Mobile Broadcasting – The Whats and Hows of Live Video as a Social Medium

Oskar Juhlin[1], Arvid Engström[1], and Erika Reponen[2]

[1]Mobile Life@Interactive Institute, Sweden
P.O. Box 1197, SE-164 26 Kista, Sweden
+46 (0)8 751 72 30
{oskarj, arvid.engstrom}@tii.se

[2]Nokia Research Tampere, Finland
Visiokatu 1, 33720 Tampere, Finkland
+358 7180 08000
erika.reponen@nokia.com

ABSTRACT

A new type of social medium, which allows users to broadcast live video from mobile devices to websites on the internet, is becoming increasingly popular. We provide a qualitative content analysis of a sample from four such services. The analysis specifically focuses on the topics presented, camerawork, and coordination, in order to investigate the possibilities and barriers to wider adoption of this new social medium. Although the services are growing in numbers of users, the study reveals an immature application area. People struggle to find interesting topics to broadcast and to manage the camera in a way that presents them in an appealing form. But there are also examples of topics such as artistic performances and tours, as well as ways to conduct live transitions and coordination, that point to a more medium-specific way of using these services. The results indicate that providing the opportunity to broadcast live video is not enough, and that there is now a need to design for amateurs' appropriation of camera handling techniques.

Categories and Subject Descriptors

H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

General Terms

Human Factors

Keywords

Content analysis, video, live broadcast, webcast, mobile, social media

INTRODUCTION

In this paper we investigate the use of a new type of applications, which we refer to as mobile broadcasting services, which make it possible to capture live video on a mobile phone and broadcast it in real time to a web page. Mobile broadcasting applications differ from earlier webcam technologies in that the cameras are wireless, which enables *capturing content from anywhere within the reach of a mobile network, and extended physical control of the device for doing camerawork*. The web application typically lets people browse through live broadcasts, access archived clips, and follow and interact with individual users

– a model familiar from other forms of social media on the internet. We have studied the topics of these videos, as well as how they are represented visually in this media format. Since the launch of the first application in 2005, these services have grown in popularity and number. There are nine services in the area to date, among which qik.com and bambuser.com are two of the most widely used. The applications emerge in the intersection of desktop video streaming, websites, and mobile video conferencing systems.

The concept has attracted some attention in research. Reponen [11] conducted a field experiment to support group interaction. Bergstrand and Landgren [1] have conducted a design investigation to explore how live video could be used in rescue operations. Extending the concept to include multiple cameras has been suggested as a way to support citizen journalism [13] as well as other use contexts [5]. In all, this research regards the concept of mobile live video broadcasts as interesting. Although such applications have been available for a couple of years, not much analysis using realistic data has yet been done of how the services are appropriated by users. The purpose of this study is therefore to investigate the pros and cons of this new medium in order to inform the design of a next generation of services.

Our study is of interest for research on user-generated content, as well as research on social media [9]. Users have found social applications for such diverse media types as text (in microblogs like twitter, chats etc); photography (e.g. in flickr); audio (e.g. myspace); and video files (e.g. youtube). It is still an open question whether mobile live video will become as successful a form of social media as these.

We examine available postings on four popular sites, bambuser.com, qik.com, flixwagon.com, and kyte.com, to investigate the contents of the videos and how they differ from webcam broadcasts. We have collected a sample corpus of 178 video clips, which have been viewed to identify aggregated content themes.

Our analysis reveals that there are many broadcasters of mobile live video who utilize the affordances of this new medium. They broadcast video from public areas in city centers, or they display outdoor activities on the ocean or in

the mountains. The mobile cameras also allow them to experiment with camera angles. At the same time, it is evident that many users struggle with both the technology and the concept. There are more people on these sites who are just testing the technology than are actually broadcasting content, and much of the latter's productions have very low production value. Although most people are already accustomed to professional live broadcasts, and in that sense have an idea of what this medium could be, it seems that taking the step to actually providing such broadcasts on one's own is very difficult.

The article is structured as follows. First we provide some background and describe previous research on commercially available live video. Next we present our methodological approach. The analysis is presented in two different sections. The first section on "Broadcast Topics" focuses on what people are recording. The second, on "Production Strategies," focuses on how this new medium is handled in terms of camera work and the giving of directions when on-air. We then discuss the findings and provide a section outlining implications for design.

