

ORIGINAL ARTICLE

Primary and secondary context in mobile video communication

Erika Reponen · Pertti Huuskonen · Kristijan Mihalic

Received: 23 December 2005 / Accepted: 31 July 2006 © Springer-Verlag London Limited 2007

Abstract The new video capabilities of mobile phones are starting to change the field of mobile communication. It is now dramatically easier to publish video in quasi-real time. We discuss how this change will affect the way people perceive video-recording, in terms of privacy, transparency, and the notion of context. We use a model of primary and secondary contexts to analyze usage situations, highlighting newly relevant research issues.

1 Introduction

Video recording in public spaces is becoming increasingly common with the ever-increasing popularity of videocams and camera-enabled mobile phones. While waiting for technical and business issues to clear, video conferencing is still a key driver for 3G mobile networks. There are new channels for publishing the video material, such as video blogs and video clip sites, such as Google video and YouTube. Because of the alwayswith characteristics of the camera phone, it is now fundamentally easier to capture and transmit video

content in everyday situations. As a consequence, there are new populations of people recording video in new kinds of situations as well as new populations of people captured to video. Easy publishing with video cameras makes it possible to distribute even live material with minor effort.

Since the whole mobile content chain, from capturing to consumption, can now be done using one device, we believe that the way we perceive video recording will fundamentally change. We discuss how the notion of context changes with the advent of ubiquitous video cameras.

Until now, little research has been done on video recordings using mobile phones. Research on mobile imaging has increased in the last few years, but since video is different from still photography, the results from still imaging cannot be directly applied to video recordings.

Our concern is with the impact of video-recordings using mobile phones as well as the importance of the social context during capturing, publishing and viewing the content. We discuss how bringing the camera to a situation changes its social context and thus, what kind of content is possibly captured and published.

E. Reponen (⋈) · P. Huuskonen Nokia Research Center, Visiokatu 1, 33720 Tampere, Finland e-mail: erika.reponen@nokia.com

P. Huuskonen e-mail: pertti.huuskonen@nokia.com

K. Mihalic ICT&S Center, University of Salzburg, Sigmund-Haffner-Gasse 18, 5020 Salzburg, Austria e-mail: kristijan.mihalic@sbg.ac.at

2 Related work

Camera phones are challenging the established models of communication. So far, communication through moving pictures has been largely the domain of broadcast media, film, and the relatively small audiences of home videos. Now, as mobile video recording is suddenly accessible everywhere, the traditional situation is changing.



However, the existing body of work in visual communications and media science still applies in this new situation. For instance, Huhtamo [6] notes that already in the nineteenth century, a camera used to have a de-humanizing effect on the person carrying it. The potential of publishing has always affected subjects.

Although there is as yet little research on videophones, the studies of imaging with mobile phones give hints to the expected developments. For instance, with mobile imaging more material is shared, and the sharing takes place relatively soon [9]. Kindberg et al. [7] classify images by their social and individual uses, helping with affectional and functional tasks. Koskinen [8] observed that people familiar to each other rely on mutual trust for controlling sharing of sensitive material. This works provides the basis for enhancing our context models (discussed later).

Baudrillard [2] discusses human notion of time. He observes two indications of the human willingness to get rid of the limitations of linear time: first, we want to capture everything and share it almost before it happened to have a feeling of living the situation; second we want to record everything to be archived as a record of our culture.

Baudrillard [1] also claims that people are so used to living in the TV reality that they always act as if they are being photographed. Ylä-Kotola [11] further states that humans are more aware of themselves as objects of recording rather than as conscious beings.

Context plays a major role in social interaction. It builds a frame for human actions in general, and is a constitutional part of every human communication process as well as human computer interaction. The importance of context has been widely recognized in communication fields for a long time [5], and has in the past years been recognized as a key issue to tackle on the road towards more effective human–computer interaction.

Dey and Abowd [3] divide contexts into primary and secondary. They identify four primary categories of context: location, identity, activity, and time. In their view, these are basic context parameters required to characterize the situation of a particular entity. On the next level, secondary context parameters can be identified, which share the common characteristic that can be indexed by primary context. For example, a phone number can be obtained by using the user's identity as an index. As most context research, their model focuses on the properties of environmental information, rather than on the user. In our present work, we emphasise the effect that context has on the users, instead of analysing the properties of context.

