

Embodied Design of Dance Visualisations

Harry Brenton
Embodied Audio-Visual
Interaction Group
Department of Computing
Goldsmiths, University of
London
London SE14 6NW, UK
h.brenton@gold.ac.uk

Andrea Kleinsmith
Virtual Experiences Research
Group
Department of Computer and
Information Science and
Engineering
University of Florida, US
alk@cise.ufl.edu

Marco Gillies
Embodied Audio-Visual
Interaction Group
Department of Computing
Goldsmiths, University of
London
London SE14 6NW, UK
m.gillies@gold.ac.uk

ABSTRACT

This paper presents the design and implementation of a software platform for creating interactive visualisations that respond to the free-form movements of a non-professional dancer. The visualisations can be trained to respond to the idiosyncratic movements of an individual dancer. This adaptive process is controlled by Interactive Machine Learning. Our approach is novel because the behaviour of the interactive visualisations is trained by a dancer dancing, rather than a computer scientist explicitly programming rules. In this way IML enables an ‘embodied’ form of design, where a dancer can design an interactive system by moving, rather than by analysing movement. This embodied design process taps into and supports our natural and embodied human understanding of movement.

We hope the process of designing an interactive experience for free form dance will help us to understand more about how to create embodied interfaces and allow us to build a general framework for embodied interaction. We would also like to create a compelling, embodied and enjoyable experience with more satisfying interactions than previous dance computer games which use pre-scripted routines where a player must repeat a sequence of moves.

The system was developed using a participatory methodology, with a software developer and an interaction designer working in partnership with users to test and refine two prototypes of the system. A third prototype has been built but not yet tested.

Categories and Subject Descriptors

I.3 COMPUTER GRAPHICS [I.3.6 Methodology and Techniques]: Interaction techniques

General Terms

Design, Human factors, Experimentation

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

MOCO 2014 Paris, France

Copyright 2014 ACM 978-1-4503-2814-2/14/06 ...\$15.00.

<http://dx.doi.org/10.1145/2617995.2618017>.

Keywords

Interactive machine learning, Bodily interaction, Dance

1. INTRODUCTION

Dance is an art form associated with a high degree of training and virtuosity, but it is also an activity undertaken by many people without formal training. For such an important social activity, untrained dancers performing spontaneous, un-choreographed and often idiosyncratic dance has been relatively unexplored in visualisation and interaction research compared to work exploring expert dancers[5].

There are many dance games designed for non-experts, but these mostly require players to perform simple choreographed sequences in a strict order. In this paper we describe the design of an interactive visualisation which responds to the free-form, embodied and non-analytical behaviour of dancers. Free-form dance can be defined as dancing with no planned choreography, memorisation or expected form or sequence. Our goal is to create a compelling and enjoyable experience which has richer and more satisfying interactions than previous systems that use pre-scripted routines. We also aim to create a technology framework for embodied interaction that is democratic, in the sense that it is accessible to ordinary people who dance using relatively cheap commodity sensing equipment.

Free-form dance suggests important new challenges for human computer interaction, because of the nature of the movement and behaviour involved. Unlike choreographed dance, there is no explicit or declarative representation of pre-planned movements and most people would struggle to explain in detail how they dance. Instead Kirsh argues that this knowledge is intimately connected to the physical act of dancing and our bodies are innately part of cognition that occurs during performance [10]. If our knowledge of dance is implicit and embodied it raises challenges to traditional approaches to interaction design which rely on explicit representations of behaviour that can be represented in a logical form which can be programmed into software. If we accept Kirsh’s view, then design processes must take into account the embodied nature of our skills by allowing people to design by active physicality. The idea of embodied design builds on Jacob’s [8] concept of Reality-based Interaction in which novel forms of interaction can make use of our existing skills learned from interacting with the real world. It also acknowledges that these skills can be very different from those traditionally used in computing. Although this work is influenced by Dourish [3], our definition of embodied

design is therefore rather different, and more specific than his definition of embodied interaction. To us it is design that allows people to use their pre-existing embodied skills for acting in the world.

