LEARNING TO DANCE WITH A HUMAN.

John McCormick, Motion.lab Deakin University. Email: john.mccormick@deakin.edu.au

Kim Vincs, Motion.lab Deakin University, Centre for Intelligent System Research. Email: kim.vincs @ deakin.edu.au

Saeid Nahavandi, Centre for Intelligent System Research Email: saeid.nahavandi@deakin.edu.au

Douglas Creighton, Centre for Intelligent System Research Email: douglas.creighton@deakin.edu.au

Abstract

Artificial neural networks are an effective means of allowing software agents to learn about and filter aspects of their domain. In this paper we explore the use of artificial neural networks in the context of dance performance. The software agent's neural network is presented with movement in the form of motion capture streams, both pre-recorded and live. Learning can be viewed as analogous to rehearsal, recognition and response to performance. The interrelationship between the software agent and dancer throughout the process is considered as a potential means of allowing the agent to function beyond its limited self-contained capability.

Keywords: Software Agent, Artificial Neural Network, Dance and Technology, Distributed Cognition, Machine Learning, Interactive Performance.

Introduction

In creating dance performances incorporating live motion capture of dancers within a projected stereo 3D environment, questions arose regarding the integration of immersive digital sound and visual environments as a component of the live dance performance. Motion data can be used as a direct source for visualisation without any analysis by the software environment. While this has led to many satisfactory results within the performances, it raised the possibility that the software environment could have a greater capacity for interpreting and responding to the dancer's movement. This would enable a more complex performing relationship between the dancer and software agent to develop. Artificial intelligence techniques have been used to visualize, sonify, and respond to dancers' movement in performance for many years [1, 2, 3, 4, 5]. We wanted to explore whether it is possible to develop a performance agent that can participate in some way within the choreographic process as well as within the performative outcome. Seminal work in this area has been done by Marc Downie [6] and OpenEndedGroup [7]. However, whereas their work did not make use of skeletal representations for the agent, we have chosen to use the same internal skeletal representation for both the dancer and agent so as to give them both a measure of equality in how they are viewed and represented. We can then use the dancer and agent's movement streams somewhat interchangeably. In this paper we have used a humanoid representation to illustrate the movement of both dancer and agent, for purposes of clarity. Both the dancer and agent are represented by similar avatars (their visual embodiment). In further artworks the representation can be markedly different.

We investigated the Artificial Neural Network(ANN) as a means of allowing the agent to learn movements from the dancer and subsequently recognise and respond using the learnt vocabulary. We modelled the performance development on a sequence of events familiar to the dancer. In our model, movement material is generated by a dancer through a process of selective improvisation. The improvised sequences are recorded and passed to the agent to learn as in a rehearsal. After the agent has learnt the material, the dancer improvises with the trained agent and the movement vocabulary is refined based on the responses from the agent. The refined movement choices become the vocabulary for the performance between dancer and agent. There were two main goals attached to choosing this performance making process. One was to try to integrate the development of the agent into a fairly typical dance developmental process. The other was to allow the experience of the dancer to support the agent as much as possible at all stages in order to maximise its capabilities in performance.

The relationship between the dancer and agent was viewed through the lens of situated cognition [8] and socially distributed cognition [9, 10] as a means of framing their interdependent relationship. Situated cognition suggests that cognition cannot be separated from the context in which it exists. Distributed cognition holds that knowledge can exist not only in individuals but also in their social and physical groupings. A cognitive ecosystem comprising two or more agents allows cognitive processes to be distributed amongst its members [11].

Designing a performing agent

Matt Carter in *Minds and Computers* writes:

... embodied experience was a necessary condition for the development of semantics, which, in turn, are necessary for having a mind. Consequently, if we want to develop an artificial intelligence it must, in the first instance, be connected to the external world in the relevant ways. In other words, it must enjoy sensory apparatus which mediate the relations between it and the external world. Furthermore, our embryonic artificial intelligence must then be able to gather a weight of experience, through which it will be conferred with mental representations. [12]

Memory, the weight of experience, is seen by Carter as a fundamental building block upon which mental representations may be constructed. Carter also introduces two other key concepts, that of embodied experience through sensory apparatus and the fundamental relationship between the agent and environment.

