CoStream: Co-construction of Shared Experiences through Mobile Live Video Sharing

Niloofer Dezuli*, Jochen Huber*, Elizabeth F. Churchill*, Max Mühlhäuser*

* Technische Universität Darmstadt *eBay Research Labs

{niloo,jhuber,max}@tk.informatik.tu-darmstadt.de *churchill@acm.org

Mobile media sharing is an increasingly popular form of social media interaction. Research has shown that *asynchronous* sharing fosters and maintains social connections and serves as a memory aid. More recently, researchers have investigated the potential for mobile media sharing as a mechanism for providing additional event-related information to spectators in a stadium. In this paper, we describe *CoStream*, a novel system for mobile live sharing of user-generated video *in-situ* during events. Developed iteratively with users, *CoStream* goes beyond prior work by providing a strong real-time coupling to the event, leveraging users' social connections to provide multiple perspectives on the ongoing action. Field trials demonstrate that real time sharing of different perspectives on the same event has the potential to provide fundamentally new experiences of same-place events, such as concerts or stadium sports. We discuss how *CoStream* enriches social interactions, increases context, social and spatial awareness, and thus encourages active spectatorship. We further contribute *key requirements* for the design of future interfaces supporting the co-construction of shared experiences during events, *in-situ*.

Mobile live video sharing, Event, Field study, Iterative design, Multimedia sharing

1. INTRODUCTION

Highly capable camera-enabled mobile phones allow users to capture and share experiences through videos virtually anywhere and anytime. Mobile video sharing has become increasingly popular for both consumers and researchers as a means for novel social media interaction [4, 15, 24]. In particular, video broadcasting platforms such as Ustream.tv allow users to broadcast live events on the Internet or for online archival. Their main purpose is to provide access to a live event even for those who are remotely located and cannot participate themselves. Obviously, bridging the physical distance and supporting the “*being there*” [10] is useful in this case.

However, sharing media between spectators participating in the very same event [2] can also enrich experiences, as Jacucci et al. [12, 13] have shown. They investigated asynchronous media sharing (i.e. taking a photo and sending it to another spectator) and its potential value. They particularly looked at large-scale events, where spectators are scattered across different sites and can only partially witness the whole event. While this is certainly helpful for fostering awareness e.g. regarding other spectators' locations, we believe that live media sharing – and user generated live video sharing, in particular – has the potential to provide a fundamentally new experience even for events where spectators share the same event *and* the same location. We further believe that this generates a novel *in-situ* experience particularly for

events happening in stadiums or concert halls, where spectators are necessarily restricted to a particular view of the ongoing action or editorially selected and projected screen views. We illustrate this kind of limitation in the following scenario derived from talking with spectators:

Alice and her parents decided to go to a soccer match. They bought tickets for the main aisle, since these seats are perfect for maintaining an overview during the match (cf. Fig. 1 left). However, a detailed view e.g. on the opposing team's goal, is only available to those with tickets in other aisles closer to the goal. Luckily, Alice discovers that Bob, a friend of hers, just checked into the stadium on Facebook. She messages him and learns that he is with friends near the opposing team's goal. Unfortunately, they cannot enjoy the match together, but Alice calls Bob through a Skype video call on her phone. Seconds later, their favorite team advances and since Bob is close by, he streams the scene to Alice and friends (cf. Fig. 1 right), who can now witness their team strike. They all cheer together by streaming videos of themselves in both directions.

The scenario illustrates that the physical restriction in such closed spaces imposes two major

Figure 1. Scene from the scenario: Main aisle in a soccer stadium. Participants are restricted to their aisle and thus also to this very point of view (Left). Bob is recording the cheering team after a goal was scored (Right).

drawbacks on spectators in traditional settings: decreased (1) viewing and (2) social experiences.

We envision mobile live video sharing services for user generated content to address these drawbacks by supporting the social co-construction of experiences during live events not only over large distances, but particularly *in-situ*. As outlined in our scenario, widespread technologies provide already a certain degree of support, yet emphasize bi-directional video streaming (cf. Skype). As to the social experience, they basically require a-priori known users. We argue that sharing is not only restricted to friends, but basically any user generated video during an event can be harnessed for experience enhancement. Broadcasting services such as Ustream.tv lack means for embedding the experience into the specific event and neglect crucial information, such as the location of potential video sources i.e. properly equipped spectators. Together, these observations led us to the following research questions: How can mobile live video sharing support the co-construction of experiences *in-situ* during events? What are the requirements for a system supporting this? And how will it affect the overall event experience?

