As If I Am There: A New Video Chat Interface Design for Richer Contextual Awareness

Jiajing Guo, Yoyo Tsung-yu Hou, Michelle Lee, Harley Mueller, Katherine Tang, Alexander Clapacs, Susan R. Fussell
Department of Information Science, Cornell University
Ithaca NY, 14850 USA
[jg2263, th588, tl428, hrm58, kat86, ac927, sfussell]@cornell.edu

ABSTRACT
Video chat for personal communication (e.g. chat with friends and family members) has been widely employed in many communication software and Apps. In spite of advanced technology, problems still exist in video communication, as well as design opportunities. In this study, we focused on video communication for personal purpose and designed a novel video chat interaction approach in order to offer richer environmental information. Our study shows that there are many potentials for the employment of 360° photo in mobile video chat when communicating with friends and family.

Author Keywords
Computer-mediated-communication; video chat; contextual Information; mobile; interface design; 360° photo.

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous; See http://acm.org/about/class/1998 for the full list of ACM classifiers. This section is required.

INTRODUCTION
Video communication has been studied for many years in area of Computer-mediate Communication. However, in spite of increasingly mature technologies, there are still many issues around the use of video as communication media [1, 3, 15]. This study looks into the issue of contextual awareness in video chat, and presents a novel video chat interface in order to offer richer environmental information.

Unexpected Events in Video Chat
Due to the lack of contextual information, Unexpected events are common in video chat. For example, a roommate who comes into the room during a video chat may be an unexpected participant. If it is in a face-to-face situation, one is very likely to know the roommate is at home. One suddenly disappears from the video seems to be an unexpected event to the remote partner, and may cause misunderstanding between the two. But the partner may not know it is because the chair is broken. In a face-to-face situation, the partner is much more likely to know the disappearance is due to the broken chair.

The feeling of unexpectedness in video chat may come from the lack of grounding knowledge [5], especially due to the lack of physical co-presence. Some of these unexpected events may result in negative consequences, such as misunderstanding, embarrassment, etc., while others may break the ice or bring humor, and facilitate the efficacy of communication.

What are the effects of unexpected events in video chat in terms of people's behavioral and affective responses? How are these responses related to event type, participants' relationship and chat content? How these findings inform the design of CMC tools? To date these questions remain unanswered.

Given this, we were interested to know how we can provide richer contextual information in video communication. To understand this potential, we created a novel video communication interface to increase contextual awareness through multiple media presentation. Following a whole HCI design process, we first adopted “unexpected event” as a starting point and conducted interviews to understand users’ experience of unexpected events in video chat, their emotions and reactions under those circumstances, and their general perspectives. The interviews showed that participants have sophisticated responses to a variety of unexpected events. But most of them reveal the demand to know more about the remote one when video chatting with family and friends. Based on these findings, we brainstormed ideas to explore the possibilities of video chat system which provide richer contextual information.

In the following part of this paper, we detail the iterative design process of our new interface and evaluation to understand how users adapt our design and how the design impacts communication among friends and family members.

USER RESEARCH

To identify design opportunities in enhancing video communication experience, we conducted one round of user interviews to understand the effects of unexpected events on users' behavioral and affective responses during a video chat. We believe user interviews were good starting points because it allowed us to understand users' pain points during video chat and help us narrow down the scope of our project.

Procedures

For the preliminary user interviews, we recruited 18 participants (14 female, 3 male). All the participants are college students. They received course credits for their participation. All the interviews were audio recorded under the permission of participants.

We interviewed the participants regarding their experiences with unexpected events during video chat, focusing on three different aspects: unexpected participants, unexpected content, and technical difficulties.

In each interview, we first asked the participant to answer a few questions regarding their video-chat habits. Then we asked the participant to recall a specific video-chat experience that involves an unexpected participant. If the user was unsure what "unexpected participant" imply, we would provide the participant with very general explanation or an example to avoid guiding their thought process. To illustrate, we would describe "unexpected participant" as something unexpected that the participant or the participant's partner may see on the screen. We did not restrict the concept of "participant" to just people. We wanted to allow participants to bring up anything they consider participants, regardless of whether it is a pet, a robot or any other things that came to their minds. Once the participant thought of a specific experience, we asked the participant to describe the specific experience and answer questions on the participant's emotion and reaction during the experience.

A few key questions from the user interview protocol are highlighted below:

- What happened at that time (when the unexpected event happened)?
- What was displayed on the screen?
- What emotions did you have? (annoyed, anxious)
- Do you think he/she had the same feeling with you? How did you tell?
- How did you behave at that time?
- What was his/her response?