In dance, embodied experience and hence memory is embedded within the morphology of the human body. Memory in dance is procedural, in the

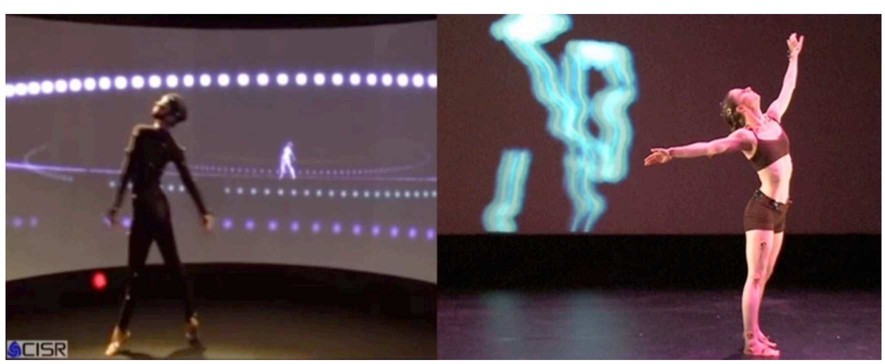

Figure 1. The agent had access to both Optitrack (left) and Motion Analysis (right) motion capture systems for sensory input in live performance environments. Movement sequences captured with these motion capture systems were presented to the agent's neural network during the learning phase. Image © John McCormick 2012

sense that, like expert movement in other elite professions, once learned, complex dance movement phrases are performed without conscious cognitive awareness [13]. Memory, in this case, is enacted only through moving one's body. Erin Manning [14] argues that dance movement is also inherently relational, proceeding from a 'pre-acceleration' that defines intentionality in relation to the world and to other people as well as trajectory. She describes dancing a duet with another person as "…not a learning by heart. It is not a choreography. It is improvising with the already-felt" [15]. Manning's argument suggests that the procedural nature of dance memory does not imply that dance performance is fixed by the past, but rather that the body memory of past movement is brought to bear on the present moment. This process is constituted in terms of felt and experienced physical morphology and structural (skeletal) organisation, because memory encompasses the sensation of movement rather than simply a linguistic encoding of the pathways of joints and limbs in space.

If our agent were to participate in a performance process, it would need access to a form of memory constituted in and by the parameters of human movement. We used full-body motion capture to provide both input and an interactive mechanism for the agent, its sensory apparatus. Two different systems were used, Motion Analysis and Optitrack, illustrating that our agent was independent of specific motion capture systems. The motion capture systems became the sensory input mechanism for the agent, providing it with a distilled view of the dancer's movement that functioned as the agent's source of experience and of sensory connection to the environment (Figure 1). Both motion capture systems are multi-camera optical systems which