Our contribution is three-fold, reflected in the structure of this paper: Following the related work section, we (1) contribute *CoStream*, a novel mobile live video sharing system and its *iterative design process*, empirically grounded in three focus group sessions. We then (2) report on the use of *CoStream* during *two field studies* that explored the aforementioned research questions. Last, based on our lessons learned, we contribute (3) *key requirements* for the design of future interfaces supporting the co-construction of shared experiences *in-situ* during events.

2. RELATED WORK

Enhancing the social experience of spectators has recently drawn the attention of both Multimedia and CHI communities [6, 24, 26]. There are a number of prior studies investigating how user generated media such as microblogs can help to understand and enrich the social experiences *around* events. TwitInfo [17] used news on Twitter to understand and describe important moments of events. Shamma et al. [28] analyzed the sentiments of tweet annotations for a presidential debate and showed that interesting events can be detected by looking at anomalies in the pulse of the sentiment signal in the event. There is a larger body of research focusing on enriching spectator experiences *during* an event, which we discuss in the following.

2.1 Providing Additional Information

Various systems [19, 20] investigated how additional information during an event can aid in fostering

awareness, as well as getting a better overview over the event. TuVista [2] supports the mobile consumption of near-to-live content (from an avg. 15 min till an avg. 30 sec), which is related to an event. This allows users either to catch up on a match remotely, or to replay scenes while at the stadium. In TuVista, one or more *professional* video editors monitor a preview of available video streams, delivered from static cameras in the stadium. In turn, the editors prepare so-called multimedia *bundles*, which consist of preselected clips from multiple angles and added links to related content such as scored goals or photos and videos *previously* captured by spectators. Spectators in the stadium can then access these bundles through the in-stadium wifi network. In sum, TuVista was used as a probe to understand what additional information spectators want to consume during a match in a stadium. The authors neither focused on active engagement, nor real-time interaction or communication between spectators. Holmquist et al. [11] evaluated an awareness device, which indicates when groups of spectators are close to each other during a rock festival. They found that their device fosters the feeling of connectedness between friends.

2.2 Active Media Creation

The aforementioned approaches focused on providing additional information to an event, therefore enriching the overall experience. There have also been efforts in involving spectators in an active media creation process during an event, such as photo taking or video recording, and investigating the effects on the spectator experience. Mäkelä et al. [18] showed that pictures, are not only used as memories of special moments of events but also as a tool for creating playful stories, expressing affection and creating art. Frohlich et al. [8] found that showing off photos taken during an event is a way of sharing experiences with others who were not participating in the event—to *tell a story*. On the contrary, they discovered that the storytelling aspect would lose its importance, if the people were co-located during the event. Peltonen et al. [21] extended large-scale event participation with user-created mobile media on a public display. They found that users are more present at events through the use of mobile cameras. Moreover, event experiences were relived and wrapped up in a fun way when users browsed through the captured videos and photos of the events afterwards together. Nilsson et al. [19] noticed that the primary interest of spectators is to experience the event as it unfolds.

2.3 Media Sharing

Jacucci et al. [13] argue that in large-scale events spectators experience the event together in other

ways than just watching [22]. They explored how capturing and then sharing experiences using mobile phones can be a participative practice to enhance the overall experience during a three-day car race and a music festival. They found that media sharing has the potential to facilitate *onsite* reporting to *offsite* spectators, coordination of group action and keeping up to date with other visitors or spectators, who are interested in different occurrences in large-scale events.

This work investigated asynchronous media sharing (i.e. taking a photo or recording a video and sending it afterwards). Live and therefore synchronous media sharing during events in real-time provides more immersive means for social interactions. For example, Sahami et al. [25] proposed to share live non-verbal opinions using mobile phones while watching a soccer match. They found that the aggregated sentiments, which correspond to important moments in the event, can be used to generate a summary of the event. Barkhuus [1] developed an application, which can distinguish different levels of audience cheering, rather than simply the presence of applause. They utilized the notion of 'reward applause to engage the audience actively without overwhelming the experience with technology during concerts. They found that although using technology augmentation with crowds can be very challenging, it provides new ways of interaction that increase the level and sense of participation among the audience. Shamma et al. [27] and Liu et al. [16] also showed evidence that online simultaneous video sharing can help people feel closer and more connected to their friends and family. There is also a variety of commercially available video broadcasting services, e.g. ComVu which was launched in 2005 to enable real-time video broadcasting from a smart phone to a public website. Other services are Livecast.com, Qik.com, Kyte.tv, Bambuser.com, Flixwagon.com, CollabraCam.com, Stickam.com, Ustream.tv and most recently color.com. These services focus on supporting remote sharing, overcoming larger physical distances.