BACKGROUND AND RELATED WORK

There is a growing interest in live video in HCI research. Shamma et al. [12] have studied desktop webcasting, and specifically the ways in which DJs use it to maintain close contact with peers, provide content related to their performances, and connect to fans through real-time streaming video. Live interaction through standard interaction techniques makes it possible for the DJ to adapt to the audience. The video communication technology in this research is standard webcam solutions that have been used for communication between people in various ways. These do, however, in effect restrict users to capturing video from fixed locations such as offices or homes. Previously, the alternative for bloggers and others who wanted to capture video from non-desktop locations to send to websites was to use offline solutions such as DV cameras or mobile phones with video cameras [6]. Mobile broadcasting also draws upon a long tradition in the telecom area of mobile video conferencing systems intended for face-to-face interaction [10]. However mobile broadcasting allows for an unrestricted audience, which might potentially be larger. Reponen's [11] field experiment with a mobile webcasting application called ComVu Pocket Caster is of special interest here. She argued that live broadcasts were beneficial for sharing contextual cues in this group of users. Such cues included information about locations, group compositions, and navigation directions.

Mobile Live Video Broadcasting Services

Mobile broadcasting services have thousands of users posting large numbers of videos. As an example, more than 4000 videos a day are shared on qik.com [14]. The first service enabling such real time video sharing from a mobile

phone to a public website, called ComVu pocket caster, was launched in 2005 [11] and later renamed Livecast. Similar services emerged in the years to come, counting another eight to date: Qik, Kyte, Bambuser, Flixwagon, Floobs, Next2Friends, Stickam, and Ustream.

Here, we present the common service features of all nine services before providing a qualitative content analysis of Qik, Kyte, Bambuser, and Flixwagon. Many of the features available in the ComVu Pocket Caster have remained in use and spread to other mobile webcasting applications. The most recognizable evolvement has been the adding of support for web 2.0 types of social interaction. All nine mobile broadcasting services share similar features and user interface formats such as:

- Sharing of live video from mobile phone to a web page

- Storage of videos to be viewed later on

- Sharing videos to other web pages, via social media services, email or embedding,

- Title descriptions

- Commenting and/or live chatting

The mobile client application is typically downloaded for free and can be installed on a range of camera phones or, more rarely, pre-installed; e.g., Qik comes bundled with high-end Nokia phones.
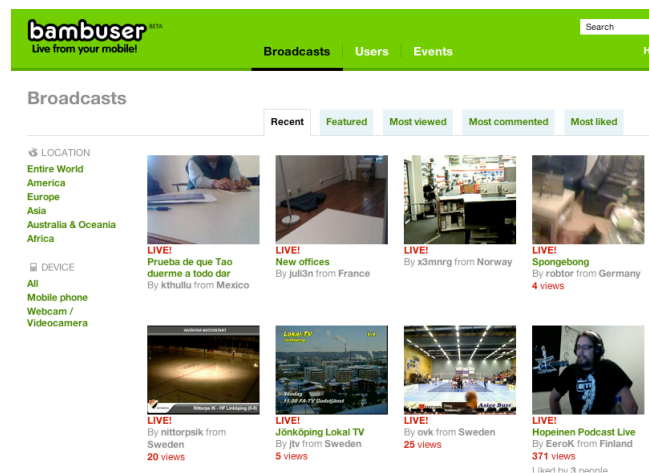

Figure 1: Interface at the Bambuser website

The web application is typically centered around a front page displaying current live feeds, much like a web-TV page displays channels. (see Figure 1). Selecting a live feed takes you to an enlarged version of that video, along with juxtaposed information such as the broadcaster's name, number of views, and recording time. Most of the services show the most recently published video or the most viewed video.

METHOD AND DATA

Qualitative content analysis is a research technique for making replicable and valid inferences from texts (or other meaningful matter) to the contexts of their use [2, 7, 15]. The term "text" is used in its wider sense to mean any media, from newspaper articles and advertisements to film and video. A key research tool in studies of mass media, content analysis involves specialized procedures for counting and coding the examined content to reveal patterns and trends in data sets [2].