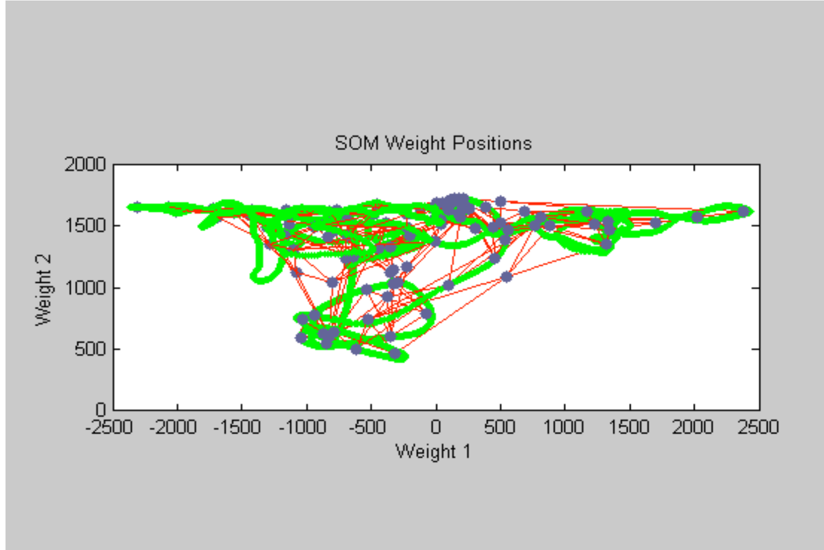

Figure 2. SOM_Weights. The green represents the input, the grey represent the neuron weights. Over time the neuron weights move to match the inputs. This a 2D representation of only the first 2 vectors of 57. In the final network there are 79 vectors or pieces of information that describe the skeletal information. Image © John McCormick 2012

captured the positions of reflective markers on the body at 120 fps. The MAC system used 40 markers, the Optitrack 34 markers. The marker positions were used to construct a skeletal representation of the dancer in order for the agent to view the dancer in terms of movement of the body and limbs. This also allowed the agent to respond through its own avatar using the skeletal movement it had learned.

Our synthetic agent was tasked with the goal of being able to recognise a live dancer's movement and responding via animating a 3D avatar with the movement vocabulary it had learnt. In our search for models of intelligence that might guide the development of an ability to synthesise movement elements into sequences, the areas of situated *[8, 16]* and distributed cognition [9, 10 11], seemed to offer an appropriate, if challenging, paradigm within which to explore the development of an intelligent agent within a live performance context. Situated cognition, with its premise of extremely tight coupling of cognitive processes to the environment, seemed potentially aligned with both the desire to more closely couple the dancer and performance environment and the development of a synthetic agent which could also respond intelligently to its environment of which the dancer is the major part. Distributed cognition has been applied to studies of social remembering and cognition, notably between couples [9,10, 11]. We were interested to see if this framework could be extended to include the agent – human relationship.

Situated cognition is broadly based on connectionist models of cognition rather than a computational model of storage and retrieval. Connectionist models favour concepts of neuronal plasticity and deep parallelism of atomistic processes to account for complexity of behaviour. For the synthetic agent, the closest model analogous to the neuronal component of cognition are Artificial Neural Networks (ANN) which have been used extensively to model cognitive processes including human gestures [17,18]. In particular, unsupervised ANNs are employed as they are currently viewed as a close representation of real neural processes [19]. Unsupervised methods of learning allow the neural network to find its own associations within the data presented to it (in this case movement encapsulated within motion captured

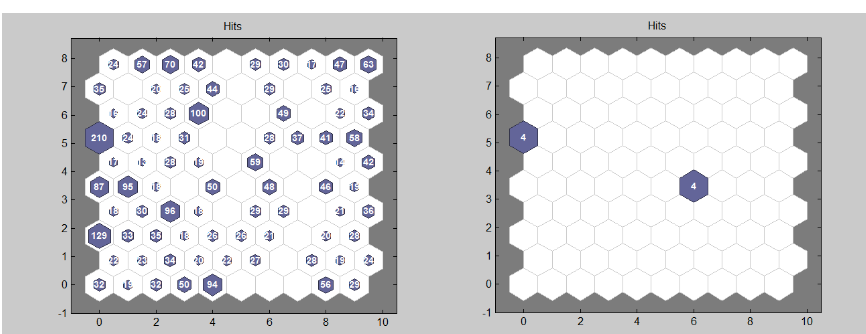

Figure 3. SOM Neuron Hits. Dance sequence as learnt by the network (left) and neurons triggered when test frames are introduced to the trained network (right). Similar postures are clustered within particular neurons as indicated by the numbers of similar postures captured by the neurons (left). Two short sequences of known postures, when introduced to the network, fired the corresponding neurons containing like postures (right). Image © John McCormick 2012