As concluded by Esbjörnsson et al., [6] "there are clear differences between this [*remote spectating and bridging larger distances*] and other types of spectating, particularly stadium-based spectating". In the latter case, the *in-situ* co-construction of shared experiences through user generated mobile live video sharing has not yet been explored in prior studies—to the best of our knowledge. Furthermore, requirements for the interaction design to support this are unclear. While artistic design guidelines do exist, scaffolding the creative process of video creation on mobiles (see Juhlin et al. [14]), we believe that media sharing in-situ and in real-time focuses more on time critical tasks than on artistic camera handling. To investigate this, we have conducted an iterative design process, which serves as an empirical basis for the development of

CoStream. In the following, we derive requirements for the interaction design and illustrate the design process.

3. COSTREAM

The design of CoStream is empirically grounded in three focus group sessions. The major goal was to elicit interface requirements, pertaining to user generated mobile video sharing in real time during same-place events. In the following, we start by briefly presenting the iterative design process. Based upon the results, we outline design implications for CoStream and present its interface design.

3.1 Iterative Design Process

We recruited 7 participants per session; 21 in total and different for all sessions. They were between 22 and 34 years old. Each focus group was equally comprised of (1) potential end users, such as passionate soccer fans and (2) interaction design researchers, who had been working at the intersection of HCI and Multimedia for 4 years in average. All had been exposed to mobile video creation before, mainly to capture precious moments; some of them during sports events particularly. Each session lasted two hours.

Discussions during the sessions had a brainstorming character, but participants were also involved in creating paper prototypes of their design suggestions. In the first session, no prototypical interface was presented, the participants were only introduced to the scenario and to the research questions. Paper prototypes generated in the first session (cf. Fig. 2 left) were then used as input for the second session. The objective there was to refine and discuss the interface concepts in detail. The refined paper prototypes were the basis for paper mock-ups with printed interface elements (cf. Fig. 2 right), which in turn were discussed in the last session.

In addition to paper prototyping, we used video-recording and photo documentation for data gathering. Both data gathering and analysis were performed iteratively. After each session we transcribed the data, selected salient quotes and coded them using an open and selective coding approach [29]. Thus, the analysis results of each session directly impacted the subsequent session.

Figure 2. Paper prototypes. Some paper prototypes, resulting from the first session (Left). Refined paper prototypes with printed interface elements used in the last session (Right).

Figure 3. The conceptual interaction design of CoStream modes: (a) overview, (b) in-situ awareness, (c) watching and (d) streaming.

3.2 Results: Design Implications

Based on this qualitative analysis, the following dimensions set the requirements for the design of CoStream.

Provide Efficient Overview and Awareness. Participants mentioned they “want to see who is in the stadium (e.g., friends) and whether a spectator is recording something or not”. In particular, the participants stressed the importance of the efficient access to this information, since they “do not want to spend too much time looking around for streams”. Indicating the orientation of the spectator was also considered important, since the participants “want to know whether a spectator is filming in the direction they are interested in”.

Support Active Engagement. While the participants generally liked the idea of being able to connect to friends close-by through video, they mentioned that they would want to “actively poll other users to stream from a certain perspective” for them. Moreover, inviting other users to their own stream was considered important, as well as feedback while streaming, e.g. as one participant commented: “something comparable to the like button in Facebook; it should be easily understandable and just communicate ‘hey! I like what I see – keep on streaming!’”.

Support Immediate Interaction and Reduce Visual Attention. Throughout our design sessions,

Figure 4. (a) Map view, (b) In-situ awareness. The arrows are color-coded: green are friends, white are strangers and black is the user herself. The arrows also contain a dot, designating the current action, whether a user is recording (red), watching (blue) or passive (black).

Figure 5. A user watches a stream in the center view while he simultaneously streams for others (displayed in the bottom right corner, picture-in-picture).

the participants underlined the fact that streaming a live situation is highly time-critical, requiring particularly careful interaction support. As one participant put it: “it must be possible to record moments quickly, without looking at the device”. They imagined this to be ideally as easy as pointing in physical space: “I just want to point in a certain direction and then see from that very perspective”.

3.3 Interface Concept

Based upon the design implications, we subdivided the interaction design conceptually into four modes: overview, in-situ awareness, watching, and streaming (see Fig. 3). In the following, we discuss CoStream accordingly. Furthermore, we present techniques to support active engagement and social interactions.

Overview and In-situ Awareness. Initially, CoStream provides an overview of the user’s current location and of nearby spectators in a ‘map’ view (see Fig. 4a). The user invokes this view by holding the device horizontally in front of herself – like a map (see Fig. 3a). Further, CoStream provides in-situ awareness through an augmented reality view. It is invoked when the device is lifted and held facing the environment like a see-through display (see Fig. 3b). In this mode, CoStream shows available streams and fellow spectators in the vicinity (see Fig. 4b). This way, CoStream fosters immediate interaction, as users are able to just point in a direction to reveal available streams for a particular perspective.