In this paper we will investigate Interactive Machine Learning (IML) as a methodology for embodied design. IML trains data based on examples rather than code. For example, a musician evaluating an IML interface for designing gestural instruments commented: “*With [an IML system called the Wekinator], it’s possible to create physical sound spaces where the connections between body and sound are the driving force behind the instrument design, and they feel right. . . . it’s very difficult to create instruments that feel embodied with explicit mapping strategies, while the whole approach of [the Wekinator] . . . is precisely to create instruments that feel embodied.*” [6, p. 249]. This indicates that IML can change the way people think about designing, from an explicit focus on features of the movement (e.g. shoulder rotation) to a holistic, embodied view of movement that taps into a kinaesthetic understanding of movement. In summary, IML allows designers to design by moving rather than by analysing movement.

We hope the process of designing an interactive experience for free-form dance will help us to understand more about how to create embodied interfaces and allow us to build a general framework for embodied interaction. This framework could then be applied to other domains, such as video games designers building encounters with interactive characters.

2. AIM AND OBJECTIVES

The aim of this research is to create a design tool for movement based interfaces that support the freeform and embodied behaviour of dancers. This can be divided into four of objectives:

- To understand what types of interaction are satisfying for dancers.
- To develop appropriate audiovisual visualisations and interactions.
- To develop embodied methods based on IML for designing movement interfaces.
- To understand how the choice of machine learning algorithm impacts upon the design process.

3. BACKGROUND

3.1 Bodily interaction

Bodily interaction covers a range of interaction styles that make greater use of body movement than traditional mouse and keyboard input. The most common kind of bodily interaction is gesture interaction, where an interface is based on recognising particular gestures, made with the hands and arms [2, 6] or the legs [12]. However, it can include broader forms of interaction such as body activated art installations [15, 5], and dance-like interfaces for controlling music[1].

Bodily interaction in video games is now a mass-market phenomenon with the launch of controllers such as the Nintendo Wii, Sony Move and particularly the Microsoft Kinect. This opens up the possibility of video games in which players interact with characters through their natural body language and other body movements. However, current Kinect

games such as Dance Masters and Dance Central adopt the same game mechanics as other dance games, requiring players to follow pre-scripted movements on screen or remember and repeat a sequence of moves.

The dance game genre started in the late 1980’s with Dance Aerobics for the Nintendo Entertainment System and the massively popular Dance Dance Revolution which was released in dozens of arcade and console versions. Most systems use floor sensors that recognise foot taps taps, although systems such as Dance Maniax and EyeToy: Groove and also recognise hand movements using optical tracking. There are four main interactions: 1) repeat a movement indicated by an on screen prompt or a flashing floor panel in time to the music; 2) remember a sequence of moves and repeat it; 3) collaborative multiplayer; 4) competitive multiplayer. Dance aerobics also has a free composition mode where you can compose your own song by moving, but this is not a common feature.

Although bodily interaction is only beginning to be explored in video games it builds on a long tradition of research into virtual characters that can engage in dialogues with humans [9], including many experiments into developing virtual characters that respond to gestures and body movements, from K. Thorisson’s early work [16] to Huang et al’s [7] sophisticated responsive listening agent.

At first glance the scripted patterns of movement, and the limited features extracted by the sensors, impose rigid constraints upon expression. However, a review of a DanceDance tournament indicates that even though the system does not permit free-form expression, dedicated players push against the rules imposed by these systems in creative and skilful ways:

“The instructions may say ‘play this game with your feet’, but through the single and doubles freestyle competitions, we saw it played with feet, hands, knees, and elbows. We saw people actually dance, and dance well as a stream of Greats and Perfects flowed up the screen. . . . Takamichi Bristol-Lee walked up to the machine. . . and proceeded to backflip and mule-kick the crowd into a state of shock” [14].

This review demonstrates that it is possible to dance expressively within the constraints of current systems, however this expression is not recognised and fed back to the system in a two way interaction.

3.2 Interactive machine learning

IML allows users to progressively refine a training model based on interactive testing. This approach gives users more control over generated behaviours compared to traditional batch approaches where you have to wait minutes or hours to see results. For example, the Crayons system [4] enables non-expert users to create image processing classifiers by iteratively drawing on images which act as training data. Fiebrink [6] has used IML to allow musicians to design gestural and bodily interfaces for performing electronic music. Her users found this form of design to be more natural and engaging than approaches that required them to analyse gestures in terms of specific features. This seems to have been because the process was more embodied, allowing users to concentrate on the physical performance of those movements, rather than analysing abstract features of the sensor data.