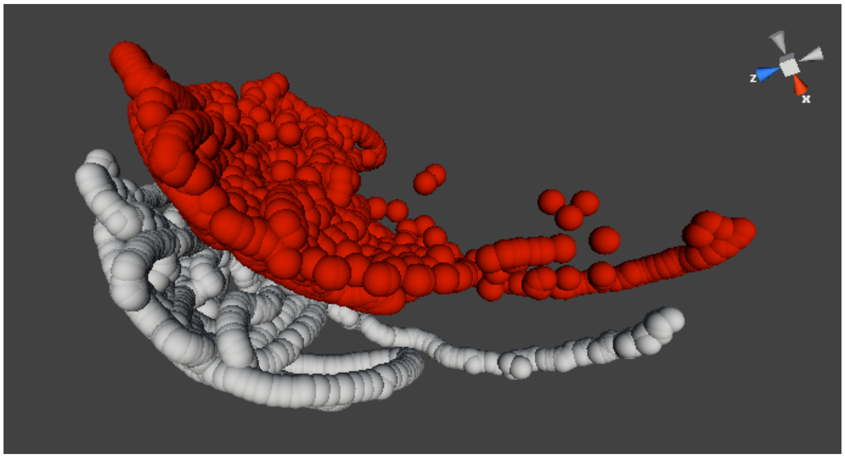

Figure 4. Visualisation of 3 dimensions out of 79, of the input data (grey) and neuronal weights (red) of a movement phrase captured with the Optitrack system and displayed in a 3D game engine. The offset is artificial for visualisation purposes, the neuron weights and input dance data are directly overlaid on top of each other. During learning the information contained in the neurons increasingly matches the movement information. Image © John McCormick 2013

sequences). This differs from supervised methods, which encourage the network to converge on an optimal solution to a known problem.

The somewhat ambiguous or latent potential offered by unsupervised learning could allow the agent greater scope for variability as opposed to known responses when synthesising movement, a feature identified by Kirsh [20] as beneficial to creativity. Kirsh suggests the high level of innovation choreographer Wayne McGregor achieves through setting his dancers tasks, arises from distributing the creative process beyond the limits of his own body and mind, allowing him to recruit ideas from a larger and more diverse pool of creative possibilities. However, this distribution of creativity is more than simply a way of sourcing external inspiration for movement invention, as in Merce Cunningham's throw of the dice [21] or Trisha Brown's alphabet cube. [22] This is not to say that Cunningham and Brown did not also engage in responsive processes with their dancers, however the Kirsh study documents these reciprocal processes in a formal study. Tasks described by Kirsch include …*imagine that their bones are made from firm rubber, or that they should imagine the feeling of being attacked. Their task is to translate those feelings into movements*. [20] In the processes, as Kirsh describes, creativity is accomplished collaboratively between McGregor and his dancers through reciprocal and responsive processes, rather than by the dancers simply providing a larger pool of ideas for McGregor to choose from.

Employing a situated approach whereby the human and synthetic protagonists become part of an extended cognitive system in our process allows the synthetic agent to be supported by the human performer's processes, and potentially allows it increased scope over the relatively rudimentary capabilities it would have as a self-contained entity. This is not unlike human dance development processes in which dancers typically learn by dancing with more experienced artists over a number of years, learning by directly experiencing the embodied knowledge of others.

To enable the agent to gather a weight of experience, a persistent memory of the dancer's movement, a type of Artificial Neural Network known as a Self-Organising Map (SOM) was used [23]. The SOM is an unsupervised ANN in that the network is presented with the movement data without any type of labelling, and finds its own associations within the data. The initial experiments were undertaken in Mathworks Matlab using the Neural Network Toolbox. Initially, a 100 neuron network was presented with a recorded sequence of dance. Over 100 iterations, the network was able to classify similar postures found within the sequence into clusters contained within specific neurons. The input sequences consisted of frames of movement defined as position and joint rotation information describing body postures. There were 99 pieces or dimensions of information to describe each frame of movement for the Optitrack system, and 161 dimensions for the Motion Analysis system. The information or weights contained in each neuron gradually changed under the influence of the input data until the neuron weights closely matched the movement inputs (Figure 2). When the network was subsequently presented with individual postures, the neuron containing similar postures was stimulated demonstrating the network's ability to learn aspects of the movement phrase and to then use this learning to recognise similar postures (Figure 3).