3 Context as a container

We model the situations where mobile video is recorded and consumed as different contexts. Here, we discuss the various factors that make up these contexts. In our model, the act of publishing (and thus the camera) defines the limit between primary and secondary contexts (Fig. 1).

Here, the primary context means the immediate surroundings; a situation where people can communicate in shared time and space without help of technical devices. The physical place and common time are identifiers of the primary context. As we add cameras into the situation, it becomes possible to publish the recorded material, and so move beyond the primary context. The situation turns to be more risky since there are more privacy concerns arising, when the cameras are present.

The secondary context is any (remote) situation where the recorded video is used. This situation may take many forms, with variations in the dimensions of time, space and subject awareness. For examples, the material may be live or time-delayed, may be saved or not, may be shared with a small or an arbitrarily large audience, or may be only retained for creators' personal uses. The people within the primary context may be unaware of the existence of the secondary context. The people in both contexts may be the same or different ones. There are no exact limits for the secondary context, it may happen in any time and any place where it is possible to consume captured material.

Publishing implies that the material is transferred in space. Telepresence seeks to break down this physical barrier of places, i.e., merge the contexts. The "hole in

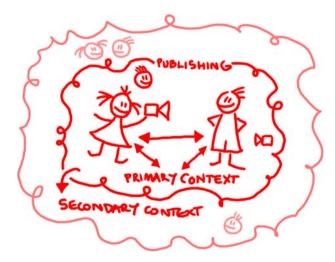


Fig. 1 Publishing defines the boundary between primary and secondary contexts of video use



space" experiment [4] is illustrative of this approach. Similarly, recording implies the aim of transferring the material in time. Archived material may only surface years (or millennia) later. Many people record autobiographical video, while some use it for reminders [7]. In general, the aim is to recover some of the primary context into the secondary context.

4 Primary context variations

Primary (immediate) context has many possible variations in terms of capturing possibilities and privacy issues. In the following, we characterize some situations for primary context. We divided the primary context into three situations based on the number of cameras: 0, 1 or N.

4.1 0: Situation without cameras

Interaction between people in an environment defines a social context with a rich set of rules and norms. One typical social phenomenon, gossip, is quite relevant to our model, as it moves on the grey areas of privacy. There are few ways of proving a piece of gossip without any recorded material from the situation. This makes a situation without a camera relatively secure for the persons present. For people absent in the primary context and relying on stories by others, the experience is not as strong as it would have been, had they experienced it with their own eyes and ears. Without a camera, the secondary context has little effect on the primary context.

4.2 1: Situation with one camera

4.2.1 Hidden camera

If a camera is hidden in the primary context, people behave rather naturally like they would in a situation with no camera present. Hidden camera may be a candid cam, a surveillance cam or a spy cam. All of them may have different effects on the people being captured, but this does not affect their behaviour since they are not aware of a camera present in the context. It is also possible that a camera is not meant to be hidden, but people just do not notice its presence; thus the behaviour of the people is the same like in the situation with a hidden camera. In these situations, people are not concerned with their privacy since they are not aware of a potential threat. The privacy concern may arise later, if they notice the camera or see the captured material elsewhere. For example, a

discussion between two friends may be unintentionally captured into the video by a stranger, who is not thought to be "dangerous". Later, the video is published on Internet page and seen by someone who recognizes the friends.

4.2.2 Visible camera

If a camera is visibly present in the primary context, people know that the situation is or may be recorded and they behave based on that knowledge. Feeling of privacy depends on the situation and the emotional closeness between the present people. The situation is different if the video is taken by one of the group members or by someone outside the group; it is different if it is taken by a trusted friend or an unfamiliar person. Example of this type of primary context is a family party captured by the father.

4.3 N: Situation with a number of cameras

In the situation with a number of visible cameras. people know that the situation is or may be recorded by many cameras, which means that the recordings are very believable. We think that believability is high because it is more difficult to deny what happened in a situation with many capturing devices. It is nowadays easy to fake videos and photographs coming from a single source, but it would be more difficult to create similarly manipulated material from many different recording sources. The question of trust for a person taking the video is similar to the aforementioned situation with one camera. Another characteristic of the situation with a number of cameras is that the material captured with many cameras is more likely to spread to the secondary context. There is less control over it compared to the material captured with only one camera.

The situation with a number of hidden cameras resembles the situation without cameras. There is a possibility for secondary context, but people in the primary context are unaware of it. That increases the privacy concerns for the people captured.