4. METHOD

The system was developed using a participatory methodology [13], with a software developer and an interaction designer working in partnership with users to test and refine two prototypes of the system. A third prototype has been built but not yet tested. Prototype 1 was an interactive dance experience in which a visualisation was influenced both by an audio analysis of the music and by the movements of a dancer. It was evaluated by three non-professional dancers who interacted with the system by dancing with music of their own choice. Participants were encouraged to articulate their thoughts as they danced and then took part in a semi-structured group. The dancing session and the interviews were videotaped. Prototype 2 aimed at investigating the use of IML for designing bodily interfaces and was tested with a longitudinal study in which a physical performer designed an interface.

5. PROTOTYPE 1 - INTERACTIVE VISUALISATION

The first prototype was a real-time interactive visualisation for dancers. There were a number of visual elements on screen that responded to features extracted from the music and the movement of a dancer. The music was analysed with a Fast Fourier Transform and the resultant frequency data was used to scale and move visual elements. The dancer's movements were tracked with a Microsoft Kinect and used to manipulate the visual elements. These manipulations were hard coded mappings from simple feature of the movement (e.g. relative positions of the hands) to simple features of the visualisation (e.g. rotation and scale). The visualisation had a number of different stages with different visual effects and interactions. The transitions between these stages were controlled by a researcher via a graphical interface.

The session begins with a transparent character with stars on the main joints (figure 1a). When the music begins it is visualised as multi-coloured stars which glow brighter in response to the bass, mid and treble of the music. There are four main modes of interaction (figure 1): b) Scale by moving arms together to create a dense ball and apart to create a dispersed star-field; b) Rotate by moving arms in a scissor motion backwards and forwards; c) Throw a gravity ball by flicking with the wrists; d) Attract from two grids by moving different joints; e) Attract to fastest limb by changing the velocity of different limbs.

P1 did not intuitively understand the connection between her movements and the scale and rotate interaction: *“what does it do, just make the stars go wide... is this connected to my body, where the stars are”*(P1). After more time she came to recognise the link between her movements and the on-screen character, however she still expressed some discontent about the latency between her movements and the movement onscreen: *“It’s quite cool but I found it quite hard to move [the stars]”*(P1). This disconnect is probably because the system was tuned to recognise wide and fast movements but not the smaller, more subtle movements which were particular to P1’s dancing style. The *Attract to fastest limb* interaction was the most enjoyable because of a clear mapping between her movements and the stars on screen which closely followed the contours of her body: *“that’s cooler, because you can see the that what you do directly reflects what you are doing... I like the last one you did when it tracked*

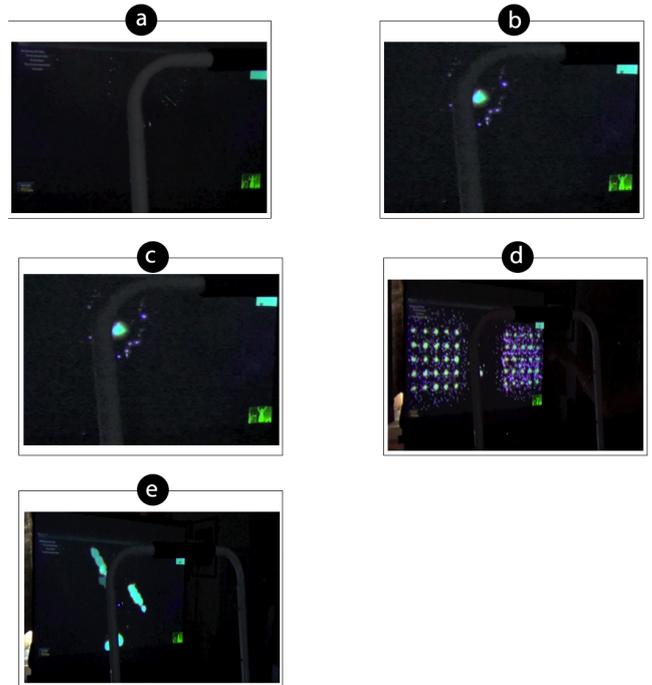


Figure 1: Images of the first prototype visualisation

me more closely, that was really fun... I could clearly see that what I was doing was having an impact... the more feedback you get the more fun it is”.

