HappyFeet: Embodiments for Joint Remote Dancing

Hesam Alizadeh*
University of Calgary

Anna Witcraft*
University of Calgary

Anthony Tang*
University of Calgary

Ehud Sharlin*
University of Calgary

ABSTRACT

We present HappyFeet, a dancing system designed for supporting the dancing experience of remotely located dance partners. HappyFeet uses 3D representations of dancers' feet in a shared virtual dance space to emphasize timing and placement of feet. It has two modes of operation: a learning mode where the user can dance with pre-recorded dance lessons, and a second mode where the system provides a shared dance floor for remotely located dancers. We evaluated our system in a laboratory study where we investigated the role of the feet embodiment by comparing its' use to a video-only condition. The feet embodiment provided our participants with a better understanding of dance moves, helped them to synchronize timing of their dance steps, and provided them with a dance space in which they could freely create dance moves with their partners.

Keywords: Remote exercise, Visual embodiment, Visual Interaction.

Index Terms: H.5.3.m Group and Organization Interfaces: Miscellaneous.

1 INTRODUCTION

Many systems intend to support remote collaboration function best when remote participants are "embodied" effectively. For instance, in shared visual workspaces (e.g. document editing systems), telepointers [3] or remote cursors represent a collaborator's presence, movement and probable focus of attention in a shared document or workspace. The presence of such telepointers facilitates interaction—for instance, by allowing collaborators to gesture or refer to parts of the document. Similarly, many multiplayer games (e.g. MMORPGs such as World of Warcraft) represent players as avatars in a game world, showing their location, view orientation and equipped weapons—again, this supports the development of common ground and shared tactics and strategies [1].

A central question that faces designers is how to design effective embodiments for systems given a novel application context. An important factor to consider is what kinds of actions, capabilities and intentions people may have within the new context, and which of these are important from the perspective of the remote partner. A "talking-head" video capture of remote participant is a common embodiment for video calling applications. Similarly, video-based embodiments encourage rich interaction, particularly for play (e.g. [5,7,20]), or for activities that demand a considerable amount of eye-gaze or gesture awareness.

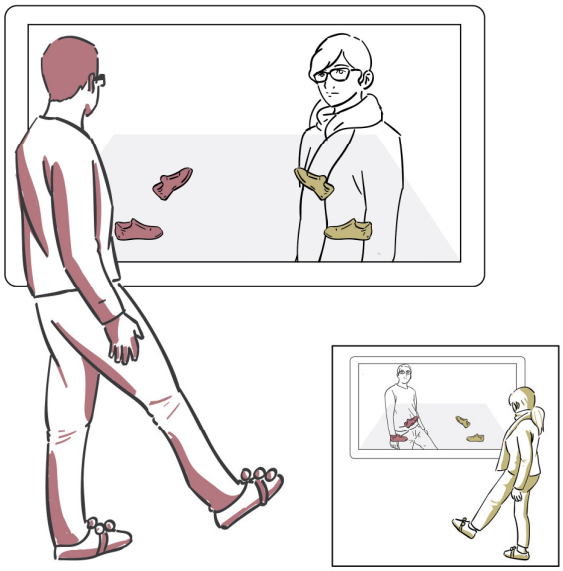

Figure 1. HappyFeet connects two remote dancers with an audio-video link, as well as a rich embodiment of their feet so they can coordinate dance movements.

The overarching goal of this work is to design a dance system that supports the dancing experience for people who are remotely located. The motivation for this system begins from the observation that having an exercise partner participating in an exertion improves motivation and the exertion experience in various ways [15]. To this end, several technologies have been designed to provide social support for remotely located exerciese peers [13,8]. However, the potential of using visual embodiments for enriching the shared experience is not well investigated. In this work, we explore the role of visual representation of remote partner on supporting dancing experience. Central to our design was trying to understand the unique characteristics of the remote dancing experience and then deriving an appropriate visual embodiment to characterize and represent the remote partner.

Our design process focused on developing a dance space where the remote dancers could easily communicate one another's dance moves. Such a space would encourage properly synchronized timing in dance steps, and allow the dancers to engage in creative play [16]. More broadly, we explore how remotely located partners ought to be represented in such a dance space, articulate the various design dimensions of such an embodiment, and evaluate how well such embodiments encourage and engage people in activity, and one another.

We introduce HappyFeet, a system that allows people located at a distance to dance in a virtual dance space together. As illustrated in Figure 1, HappyFeet shows a 3D representation of the remote

* hesam.alizadeh@ucalgary.ca
* anna.witcraft@ucalgary.ca
* tonyt@ucalgary.ca
* ehud@cpsc.ucalgary.ca

dancer's feet in a virtual dance space to emphasize timing and placement of feet during joint dancing. During our design process, we prototyped various configurations of HappyFeet that can support or discourage engagement between partners in a remote dancing scenario. To evaluate our prototype, we conducted a study to understand how people would appropriate different variations of embodiment types. Our findings suggest that HappyFeet enables a far richer experience for participants compared to a video-only condition, and that different orientations of remote feet can support mutual learning and creativity. These findings, while focused on remote dance, suggest a preference for designing activity-centric embodiments that target and are designed specifically for aspects of the activity which require or demand coordination.

This work makes two contributions. First, we describe the design of HappyFeet, a novel system that supports dancing between remote participants, and second, we illuminate considerations for designers of future systems intended for remote activity based on findings from our study of HappyFeet.

2 Related Work

To set the stage for our research, we review prior work on visual embodiment in remote collaboration—generally in terms of how to design shared experiences, then with a domain-specific focus on dance. We close by exploring how researchers have considered remote embodiment as a way of facilitating deep shared experiences.