5 Secondary context variations created by publishing

Publishing is the border between primary and secondary contexts. Publishing may be done by digital or analogue means, for oneself or for the others. Other people may have been a part of primary context. They may be persons who only have access to the content in the secondary context. Publishing is a restrictive



security and privacy line between two contexts but also an enabler between those. Kindberg et al. [7] presented a division of images into affective versus functional and social versus individual categories. While they are not using the terms primary and secondary context, they talk about the same issue: about recording situations, what is considered publishing, and to whom. Their results suggest that people are taking mostly images that are affective. Those images are used for both social and individual needs. Since Kindberg et al. [7], present that most of the functional images are representing some "specific things" which are not sentientbeings, our research is concerned mostly with their affective category. However, there is some use for functional videos from sentient beings as well, like documentaries, crime proves or paparazzi videos.

We divide the publishing possibilities for secondary (remote) context based on the time and extent of publishing. Material can be shared live or later, with many people or only between two people. In the following, we characterize the differences between these situations.

5.1 The event is shared live, ad hoc, immediately

If sharing is done with numerous people, the spreading happens immediately among a large audience. An example of this kind of publishing to secondary context is the telepresence experiment "hole in space" by Kit and Sherrie [4]. This experiment was not mobile but happened in public environment. Basic mobile example of this is a video conference call made via mobile devices. An interesting example of live sharing of recorded material is Steve Mann (http://www.genesis.eecg.toronto.edu/) who wears an EyeTap web camera to share audiovisual environment he senses live and also to create his personal records for later use. Illusion of real time plays an important role in our research of mobile video communication, which makes Mann's experiment very interesting from our point of view. Mann seems to be fine with the fact that friends and unfamiliar people see the world through his eyes and experience the same things as he does all the time. However, people are not yet accustomed to be observed all the time and it may take them some time to get used to it.

Live sharing between two persons basically makes spreading to a large audience possible, but is more difficult and slower, since each person needs to spread the content to their individual recipients recursively. The speed of spreading depends on how interesting the material is for the persons who receive it. Also the relationships between the people affect the spreading

speed. Example of this kind of live sharing with one person is a video call between two people. The video may be saved to memory and forwarded to others as well.

5.2 The event is shared later

A study showed that 89% of digitally taken photos are shared at least once, and that most of the sharing is done within 3 days of taking [9]. This means that the timeframe in which digital images or videos are shared is relatively small when compared to the paper photographs. Paper photographs are processed in larger batches, which delays sharing by weeks or months after the original event. With photograph printouts live sharing was not very common: people shared them when they eventually met in their spare time.

We divide the material that is not shared live into three categories: (1) preserving for own use (2) sharing with people who were present in the primary context, and (3) sharing with people who were not present in the primary context.

- (1) The material that is preserved for personal use creates a secondary context even if viewed only by the recorder self. Example of this kind of secondary context is reviewing a self-recorded video, or browsing a photo collection. The same situation can be experienced again and again by reviewing the material.
- The material is often shared with people who were present in the same primary context. Secondary context like this is rather safe, because each person has already experienced the event as part of the primary context. It may however, create some privacy concerns, since the captured material is stronger evidence than material that is stored only mentally. An example of this type of secondary context is sharing a holiday trip video with the person who took part in the trip. Some discussions may arise on what has been done or said between the parties, which the recording would undeniably prove. There may have been persons present in primary context without knowing that they were being captured. For those people, this situation of secondary context raises big privacy concerns because they may have a feeling of being captured covertly, even if it was unintentional by the recorder.
- (3) If the material is shared with people who were not part of the primary context, the spreading has quite high security risks concerning peoples' privacy. Spreading is wide and may reach persons



who were not meant to receive the material. This is because not all people are loyal to the same extent. With the increasing number of recipients, the sharing becomes less controllable. An example of this secondary context is sharing of (stolen) private home videos on the Internet. If the material is shared with one person at first, the spreading is slower and more uncertain than when shared directly to many people. Spreading depends on many factors like how interesting the material is and how close is the relationship that exists between the sender and the recipient of the material. An example of this is sharing with someone a video of a boyfriend. If the material is shared with a good friend and the author does not wish further sharing, the spreading probably stops. However, if the person with whom the material is shared is not reliable, the loyalty is low and material may further be shared, even though this is not the wish of the author.