In this study, the first step was to build a sample corpus of videos from these broadcasting services. A systematic sample was assembled to give us a general and initial understanding of the use of these services. For practical reasons, we could not sample videos from all services, but selected four of them, ranging from large (Qik) to small (Kyte) in terms of number of mobile users.

Since the broadcasts are live, typically appear without previous notice, and sometimes occur simultaneously, we carried out a systematic sampling based on time slots rather than the order of broadcasts, which would have been the standard procedure with prerecorded material. For each web service, video streams were manually sampled for ten minutes on eight predetermined occasions over a 24-hour period. In order to practically manage the sampling, we could only record two services per day, extending the sampling of a 24-hour period to two days. With this technique we intend to cancel out broadcast variations due to global time differences. Of course, such a sample will not capture long-term variations over days or weeks. We have secured permission to use material from the websites.

In all, we recorded 254 video streams from the four web services, of which 178 were broadcast from mobile devices, 133 on Qik, 7 on Kyte, 16 on Bambuser, and 22 on Flixwagon. The varying numbers reflect the relative popularity of the services. However, this is less of a problem, since the services are conceptually very similar and we have not intended to make comparisons between them. The recordings were made with video and audio capture software. The video capture program Jing has a maximum limit of five minutes per capture, which limited our samples of broadcasts to this length. There were as many as 55 samples that exceeded this limit, but only eight of these were recorded on mobile devices. Thus, this limitation only causes a minimal restriction of what we can see of the broadcasts from our selected time slot. We then examined each video file and described its content, drawing on what is empirically available in the material: the video, audio, and title. A set of aggregated content categories that bring out salient characteristics in the videos was developed. We inferred some patterns in types of usage of live video as a medium in these services. The aggregation was conducted following an inductive approach [15], whereby a classification scheme was developed by studying the individual clips one by one, and several times. This resulted in a category scheme that covers the entire corpus. In the following, we present the concepts we suggest are most relevant to understanding this emerging medium. Importantly, these concepts are not mutually exclusive, and they might not even be the quantitatively most salient aspects.

CONTENT ANALYSIS OF BROADCAST TOPICS

In the following, we present our category scheme, which describes what people have filmed. We illustrate each category with examples from the recorded video streams.

Tests and Demonstrations

Technology tests are by far the most common theme in the recorded broadcasts (71 out of 178). Tests are characterized by arbitrary and unsteady framing, indicating that attention is being placed on the handset interface rather than the broadcast content, and by irregular camera movement, particularly at the start and end points. They are typically short in length and lack commentary and other people than the user. Most test broadcasts are surprisingly similar in visual appearance, commonly displaying interiors, desks, users' feet and hands, and computer screens. A typical test is shown in Figure 2, a seven-second sequence first framing an apartment interior, followed by a quick pan stopping for a moment at a TV screen, then sweeping back to frame a tilted shot of a computer in the user's lap. The frequency of this category could be explained by the fact that the technology is fairly new to and unexplored by the general public, and the number of users is constantly rising, which means the share of users trying it out for the first time or familiarizing themselves with the service is quite high at any given time.
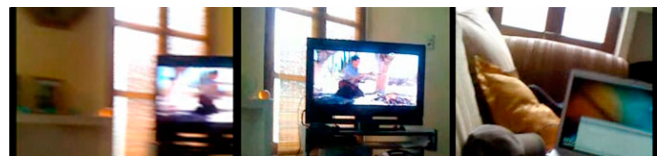

Figure 2. Test broadcast sequence

A related but less common theme consists of videos where users demonstrate the technology to family members, friends, or colleagues. All eight *demonstrations* in our analysis are done for people within the same physical space. Here users have been trying out the technology and grown accustomed to it, and are now taking the role of an authority explaining how the technology works by commenting on its handling while on air.