Sharing experiences, parallel experiences and shared experiences. Considerable prior work has designed and built different kinds of "shared experiences," where remote people can engage in an activity together, in real-time. We distinguish here three different ways of sharing an activity according to the extent of people's engagement in such activities, in order to help clarify the contribution and intent of our work. First, many systems allow for "sharing experiences," wherein one participant can share (e.g. via a video stream) an activity or an event. For instance, Inkpen et al. [6] report on a series of camera and video-streaming prototypes that allow a remote party to watch a live event (e.g. soccer match) that a loved one is participating in. Here, the remote party is brought in by a local party: the local party is "sharing the experience," but the remote party is restricted to viewing the event without meaningful participation beyond conversation. Current video chat technologies (e.g. FaceTime, Skype) are designed to support this kind of activity, where the technology acts as a portal for one party to share the activity in one space with a remote viewer. "Parallel experiences" are those where two parties are connected via video chat, and are simultaneously, but effectively independently engaged in the same activity. Procyk et al. [11] offer an example of a parallel experience where people engage in a real-world treasure hunt, and are connected via a video chat application. But, rather than being collocated, each person engages in the treasure hunt remote from one another. O'Brien et al. [10], introduce another instance of a parallel experience where two remote jogging partners are connected using an audio interface serving as a communication link. Similarly, in this exercise the jogging partners' performance doesn't affect the remote partner [8]. Thus, the participants engage in a parallel experience—they are both engaged in the same activity, but where one's actions do not meaningfully affect the other's experience or engagement. From these, a "shared experience" can be distinguished where remote participants are actively engaging in the same activity together, and one's participation meaningfully affects the other's engagement with the activity [9]. Many online video games are designed this way, for instance, in competitive games, where each party controls an opposing team, or are collaborating together toward a shared goal. Many of the remote systems that support exertion interactions, are another examples of shared experience where exercise partners remotely participate in a shared activity where in most of them the goal is to defeat the opponent [9].

Our approach with HappyFeet was to design a shared dancing system where participants would have a "shared experience". Yet, how to do this is not clear—what are the unique characteristics of dancing that can make it a rich experience for remote participants? Furthermore, is a video connection sufficient to enable a rich experience that dancers expect, or would it restrict the engagement, making it more akin to a parallel experience?

Systems for supporting dancing experience. There is a body of research focused on designing systems to support dance. Most of these works have explored different ways of teaching people to dance; others proposed designs allowing people to dance with others. Of particular interest to us is how they provide dance instructions to the dancers and how they connect remote dance partners.

Most systems in this space use a mix of visual and auditory feedback for dance instruction. For instance, in DL-BUS [18], dancers wear a special suit so that their movements can be tracked full-body motion capture system. A 3D avatar on a wall display is used to demonstrate the dance instructions to the user. The dancer follows the 3D avatars lead and receives a performance score at the end of each dance lesson. Tang et al. [14] evaluate a similarly designed system (with a virtual 3D avatar as feedback), and show that this avatar was not only effective in representing the movements of the dance, but that it was effective in motivating people to participate in the dancing experience. One drawback of these approaches is that dancers need to wear a full-body suit for effective tracking.

In variance to the visual feedback approach, Saltate! [2], aims to provide instruction through the auditory channel. Here, dancers wear force sensors on the soles of their shoes, allowing the system to detect steps. The timing of these steps (in comparison the system's understanding of the dance) changes the loudness (and emphasis) of the musical beats in the music that is played. This helps a dancer to stay in sync with the music.

Related to this, several authors have explored how to design immersive experiences with dance through various augmentations with pre-recorded content. For instance, OutsideMe [17] is a mixed reality dance teaching system that enables dancers see their body movements as external observers along with a virtual character through a head-mounted display (HMD) device. This system captures dancer's posture and blends it into scenes from the dancer's original field of view. It uses an augmented virtual dancer as an instructor, which is added into the dancers' view to increase training motivation. This blending approach has also been used to (virtually) place a dancer into an existing dance/music video: VRMixer [4] blends real-time captured video of the dancer, segmenting him/her, and placing him/her within the context of a pre-recorded dance/music video.

These systems aim to augment the dancing experience with computation. Thematically, a recurring theme is to employ visual representation of oneself in a virtual space—a general approach that we appropriated in HappyFeet.

Remote embodiment. Researchers have looked at remote embodiment as a way of connecting people for remote, joint activity. Embodiment in its most basic form could be used for communicating location, movement, gesture, etc. Telepointers [3] (i.e. mouse cursors that represent a remote participant's mouse cursor) are an example of a simple embodiment that is used in real time groupware systems. Telepointers can communicate location, movement, and focus of attention working on a shared document.

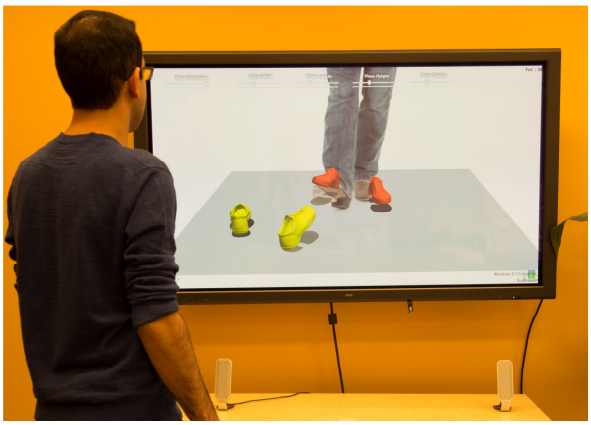

Figure 2. HappyFeet represents the feet of both local dancer (in yellow) and remote dancer (in red). This is superimposed on a video of the remote dancer, or instructional video.

