

Visual Awareness in the Home

J M Thorne†

†University College London and BT Group CTO

Abstract: This paper discusses the use of Video Snapshots as a compromise between the stillness of a photo and the liveness of real-time video, for providing visual awareness to support connectedness in and between homes. It is hoped that Video Snapshots will be more emotionally engaging than still photos, yet less intrusive than live video.

1 Introduction.

With the rise in single and lone parent households [1], individuals are increasingly geographically disparate. Humans have basic needs to belong and form relationships, as such, an increasing amount of communication will occur across distances or in brief face to face encounters.

When people are sharing a physical space it is easy to slip in and out of conversation. The awareness of each others' activities and moods implicitly keeps the conversational channel open such that greetings and farewells are not required for each brief fragment of explicit communication [2]. As people become physically separated this awareness is lost and communication channels become much more binary - either open or closed. When a shared meal ends or the telephone is put down, no further communication is possible until the next meeting occurs or a new call is made.

A sense of connectedness remains between the parties, but this will fade with time if they cannot judge each other's availability for a new contact, cannot demonstrate their willingness to continue the relationship and cannot engage in shared activities or physical touch [3].

Always on, background video based displays might be used to restore a sense of awareness, such that some implicit communication remains, activities are seen as shared and connectedness is maintained. Such displays might also make judgements of availability easier and help users gain familiarity with video as a medium for communication.

This paper concerns the use of visual awareness to support connectedness and informality of visual communications in the home. Two extremes of video display that will be considered are the static photograph and continuous live video. An argument for a third form that lies between these extremes will be made.

2 Video Snapshots

Could some compromise between a static photo and live video improve support for emotional connectedness, but without becoming too distracting or suggesting capabilities that do not exist? This is what we are trying to achieve with Video Snapshots. Video Snapshots are short fragments of video captured automatically at regular intervals. The power of a photograph that embodies some of the qualities of live video is exemplified by the magical photographs in the Harry Potter films [4]. These 'video snapshots' vividly capture a moment in time.

In our initial implementation [5] one second of video was captured every 10 seconds and then looped backwards and forwards in the remote display. This looping avoided a jump in the display when the clip was restarted. However recent research [8] suggests that fidelity of motion is important in the perception of emotion and consequently reversing video clips may have an effect on this perception. A second implementation presents the first frame from the clip as a static image for a few seconds then ramps up to real frame rate through the rest of the snapshot, slows down and holds the final image static for the remaining duration of the interval between snapshots. A smooth blend is used to transition between the last frame of one and the first frame of the next snapshot.

If this interface is to be presented in the periphery and only glanced at occasionally, it must be designed to provide all the information needed in that quick glance. So in conjunction with the 'live' snapshot we can also present a history of behaviour (See Figure 1). For example, on glancing at the live display I might see an empty room, but in the image taken a minute ago I can see you sitting on the sofa, I can thus surmise that you've just left and may come back shortly.



Figure 1: Model of a ‘smart’ picture frame displaying Video Snapshots. Overlaid, spiral history of previous snapshots, largest image is most recent - 10 seconds ago, 1 minute ago, 5 minutes ago, 10 minutes ago.

3 Natural Video

If these displays present natural video, that is video images direct from a camera, then any subtleties of expression and posture necessary for understanding the mood and emotion of that user should be preserved. It could also be argued that humans need natural displays, to quote a recent blog entry

“We never had to learn to process body language, facial expressions, and tone of voice. We evolved this capability...it’s innate. no matter how much we practice communicating through text, the brain still finds it stressful.” [6].

This ease of perception is important if we are to create calm technology [7]. That is, instead of creating yet more devices that clamour for our attention, we try to design non-distracting interfaces that are just there, when we need them, and that easily slip between the periphery and the centre of our attention.

The static photograph is easily ignored and yet is sufficient to conjure up the memory of a person, a past shared activity and associated feelings of connectedness. If the moment at which a photo is taken is chosen with care, then the giving of that photo also allows a person to demonstrate their commitment to the relationship. The single photo cannot though give any sense of current availability and the perceived intensity of emotion is weaker than that from video [8]. However, continuous video can be very distracting - a TV in the corner of the room can easily capture attention. A Video Snapshot might find a balance between the two, partially static so less distracting than video, but containing some motion so that the emotional content is perceived as more intense than a photograph and frequently updated to give a sense of what’s happening now.

4 Privacy

Unfortunately, unfiltered video can expose more than we would like. Not all facial expression is voluntary, and much of our habitual behaviour is unconscious. This unintended behaviour would be transmitted as faithfully as any deliberately exposed actions. The home is also seen as a very private location. Away from the perceptions of neighbours and colleagues the home allows people to drop their public mask and be themselves. If part of that space is in the field of view of a camera that is continuously capturing and transmitting, then that area has essentially become public again.

Other users of the space must also be considered - for example a tele-worker might tolerate a connection to the office for the benefit of feeling part of the group, but the tele-worker’s spouse is not part of that relationship, gains little benefit from the connection but experiences the same invasion of privacy and so might feel awkward using the same space [9].

It is worth noting that attitudes to privacy are culturally dependent and changing. Young people, growing up with CCTV and the internet expect others to know about them, and are most concerned with how they can manage their image [10].

A key concept in mitigating privacy concerns is that of reciprocity - if you can see me, I can see you. In a trusting relationship, this mutual waiver of a right to privacy might be tolerated.

Neustaedter [9] details a number of design guidelines to protect user privacy, including dedicated physical

controls on the level of privacy and visual feedback of the current state.

