Interface 2.0 in 2010. [8] It provides a way to sense the participant's upper body movements and send the information about which part of the body moved into the MUVE. There, an avatar animation is activated accordingly (see Figures 2 and 3). By doing so, the webcam interface contributes to connect the participants to their avatars and to each other, creating a hybrid embodied improvisation environment.

Since 2010, this interface has been generally used in all SP performances and by anyone in SL with access to a webcam, spreading this experience worldwide.

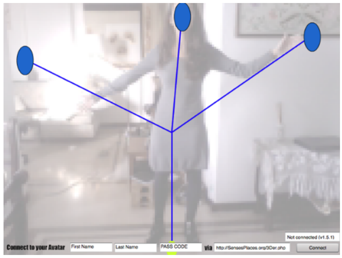

Fig 2. *The webcam interface developed in 2010,* 2015, Joana Martins, original snapshot.

Developing the Webcam Interface

The webcam interface is implemented on a webpage in an embedded Adobe® Flash® component. This provides access to the interface from any web browser on the Internet. Participants after entering in a SP session are offered, from within the 3DVW, a script that is an object to be attached to their avatar as a heads up display (HUD). When the HUD is attached to their avatar, they receive the corresponding instructions of how to login to the webcam interface via a webpage. After permission to use the webcam is given, the webpage displays a video of the webcam view with a superimposed graphical abstraction of a "skeleton" (as in Figure 2) with circles connected by lines, representing a possible location for detected head, hands and torso in movement.

The interface is treated as a game controller that causes action in the 3DVW. All detected movements by the webcam are naively treated as anthropic movement. People using the interface tend to adjust their position to match the layout of the "skeleton", or alter this correspondence through improvisation, by moving out of the vertical axis and approaching or distancing from the webcam (see Figure 3). Yet another mode, disguising the "skeleton", includes several people moving in space within the range of the webcam triggering more animation on sole avatar.

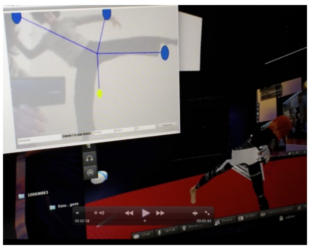

Fig 3. *A dancer interacting with an avatar through the webcam interface,* 2010, Isabel Valverde, original snapshot.

The data supplied to the 3DVW with SP's 2010 webcam interface (see Figure 2), while sufficient for activating some upper body movement, were missing details about the upper body, and no data about the lower body were provided. Hence, in the triggered animations, detailed movement of the upper body and all of the lower body in the 3DVW are created as an optional choreographic "effect" rather than matching control by participants.

Improvement in the Webcam Interface

The webcam interface is being developed further using a frame difference method with HTML5 as the implementation platform. While the platform continues to provide access to the interface through a webpage (as described above), it also has now the potential to run on any device running an advanced web browser. [9] Participant's movement detection presently includes the lower body pelvis and members, leading to more complete full body data in the 3DVW. We started experimenting with eleven body parts (head, hands, feet, knees, elbows, pelvis and chest), but as the image got too chaotic with the simultaneously movement of so many nodes, we decided to eliminate the nodes corresponding to elbows and knees, achieving the final representation shown in Figure 4.

Fig 4. *SP webcam interface with seven nodes mixed with streaming, during a performance,* 2015, Isabel Valverde, original snapshot.

The algorithm created to recognize and track the movement is based on the frame difference method. This technique starts from the assumption that the background is static and compares the pixels of the frames captured in very small time intervals (Δt). [10] If the absolute difference comes to a value above the defined threshold (Γ), it means that the pixel has changed and therefore there was movement.

This method can be described by the following formula:

$$D_k = \begin{cases} 1, \text{ if } |f_k - f_{k-1}| \geq \Gamma \\ 0, \text{ if } |f_k - f_{k-1}| < \Gamma \end{cases}, \qquad (1)$$

where D_k represents the binary result of the absolute difference of a pixel between two consecutive frames (f_k and f_{k-1}). The difference of all frame pixels gives us a collection of binary numbers that, ultimately, will translate the movement (see Figure 5).

Fig 5. *Binary image of a hand movement*, 2015, Joana Martins, original image.

This technique allows a good recognition of the movement, but it was also important for our final objective that it could also distinguish which different body parts are moving. However, this has proven to be a very difficult task, involving complex algorithms that are not suitable for real time online performances. [11], [12] Therefore, presently the only actual body part that we are able to distinctively detect with an acceptable computational efficiency is the head. The human face has unique features that allow its quick recognition, besides it does not change its shape drastically, as e.g. hands do. In order to implement the head detection, we used the well-known object detection method developed by Viola and Jones. [13] When the face is obstructed or it cannot be recognized for some reason, the interface resumes by applying the framed difference algorithm. Attempting to distinguish somehow the movement of the other body parts, we divided the overall frame into different regions, as can be seen in Figure 6. Some areas are shared by different body parts, and in those overlapping areas the node that is responsible for the movement is the one that is closest to that region. Moreover, from a region frame to another (Δt), there is a maximum distance that a given part of the skeleton can move.