Yet other kinds of domains place different demands on these embodiments. In many video games, players are represented by avatars that show their location and orientation within the game world [1]. Similarly, for many conversation-focused interactions, a face-to-face video chat embodiment is desirable, even this is frequently not sufficient depending on the specific needs of the activity.

For instance, ShareTable [21] employs a "shared desk" metaphor to connect two remote locations for supporting interaction between a child and parent. Here, beyond a simple video chat, the system also embodies remote participants through video capture and projection of their arms as they work over the workspace. This embodiment allows family members to see one another's interactions with the shared desk space (e.g. drawing, gestures, and so forth). In Family Story Play [12], the researchers designed a sophisticated shared book reading experience to allow grandparents to read with their grandchildren. Here, the book is augmented with an integrated video chat application, and beyond this, the system allows the grandparent to see what page in the book the grandchild is viewing (and whether a page is being turned). This latter piece of information helps simulate routine reading patterns such as page turning to encourage child participation in the reading experience. This prototype improves child engagement in remote communication and creates a collaborative shared activity for distant grandparents and their grandchildren. Thus, we see that the particular demands of the activity can influence what is important (and what is not) in a shared experience for that domain.

With specific focus on a shared dance system, Yang et al. [19] present a tele-immersive dancing system where remotely located professional dancers dance in a shared virtual space. They used multiple 3D cameras to capture dancers' movements and a multi-display 3D rendering system. The multi-display system allows the dancers to watch their remote partner represented in a 3D space from different views simultaneously.

We propose a remote dancing system—aimed at non-professional dancers—in which visual embodiment of the dancers' feet is used as a fun mediator in a minimal way. The goal of HappyFeet is to provide the dancers with a shared virtual environment in which they can playfully create dance moves with their remote partner.

3 HappyFeet: Design

HappyFeet (illustrated in Figure 2) is the result of an iterative design process where we began with the idea that we would design a system for shared dance, and iteratively worked to address challenges we faced in terms of the representation of both the local and remote participant. Using feet for representing the dancers was motivated by the fact that our aim was to focus on a dance type which does not involve touching the dance partner as the dancers are remotely located and they can't touch each other. So, we decided to choose line dancing, a dance type mostly focused on foot movement.

Next, we describe the design that was used in our study, the implementation, and the design rationale for HappyFeet that captures important design decisions made during our process germane to supporting shared dance activity.

3.1 Design

As illustrated in Figure 2, HappyFeet embodies and represents participants through a connected audio-video channel, as well as a shared virtual dance floor. The shared dance floor shows 3D rendered shoes, whose positions are mapped based on the tracked positions of the actual shoes worn by participants. This space affords a limited range of customizability: the orientation of remote feet can be changed; local feet can be turned on or off; the perspective of the dance floor can be changed; opacity can be manipulated, and the feet appearance and behavior models can be customized.

HappyFeet enables two major modes of operation: a learning mode, where the system can play pre-recorded videos (along with rendered feet to represent the feet of the dancers in the videos), and a shared dance space mode, where the system can connect two remote dancers into a shared dance space, allowing them to dance, speak and interact with one another.

Our design allows people to dance from "home" with others through a shared visual interaction system. The embodiment of the feet focuses and emphasizes the timing of movement and dance, while the video connection allows people to see and converse with one another.

3.2 Prototype Considerations

HappyFeet is a custom C# application written using the Windows Presentation Framework, and Helix 3D graphics toolkit. The client application connects through a custom node.js server that handles synchronization across instances. Video is handled through a consumer-grade video chat application running in the background.

Our implementation relies on the VICON tracking system, which tracks the movement of participants' feet in a marked dance space. A marker system is affixed to slippers worn by participants to capture position, orientation, yaw, and pitch of the shoes. In principle, consumer grade depth cameras (e.g. Microsoft Kinect; Intel RealSense) might be deployed to similar effect; however, we were interested in developing our sketch using "best available" technology rather than concerning ourselves with deficiencies in the capture system. In time, such depth cameras will reach the accuracy required for our application.

3.3 Design Rationale

We document here several design decisions we made, and the rationale that we followed.

People Space vs. Activity Space. In a standard video chat application, the focus is on "People Space"—an audio/video connection that allows people to make eye contact, and talk with one another. In HappyFeet, we realized that in addition to needing to move the "people space" camera back further to capture more than just the "talking head" view of participants, we also needed to do more to capture the particulars of the activity. Specifically, it

became clear in early trials that simple video capturing the feet of participants in addition to the head provided insufficient emphasis on the timing and placement of dancers' feet. Thus, beyond the conventional audio/video connection, we added a separate facility to track and render the feet of dancers in a shared space.

Dancing with vs. Dancing next to. Our early experiences revealed two fairly different "modes" of dancing with others that people were interested in engaging in. In early iterations, we placed the remote partner's feet in the shared dance space such that they faced the local dancer—this complemented the video-based capture of the remote dancer well as left-to-right conventions were maintained for both the video view and "feet view" of the remote dancer. Yet, it became clear that this view, although intended as "dancing with" perspective, made for a challenging experience because when *teaching* a dance step, it was impossible to stand "side-by-side" with the learner: it was a little too challenging to accurately read timing and positional information from this perspective. Consequently, we added a toggle to HappyFeet that allows dancers to dance alongside (i.e. "next to") a remote dancer. While this breaks the left/right conventions of the spaces, it allows dancers to dance together, and to watch one another's motions.