Results

The results using pre-recorded data were very promising. The SOM could learn to cluster similar movement postures into groups within specific neurons within the network, and these neurons responded to incoming movement postures, allowing the remembered movements contained therein to have focus.

The next stage positioned this capability within an agent running in a 3D game environment with live input from a human dancer. We used the Unity game engine and developed the agent and data streaming components in C# and C++. The data stream was reduced to 79 di-

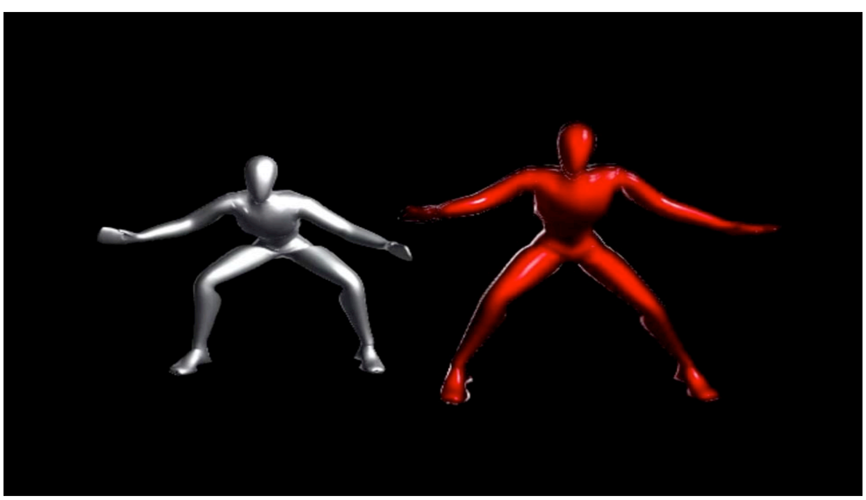

Figure 5. Neural Network Agent (red) responding to the live dancer (silver) with the closest match from its learnt memory. Image © John McCormick 2013

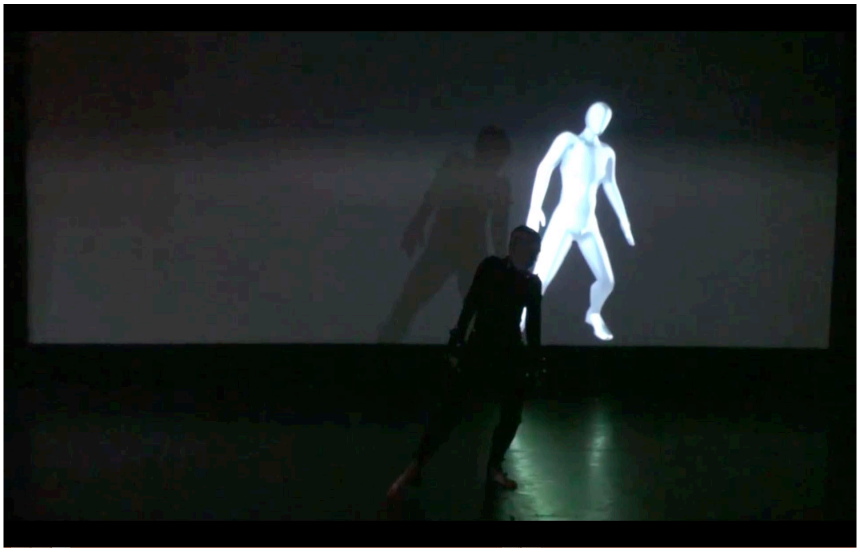

Figure 6. Neural Network Agent responding to live dancer with learnt movement from memory. Image © John McCormick 2013