We mirror the secondary context to already existing phenomena like gossip and paparazzi. It has already been possible for long time to use a technology for capturing audio or video material and present it later. For example, paper photographs have been showed to friends very commonly when telling about past events. Secret situations between celebrities have been photographed by paparazzi and published in magazines. Famous people are already used to cameras shooting everything they do. For ordinary people not interesting to the media, this kind of public sharing of captured material has not been happening a lot. The spreading of non-digital or non-mobile material is slower and requires more effort than spreading of the digital mobile content. The content captured with a mobile video cam is entirely digital and even optimized for small mobile device. This makes the non-digital material "safer" and digital mobile content potentially more "risky" when thinking of privacy and security.

6 Overlapping contexts

The division into primary and secondary context may not always be clear, although the distinction itself is easier to make by the use of technology. There are situations wherein a person may be a part of both the primary as well as the secondary context as the same time. Every time the person who records the material and watches it simultaneously via the viewfinder of the camera is physically in the primary context: feeling and hearing the situation live in real physical world. At the

same time, he/she sees the situation in the secondary context, by means of the viewfinder or display of the camera. A person other than the recorder may also be in the primary and secondary contexts simultaneously. While being physically in the place and time of the primary context, a person may be receiving a live video call from a friend who is recording the material in the same primary context.

6.1 Mobile shared shooting and mobile observing

Mobile video recording can enable new genres of visual media. In the era of reality television, one possible genre is a mobile first-person reality television show. Reality programs are created by the private persons with their mobile video-recording devices. It is also consumed in mobile context by the people with their mobile device able to receive video material or mobile television. The content of this kind is recorded as firstperson view, since the recording person is carrying the device. This enables the audience to see the program "with the eyes of the recorder", to see the world as the recorder sees it. If there are many recorders in the same primary context, every one of those is both a subject and an object of recording. Also the audience can affect the program by interacting with the recorders, for example by showing up in the camera, talking to the tape, or more quietly manipulating the recorder. People can also interact via devices thus affecting the show (Fig. 2).

The technical enablers for this kind of mobile reality program are rapidly being established. You-Tube (http://www.youtube.com), a web site that collects mobile video-recordings contributed by private persons, has been one of the most rapidly growing sites of Internet in the past years. Google has also added user-contributed video clips to their sites. Besides the abundant clips from commercial material, the user-created videos range from non-edited real life pieces to almost professional quality material. The motifs are everything that can be imagined from reality to fiction.

In general, video has become a practical data type in the Internet. At the same time, the various audio distribution methods such as net radio and podcasting have become a popular form of communication. We assume them to be an indication for the forthcoming success of user-contributed video broadcasting. Video is a powerful way to convey experiences and feelings. By receiving audio and video material from some situation via some device, the audience in secondary context is getting closer to the primary context experience than it would be otherwise.



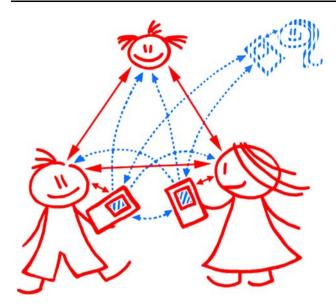


Fig. 2 Mobile shared shooting and mobile observing

7 Subject awareness as function of context

Camera phones make publishing easier for many everyday situations in which cameras were previously not present. Easy, even live-publishing enables contexts to spread from primary to secondary. It therefore means that secondary context can now have a greater effect on many new situations. This changes the way people behave in such situations.

Different situations can present different risk levels for a subject of a recording. This depends on the person's awareness of the recording, on the possibilities of publishing, and on how natural the situation may be.

Huhtamo [6] discusses personal photography in his article on mobile media archaeology. He points out that privacy concerning personal photography is not a completely new issue; already in the nineteenth century there were frequent complaints about amateur photographers transgressing existing social rules related to privacy and decency. The threat of anyone being a target of a snapshot was noticed already then and a camera used to have a de-humanizing effect on the person carrying it. There were candid camera snapshots taken, too. Huhtamo compares the situation to the discussions now taking place with the mobile phones and more specifically to the emerging habit of using camera phones.

Figure 3 shows a chart comparing some well-known situations where people behave naturally versus artificially, related to their awareness of cameras. The more aware the subject is of cameras and the more artificially the person is behaving, the safer the situation is for her.

Similarly, the more aware the subject is of the secondary context, the safer the situation.