Saliency of Coordination-Specific Features of Dance. This episode highlighted for us the importance of identifying and making extremely salient aspects of the activity that people rely on for coordinated activity. In this particular case, it was *not* the entire view of the remote person—instead, it was the movement of the feet, the timing of the steps, and the positional information. Thus, beyond simply tracking positional information of where the dancers' feet are with respect to the ground, we also track and render the subtle movements of feet—how they are tilted (i.e. pitch/yaw/roll), or their height in relation to the ground.

Open Experience for Expression and Engagement. One of our principal interests was to design a space that allowed people to engage with one another through the dance activity. Rather than constrain their engagement through a specific song or set of motions, we wanted to allow people to freely use the shared space, dancing to the songs they wanted to, and so forth. Nevertheless, we imagined scenarios where people might have difficulty finding dance partners, and so included a "learning to dance" mode. We created a small set of dance videos to allow people to dance alongside the dancers in the video in the shared "dance floor" space.

4 Study

To evaluate HappyFeet, we conducted an observational lab experiment with pairs of participants. Our interest broadly was to understand how people would appropriate the dance space, and how they would interact with one another. Specifically, we were interested in how the embodiment strategies (i.e. the shared dance space) influenced activity. We compared variations of HappyFeet with a standard audio/video connection for both learning and creative dance. We were interested in addressing the following research questions about HappyFeet:

- What is the role and impact of the feet embodiment compared to a video-only condition?
- What is the impact of *feet-aligned* (the remote partner's feet appear to be dancing next to one's own feet) vs. *feet-towards* (the remote partner's feet face one's own feet in the dance space) perspectives on the dance space? How do embodiment needs change given different kinds of activities (e.g. dance learning vs. dance creation)?

4.1 Design and Method

Our study had two phases: a dance-learning phase, and a dance-creation phase. The dance-learning phase was completed individually, and the purpose is to allow participants to explore each embodiment style in turn as part of a learning activity (participants need to learn basic dance steps). Pairs of participants then complete the dance-creation phase together, where they were connected via a video-based connection (and, depending on the condition, a given HappyFeet embodiment as well). Here, the pair was responsible for creating a dance together, and demonstrating it to the experimenters.

Dance-learning phase. Participants watched an instructional dance video twice, and had the opportunity to mimic/learn the dance being taught. This phase was completed alone, and each participant experienced three conditions depicted in Figure 3: *video-only* (equivalent to watching the video at home, with no embodiments); *feet-aligned* (both learner and instructor's feet are embodied in the space, and pointing in the same direction—i.e. a simple view of the space), and *feet-towards* (learner and instructor's feet are embodied, but instructor's feet face the participant—mimicking the perspective of the instructional video, where the teacher's shoes face the learner). After each condition, participants completed a short questionnaire that asked them about their experience with the embodiment. They were asked to rate the difficulty level of each condition and to list what they liked/disliked about each condition. The presentation order of the embodiment types was counterbalanced across participants, and participants watched three separate dance videos. At the end of this phase, we handed our participants another questionnaire asking them about their preferred condition and how they perceived the effectiveness of the virtual feet.

Dance-creation phase. Participants danced together as a pair across distance, connected via HappyFeet. They were asked to

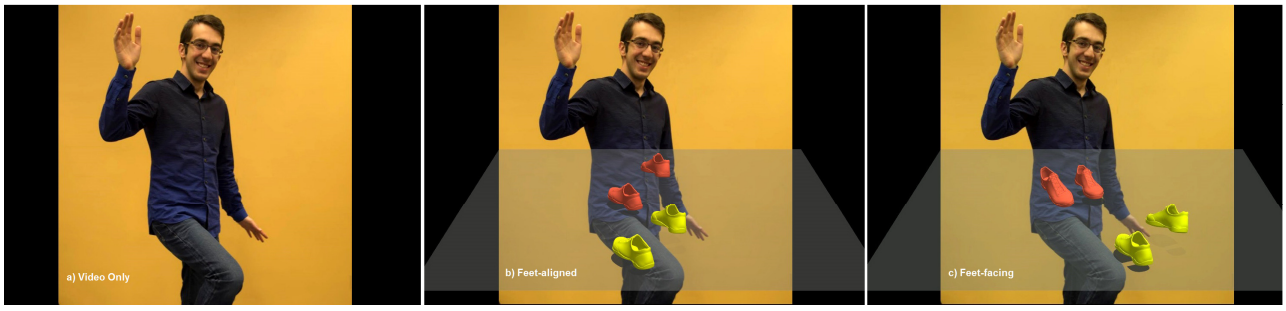

Figure 3. Our study compared three different embodiment conditions: a) video-only, b) feet-aligned, and c) feet-towards. Here, Larry's feet are represented by the red shoes (he is remote and waving) while the local participant's feet are represented by yellow shoes.

construct a dance for a one minute music clip. They were given a total of 9 minutes to practice, and then asked to perform the dance for the experimenters in the remaining minute. This was repeated twice: in the first trial, participants got to experience the three conditions in a random order. In the second trial, participants were allowed to choose an embodiment condition.

At the end of the second trial, participants completed a questionnaire that asks about their experiences—in particular, their preferences, and a reflection on the role of the embodiment.

Data Collection. We collected questionnaire data, and video recorded each session for later analysis. We also collected field notes of interesting occurences throughout the study.

4.2 Materials

Our study made use of line dancing music—a type of common folk music from our locale (typically accompanied by Western Country-style music). Line dancing is a form of choreographed dance with repeated sequences of steps. In this form of dancing, people can dance in multiple lines/rows, sometimes facing one another, and sometimes in opposite direction. Dancers execute steps at the same time. For this type of dance, timing and synchronicity of steps between dancers is very important.