6. PROTOTYPE 2 - EMBODIED DESIGN OF MOVEMENT INTERFACES

The evaluation of prototype 1 highlighted the limitations of hardwired mappings and suggested the need for a more flexible system that could accommodate free-form idiosyncratic movements. The second prototype addressed these limitations using IML. This prototype was not aimed at dancing and used a different visualisation from prototype 1, with responses performed by an animated character (figure 2). IML allowed two participants to perform examples of interactions which trained a dataset, without any technical knowledge of programming. We performed a longitudinal study with a physical performer to understand how he would use the prototype. An initial phase aimed at understanding his working practices and revealed that his approach was to design specific actions by the user that would create specific reactions. This information was used to design the prototype. The second phase involved the performer working with our prototype to design and test an interface together with members of the public.

There were a number of positive outcomes of these studies. The performer and other participants were able to follow the process and found the interface quick to learn, as the performer commented: *“technically it is quite easy to learn, you learn it fast. All the people that came in after a few seconds everybody was able to add clips, add classes”.* However, overall, the performance of the interactive character was disappointing. While some actions were recognised reliably others were not. In many cases there was a single action that was recognised for the majority of the time.



Figure 2: Capturing examples of interaction during embodied design.

As one participant describes “about 60% of the time it did what I wanted it to do. The other 40% of the time it was insulted”. Overall, the performer was more negative about the performance, saying that the character performed the correct action only 20% of the time.

This shows that, while the performer and other participants were able to grasp the mechanics of the software they were not able to use it effectively to create reliable interactive characters. This was most clear in the difficulties that they had in debugging the model when it did not work correctly, a key issue for IML as noted by [11]. Debugging requires users to diagnose the causes of a problem and to formulate an appropriate solution. Our interviews showed that the performer and the other participants were not able to clearly identify specific causes of problems (“about the clips I’m still not clear what . . . the problem [is]”) and therefore any proposed solutions were vague and often inappropriate. This lack of concrete ideas of how to improve the model is shown in the fact that he was well aware that the performance was poor but was not able to improve it beyond his rating of 20% correctness.

The participant did not shape his movements to fit the constraints of the system. However, we see this as a limitation of the prototype rather than the performer, because ideally the system should have been about to accommodate his behaviour.

These problems seem to indicate that the software did not support our participants in understanding what was happening in the system, particularly when it went wrong, as shown by the quote: “about the clips I’m still not clear what . . . the problem [is]”. Our hypothesis is that this is due to the lack of immediate visual feedback about how the classifications were being made. Our users were not able to inspect the relationship between the training data and the classifications.

Another problem was the quantity of training data. Creating and labelling the data is an intensive process. While the participants understood the need to provide multiple examples for each action, the number provided was small, between 2 and 4 for each action. This is tiny compared with a traditional machine learning method which requires hundreds or thousands of examples. This suggests that they may have been following the strategy that [6] identify in the participants in their IML study: focusing on selecting a small number of “good”, representative examples rather than on producing a large number. This approach caused considerable difficulties when used with many learning algorithms, including the decision tree algorithms we used. Decision tree learning analyses all examples to determine which features best discriminate between classes. If there are a large number of examples with considerable variation between them, only the features that are truly representative of particular classes will discriminate between all examples of those classes. However, if the number of examples is small, it is more likely that there will be spurious features that will discriminate them. For example, a participant might happen to have his or her head at an angle in all three examples of a waving action. This would make head angle a good feature to discriminate waving on those examples, but it would not generalise well to other people waving. Another requirement of our classifier, is therefore that it can work effectively with a small number of carefully chosen examples, rather than a large data set, with randomly sampled examples.

A final problem emerged from the requirements identified in Prototype 1. The learning method used a traditional classifier which recognises discrete actions and produces discrete responses. The interaction is therefore highly quantised, to a small number of actions response pairs. This makes it impossible to produce the kind of tight coupling between movement and visuals that we identified as an important requirement of our system following our studies with prototype 1.

7. CURRENT PROTOTYPE

This section will describe the design of a third prototype that uses interactive machine learning based on the results of prototype 2 as a tool for designing motion to visual mappings similar to those used in prototype 1. This new prototype uses an Oculus Rift Virtual Reality head-set to create a more immersive experience of dancing *with* the visualisation. This greater immersion is likely put greater demands on the immediacy of the interaction and feedback.

This section will discuss some of the changes made to the interactive machine learning approach before discussing how it is applied.