mensions, optimised for the live context. The dancer created some improvised movement sequences while considering that the agent would use them to learn a shared movement vocabulary, use the learnt movement to recognise what she was performing and follow her, and use the learnt material to generate movement sequences in response to movement "seeds" (postures) that she provided. The agent's neural network was first exposed to the recorded sequences of movement, akin to a rehearsal stage in a creative context, and allowed to learn to classify individual movements contained in the dance (Figure 4). During the learning process, a second map was introduced to capture temporal information in the form of links between neurons as they were stimulated. This created pathways through the neural network that linked movements over time and provided a basis for the agent to navigate its learnt movement memories to synthesise movement responses. Some neurons would accrue multiple possible future pathways and the initial tendency was to follow the last known good connection.

Once the network had finished the learning process, the dancer improvised with the agent to both reacquaint herself with the movement vocabulary (it was improvised not set), and to discover how the agent responded to her movements and the kinds of choices she could make in response. As a trial performance, we tasked the agent to firstly recognise the dancer's improvised movement as best it could and respond with the closest movements it had learnt (Figure 5, 6). Next the dancer could improvise and at any time provide a seed movement which the agent could use as a beginning point for a newly generated movement sequence, until the next seed (Figure 7). The first study tested the agent's ability to continually recognise particular movement postures and produce a reasonable response. The second study tested the agent's ability to use the learnt vocabulary to create appropriate movement responses to a dancer's movement cue. The agent and dancer were confined within a relatively known vocabulary emerging from a semi-improvised structure, however a typical performance might also be confined to a particular, finite movement vocabulary.

The SOM chosen for the agent's neural network proved robust in engaging with a dancer in a live performance context. While the SOM is a relatively simple neural network, the results indicate the neural network approach to learning and creating movement sequences in response to a live dancer has definite potential.

Figure 7. Neural Network Agent creating a movement sequence from a seed movement supplied by the live dancer. Image © John McCormick 2013

Agent as collaborator

Watching the agent's avatar dance with a live performer was like watching a young dancer attempting to learn from a more experienced performer. The avatar's movement closely resembled that of its 'teacher', but with subtle variations. The avatar movement seemed conceptually related to that of the human performer (in the sense that 'conceptual', here, is defined in three-dimensional, spatio-dynamic terms rather than in cultural or gestural ones), but not identical. The agent seemed to be effecting an iteration rather than a copy of the human performer's style.

Having designed the agent with the knowledge that it could draw on the intelligence of the dancer at all stages enabled us to use relatively simple processes to reach a significant outcome. The agent was able to recognise and respond in an appropriate manner to the dancer in a performance setting. The dancer was also able to proceed on a familiar creative trajectory with an understanding of how her creative work would be embedded within the relationship between herself and the agent.

Conclusion

In this project, we designed a performance agent that can become part of a collaborative, creative process, as it does a typical workflow from inception, through rehearsal to performance. Applying principles of situated cognition prompted us to view the learning acquisition of the agent in terms of increasing effectiveness in performance rather than self-contained accumulation of knowledge. The relationship with the dancer was seen as part of this effective capacity and allowed the agent to become part of a cognitive ecosystem that vastly aided its capabilities in terms of being an effective performer.

One of the possibilities for future research is the ability for the agent to better synthesise movement. While the generated movement is confined within a finite vocabulary, this is not unusual for any one performance. The vocabulary can be easily extended by the dancer recording thematic improvisations and passing them to the agent to learn. New methods of traversing the neural network to create variations on the movement vocabulary are currently under investigation.

References

1. J. Birringer, *Performance, Science and Technology* (New York: PAJ Publications, 2008).

2. S. Dixon, *Digital Performance: a history of new media in theater, dance, performance art, and installation* (Cambridge, Mass, London : MIT Press, 2007).

3. S. deLahunta and F. Bevilacqua, "Sharing descriptions of movement," *International Journal of Performance Arts and Digital Media*, 2007 3, no. 1 (2007) pp. 3–16.