We selected pre-recorded instructional videos of roughly equal difficulty (i.e. three different steps) and length (~ 2:00 minutes). We recorded the feet of the second author, an avid line dancer, mimicking the steps and timing of the instructors of each of the videos. Using HappyFeet, we played back the recording of her feet atop the instructional video to create the illusion that it was the instructor's feet that were being embodied.

4.3 Participants

We recruited 12 pairs of participants (17 females and 7 males) through physical postings, targeted emails to mailing lists, and word of mouth. Participants were recruited as pairs, and knew each other coming into the study. All participants were university students, and were young adults (22-34 years).

Of these participants, two had prior dance experience (defined as formal training or regular attendance at discos), while thirteen had prior musical training background.

4.4 Findings and Observations

Observations. Participants were generally quite engaged with the prototype system, and enjoyed learning how to dance, as well as interacting with their friends through the system. 10 out of 12 groups danced to the time limit. No participant felt that the system resulted in their performing exercise, even though they were clearly engaged in physical activity (albeit low-intensity).

For most participants, this was their first encounter with line dancing, so the dance-learning phase was crucial to helping them develop an understanding of the basic steps. Many indicated that the dancing tutorials were enjoyable:

"It was fun and I felt like I knew what I was doing. I liked the teacher, he was clear." –P7

"It was interesting and the guy was explaining dance moves slow enough for me to follow." –P5

Beyond this, many participants felt that this kind of system would allow them to engage and learn altogether new dance steps and routines:

"I enjoyed the dancing moves and it has motivated me to look for similar video and practice at home." –P18

Of course, the embodiment of the instructor is not the same as having a real-life instructor to guide one's movements. Instead, the shared dance space design forces a dancer to carefully evaluate his/her own movements in relation to the instructor's, rather than an instructor's verbal guidance, or system-generated feedback:

"I didn't feel much engaged since I didn't get any real-time feedback letting me know how correctly I am following the moves." –P14

"It was helpful to see the [instuctor's] feet because I could see his feet from different angles." –P5

In the remote dancing phase, we observed multiple instances of groups laughing as they put together their own dance routine. Much of this was due to the participants simply playing with one another's feet embodiments—for example, miming stomping on one another, or playfully moving their feet. The embodiments in the shared space allowed participants to "play" with one another without the constraints of the physical world (e.g. deliberately walking over one another's shoes; stomping on one another; following by "stepping on one another's shoes", etc.).

In this phase, many groups added creative steps to their routines that were not introduced in the instructional videos. Of these, group 5 (P9 and P10), produced a memorable sequence where they added "cha cha" steps from Salsa (a fairly unrelated dance type) to their dance. Other groups reported enjoying the open design of the tool (i.e. that it does not force a particular style of interaction):

"I felt more comfortable and enjoyed it more with my friend. We were able to laugh together at our struggles." –P8

Most participants (18 of 24) found the joint dancing activity engaging, as it allowed them to connect with their friends in fundamentally new/fun ways:

"Dancing with a remote partner was more fun, and didn't feel like I am doing it in front of a TV." –P7

"I preferred remote dancing as I could create something new and different." –P21

The system's design allowed participants to engage in creative, free play, engaging them with one another through the virtual shared space.

Leader/Follower. Many groups adopted a "leader/follower" style of interaction during the creative dance phase. For instance, one partner would alternate between dictating the dance steps, and stopping to ensure the message was well understood. If a step was not understood, the leader would perform the actual steps. At this point, the follower would replicate the dance steps, and the cycle would repeat. In some groups, partners would alternate turns (i.e. each introducing their own dance step, as they liked).

Four of the twelve groups used counting aloud as a means of synchronization. That is, each dance step would take a certain number of beats that were counted out as they performed them. For instance, once partners had determined the sequences of steps, one of them would lead by counting out beats, and then would pace the dance by counting aloud.

This simple style of interaction was not without difficulty: it was important, for instance, that the follower be paying close attention to the right part of the dance floor/looking at the correct feet, and so forth.

Role of Embodiment. We observed a high level of engagement in both phases (learning/dancing alone and remote dancing)—yet, what is it that HappyFeet provides over a typical dancing video one might find? It seems that the feet embodiment provide a concrete means for participants to concentrate on the positioning, orientation and movement of the feet—regardless of what is happening in the video. The video—particularly if it is oriented to remote participant's face—mainly provides a concrete means to gauge attention and understanding in relation to conversation without sufficient emphasis on the dance steps themselves.

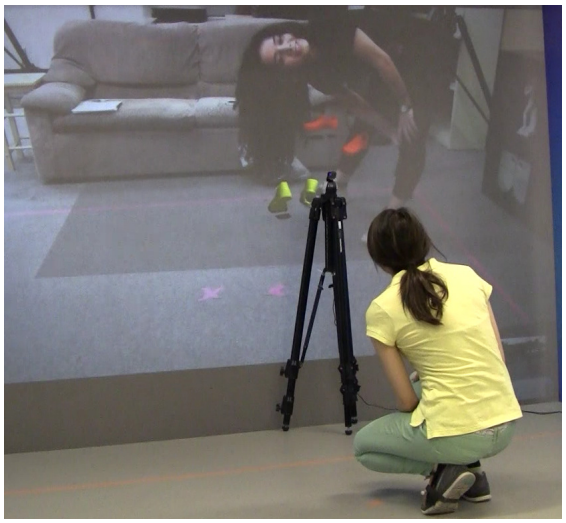

Figure 4. P10 and P11 trying to communicate by squatting so they can see one another's faces as a sign of attention.