7.1 Algorithms for small data

One key issue identified with prototype 2 as the quantity of data provided. An algorithm that creates a model of the data is reliant on a large data set to generalise effectively, but with sufficient data, can avoid overfitting, i.e. taking excessive account of atypical data items. However, our initial study and Fiebrink [6] have shown that users of interactive machine learning systems do not approach the task by providing large amounts of mixed quality data, as would be typical in a batch system. They tend to select a small number of data items, often discarding many items [6], but put a lot of effort into ensuring that the examples they do provide

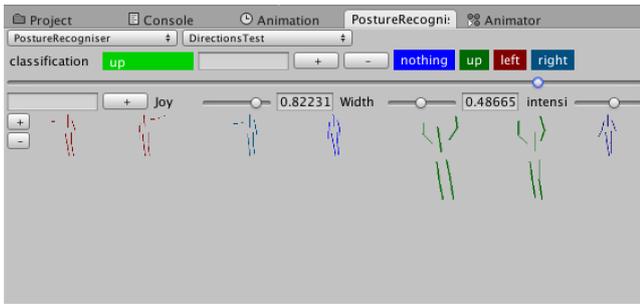


Figure 3: The dynamic visualisation of the nearest neighbour algorithm. Each data item is displayed as a stick figure. The figures are scaled in proportion to the amount of probability each contributes to the classification of the current posture.

are of a high quality: highly representative of the actions they are trying to perform. This means that the user naturally takes on the task of avoiding overfitting and outlying data items, but they do not provide sufficient data for many learning algorithms. We chose to use a weighted nearest neighbour algorithm for our second prototype, as it makes maximal use of all data items by using them directly. Pitfalls of overfitting are likely to be avoided by users' selection of examples.

7.2 Dynamic Visualisation

A key part of the redesign of the learning system was the creation of a dynamic visualisation that allows immediate feedback about the working on the system and why it is making the classifications it does. The interface must show how the classifier represents the data, but it must also show how it uses it to classify poses. In order to understand why the classifier makes mistakes, users must have feedback about what the classifier is doing when it is making decisions. An effective visualisation must, therefore, show this dynamic behaviour.

Figure 3 shows the visualisation used in the final prototype. Each training example is represented using the same stick figure representation as is used for the live and recorded Kinect data. The examples are colour coded to represent classes (the same colours are used in the class buttons shown near the top of figure 3). In order to give real time feedback about the system, the visualisation needs to represent the reasons why the classifier is classifying particular instances the way it does. The classifier weights the training examples based on their closeness to the current pose when calculating the classification. The state of the system can therefore be represented by the weights contributed by each example. The visualisation shown in Figure 3 scales each example in proportion to the probability contributed by that example (plus a constant to ensure that even components that contribute zero probability are still visible). This scaling is done during live testing, but also on the labelling screen, so users can see how the classification changes as they scrub through their data.

7.3 Continuous and Discrete Interaction

Our investigations with prototype 1 revealed the importance of continuous, tight mappings between movement and visuals. However, the classifier in prototype 2 only sup-

ported discrete classifications which result in a small number of actions and responses. We have therefore augmented the classifier to support continuous responses as well as discrete classifications. Our interface allows users to define a number of continuous parameters (shown above the stick figures in figure 3). Each example pose can now be given values for these parameters as well as its discrete label. Once a pose has been classified with a certain label a value for each parameter is calculated based on the example poses with that label. Each parameter is calculated as a weighted sum of the values of the example poses using the same parameters that are used to calculate the classification.

7.4 Designing interactive visualisations

This interactive machine learning interface can be used to design interactive visualisations. The visualisations used in our prototype consist of a number of discrete stages which have their own visual elements and forms of interaction. Each of these stages has one or more continuous interactions in which dancers movements are mapped onto visual features. The two aspects of our learning algorithms are used to control both. Discrete classifications are used to select particular stages and the continuous parameters are used to control the mapping of movement onto visual features.

The design process is similar to that used in prototype 2. Dancers are recorded while dancing to music. The resulting movement can then be played back in the IML interface. A number of discrete labels are defined for the visualisation, each one corresponding to a particular phase of the visualisation. Each of these phases has a number of continuous parameters, which are also represented in the interface. Users select a number of representative examples poses. For each of these poses they select a label corresponding to the phase that should be triggered by that pose and set values for the parameters to determine how the visualisation should look when they are in that pose. This provides example data to the classifier which will then control the visualisation by changing its parameters based on the similarity of the dancers poses to the examples. To test the result users can dance again using the resulting visualisation. If they are not happy with the result that can add more examples. During testing and labelling they can watch the visualisation shown in figure 3 to support them in understanding and debugging the behaviour of the system.