A potential problem with physical controls that allow users to increase the level of privacy is that they may forget to later reduce the privacy level so that awareness is restored. If the benefits of awareness need time to take effect, then learning to restore broken awareness may be difficult.

An example of visual feedback would be to make the field of view of active cameras visible, so that it would be possible to tell if a remote viewer would be able to see a person or not. This would allow users to make the choice to step around the zone if they did not want to be seen.

Once a static photograph has been taken and shared with a closed group of people, no further loss of privacy occurs. In contrast, continuous video will constantly be revealing new facts. Video Snapshots might find a balance by only sampling a fraction of the time, however the instant at which that sample is taken lies outside of user control and although the chances of capturing a revealing moment is decreased, if one is captured then it will be displayed for longer.

5 Affordances

A recent paper on the use of reciprocal background visual awareness [11] raises some interesting issues. In their experiment participants watched a football match simultaneously but in different locations, and were presented with visual awareness of each other in the form of a projection onto the wall behind the TV.

Watching football is a shared activity with a fair amount of interaction - “did you see that?”, “my team’s winning!” etc. and many of the participants tried to use the visual awareness for active communication. This worked ‘OK’ when the visualisation was full colour real time video, when gestures - e.g. waving, raising a glass, were adequately conveyed. However the medium was designed to be a calm, background visualisation transmitting subtle non verbal cues from non-explicit communication. For instance audio had been deliberately left out. As such it was not a good medium for explicit communication and required gross exaggerated gestures to communicate. In a second experimental condition, the visualisation presented a black and white frame difference so only motion was visible, here gestures were often missed breaking the normal action-response of explicit communication, potentially leading to frustration.

The insistence of users to try and misuse technology supports the notion that people have a strong desire to communicate that is not met adequately by domestic communications equipment. It also raises the question, how do we correctly suggest the affordances of a communications system? If a system has been designed to provide background awareness (requiring low bandwidth, power and cost technology) but is not capable of supporting low latency, high bandwidth explicit communication, how do we indicate this to guide users to choose more appropriate mechanisms?

The traditional static photo is clearly understood to have no power to support communication. Nor do we attempt to communicate with presenters on broadcast TV. Perhaps we quickly learn that characters on TV are completely unresponsive to our behaviour. The samples that make up a Video Snapshot are captured over a period of time and cannot be displayed until the next period starts. As such there is considerable delay between the capture and display of a snapshot. It is hoped that these elements of stillness and asynchrony in a Video Snapshot may begin to reflect some of the properties of a picture or TV.

6 Experiment

An experiment is planned to assess the affective benefits and costs [12] of Video Snapshots against photos and live video for providing visual awareness. Pairs of close friends will be placed in separate locations and engage in a shared, but passive activity such as watching a film, while provided with different forms of reciprocal visual awareness. On completion they will be asked to assess their experience using formal questionnaires.

7 Conclusions

We have presented Video Snapshots for use in the home to provide awareness and support connectedness to geographically separated others. We have contrasted their theoretical properties for emotional connection and privacy protection with both static photos and live video connections. The suggestion is that Video Snapshots will be more emotionally engaging than still photos, but less invasive and distracting than live video.

Acknowledgments.

This work is supported by BT, the EPSRC and UCL.

References

- [1] Social Trends, National Statistics, The Stationery Office, London. No. 34, 2004 edition
http://www.statistics.gov.uk/downloads/theme_social/Social_Trends34/Social_Trends34.pdf
- [2] Frohlich, D., Whittaker, S., Daly-Jones, O. (1994) Informal Workplace Communication: What Is It Like And How Might We Support It? CHI'94.
- [3] Nardi, B. A. (2005) Beyond Bandwidth: Dimensions of Connection in Interpersonal Communication, CSCW 2005.
- [4] Portrait (Harry Potter), From Wikipedia, the free encyclopedia, (as of 27 th June 2006)
http://en.wikipedia.org/wiki/Portrait_%28Harry_Potter%29
- [5] Thorne, J. (September 2005) Supporting Informal Communication and Closeness Through Video Snapshots, London Communications Symposium 2005.
- [6] Sierra, K.(April 13, 2006) Why face-to-face still matters! Creating Passionate Users,
http://headrush.typepad.com/creating_passionate_users/2006/04/why_facetoface_.html
- [7] Weiser, M. and Brown, J. S. (October 5, 1996) The Coming Age of Calm Technology, Xerox PARC
<http://www.ubiq.com/hypertext/weiser/acmfuture2endnote.htm>
- [8] Ehrlich, S.M., Schiano, D.J., and Sheridan, K. (2000) Communicating facial affect: It's not the realism, it's the motion. CHI 2000.
- [9] Neustaedter, C. and Greenberg, S. (2003) The Design of a Context-Aware Home Media Space for Balancing Privacy and Awareness. UbiComp 2003.
- [10] Coleman, S. (8.35am May 16, 2006) Transcription from radio interview The Today Programme, BBC Radio 4.
- [11] Markopoulos, P., IJsselsteijn, W.A., Huijnen, C., and de Ruyter, B. (2005). Sharing experiences through awareness systems in the home. *Interacting with Computers* 17, 506-521.
- [12] van Baren, J., IJsselsteijn, W.A., Markopoulos, P., Romero, N., de Ruyter, B. (2004). Measuring Affective Benefits and Costs of Awareness Systems Supporting Intimate Social Networks. In: Nijholt, A. & Nishida, T. (eds.), Proc. 3rd workshop on social intelligence design. CTIT Workshop Proceedings Series WP04-02, 13-19.