For the dance-learning phase, the feet embodiments were useful to follow and learn the steps for most of our participants. Many tutorial videos were captured from multiple cameras, and the changes in view meant that while they were inherently interesting to watch, they were challenging to understand. The feet embodiments provided a consistent view both of the teacher (i.e. the person in the dance video), and the participant.

"I think red shoes were pretty effective and engaging. It let me follow the moves more accurately." –P23

"It was useful to correct the movements and it gave me insight to do the movements in the correct way." –P9

This suggests that feet embodiment provide a rich sense for the actions of others (in this case, the instructor), and that the participants enjoyed this additional awareness information even for a non-live partner (as in the dance-learning phase).

Participants found that different orientations of the feet were useful at different times. When the virtual shoes were next to one another (feet-aligned condition), people felt this was useful because it was easier to match the movements of one's own feet with the other (whether it was a remote partner, or a pre-recorded dancer).

"It (feet-aligned condition) was effective as I could compare my moves with the (virtual feet)." –P8

"It (feet-aligned condition) was much easier to follow the shoes because they were parallel to my feet." –P11

On the other hand, when the partner's feet were facing the participant (feet-towards condition), it was somewhat easier to interpret for participants, because it matched the orientation of the feet of the remote participant in the video.

"Following my partner's visual shoes were easier in this way compared to feet-aligned condition." –P9

"Showing instructor's shoes facing me (feet-towards) made it easy to understand the dance step." –P16

At the same time, the video connection provided an important information resource. As others have argued, the video connection (particularly of the remote participant's face) is important in establishing shared attention. While participants might, for instance, glance and watch their partner's shoe embodiments to understand what they were doing, they would frequently glance back up to look at their partners' face. They might do this, for instance, to ensure that an instruction had been understood, or when they were trying to get their partner's attention. In the following vignette, the participants (Group 5) requested the video be pointed at each other's feet rather than their face as the camera could not cover whole body of the dancers, and they wanted to be able to see their partner's lower body movement:

"Positioning the camera in a way that both partners feel they are in a same room makes it more real." –P9

As illustrated in Figure 4, this frequently resulted in bizarre sequences where they would bend over to "look under the fence" to ensure that a verbal instruction had been understood. The absence here of a video connection for seeing one another's faces and reactions was extremely evident. The following vignette shows a sample conversation between them:

Time	Verbal	Action
30:57	P10: Hey, look at here.	P10 squats down facing the camera to get P11's attention.
30:59	P10: First, you do the side. Then, kick. Then, triple step.	P11 squats. P10 stands back up to demonstrate the dance movements.

Table 1. Participants' preferred conditions by study phase.

	Feet-aligned	Feet-towards	Video only
Dance Learning	10	5	9
Dance Creation	7	9	8

Preferences between Embodiment Conditions. At the end of each phase of the study, we asked participants to indicate their preferred condition (video-only, feet-towards or feet-aligned). Table 1 summarizes these results for each study phase.

In both phases about two thirds of the participants preferred seeing the feet embodiment on the screen (15 out of 24 for dance-learning and 16 out of 24 for the dance-creation phase) over the video-only condition.

On balance, no one condition was a clear winner. Each configuration had its respective strengths and weaknesses, so depending on how a participant used or thought about the activity, the embodiment might suit the activity better or worse.

Some participants found the shared dance space to be overwhelming (i.e. seeing feet embodiment in addition to the video), and instead preferred to simply focus on the video itself. This issue was mitigated when the participants got familiarized with dance moves and learned how to use the visual embodiment.

"Preferred to follow the video rather that the red shoes, following [the feet embodiments] needs practice." –P12

"I would say—for the beginning—it is better for me not to see my feet. However, when you learn the moves, seeing your feet could help and be effective for proficiency." –P10

"The feet was more effective when the moves were easier."-P20

Finally, other participants preferred the feet-toward condition because the video of the instructor's feet would match his/her feet embodiment:

"It was really fun although I couldn't follow the dance moves very well. Seeing the instructor's feet facing me was effective in understanding the moves better." –P17

With the dance-creation phase, many participants would rely on the feet embodiments for demonstrating the dance sequence rather than for learning. As such, the "task demands" were much lower, and instead, participants were more interested in getting the dance sequence and the timing right.

"Watching my partner's feet helped me ensure we are in sync." –P2

"Feet-aligned gave me a real feeling about my partner, and you would be able to do the exact dance steps. Feet-facing would be more appropriate if you had some previous experience" –P4

Here, more participants indicated a preference for the Feet-towards condition, again, because it reduced the dissonance between the orientation/movement of the remote partner's feet in the video and the feet embodiment.

"I prefer dancing while facing my partner. Otherwise, I would be kind of exercising." –P9

Challenges with HappyFeet Embodiment. In observing how participants used HappyFeet, we identified three major challenges with the design. First, the lack of a temporal "trace" meant that the feet were only of limited value to illustrate the historical movement of other dancers' feet over time. Second, the shoe embodiments sometimes seemed to add too much information for dancers to take in. Finally, that when the orientation of the shoes did not match that of the remote participants' feet in the video, some participants would confuse left and right. We discuss each of the challenges in turn.