Nearby spectators are visualized in both views using small icons that double as arrows (cf. Figure 4). The icons show the social relationship to the spectator (friend or stranger) and are oriented according to the direction of the camera. In accordance with the iterative design sessions, this design aims at conveying the direction a spectator (or rather, her device) is currently looking at. Furthermore, icon decorations reveal whether a user is currently recording, watching or passive.

Watching and Streaming. Once a stream has been located, users can immediately start watching that stream by simply rotating the device into landscape mode (see Fig. 3c). If multiple streams are available in the considered direction, a thumbnail grid with the latest video frame of each stream is provided. Users

can then select a desired perspective by tapping onto the thumbnail. CoStream also supports replaying scenes: playback can be rewound by 30s by tapping onto the circular icon on the left. Tapping again resumes live playback.

To start streaming (cf. Fig. 3d), users can tap and hold down with two fingers anywhere until the camera is ready. This allows users to concentrate their visual attention on the event and just use the device to point at an important scene and immediately start streaming. Tapping again with two fingers ends streaming and allows for an efficient mode switch. If the user is already watching a stream, the playback will be continued and the camera is shown in a picture-in-picture mode (see Fig. 5).

Active Engagement and Social Interaction. CoStream allows users to actively draw their friends' attention to what they are doing: for this purpose, the interface provides an overview of friends in the vicinity in a sidebar for both watch and stream mode. By tapping onto the current video and dragging it to a friend (*push mode*), the friend is invited to watch the same stream as the user. By tapping onto a friend's icon and dragging it into the video screen (*pull mode*), the user switches to the same video the friend is currently watching or streaming. In addition to visual notifications, CoStream signalizes invitations through vibration. The interface also provides a like and a dislike button, allowing users to rate streams. The number of likes/dislikes and the current amount of viewers is shown in the top right corner (see Fig. 5).

4. FIELD STUDY I

We conducted a field study to explore (1) how CoStream is actually used in real-world settings, (2) whether it supports the in-situ co-construction of shared experiences during events, and (3) how it affects social interactions. Furthermore, we wanted to get insights into CoStream's usability and the overall user experience.

4.1 Study Design

The study took place during a mid-scale soccer match with 8190 spectators in Germany. We used a snowball sampling technique to recruit 2 groups of 4 friends (P1-P4 and Q1-Q4; 7m, 1f; avg. 25 years). The groups did not know each other. All of them but two were regular attendees of soccer matches. The two, however, were regular attendees of ice hockey and basketball matches. Four participants (2 in each group) confirmed to use live video sharing services such as Ustream.tv. The participants were introduced to CoStream upfront. The session took about 4 hours. The participants were paid the entrance fee and a mobile data plan.

4.2 Apparatus and Methodology

CoStream is implemented as a client-server architecture. The mobile clients are Android-based and use RTP streaming. The server uses VLC to distribute the videos via HTTP. Thus we achieve a delay of less than 1s over HSUPA. All videos are stored on the server-side. CoStream uses Facebook to manage social relationships between users.

As data gathering methodologies, we used interaction logs (video and usage), semi-structured interviews (before, after the match) and observation. Two of the authors engaged with the participants as participant-observers but did not use CoStream at all throughout the study. Figure 6 shows the participants' and observers' location during the two halves of both matches. Neither of them knew their actual location upfront. This arrangement enabled us to observe the behavior of the participants between co-located friends, strangers and distributed individuals during the match.

4.3 Data Analysis

Interviews and observations were transcribed and analyzed using an open coding approach [29]. To analyze the time-coded interaction logs, we have implemented an analysis tool. It supports the analysis of the recorded videos of the participants with respect to geo-location, social relationship and sharing behavior. Furthermore, all data is synchronized and aligned with respect to a unique timeline. Although there exists a variety of tools for this purpose [7, 9, 31], these tools do not support extensible presentation and navigation of multimodal time-based data from different sources (in our case the recorded videos and related interactions). Furthermore, they do not support the visualization of social relationships between participants and their interaction respectively, which is essential for investigating our research questions.

Figure 6. Participant locations during the first study. Unfortunately, the south aisle was under construction and no participants could be seated there.

Figure 7. (a) The analysis tool has three main views: video, geo-location and timeline with annotations. The video view shows an overview over the recorded videos at the selected moment in time. (b) The geographic position and the relationship of the participants are displayed on a map; anonymized. (c) The timeline view visualizes the interaction logs.