Video Logs

There are a number of video logs, or Vlogs, in the data (6 out of 178), characterized by aesthetics familiar from webcam video chats. In its most basic form, the camera is fixed, typically framing the broadcaster in *third-person view* sitting in front of a computer. This view sets Vlogs apart from other categories that are almost exclusively in
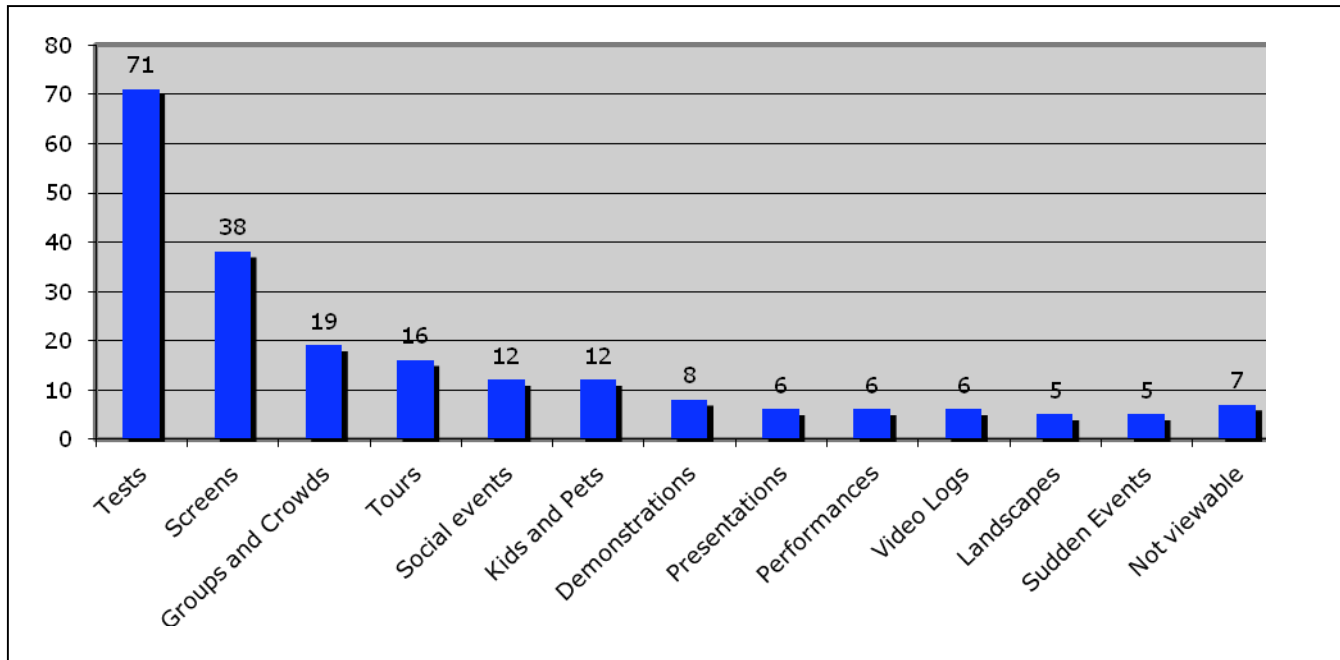
Figure 3. Topic categories and number of occurrences

first-person view; i.e., the viewer sees what the broadcaster sees and frames.

While most of the chats (four out of six) are conducted from a desktop environment, essentially replicating the webcam format, two of them take place in more mobile contexts. These broadcasts then border on the next category (tours), but maintain the characteristic third-person view of the broadcaster. For example "djdlo" reports to his Vlog while walking the streets of Toronto (see Figure 4). The broadcaster is talking into the camera phone in his hand, framing himself from below. He frames himself in this

Figure 4. Mobile Vlog broadcast sequence

handheld third-person view throughout the 1:15 minute sequence, except for two occasions when he pans in a controlled fashion to display images of what he is reporting about: first the square in central Toronto which is the scene of the broadcast; and then his friend, who is presented in the middle of the sequence. The broadcaster refers to both the location and to his friend in turn as they come into the frame. The setting also remains visible in the background of the broadcast as the two people move through the square, over a crossing visible by the traffic lights ahead, and onto a sidewalk, where we see the wall of a building to one side.

Tours

Tours and sightseeing are a major category of usage outside of the desktop setting (16 out of 178). Sixteen sequences adopt a first-person view and focus on the local environment. They are typically physically mobile, descriptive in framing, and commented live by the broadcaster; e.g., the artist Kerli takes the audience on a tour of an old Estonian town on Kyte. He is moving the camera and alternating between framing things that could be of interest to the audience and close-ups of himself and a friend.