4. S. delahunta, "The choreographic language agent", in *World Dance Alliance Conference*, Brisbane, Australia, 2008. <http://ausdance.org.au/articles/details/the-choreographic-language-agent> accessed July 1 2013.

5. S. Broadhurst, *Digital Practices: Aesthetic and neuroesthetic approaches to performance and technology* (Basingstoke, UK; New York, N.Y. : Palgrave Macmillan, 2007).

6. M Downie, *Choreographing the extended agent : performance graphics for dance theatre,* PhD Thesis, MIT 2005. <http://dspace.mit.edu/handle/1721.1/33875> accessed July 1 2013.

7. OpenEndedGroup, *how long does the subject linger on the edge of the volume*, 2005 Artwork <http://openendedgroup.com/artworks/howlong.html> accessed July 1 2013.

8. Robbins, P.; Murat Aydede, e., *The Cambridge Handbook of Situated Cognition*. Cambridge University Press: 2008.

9. Barnier, A. J.; Sutton, J.; Harris, C. B.; Wilson, R. A., A conceptual and empirical framework for the social distribution of cognition: The case of memory. *Cognitive Systems Research* 2008, *9* (1-2), pp. 33-51;

10. Harris, C. B.; Keil, P. G.; Sutton, J.; Barnier, A. J.; McIlwain, D. J. F., We remember, we forget: Collaborative remembering in older couples. *Discourse Processes* 2011, *48* (4), pp. 267-303;

11. Sutton, J.; Harris, C. B.; Keil, P. G.; Barnier, A. J., The psychology of memory, extended cognition, and socially distributed remembering. *Phenomenology and the Cognitive Sciences* 2010, *9* (4), pp. 521-560.

12. Carter, M, *Minds and Computers : An Introduction to the Philosophy of Artificial Intelligence*, Edinburgh University Press 2007. p. 206

13. C. Stevens & S. McKechnie, "Thinking in action: thought made visible in contemporary dance," *Cognitive Processing: International Quarterly of Cognitive Science*, Vol. 6, 243–252 (2005).

14. E. Manning, *Relationscapes: Movement, art, philosophy* (Cambridge, Massachusetts & London: MIT Press, eBook, 2009).

15. Manning [14] p. 380.

16. A.Clark, D. Chalmers, The Extended Mind. *Analysis* 1998, *58* (1), pp. 7-19.

17. G. Caridakis, K. Karpouzis, A. Drosopoulos & S. Kollias, 2010, 'SOMM: Self organizing Markov map for gesture recognition', *Pattern Recognition Letters*, vol. 31, no. 1, pp. 52-59.

18. A. Shimada, 'Gesture recognition using sparse code of Hierarchical SOM ', paper presented to ICPR 2008. 19th International Conference on Pattern Recognition, Tampa, FL 8-11 December 2008.

19. T. Kohonen , *Self-organizing maps*, 2nd ed, Springer, Berlin, New York, 1997.

20. D. Kirsh, 'Creative Cognition in Choreography', in *Proceedings of 2nd International Conference on Computational Creativity*, Mexico City, 2011.

21. S. Banes, '*Writing Dancing in the Age of Postmodernism*', Wesleyan University Press, ebook, 2011, p. 103.

22. M. Sulzman, *Choice/Form in Trisha Brown's "Locus": A View from inside the Cube*, Dance Chronicle, Vol. 2, No. 2, 1978, pp. 117-130.

23. T. Kohonen, *Self-organization and associative memory*, 3rd ed, Springer-Verlag, Berlin, New York, 1989.

Acknowledgements

The authors would like to thank the following people for their much appreciated assistance:
Steph Hutchison, Peter Divers and Daniel Skovli from the Deakin Motion.lab, Chee Peng Lim and Matthew Watson from the Centre for Intelligent System Research.

This research was supported under Australia Research Council's Discovery Projects funding scheme (project numbers DP0987101 and DP120101695)