The analysis tool has three main views (see Fig. 7a). The analysis tool has three main views (see Fig. 7a). The geographic position of the participants is displayed with a pin on a map (see Fig. 7b). The social relationship is visualized through the color of the pin: pins with the same color are friends (i.e. a group during the study). Furthermore, the icon of the pin reflects the current state, whether the respective participant was streaming, watching or passive. The timeline view at the bottom (see Fig. 7c) shows histograms of the participants' sharing behavior (streaming, watching and a combined view on top). The timeline view at the bottom (see Fig. 7c) shows histograms of the participants' sharing behavior (streaming, watching and a combined view on top). Furthermore, it allows for coding of discrete moments in time or regions.

4.4 Results

The analysis yielded four categories. We present the results within these below.

Production vs. Consumption. The participants recorded a total of 96 videos, which we classified into 4 categories: 45 streams recorded the match, 26 showed communications, 16 the surroundings (such as side-events happening at the aisles) and 9 were recorded outside the stadium. A total of 43 streams out of the recorded 96 were actually watched. These 43 streams were accessed 85 times, 56,47% from friends; the rest from strangers. Thus the participants focused more on producing,

than on consuming.

Figure 8 shows the combined histograms for both streaming and watching behavior. Additionally, the graph shows important moments in the match as reported by an after match summary of the match by a local newspaper. These moments correspond to peaks in the histogram. Interestingly, the CoStream usage drops after every peak. Furthermore, CoStream has been used nearly throughout the whole match. While the novelty factor of CoStream probably accounts for a high application usage, the results from the interview provide evidence as to why participants streamed in certain situations. They commented that they either anticipated an exciting situation in the match or wanted to show their view to their friends in the stadium.

Occasionally, we observed that participants did not stop streaming. P3 explained that “streaming generated the pressure to continue streaming for others, even strangers”, since he did not want to “annoy the viewers”. We also observed that participants demanded to know whether others liked their stream, since it is “rewarding” and “an incentive” (Q1); P1 commented further: “When I share what I’m recording, I need to know if any one follows my stream”.

Complementing Spatial Awareness and Event Context. At the beginning of the study, we observed participants to first use CoStream in order to get an overview over nearby friends. P1 commented in the post-match interviews: “Once I knew where my friends were, they became a shortcut to their

Figure 8. Usage histogram during the first field study, sentiments correspond to the important incidents of the match.

Figure 9. P1 asked P2 and P3 to stream. They instantly recorded themselves and cheered with their friend.

location. I pointed toward my friends' direction whenever I wanted to see the match from their perspective. [...] I had the feeling my friends were in reach".

The participants repeatedly stressed that peeking at the other perspectives through the live videos fosters awareness for things happening in other parts of the stadium. Q2 commented: "You get so much more context [...] because I didn't know why spectators standing at the other side of the stadium are shouting angrily". Moreover, P3 stated: "The application helps me to know what is going on, on one side of the stadium at the same time of when I'm standing in an opposite side. I could stay connected to the match, even if I left the game for a while to grab something to eat".

Extending Experiences Beyond the Match. The participants commented that they typically switched to a stream when they could not get a good view and hoped for a better perspective. Apart from that, the participants also used CoStream for recordings not directly focused on the match. This is reflected in the histogram, which shows several aggregated sentiments, not corresponding to the important moments of the match. On the one hand, the participants used CoStream for sharing side-events, such as singing bystanders. On the other hand, they engaged in playful elements when the match was boring: the situation shown in Fig. 10 took place during the last 7 minutes of the match (cf. Fig. 8), where there was literally nothing exciting happening in the match.

Social Interaction and Communication. CoStream was often used for communication between friends (see Fig. 9) and also to communicate with viewers of a participant's own stream, i.e. with strangers in particular. For instance when P2 started streaming, he always recorded himself from time to time and waved; he commented: "I felt that I had to show that I am recording the stream". Participants considered streaming "something public" (Q2), since "everybody is also present in the same event" (Q3). Thus the participants did not care whether a stranger or a friend watched their stream, since "even bystanders can see what I record" (P4).

In the interviews, the participants noted that both push and pull features allow them to immediately interact with others, as Q3 stated: "Inviting friends to

a stream is absolutely important for immediately communicating an incident, be it event-related or not, among CoStream users". The participants also polled others to stream their surroundings, as Q4 commented: "I want to see my friends and the reactions of the fans around them". However, participants also struggled to focus their attention to the event while interacting with CoStream, as P3 described it: "at one point, I had the feeling that I did not participate in the event". We even observed P1 missing the first goal.

Interface Usability. Although the participants generally appreciated using CoStream, there were some concerns with respect to the application user interface and interaction techniques. While they stressed the importance of the conceptual subdivision into the four interaction modes, they mainly criticized that they had to switch between the modes via implicit embodied gestures, such as turning the device. They argued that e.g. turning the device seems appropriate in controlled environments; in a stadium however, other interactions such as cheering with the device in hand accidentally lead to a mode switch. Moreover, they demanded a dedicated recording button instead of the two-finger touch gesture to immediately trigger streaming.