For instance, a staged movie is safe, since it is assumed that the material will be carefully edited and published in a very controlled manner. On the contrary, a candid camera can be risky, since the subject is acting naturally without the knowledge of a recording. Most practical situations fall in between these extremes. Reality TV is stretching the perceptions of naturalness and awareness; for instance, in the TV show "Big Brother", the subjects are quite aware of the cameras and the artificial situation, but their behaviour still tends to become more natural over time.

The most risky area for recording is the situation in which the subject is completely unaware of the camera and is behaving naturally. In the safest recording situation the subject is completely aware of the presence of the camera and is behaving artificially. Example of safe situation is an organized movie-capturing event with scripted (acted) behaviour. A situation where person is completely aware of the camera but is acting naturally is semi-safe as well as the situation where a person is completely unaware of the camera but is behaving artificially (acting).

Awareness of the camera and artificiality of the situation in primary context (*x* and *y*-axis of the box) affect people's willingness to pose, record and publish video. This directly affects the amount and nature of material available in the secondary context (inside the

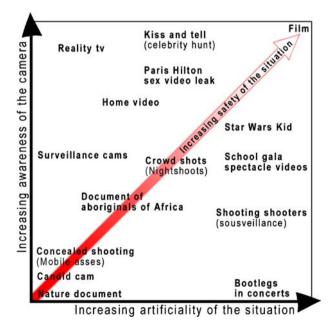


Fig. 3 The "safety" of a recording situation varies as a function of the subject's awareness of the camera and the artificiality of the situation



box) if the recorded material is published. Since mobile video recording and publishing is still a fairly new phenomenon, we cannot yet say what kind of content and format has the potential to become most popular among the audience, but we think that it is inevitable that novel program formats will surface.

"Scare tactics" is a television show by SciFi.com television channel. The show presents films that are partly hidden camera and partly written, the "victim" being a part of the show without knowing the situation is staged or other people acting. Those victims are being scared as heavily as possible and the end result, the mini movie, looks more or less like a professionally handwritten, acted, and shot movie. Similar format where one person is victim and the others are acting, but in more normal (not movie-like) situations is presented in "Punk'd" television show running on MTV. These two programs are variations of generally known hidden camera programs, which usually present short occasions, like jokes, where people make fun with the "victims" of the shows. These examples are mixtures of safe and unsafe shooting situations, because some of the people in primary context are acting and thus know very clearly that cameras are present. However, one person is completely unaware of cameras and of others acting. For the audience, this format gives both the joy of well-written drama and the joy of watching a nature document kind of capture of the fellow creature, peeping the human behaviour in various situations.

8 Mobile content

8.1 Characteristics of mobile content

In our view, there are some inherent characteristics of mobile content that make it worthwhile to differentiate it from non-mobile use. These characteristics enable an easy mobile content chain and thus an easy leap from primary to secondary context.

The content that is created with a mobile phone video camera is in the appropriate format because it is entirely digital throughout the production–consumption chain. It is also in the right size for the consumption with mobile devices, both in terms of resolution as well as file size. This is due to the fact that it is captured with a mobile device, which has equivalent capabilities as the consuming device. Mobile content is accessible at the right time because it is sharable live or at least within relatively short time frames.

Capturing and consuming in the mobile context enables new program formats or genres, like mobile

reality television. It can be done by the users themselves because cameras and receivers are always with, being a part of a mobile phone. The content is often self-created, but can also include professionally manufactured material. It is immediately editable with mobile video editors and quickly publishable via various channels, possibly to large audiences, thus being particularly interesting to advertisers and sponsors. Mobile content is usually recorded from a first-person camera perspective, which gives both a telepresence-effect and a homemade effect by enabling the consumer to see with the eyes of the producer. In addition, the ability to integrate environmental audio and narratives further helps capturing the atmosphere of the situation.

8.2 Mobile content chain

In the mobile domain it is already possible to cover the entire end-to-end chain from creating the content to consuming it. There is equipment available for every needed part of the chain; all required applications are available on a single device such as a mobile phone.

Parts of the mobile content chain are video capturing, editing, sharing and consuming. In addition, a feedback channel is available as well.