After interviewing participants regarding their video chat experience involving "unexpected participants", the same steps were repeated for "unexpected content" and "technical difficulties".

Results

The interviews showed a variety of diverse video calling experiences from the participants falling under our three categories of unexpected video calling incidents. During these video calls, the majority of participants were calling family and friends. The other most common video call setting for the appearance of unexpected events was during interviews, professional and casual.

Each event elicited very unique and different reactions compared to another event from both the participant and their call partner depending on the current situation, setting, and the nature of their video call. For example, two participants both had their own past experience during a professional interview during which an unexpected participant showed up in the middle of their call. For one participant, having the unexpected participant made her feel less nervous and broke the ice between herself and the interviewer.

P7: His colleague came in and said, "Hey Scott! … Oh sorry you are in … in an interview. I wanna go now." And he closed the door. It kinda broke the ice. He started laughing, I started laughing.

However, for another participant, the appearance of the unexpected participant annoyed her.

From the results of the interviews and the various past unexpected video call incidences participants experienced, we then categorized each event based on its context and discovered many were related to "environment context", most often due to either the participant or their caller not knowing about the other's physical surroundings or presence of others not directly involved in the video conversation. Other notable categories included "informational context" (participants or their partners were missing conversational information), "technical issues" (problems with the video calling platform or technology occurred), and "randomness" (sudden changes in the video call conversation itself).

Based on these results, we found contextual information is an essential factor that would impact the experience of video chat. And there are many potentials for the improvement of video communication with family and friends.

VIDEO COMMUNICATION IN PERSONAL LIFE

In recent years, research focus has moved from professional life to personal life [1, 3, 13, 15, 16]. Different from video communication in workspace, video chat with friends and family members focus more on building and maintaining connectedness, trust and love [1, 2].

Shared Experience Instead of Talking Heads

Shared Activities

Recent research finds that in personal domain, individuals tend to utilize video communication beyond the desk, moving from conversation to shared activities. Video chat

among friends and family does not have a particular conversational task but more a shared everyday experience of mundane activities together [2]. For example, children show things and talk to grandparents [8], long-distance couples share everyday activities [16], children and teenagers play and study together [4, 20].

Presentation in video chat

Talking head concept describes the scenario that simulates face-to-face conversation between two seated individuals. Different from video conference which usually happens in the office via desktop devices, video chat among family and friends may happen at home or outdoor via both desktop and mobile devices. Recent research showed that objects, mediated presentation, and activities may play a more significant role in personal video chat [1].

Locations and Atmosphere

Shared activities often happen not only a fixed place at home, also outdoor, unexpected places and moving around.

The most common shared activities are to show the remote location, but users' appropriations extend the reach of webcams into a variety of places such as living rooms, kitchens, and gardens. Individuals tend to regard video communication in personal life as a result of relocation [3]. Sometimes people want to show some specific objects or give a "tour" of the house by moving the camera. It suggests the potential of configurable and portable video communication devices in personal domain [13].

In a regular video chat, information is two-way and equal for both individuals. While in some circumstance, the atmosphere in two locations is imbalanced, and one side may need more information than the other. Recent research shows that locations with different atmosphere should also be taken into account [15]. For example, grandparents attend child's graduation and wedding via Skype. In this situation, grandparents may want to know and feel more about the remote location.

Technologies for Domestic Video Communication

There has been a large amount of research on video conferencing in workspace, along with much design and deployment of video communication technologies at home in the recent years.

HomeProxy is an integrated, no-touch video messaging prototype that allows shared activities for distributed family members. Users can operate the system via gesture control, watch recorded or live video messages, and interact with the remote partner.

Family window [9] is a domestic media system that connects family members over distance. Via the always-on video, family members can see their loved ones on a regular basis. They are able to see live things and everyday domestic activities, which largely increases the sense of believing and connectedness.

ShareTable [21] is a shared tabletop task space and an easy-to-initiate video chat system which project a remote table surface and allow remote parents to do activities with children. ShareTable allows the communication to focus more on shared activities rather than conversation, which created a great sense of closeness and connection.

Technologies above are domestic video communication systems which are deployed in a fixed place at home. Many of them require multiple cameras [6, 14], projectors [20, 21], and display screen [9, 19]. However, few research looks at portable devices for video communication among family and friends that is able to support shared activities in various locations.