8. CONCLUDING REMARKS

This paper has described the design of an system for embodied design of dance visualisations for free-form, non-professional dance. The system uses IML to allow users to design interfaces by giving examples of movements rather than by analysing their movements. This design was based on experiences with users interacting with two preliminary prototypes. These prototypes lead to a number of important insights, such as the importance of continuous interaction and the need for easily understandable machine learning methods.

We now have a prototype of the complete system. The next step is further user testing to understand how ordinary dancers use embodied design to create tailored interactive experiences.

9. REFERENCES

- [1] A. N. Antle, G. Corness, and M. Droumeva. What the body knows: Exploring the benefits of embodied metaphors in hybrid physical digital environments. *Interacting with Computers*, 21(1-2):66–75, 2009.
- [2] F. Bevilacqua, B. Zamborlin, A. Sypniewski, N. Schnell, F. Guédy, and N. Rasamimanana. Continuous realtime gesture following and recognition. *Gesture in Embodied Communication and Human-Computer Interaction*, pages 73–84, 2010.
- [3] P. Dourish. *Where The Action Is: The Foundations Of Embodied Interaction*. MIT Press, 2001.
- [4] J. A. Fails and D. R. Olsen, Jr. Interactive machine learning. In *Proceedings of the 8th international conference on Intelligent user interfaces*, IUI '03, pages 39–45, New York, NY, USA, 2003. ACM.
- [5] S. Fdili Alaoui, C. Jacquemin, and F. Bevilacqua. Chiseling bodies: An augmented dance performance. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '13, pages 2915–2918, New York, NY, USA, 2013. ACM.
- [6] R. Fiebrink. *Real-time Human Interaction with Supervised Learning Algorithms for Music Composition and Performance*. PhD thesis, Princeton University, Princeton, NJ, USA, January 2011.
- [7] L. Huang, L.-P. Morency, and J. Gratch. Learning backchannel prediction model from parasocial consensus sampling: a subjective evaluation. In *Proceedings of the 10th international conference on Intelligent virtual agents*, IVA'10, pages 159–172, Berlin, Heidelberg, 2010. Springer-Verlag.
- [8] R. J. Jacob, A. Girouard, L. M. Hirshfield, M. S. Horn, O. Shaer, E. T. Solovey, and J. Zigelbaum. Reality-based interaction: a framework for post-wimp interfaces. In *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, pages 201–210, New York, NY, USA, 2008. ACM.
- [9] Y. Jung, A. Kuijper, D. Fellner, M. Kipp, J. Miksatko, J. Gratch, and D. Thalmann. Believable virtual characters in human-computer dialogs: State of the art report. In *32nd Annual Conference of the European Association for Computer Graphics*. ACM Press, 2011.
- [10] D. Kirsh. Embodied cognition and the magical future of interaction design. *ACM Trans. Comput.-Hum. Interact.*, 20(1):3:1–3:30, Apr. 2013.
- [11] T. Kulesza, S. Stumpf, W.-K. Wong, M. M. Burnett, S. Perona, A. Ko, and I. Oberst. Why-oriented end-user debugging of naive bayes text classification. *ACM Trans. Interact. Intell. Syst.*, 1(1):2:1–2:31, Oct. 2011.
- [12] J. Moen. From hand-held to body-worn: Embodied experiences of the design and use of a wearable movement-based interaction concept. In *Proceedings of the 1st International Conference on Tangible and Embedded Interaction*, TEI '07, pages 251–258, New York, NY, USA, 2007. ACM.
- [13] M. J. Muller and S. Kuhn. Participatory design. *Commun. ACM*, 36(6):24–28, June 1993.
- [14] D. Smith. Dance dance revolution tournament report. *IGN website*, 2000.
- [15] S. S. Snibbe and H. S. Raffle. Social immersive media: pursuing best practices for multi-user interactive camera/projector exhibits. In *Proceedings of the 27th international conference on Human factors in computing systems*, CHI '09, pages 1447–1456, New York, NY, USA, 2009. ACM.
- [16] K. Thórisson. Real-time decision making in multimodal face-to-face communication. In *second ACM international conference on autonomous agents*, pages 16–23, 1998.