- 1. Video capturing with mobile camcorder: Portable video recorders (e.g., VHS, DV) have been available to the public for a long time, although solid-state storage was introduced only recently in such devices as Panasonic's SV-AV100. High-end mobile phones, such as Nokia N93, record near-DVD quality video on memory cards.
- 2. Editing with mobile video editor: There are already some video editors available on mobile phones. For example, new Nokia models have following editing possibilities: "MMS-cutting", a process in which a video is cut into MMS-sized chunks and the video size is counted automatically; making regular video cutting; adding and cutting audio; adding transitions and fades; adding effects like slow motion, black/white and sepia colouring; adding images and titles in between the video frames; transcoding to various sizes and formats, such as MPEG4 CIF or H.263 QCIF; Sending videos via Bluetooth, MMS, or email; Recording audio comments.
- 3. *Sharing*: Technologies and services for mobile content publishing include MMS, P2P (Infrared, Bluetooth, WLAN), mobile blogs, video blogs (lifeblog), email, mobile Internet, podcasting (for audio content), broadcasting, etc.



- 4. Consuming on a mobile phone: Technologies are similar in sharing and include MMS, P2P (Infrared, Bluetooth, WLAN), email, mobile Internet, mobile television (e.g., DVB-H), RSS, etc.
- It is also possible to have a reply channel for feedback and interaction offered by the mobile consumption services.

9 Conclusion and outlook

In our view, the rise of mobile video will change the way people behave in many everyday situations. In other words, the existence of potential secondary context affects people in the primary context. For instance, the Star Wars kid [10] might never have recorded his exercises, had he known that the material would spread widely on the net.

Easier recording and easier publishing implies a greater perceived level of risk when it comes to security and privacy. How significant will this be for our everyday lives? While research does not yet give answers, we see telltale signs related to this issue. Some nightclubs do not allow use of camera phones. In Japan, it is mandatory to have shutter sounds in digital cameras. On the other hand, the slightly doubtful website http://www.mobileasses.com only became popular with camera phones.

In general, digitally recorded material can be considered more risky, since it can be more easily published to larger audiences. With mobile communications, digital content can also be published almost instantaneously even on the move.

One factor affecting the perceived risk level is the familiarity of the subject with the content author. We trust people we know well, and assume that they will better control the distribution of the recorded material. Strangers recording strangers have less moral obligations to restrict publishing [8]. With increases in mobile imaging, this issue is becoming more relevant. It is possible that people will give up control of their privacy, given sufficiently wide spread but socially accepted use of recording. This is evident with surveillance cameras—what are then the mobile video applications that will win the public acceptance?

There are privacy issues connected to publishing audio and video material. Is there a need for intended restrictions in order to make sharing more difficult? Restrictions may be connected to privacy issues such as encrypting or record preventions. System design has to take a broad view on how technologies are used in our daily environments. This is especially important with video capturing and sharing, which is a highly social activity involving numerous parties.

We presented a theory of primary and secondary context, publishing being an enabler between these. The division is made based on the technology used for sharing but is more social in its form. We are now analysing various use situations in the space of awareness versus artificiality of the primary and secondary context model. We believe this analysis will give fresh theoretical tools to discuss mobile video communication and its impact.

Acknowledgments Parts of this work have been carried out under the EU project Mimosa (contract IST-2002-507045).

References

- Baudrillard J (1981) Simulacres et simulation. Galilée, Paris, pp 48–56
- Baudrillard J (1995) Lopun illuusio. Galilée, Gaudeamus Kirja Helsinki, p 25
- Dey AK, Abowd GD (1999) Towards a better understanding of context and context-awareness. Proceedings of HUC 99, Springer, pp 304–307
- 4. Galloway K, Rabinowitz S (1980). Hole in space. Available at http://www.ecafe.com/getty/HIS/index.html
- 5. Hall ET (1976). Beyond culture. Doubleday, New York
- Huhtamo E (2004) Pockets of plenty: an archaeology of mobile media. Proceedings of ISEA 2004. http://www. isea2004.net/
- Kindberg T, Spasojevic M, Fleck R, Sellen A (2005) I saw this and thought of you: some social uses of camera phones. Proceedings of CHI 2005, ACM, New York, pp 1545–1548
- 8. Koskinen I, Kurvinen E, Lehtonen T-K (2001) Mobiili kuva. Edita/IT, Helsinki, pp 80–81
- Sarvas R, Oulasvirta A, Jacucci G (2005) Building social discourse around mobile photos—a system perspective. Proceedings of mobile HCI 05, ACM, New York, pp 31–38
- 10. Star Wars kid (2002) Available at http://www.jedimaster.net
- 11. Ylä-Kotola, M (1997) Evoluutio vai nollapiste? Mediatieteen bioparadigma ja Hermann Friedmann. In: Hintikka KA, Kuivakari S (eds) Mediaevoluutioita, pp 93–94