5. INTERIM DISCUSSION

The observed phenomena provide evidence that CoStream supports the in-situ co-construction of shared experiences in various ways. First, CoStream *enriches social and spatial awareness*: the participants built a cognitive map of the stadium with their friends being landmarks, therefore serving as quick access shortcuts to different perspectives. This in turn also helped them to overcome social and physical restrictions, since they felt near to their friends. Second, CoStream *encourages active spectatorship*: the participants engaged with CoStream throughout the match to (1) record and also watch other streams and (2) to point their friends' attention to interesting streams they were either watching or recording. The latter also underlines that CoStream *enriches social interactions*: apart from sharing streams, the participants communicated with friends over

Figure 10. P3 and P4 engaged playfully because the "match was boring at that moment". P4 recorded P3 while he was watching P4's stream, generating a chained recording.

distance through video or even with the whole audience of their stream (as in the case of P2). Thus, CoStream has potential to *enhance and complement the overall event experience*: the participants did this deliberately either by polling others to stream their perspective or by just peeking at other perspectives, therefore gaining a richer context.

In addition to these benefits, study I also revealed a *tension* between the conventional *physical experience* of the event and the CoStream-based *digital experience* of the event. We could not classify this tension as either positive or negative. The more users connected to other participants through CoStream, the more they were distracted from the physical experience – and vice versa. This is also underlined by the results from the histogram analysis. On the one hand, important moments match peaks in the histogram. This shows that participants chose the event as a topic for their recordings, therefore connecting to the event through CoStream. On the other hand, we observed a drop in the CoStream usage right after each of those moments. In the post-match interviews, the majority of the participants stated that in such moments, reactions from the audience focused their attention away from the device back to the immediate perception. However, a contrary effect could be observed when surroundings and match became less interesting: participants intentionally ‘disconnected’ from the actual event in such cases, as described in the scene of figure 10.

In summary, the novel digital experience with CoStream ‘competes’ with the real-world experience. After study I, it was still unclear whether this tension had an impact on the overall event experience. We decided to conduct a second study to examine the tension in more detail. Additionally, we refined CoStream's user interface in accordance to the usability results.

6. FIELD STUDY II

We conducted a second field study during a mid-scale soccer match with 4500 spectators to address the aforementioned open research question. The study design was analogous to study 1. We recruited another 2 groups of 4 friends (P1-P4 and Q1-Q4; 7m, 1f; avg. 27 years), who did not participate in the first study, using a snowball sampling technique. All of them were regular attendees of soccer matches. They were introduced to CoStream upfront in a hands-on session and were paid the entrance fee and a mobile data plan. Overall, the session lasted about 4.5 hours.

6.1 Apparatus Refinement and Methodology

The results of the first field study showed that the concrete implementation of switching between

Figure 11. Revisited views, left: map view, showing spectator's Facebook avatars instead of the original color coding. All views now have a dedicated record button. Tapping it switches to stream mode, explicitly. Right: stream mode, notification area is highlighted.

CoStream's four conceptual modes was inappropriate and not explicit enough. To address this issue, we introduced a dedicated record button to every view, once being tapped, switches to the stream mode explicitly (see Fig. 11 left). Furthermore, the results from our study showed that active spectators, who are currently streaming a video, require awareness over users watching their very stream. We therefore added a notification area at the top of the interface, showing the current viewers (cf. Fig. 11 right).

We employed the same data gathering and analysis methodologies as in our first study. Figure 12 shows the participants' and observers' location during the two halves.

6.2 Results

The participants recorded a total of 106 videos during the match. Similar to the results of the first study, the videos' contents can be classified into three different categories: 54 videos recorded the moments related to the match, 47 showed social communication and 12 the surroundings. Note that several videos are classified in more than one category. A total of 58 streams out of the recorded 106 were actually watched. These 43 streams were totally accessed 80 times, 53.33% from friends. Figure 13 shows the accumulated usage histogram for both streaming and watching behavior, which

Figure 12. Participant locations during the first study.

Figure 13. Usage during the second field study. The sentiments correspond to the important incidents of the match.

indicates that the important moments of the event such as goal chances or fouls again match the peaks. This is in line with our qualitative findings, as Q3 stated: “I want to stream and share the match-winning scenes” and confirms the results from the first study. We present the results from the second study below.

Required Attention While Recording. Throughout the study, we observed that participants easily streamed video for a long period of time, while preserving their attention to the actual event (cf. Fig 15). P2 commented: “I was continuously streaming video during the second half, since the players were frequently approaching the goal and I hoped to stream a strike”. P1 added: “I put the device into my chest pocket while streaming and simultaneously clapped hands to applaud my team” (cf. Fig. 14).