MOBILE VIDEO CONFERENCES

Commercial video applications such as Messenger, Skype, and FaceTime have been widely employed. And some studies have investigated the use of mobile video chat when sharing activities with family and friends. In spite of the convenience brought from current advanced technology, some issues still exist and impact experience during mobile video chat. On the one hand, as the camera handler, users always try to provide a good view for the remote partner but usually hold the phone in awkward ways, either too close to the body or to the chin. As a result, they become too engaged in the video chat to notice the surroundings [7]. Some technology probe such as Shared Geocaching [17] uses camera on glasses as a video input, which solved the problem but also raise more issues such as video quality, camera orientation, continuous moving camera etc.

On the other hand, as the remote partner, users find the limited field of view and lack of camera control would result in frustrating experience [7]. Previous research prototype also emphasized the importance of spatial context. Kim et al.'s prototype provides three types of contextual information in mobile video communication: map, live video, and a high-resolution image. The image enables users to review what happened without affecting the live conversation. It indicates that video is not always the best form of contextual information [12].

Also, users are in need of the ability to see both the physical environment in front of the handler and the face at the same time [6, 17]. Current mobile video chat system seldom satisfies this need. Users either turn on front or rear camera to show the face or the scenario in front of them.

Several prototypes have tried to give more environmental information by utilizing 360° camera to create an immersive experience. For example, "JackIn Head" [10, 11] includes a head mounted omnidirectional camera and image stabilization. The remote partner is able to see a first-person omnidirectional video via an immersive viewing device. Tang et al.'s [18] prototype employed a fixed 360° camera and audio communication. They found that remote users are able to explore the environment independently and assist the local partner in navigation tasks. While participants also

have difficulties in communicating location and orientation information. These prototypes emphasize agency of remote participants and offer control of views. But they require specialized devices and image processing technologies, which may not be suitable for everyday activities.

DESIGN

Informed by our user study and previous literature, and existing technologies, we ran several brainstorming sessions to generate potential designs to improve current interface for video chatting. We then created some prototypes to get an initial sense of how these ideas might work in real-world scenarios. In the end, we narrowed down our scope to one idea and came up with an innovative way to test our idea without having to implement a new system for video chatting.

Concept Formation

We centered our brainstorming session around the concept of contextual awareness, because it was the emerging common theme that we found from our interviews on unexpected events. We found participants had got both positive and negative responses to the lack of contextual awareness. In our brainstorm session we focused not on "solving the problem". Instead, we tried to be creative and explorative in investigating the opportunities of contextual awareness in video chat (Figure 1).

Figure 1. We try to be creative and explorative, rather than just "solving the problem", in our brainstorming sessions.

Here are some of our favorite ideas (Figure 2).

- Notify users when other people enter the room where video chat takes place or are in a nearby location.
- Display of panorama / 360° photo of users' environments.
- Offer fun ways to pause video chat, such as allow users to "close curtain on screen".
- AR projection of partner over chosen camera view / Overlay users over views captured by rear camera or screen sharing.

Figure 2. The "Black horse" ideas from brainstorming.

Idea Testing and Consolidation

Prototype

We used our design sketches to collect opinions about these ideas through informal interviews. Based on feedback and prototype feasibility, we decide to combine the two ideas, panorama and AR projection, both are about showing the environment during a video chat.

After we made a decision on what idea to test, we start another round of discussion and brainstorming on how this idea can be implemented, both for testing phase and for the final product phase. Here are some visual mockups of our ideas.

Augmenting Current Video chat Interface

The simplest way to incorporate environmental information into a video chat interface is to augment a current one (Figure 3). In this design, people will be asked to take a picture of their surroundings when they make (or receive) a phone call, and the picture will be the background during the whole conversation. This requires minimal change to the current interface. However, the picture is static and does not provide enough information of the environment.

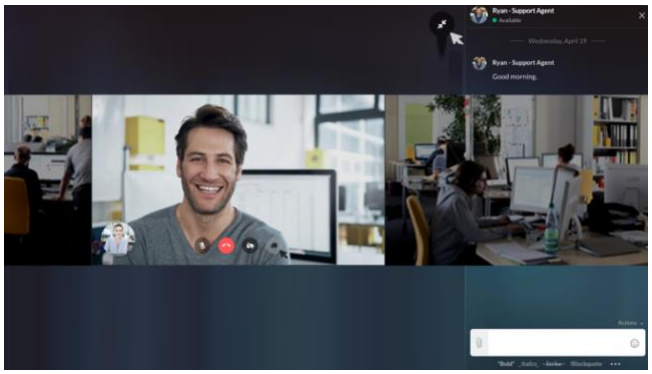

Figure 3. Augmenting current interface is the easiest way to incorporate environmental information into video chat.

Overlay Users Over Panorama Views

In this design, while talking on the phone, the image of distant speakers will be overlayed onto a background, which is a panorama view of their environment, captured by the rear camera. Users is able to see the whole environment by simply tilting and rotating the phone (Figure 4).