Performances and Presentations

There are six artistic performances in the material. The examples include kids singing or playing, someone playing the piano, and live street art. In these videos, the broadcaster generally takes an active role in directing people in front of the camera, and several videos begin with the broadcaster making preparations by adjusting the camera and arranging the framing, probably with an aim to create a more compact framing of the performance. There are also an equal number of *presentations* in our corpus, including lectures and seminars in settings ranging from workplaces to lecture halls. These are typically shot using a fixed camera and wide framing.

Social Events, Groups, and Family

This category includes broadcasts where people seem to be the main topic. There are twelve broadcasts that feature social events of various kinds. This is, for example, obvious in a broadcast from Uppsala University, with a title stating that they are having an "open house." The actual broadcast shows an empty corridor at an office. Nineteen broadcasts

display groups and crowds. Many of these groups are taking part in events such as dinner parties, or visiting nightclubs. There are twelve broadcasts displaying children or pets. These broadcasts often stay on a single subject, such as a newborn baby or a dog running around, and are more closely framed than the average shots of people, who are typically framed in full figure or from a distance.

Landscapes

The five broadcasts in this category include sequences presenting landscapes and nature motifs, such as distant mountains or views from hotel windows, where people and social events are not the primary focus. They are widely framed, often shot from a distance, and share other common characteristics such as a lack of variation in framing and an extended duration. They are typically devoid of action and slow in tempo.

TV and Computer Screens

Computer and TV screens frequently occur in the image content of the broadcasts; just over 20 percent, i.e., 38 out of 178 videos, contain one or more screens, making them more commonplace than any other subject, including people. This might seem a little odd given the pledge of mobile broadcasting services to take live video out into truly mobile and physical settings. The explanation may be that they occur as part of tests and demonstrations, where the screen activities give direct feedback on broadcast quality and delay. Here, the fact that the technology is divided into a mobile and a web interface that both need to be monitored during a test, seems to restrict that mobility, at least initially. Another explanation could be that screens are "where things happen" in many homes and workplaces, and therefore draw a lot of attention in the videos. In comparison to armchairs and tables, screens may be seen as more visually interesting, and sitting in front of the computer is a common activity.

Sudden Situations

There were five videos in the sample where the broadcast was clearly a response to events suddenly unfolding around the broadcaster. These broadcasts typically do not have a marked start, but rather take us straight into the event. The situations framed are out of the ordinary and clearly distinguishable from, e.g., the more mundane videos of the social events category.

All of these show what is either a fire drill or an emergency, with people running and walking out of a house and gathering outside, to the sound of an alarm. This category is potentially larger than the small amount of videos attributed to it, since a number of other videos may have been prompted by sudden events before the broadcast's start, but in the cases that this is not clearly evident in the video material these have not been included here. Still the few

live feeds seen in this category are interesting in that they represent situations where we can clearly see that people have spontaneously responded to live events by grabbing their mobile phones and sharing the event in real time using a mobile broadcasting service.

ANALYSIS OF PRODUCTION STRATEGIES

In the following we will analyze the corpus to tease out the ways in which people manage this new format, that is, how they handle the production of live video. First we focus on their camerawork and how they vary the framing of objects in view, which is often referred to as transitions. Second, we focus on how the camera operator talks to people in view as a way to manage their appearance. The content is then discussed with reference to theories from professional filmmaking.

Transitions in Live Video Work Using a Single Camera

The ways in which people make camera movements, either to follow a particular topic or to change the view from one topic to another during the broadcast, is a salient feature in the corpus. Broadcasting live from a handheld device has a very important restriction in that *editing,* in the traditional sense, is disabled. Editing is the standard way of combining image sequences shot at different times and/or places into a continuous narrative, in order to more effectively tell a story [8]. It is used among other things to condense elements of a story that take place over time and, which is of relevance here, to make *transitions* from one framing of a topic to another. Transitions are important in any image sequence containing more than one topic, and can be more or less deliberate and well executed. Making transitions work through camera movements in mobile live streaming is particularly demanding, as we will explore in this section.