Coupling Between Physical and Digital Experiences. Our observations revealed watching video (as a digital experience) required more attention to the device. The participants repeatedly stressed that it is important for the stream to be live, since “the stream itself fosters awareness over the current situation in the stadium” (P3). In line with this is Q3 commenting: “It was great that the stream was live. So thus what I watched on the device, matched what I heard from the atmosphere and other spectators around me”. However, in the post-match interviews, the majority of the participants stated that reactions from the audience focus their attention away from the device back to the match (as a physical experience). P4 noted that “when the spectators are cheering, my attention draws back into the match”. Occasionally, we observed the participants use CoStream’s replay functionality to replay certain scenes. Q1 said: “during a match in stadium without CoStream, there is no chance to replay scenes, particularly those of other spectators around”. However, P4 added: “I imagine that if a match were really intense with lot of action, I would

not use the replay function, to not miss anything”.

Synchronous Communication. The participants did not want to use any additional communication such as audio, as P2 stated: “A soccer match is too loud. This makes audio communication almost impossible. And texting is certainly not an option, it requires too much attention”. Q4 elaborated on this “and if I sent my friends a message related to the match, they might not read it immediately and later on, wouldn’t understand it”.

7. DISCUSSION AND SUMMARY

As outlined above, the first study revealed a tension between the actual event and the application usage. As a consequence, spectators might become disconnected from the event and miss important scenes. The results from the second study highlight two important phenomena: (1) CoStream contributes to the event through a strong real-time coupling between physical and digital experiences and therefore (2) the tension between CoStream and the actual event can be characterized as an interplay of both experiences.

We propose to conceptualize this interplay as a *focus+context* [3] approach to experiencing events with ubiquitous live multimedia sharing. While the participants were watching a live stream, CoStream was their *focus*. The actual *context* with respect to the event was preserved through both listening to the atmosphere in the stadium (e.g. cheering sounds) and peripheral vision (e.g. immediate reactions of bystanders). Vice versa, while streaming, their attention switched away from CoStream (cf. Figures 14 and 15), the event became the actual *focus*; CoStream was still in the peripheral *context* (e.g. visually or vibrotactile). Moreover, both event and CoStream provide means to support the fluid transition between focus and context. On the one hand, the atmosphere can draw

Figure 14. P1 put his device into his chest pocket, continued streaming and cheered for his team.

Figure 15. P2 is streaming and not focusing on the display, but on the actual event.

a user's attention to the event. On the other hand, CoStream supports interactions such as pushing and pulling video streams to users, therefore demanding their attention. Our studies show that CoStream enhances and intertwines with the overall event. Key to this is that

- both physical and digital experiences reflect the peculiarities of activities in the very same event
- the system provides a strong real-time coupling to the event, the spectators' locations and their social relation
- the UI enables the fluid transition between focus and context through (a) providing efficient overview and awareness, supporting (b) active engagement, (c) immediate interaction and (d) reducing visual attention.

Otherwise physical and digital experiences will be decoupled, leading to disconnections (*sensu* Turkle [30]). These findings go beyond prior work, focused on either active media creation as a memory aid to relive experiences later on, asynchronous sharing of multimedia over distance or providing additional information. In conclusion, we believe that future systems, adhering to the design guidelines of CoStream, will pave the way for new possibilities to co-construct shared experiences *in-situ* through the ubiquitous sharing of multimedia intertwined with physical event experiences in *real-time*.