Temporality - Because the shoe embodiments track only the live position of a dancer's feet, it can be challenging to explain a series of dance steps. Deictic references (e.g. saying, "You put your left foot here", while placing one's foot in the right position) need to happen in the moment; if the remote dancer is not paying attention, then this reference is completely lost. This lack of temporality also causes problems when people are trying to explain what the other person is doing incorrectly. That is, any reference needs to be made at the moment, as recall of false steps/poor positioning/etc. will necessarily be lost in time. For example, the following vignette shows a situation where the one of the participants is trying to come up with a new dance move, but has trouble explaining the movement path to her remote partner:

Time	Verbal
9:30	P19: We can go to this direction
9:37	P20: I am just confused about the directions
9:41	P19: I am saying that move in a kind of crossway

The problem here is that the dancers do not have an easy way to refer to previous dance steps in the routine, or previous moments in the sequence of dance moves—specifically, the embodiments provide a means to understand the position of another person's feet, but only in the moment—not in the past.

Visual overload - In designing HappyFeet, we deliberately overlaid the embodiment of the feet and the virtual shared dance space atop the video of the remote dancer. This makes both the video and the embodiments more challenging to see and interpret. Several participants—particularly for the Learning phase, did not like the embodiments, as it added too much information that needed to be interpreted. This made learning an altogether new dance very challenging. Nevertheless, it seemed as though this was a challenge that could be overcome with practice.

"Following the shoes and the feet at the same time was kind of distracting and it distracted my focus from dancing to focusing on what happening" –P11

"At first I was confused. Then, used the feet, then used both. It was pretty easy (after I learnt) to use both" –P8

Joint Orientation - As discussed earlier, many participants had challenges interpreting and understanding the shoes in the feet-aligned condition. This causes problems in two different ways. First, in people's implicit assumptions about which foot and which direction they should raise/move. Second, when people discuss "left" and "right" verbally, this makes sense until the video seems to suggest they are going the wrong way. Many participants felt that this was merely something that they could get used to over time, too.

5 Discussion

We designed our study to address two questions about embodiment design within this specific dance context: what impact do the embodiments have on the interaction compared to video-only, and how do the various embodiment conditions compare to one another in terms of how they are perceived or used? Our results indicate that using feet embodiment empowered the dancing experience in several ways:

Feet embodiments play different roles when used in different dance conditions. We found that people perceive the virtual feet differently when learning dance moves and when actually dancing with a partner. When people were learning new moves the feet embodiment were used as a reflective tool, helping the dancer understand the nuances of the dance steps, and provided them with a way to compare their feet movements with those of the teachers. As a result, many found feet-aligned more useful as they could see their feet side-by-side with the teacher's feet, and easily mimic their dance moves. On the other hand, when dancing with a remote partner, feet embodiment encouraged our participants to playfully dance with their partners (e.g. stomping on partner's virtual feet). It helped them to demonstrate their desired dance steps to their partner more easily, and to synchronize their dance steps more effectively. People perceived the virtual shoes as a shared connection or link from themselves to their partners, and it made more sense for them to see the shoes facing towards them.

Role of video. Nevertheless, video remained an important mechanism through which the partners maintained contact. We observed that the participants used video to follow the body parts that were not tracked, and to understand their partner's reactions to their movements—specifically, being able to gaze at one another's faces provided an easy mechanism to gauge attention (and inattention), as well as see one another's reactions to jokes and bodily play.

5.1 Limitations

Our goal was to highlight and bring dancers' attention to the feet—that is, the dance steps for line dancing. Nevertheless, we do acknowledge several weaknesses in this work.

Feet-focused. The embodiment places specific focus on the position and movement of the feet. And, while this is perhaps appropriate for the kind of music and dance that we were working with (i.e. country line dancing), we still saw instances where this broke down: for instance, when dance steps involved other body parts, such as hips, or when the dancer was required to turn his/her body in such a way that viewing the screen would be difficult. As one participant acknowledged, "Simulating hip and hand would improve the experience," [P9]—particularly for different types of dance.

Multiple points of visual focus. For some participants it was challenging to focus on them and on the partner's body at the same time. This detachment was a part of our design in which we represent the dancers in a shared space using visual embodiment of their feet. That being said, an alternative design can address this issue through tracking the feet and correctly superimposing the feet embodiment to, or adjacent to, the actual feet.

Camera Placement. We used off-the-shelf limited FOV webcams for video streaming in our study. These cameras could not cover the dancers' whole body while still providing enough level of details. As a result, the participants had to choose whether they want to see their partner's feet or upper body. Wider FOV cameras could improve this by providing a full body view of the remote dancer. Some participants suggested to show the dancers video feed side-by-side so that they can easily compare their moves: "I wonder

if it would reduce confusion if the video feed were positioned so that it was like my friend was standing beside me, and the feet guide was also like that“ –P8

Other modalities. In this work, we relied specifically on projected, visual embodiments. We leave open the possibility of considering embodiment that makes use of other modalities (e.g. auditory, haptic, and so forth).

Sample population. The sample population for our study was strictly made of graduate students—very few of whom had dancing background/knowledge. It is unclear how well these findings generalize to other populations, or for dancing purposes other than casual dancing.

Beyond a pair. It is also unclear how this type of solution scales to dancing groups that may be larger than two participants. Specifically with country line dancing, this is an activity that is frequently performed with large groups of dancers (e.g. 8 to 24). Undoubtedly, it would be difficult to near impossible for a dancer to make sense of this many shoes on the screen at once. Of course, it begs the question of whether it is important to actually see all these feet simultaneously to have an engaging, shared experience.