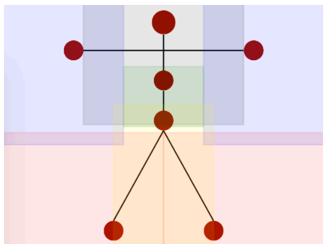

Fig 6. *Regions that detect the different body parts,* 2015, Joana Martins, original image.

As a result, this new version of the webcam interface makes a more reliable recognition of the human movement, which grants more freedom to the participants, allowing them to express themselves candidly.

Improvement of the Virtual World Scripts

The connection to the 3DVW of the new version of the interface is made almost in the same way as for the first webcam interface. When a movement is detected, the information about which node moved is sent to a PHP script on a web server. With this new version, we also send the coordinates to where the node moved to the PHP. Then, the script in the virtual world pulls this information to animate the avatar.

Firstly, together with the first implementation of the interface, the 3DVW received information about what body part moved and then it randomly selected between four animations pre-set to be connected to that node. Now, we implemented two different kinds of 3DVW scripts. One script is based on the previous one, and it's only concerned about which node has moved. After that recognition it selects arbitrarily one of the four animations pre-selected for that boy part. The second script also detects which node has moved, but afterwards triggers a related animation to the corresponding spatial location of the node. We also created animations in the biovision hierarchy (BVH) format to better establish a correspondence between where the nodes are and what animation the avatar performs.

As a result, the new webcam interface, providing data with a much more detailed level, allows participants to have a finer degree of control on avatars' animations, which may provide more engagement and a stronger sense of immersion (see Figure 7).

Fig 7. *SP interface with seven nodes in connection with SL, in a SP' performance,* 2016, Joana Martins, original snapshot.

What's next?

In the near future, we plan to create more animations for the SL platform to better integrate the participants' movements with their avatar. With further use of motion capture, we will be able to create better animations that more accurately suit our somatic-technological choreographic approach.

A new improvement that we hope to achieve soon is to give more freedom to the avatar. Presently, there are interfaces that trigger different animations (using a webcam) and there is one that controls the avatar's movement in space (using a Wii Remote). The ideal interface would join both these strengths into one. By doing so we could more easily perceive the avatar as a dance partner that moves along with and in relation to us, and to other avatars through space and time, creating a richer sensitive and emergent environment with more choreographic options.

This project innovates in many ways, both artistic and technological. The system developed for the new interface can also be applied in engineering, to act like a sensor, to control and to simulate e.g. different scenarios improving computer vision. These functional extrapolations bring agency to the influence of SP's inventive and original nature.

References

1. Isabel Valverde, "Senses Places: developing a somatic dance-technology approach", unpublished paper. Accessed November 10, 2015, https://www.researchgate.net/profile/Isabel_Valverde/contributions
2. Katherine Hayles, *How we Became Posthuman: Virtual Bodies in Cybernetics, Literature and Informatics,* (Chicago: The University of Chicago Press, 1999).
3. Susan Leigh Foster (editor), *Corporealities: Dancing Knowledge, Culture and Power*, (Routledge, 1995).
4. Bonnie Bainbridge Cohen, *Sensing, Feeling, and Action: The Experiential Anatomy of Body-Mind Centering* (2nd ed. Contact Editions, 1993).
5. Irmgard Bartenieff, *Body Movement: Coping with the Environment* (Routledge, 1980).
6. Isabel Valverde and Todd Cochrane, "Weathering In/Com Tempo: An intervention towards embodying multimodal self-organizing environments," *Proceedings of the SLACTIONS 2009 International Conference - Life, imagination, and work using metaverse platforms: 24-25 September (SLACTIONS* 2009).
7. Isabel Valverde and Todd Cochrane, "Innovative Dance-Technology Educational Practices within Senses Places", *SLACTIONS 2013 Research conference on virtual worlds - Learning with simulations: 21-23 November*. Accessed December 10, 2015, http://www.sciencedirect.com/science/article/pii/S2212017314000267
8. Isabel Valverde and Todd Cochrane, "Physical-Virtual Body-Game Interfaces", *Interface 2 (2010)*, Dance Tech website. Accessed January 5, 2016, http://www.dance-tech.net/profiles/blogs/interface-20-second-life
9. W3C, "HTML5", W3C website. Accessed January 5, 2016, http://www.w3.org/TR/html5/
10. Nishu Singla, "Motion Detection Based on Frame Difference Method, *International Journal of Information & Computation Technology",* Vol. 4, No. 15. Accessed November 24, 2015, http://www.irphouse.com/ijict_spl/ijictv4n15spl_10.pdf
11. Ralf Plaenkers and Pascal Fua, "Model-Based Silhouette Extraction for Accurate People Tracking", in *Lecture Notes in Computer Science,* (Berlin and Heidelberg: Springer-Verlag 2002), 2351.
12. Chen Wei Chao, "Real-Time Palm Tracking and Hand Gesture Estimation Based on Fore-Arm Contour", (Master Diss., National Taiwan University, Department of Computer Science and Information Engineering).
13. Paul Viola and Michael Jones, "Rapid object detection using a boosted cascade of simples features," (paper based on a talk presented in 2001). *IEEE Computer Society Conference on Computer Vision and Pattern Recognition.*