8. REFERENCES

- [1] Barkhuus, L., Jørgensen, T. Engaging the crowd: studies of audience-performer interaction. In Proc. CHI EA '08, ACM, 2925-2930.
- [2] Bentley, F., Groble M., TuVista: meeting the multimedia needs of mobile sports fans. In MM'09, ACM, 471-480.
- [3] Card, S.K., Mackinlay, J.D., and Shneiderman, B. (Eds.) Readings in Information Visualization: Using Vision to Think, pp.,1999, Morgan Kaufmann, 1-34.
- [4] Cesar, P., Bulterman, D.C.A., Jansen, J., Geerts, D., Knoche, H., Seager, W. Fragment, tag, enrich, and send: Enhancing social sharing of video. ACM Trans. Multimedia Comput. Commun'09. Appl. 5, 3, 27 p.
- [5] Counts, S.J., Fellheimer, E. Supporting social presence through lightweight photo sharing on and off the desktop. In Proc. CHI '04. ACM, 599-606.
- [6] Esbjörnsson, M., Brown, B., Juhlin, O., Normark, D., Östergren, M., Laurier, E. Watching the cars go round and round: designing for active spectating. In CHI '06.
- [7] Fouse, A., Weibel, N., Huchins, E., Hollan, J.D. ChronoViz: a system for supporting navigation of time-coded data. In Proc. CHI EA '11. ACM, 299-304.
- [8] Frohlich, D., Kuchinsky, A., Pering, C., Don, A., Ariss, S. Requirements for photoware. In CSCW '02. 166-175.
- [9] Hagedorn, J., Hailpern, J., Karahalios, K. Vcode and VData: Illustrating a new framework for supporting the video annotation workflow. In Proc. AVI '08. ACM.
- [10] Hollan, J., Scott, S. Beyond Being There. In Proc. CHI '92, Penny Bauersfeld, John Bennett, and Gene Lynch (Eds.). ACM, 119-125.
- [11] Holmquist, L.E., Falk, J., and Wigström, J. Supporting group collaboration with interpersonal awareness devices. In Proc. Personal, Ubiquitous Computing'99.
- [12] Jacucci, G., Oulasvirta, A., Ilmonen, T., Evans, J., Salovaara, A. CoMedia: Mobile Group Media for Active Spectatorship. In Proc. CHI '07. ACM, 1273.
- [13] Jacucci, G., Oulasvirta, A., Salovaara, A. Active construction of experience through multimedia: a field study with implications for recording and sharing. In Proc. Personal Ubiquitous Comput'07, 215-234.
- [14] Julhin, O., Engström, A., Reponen E., Mobile broadcasting: the whats and hows of live video as a social medium. In proc. MobileHCI'10, ACM, 34-44.
- [15] Kindberg, T., Spasojevic, M., Fleck, R., Sellen, A. The ubiquitous camera: an in-depth study of camera phone use. IEEE Pervasive Computing '05.
- [16] Liu, Y., Shafton, P., Shamma, D., Yang, J. Zync: the design of synchronized video sharing. In Proc. DUX '07. ACM.
- [17] Marcus, A., Bernstein, M. S., Badar, O., Karger, D. R., Madden, S., Miller, R. C. TwitInfo: aggregating and visualizing microblogs for event exploration. In Proc. CHI '11. ACM, 227-236.
- [18] Mäkelä, A., Giller, V., Tscheligi, M., Sefelin, R. Joking, storytelling, art sharing, expressing affection: a field trial of how children and their social network communicate with digital images in leisure time. In CHI '00., 548-555.
- [19] Nilsson, A. Using IT to make place in space: evaluating mobile technology support for sport spectators. In Proc. ECIS 2004.
- [20] Nilsson, A., Nulden, U., Olsson, D. Spectator information support: exploring the context of distributed events. The interaction society: practice, theories and supportive technologies, IDEA 2004.
- [21] Peltonen, P., Salovaara, A., Jaccucci, G., Ilmonen T., T., Ardito, C., Saarikko, P., Batra V. Extending Large-scale Event Participants with User-created Mobile Media on Public Display. In Proc. MUM '07. ACM, 131.
- [22] Reeves, S., Benford, S., O'Malley, C., Fraser, M. Designing the spectator experience. In Proc. CHI '05. ACM, 741-750.
- [23] Reponen, E. Live @ Dublin- Mobile Phone Live Video Group Communication Experiment. In Proc. EUROITV '08, Springer-Verlag, 133-142.
- [24] Reponen, E., Lehtikainen, J., Impiö, J. Mobile Phone Video Camera in Social Context. In Proc. HCI'07, Julie A. Jacko (Ed.). Springer-Verlag, 460-469.
- [25] Sahami Shirazi, A., Rohs, M., Schleicher, R. Real-Time Nonverbal Opinion Sharing through Mobile Phones during Sport Events. In Proc. CHI '11. ACM.
- [26] Salovaara, A., Jacucci, G., Oulasvirta, A., et al., S. Collective creation and sense-making of mobile media. In Proc. CHI '06, ACM, 1211-1220.
- [27] Shamma, D. A., and Liu, Y. Zync with Me: Synchronized Sharing of Video through Instant Messaging. Social Interactive Television: Immersive Experiences and Perspectives '09.
- [28] Shamma, D.A., Kennedy, L., and Churchill, E.F. Tweet the debates: understanding community annotation of uncollected sources. In Proc. WSM '09. ACM, 3-10.
- [29] Strauss, A., Corbin, J. Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory. Sage, 2008.
- [30] Turkle, S., Alone Together: why we expect more from technology and less from each other, Basic books 2011.
- [31] Wittenburg, P. Brugman, H., Russel, A., Klassmann, A., Sloetjes H., ELAN: A professional framework for multi-modality research. In Proc. LREC '06.