Figure 4. Overlay Users Over Panorama Views. Viewers can rotate their phone to see the whole environment.

In our working prototype for user testing, we used Skype, which provides a floating video chat window on Android Phones, and Messenger, which affords easy creation and sharing of a 360° picture on IOS system. (Figure 5). We therefore can test our idea of a new interface without having to implement a new system.

Figure 5. Working prototype for testing. An Skype floating window overlays a 360° picture on Facebook app.

PROTOTYPE EVALUATION

After solidifying a functional prototype, the next stage was to evaluate the prototype in a live setting with research participants. The purpose of this evaluation was to examine how our conclusions and design choices from the prototyping phase would be received by users.

Procedure

The evaluation phase of the project was an approximately 30-minute observational study of a participant using the prototype. This was then followed by an online post-task survey, in which the user reported on their experiences during the test and with the prototype. The results of this survey were reviewed and discussed between the researcher and the participant to conclude the session.

In each interview, we first asked the participant to select a "partner" from their Facebook Messenger contacts for the study. The participant was instructed not to inform the partner that they were taking part in a study. The participant would then ask for the partner to take a 360° photo of their surroundings, and to send this photo back to the participant via Messenger. Then, the call began, using the Messenger video call feature and the 360° photo as described in the prototype section.

The participant was instructed to have the call last approximately 10 minutes, and to discuss normal topics that they might discuss with the partner in ordinary call situations. Pictures were taken of the subject, the environment, and the prototype in action. During the call itself, the researcher observed key points of the call, such as the positioning of the call window, the amount of use of the 360°photo by the participant, and the partner's reaction to the participant's movement of the phone. Finally, the participant was asked to take a screenshot of their phone while in the call. Once the call has finished, the participant completed the survey for review with the researcher. The survey asked for the participant to describe their experience through a series of questions focusing on co-presence, psychological involvement, and behavioral engagement.

The review of the survey involved discussing the participant's answers to the survey questions. For example, if the participant had a low response about co-presence, then the researcher inquired further with questions such as; "Why did you feel alone?", "Why did you feel easily distracted?", "Why did you think your partner did not pay attention to you?". For the survey questions comparing the study experience to other video calls, the researcher asked questions such as; "What aspects of this call was better/worse? Why?", "Are you willing to share your own environment through a 360° photo? Why?". The conversation between participants and the researcher were audio recorded for further analysis.

Results

Due to time constraints, we recruited 3 participants for the user interviews. The interviews all concluded that the

increased contextual information was very beneficial for the users. Through the survey results, we found that the users had an overall high sense of mutual understanding with their partners during the video chat. The prototype was designed to increase conversational grounding by placing each participant in the other's physical location. With this in mind, users can cite objects in their partner's background and notice things they would not have known were there. One participant was speaking with their mother, but because of the 360° photo became aware that their father and dog were also in the same room.

Participants explained that they enjoyed sharing their own environment, especially interesting environments and with people they have intimate relationships with. They also expressed interest in being able to see their partner's environment to know where they are and see what they are seeing.

Users also claimed to feel less isolated and more focused in on their partner in the post-interview survey. One participant, in particular, explained that having the 360° picture of their partner's environment there made her feel more inclined to look at the picture rather than looking around her own physical environment. The participant explained that she typically gets distracted by her laptop or other things in her physical environment or feels inclined to multitask during video chats. However, the 360° photo made her feel like she was within her partner's environment.

CONCLUSION

This paper presents the process of designing video communication interface for personal life with the purpose of providing richer contextual awareness. Our user evaluation shows positive feedback on our design. But the data collected is not enough to generalize our conclusion. In the future, we plan to collect more user feedback and iterate our prototype. We also want to see the effect of this prototype in different situations, such as emergency call and major events.

ACKNOWLEDGMENTS

We thank all the participants and teaching staff in this course.

REFERENCES

[1] Ames, M.G., Go, J., Kaye, J. "Jofish" and Spasojevic, M. 2010. Making love in the network closet: the benefits and work of family videochat. *Proceedings of the 2010 ACM conference on Computer supported cooperative work - CSCW '10* (2010), 145.

[2] Brubaker, J.R., Venolia, G. and Tang, J.C. 2012. Focusing on shared experiences. *Proceedings of the Designing Interactive Systems Conference on - DIS '12* (2012), 96.

[3] Brubaker, J.R., Venolia, G. and Tang, J.C. 2012. Focusing on shared experiences: moving beyond the camera in video communication. *Proceedings of the Designing Interactive Systems Conference on - DIS '12* (2012), 96.