5.2 Embodiment Design Beyond the Dance Floor

We view HappyFeet as a specific case study in embodiment design that sheds light on the question of how to design embodiments generally for shared activities at a distance. In contrast to the approach by Yang et al. [19], where the authors perform a complete 3D scan of the dancer in real-time, HappyFeet takes a reductionist approach. This approach necessarily means focusing on some narrow characteristic or aspect of the dancer—in this case, the position of the dancer's feet. Given the style of dancing we were designing for (i.e. country line dancing), this focus on feet was appropriate; however, for other types of dance (e.g. jazz/hip-hop), the focus might need to be on different aspects of the dancer's body.

For us, the core insight was to focus on aspects of the embodiment that would be important for dancers to feel that they were having a meaningful, shared experience. Within the context of line dancing, this meant focusing on aspects of the activity that had demanded coordination.

Yet our final approach (i.e. rendering shoes) leaves several unanswered questions. Could we have gotten away with even less? —that is, what if rather than capturing all aspects of the dancer's feet (roll, pitch, yaw, height), we only captured height? Alternatively, what would happen if the representation was a set of points rather than a shoe? These questions are important both within this specific context (i.e. can we use less capture infrastructure?), and more broadly (i.e. minimally, what needs to be captured for effective embodiment?).

6 Conclusions and Future Work

Current designs of dance systems are mainly focused on different ways to teach dancing. Little work has been done on understanding the main characteristics of remote dancing experience and finding appropriate representation for those characteristics. HappyFeet explores the role of feet embodiment in supporting dance training and remote dancing. We found that our system helps participants to be engaged in the dancing experience. The feet embodiment played a different role in different dancing conditions. While learning new dance steps, the feet embodiment provided a better understanding of dance steps to our participants and they used it to compare their moves with the ones of the teacher. In the remote dancing situation, they used feet embodiment to demonstrate dance moves to their partner and to synchronize their moves. Based on these observations, we have outlined implications and challenges for designing remote dancing systems in the future. Next steps in this space will be to engage in more broadly testing the system, and then designing a system robust enough to be deployed and studied long-term.

Acknowledgements

The authors wish to thank the Frank and Sarah Meyer Foundation CMG Collaboration Centre for setup of the remote participation experiment, NSERC and AITF for funding, and Melissa De Smet, David Ledo and Jennifer Payne for their kind support.

References

[1] Benford, S., Bowers, J., et al. (1995). User embodiment in collaborative virtual environments. In Proc SIGCHI 1995, 242–249.

[2] Drobny, D., Weiss, M., and Borchers, J., (2009). Saltate!: A Sensor-based System to Support Dance Beginners. In CHI EA 2009, 3943-3948.

[3] Greenberg, S., Gutwin, C., and Roseman, M., (1996). Semantic telepointers for groupware. In Proc OzCHI 1996, 54–61.

[4] Hirai, T., Nakamura, S., et al. (2014). VRMixer: Mixing Video and Real World with Video Segmentation. In Proc ACE 2014, 1–7.

[5] Hunter, S., Maes, P., et al. (2014). WaaZam! Supporting creative play at a distance in customized video environments. In Proc CHI 2014, 1197–1206.

[6] Inkpen, K., Taylor, B., et al. (2013). Experiences2Go: sharing kids' activities outside the home with remote family members. In Proc CSCW 2013, 1329-1340.

[7] Ledo, D., Aseniero, B. A., et al. (2013). OneSpace: shared depth corrected video interaction. In CHI EA 2013, 997-1002.

[8] Mueller, F. F., O'Brien, S., and Thorogood, A., (2007). Jogging over a distance: supporting a jogging together experience although being apart. In CHI EA 2007, 1989-1994.

[9] Mueller, F. F., Gibbs, M., and Vetere, F., (2013). Reflections on designing networked exertion games. In Proc IE 2013, 5:1–5:8.

[10] O'Brien, S., and Mueller, F. F., (2007). Jogging the distance. In Proc CHI 2007, p. 523.

[11] Procyk, J., Neustaedter, C., et al. (2011). Shared geocaching over distance with mobile video streaming. In Proc CSCW Companion 2014, 293–296.

[12] Raffle, H., Spasojevic, M., et al. (2010). Family story play. In Proc Proc CHI 2010, 1583-1592.

[13] Stevens, G., Thorogood, A., O'Brien, S., & Wulf, V. (2007). Sports over a Distance. Personal and Ubiquitous Computing, 11(8), 633-645.

[14] Tang, J. K. T., Chan, J. C. P., and Leung, H. (2011). Interactive Dancing Game with Real-time Recognition of Continuous Dance Moves from 3D Human Motion Capture. In Proc ICUIMC 2011, 50.

[15] Weinberg, R. S., & Gould, D. (2014). Foundations of Sport and Exercise Psychology, 6E. Human Kinetics.

[16] Wozniewski, M., Bouillot, N., et al. (2008). Large-Scale Mobile Audio Environments for Collaborative Musical Interaction. In Proc NIME 2008, 13-18.

[17] Yan, S., Ding, et al. (2015). OutsideMe: Augmenting Dancer's External Self-Image by Using A Mixed Reality System. In CHI EA 2013, 965–970.

[18] Yang, Y., Leung, H., et al. (2013). Generating a Two-phase Lesson for Guiding Beginners to Learn Basic Dance Movements. J. of Computers & Education. 61 (1): 1–20.

[19] Yang, Z., Yu, B., et al. (2006). A Study of Collaborative Dancing in Tele-immersive Environments. In Proc ISM 2006, 177–184.

[20] Yarosh, S., Inkpen, K. M., and Brush, A. J. (2010). Video playdate: toward free play across distance. In Proc CHI 2010, 1251-1260.

[21] Yarosh, S., Tang, A., et al. (2013). "Almost touching": parent-child remote communication using the sharetable system. In Proc CSCW 2013, 181-192.