[4] Buhler, T., Neustaedter, C. and Hillman, S. 2013. How and why teenagers use video chat. *Proceedings of the 2013 conference on Computer supported cooperative work - CSCW '13* (2013), 759.

[5] Clark, H.H. and Brennan, S.E. 1991. *Grounding in communication*.

[6] Inkpen, K., Taylor, B., Junuzovic, S., Tang, J. and Venolia, G. 2013. Experiences2Go: sharing kids' activities outside the home with remote family members. *Proceedings of the 2013 conference on Computer supported cooperative work - CSCW '13* (New York, New York, USA, 2013), 1329.

[7] Jones, B., Witcraft, A., Bateman, S., Neustaedter, C. and Tang, A. 2015. Mechanics of Camera Work in Mobile Video Collaboration. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15* (New York, New York, USA, 2015), 957–966.

[8] Judge, T.K. and Neustaedter, C. 2010. Sharing conversation and sharing life: video conferencing in the home. *Proceedings of the 28th international conference on Human factors in computing systems - CHI '10* (New York, New York, USA, 2010), 655.

[9] Judge, T.K., Neustaedter, C. and Kurtz, A.F. 2010. The family window: the design and evaluation of a domestic media space. *Proceedings of the 28th international conference on Human factors in computing systems - CHI '10* (New York, New York, USA, 2010), 2361.

[10] Kasahara, S., Nagai, S. and Rekimoto, J. 2015. First Person Omnidirectional Video: System Design and Implications for Immersive Experience. *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video - TVX '15* (New York, New York, USA, 2015), 33–42.

[11] Kasahara, S. and Rekimoto, J. 2015. JackIn head: immersive visual telepresence system with omnidirectional wearable camera for remote collaboration. *Proceedings of the 21st ACM Symposium on Virtual Reality Software and Technology - VRST '15* (New York, New York, USA, 2015), 217–225.

[12] Kim, S., Junuzovic, S. and Inkpen, K. 2014. The Nomad and the Couch Potato: Enriching Mobile Shared Experiences with Contextual Information. *Proceedings of the 18th International Conference*

on Supporting Group Work - GROUP '14 (New York, New York, USA, 2014), 167–177.

[13] Kirk, D.S., Sellen, A. and Cao, X. 2010. Home video communication: mediating "closeness." *Proceedings of the 2010 ACM conference on Computer supported cooperative work - CSCW '10* (New York, New York, USA, 2010), 135.

[14] MacCormick, J. and John 2013. Video chat with multiple cameras. *Proceedings of the 2013 conference on Computer supported cooperative work companion - CSCW '13* (New York, New York, USA, 2013), 195.

[15] Massimi, M. and Neustaedter, C. 2014. Moving from talking heads to newlyweds: exploring video chat use during major life events. *Proceedings of the 2014 conference on Designing interactive systems - DIS '14* (2014), 43–52.

[16] Neustaedter, C. and Greenberg, S. 2012. Intimacy in long-distance relationships over video chat. *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems - CHI '12* (New York, New York, USA, 2012), 753.

[17] Procyk, J., Neustaedter, C., Pang, C., Tang, A., Judge, T.K., Procyk, J., Neustaedter, C., Pang, C., Tang, A. and Judge, T.K. 2014. Exploring video streaming in public settings: shared geocaching over distance using mobile video chat. *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14* (New York, New York, USA, 2014), 2163–2172.

[18] Tang, A., Fakourfar, O., Neustaedter, C. and Bateman, S. 2017. Collaboration with 360° Videochat: Challenges and Opportunities. *Proceedings of the 2017 Conference on Designing Interactive Systems - DIS '17* (New York, New York, USA, 2017), 1327–1339.

[19] Tang, J.C., Xiao, R., Hoff, A., Venolia, G., Therien, P. and Roseway, A. 2013. HomeProxy: exploring a physical proxy for video communication in the home. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13* (New York, New York, USA, 2013), 1339.

[20] Yarosh, S., Inkpen, K.M. and Brush, A.J.B. 2010. Video playdate: toward free play across distance. *Proceedings of the 28th international conference on Human factors in computing systems - CHI '10*. (2010), 1251–1260. DOI:https://doi.org/10.1145/1753326.1753514.

[21] Yarosh, S., Tang, A., Mokashi, S. and Abowd, G.D. 2013. Almost touching: Parent-child remote communication using the sharetable system. *2013 2nd ACM Conference on Computer Supported Cooperative Work, CSCW 2013*. (2013), 181–192. DOI:https://doi.org/10.1145/2441776.2441798.