We have identified two basic skill levels in managing transitions, hereafter referred to as "controlled movements" and "spray paint." The first reflects some familiarity with camera operation techniques used to produce quality footage, the latter a less trained approach, simply pointing the camera at whatever is interesting at the moment. As opposed to the previous content analysis, we have not conducted a detailed quantification of the prevalence of these content categories. Such an analysis would be too cumbersome, given the overall abundant availability of these phenomena and the size of the corpus.

Controlled Movements

There are a number of examples displaying what appear to be controlled camera movements in one basic plane at a time during broadcast – i.e., with the camera moving in a single plane only while remaining fixed in the other two – thereby resembling well-established professional techniques in filmmaking.

Figure 5. Broadcast sequence using controlled camera movements and variations in framing to report from a break during a motorcycle trip.

A broadcast on Qik, the 1st of June, 10:30–10:40 GMT, is a good example. The video shows a group of motorcyclists taking a break to look at the view from a mountain road in Italy (see Figure 5). The first-person-view broadcast starts with a medium close-up of a man in a motorcycle outfit and helmet, shot from a straight-on angle. The camera then *pans* to the left and reveals the view from the mountain, stays on the view for a couple of seconds, and then *tilts* up and down, showing the depth of the valley. The camera then continues to pan to the left, stops for a moment and zooms in to a close-up of a mountain peak, and then comes to a stop on a long shot of the road with motorcyclists standing in a row. The broadcaster then begins *tracking* past the people, one after another, framing steadily while the passing people appear in close-ups, medium shots, and extreme long shots. The subjects are waving in response to the comment that they are being streamed live on the internet.

The types of camera movement discussed above (pan, tilt, and tracking shot), correspond to well-established transitions in filmmaking. A *pan* is a camera movement where the camera turns right or left, rotating on its axis, producing a mobile framing that scans the space horizontally. Analogously, a *tilt* is a movement up or down on a fixed axis. A *tracking shot,* on the other hand, is a movement where the camera position is changed, causing a mobile framing that travels through space forward, backward, or laterally [4].

Looking at the video broadcast as a whole, the production resembles a classic hand-held single camera video reportage, with the reporter commenting on things in front of the camera and directing the people present. Instead of using camera movements to follow people or objects, the broadcaster constantly moves the camera to put emphasis on new content onscreen. Using this technique, the framing, actions, and camera movements look relatively well planned, to the extent they can be so at an ad-hoc event like this.

Uncontrolled Movements
Amateur camerawork tends to be a mixture of the movements discussed above. Moving the camera in more than a single plane at the same time, often without a distinct beginning and end of movements, is often referred to as "spray painting," likening untrained camerawork to holding a spray can. There are many examples of such production strategies in the corpus, such as in the following example

on Bambuser, the 6th of May, 11:00–11:10 (GMT+1), from a private home in Russia.

The broadcast starts with an extreme close-up of a TV screen. After about two minutes the camera moves to the left, scanning the wall in the room, and then holds on the floor where a carpet and a heap of clothes are visible. The image lingers on the floor for a few seconds until the camera slides back, more slowly this time, to a close-up of the TV screen. The image stays on the TV screen very briefly, and then moves to the left one more time – returning to the floor with the carpet and clothes, and then to the entrance of the room in the back where a cat is walking in. The camera follows the cat coming closer and skips down as it walks over the carpet. Then the camera pans to the right, passing the TV screen to a close-up of a computer screen showing the Qik site with the live streaming broadcast. The camera moves to the left again, back to the TV screen and at the same time zooms in to an extreme close-up. After that, the camera moves indistinctly between a close-up of the face and fur of the cat, then back and forth over the walls, as the cat moves out of the picture, first showing a poster, then moving back to the right, ending up on a close-up of the TV screen.

This episode contains several characteristics of inexperienced camerawork: unsteady movements, zooming while panning, and indistinct beginnings and endings of movements. The broadcaster skips swiftly between subjects (the cat and the TV) and reacts to them rather than maintaining a steady framing. Quality judgments aside, this leaves fewer opportunities for distinct transitions between subjects of interest, and makes it harder for the viewer to concentrate on any intended topic in the video, since the camera is constantly moving.

Starts and Endings
The characteristic look of the starts and endings of broadcasts is something of an inadvertent transition feature deriving from the user interfaces of mobile phones and video services. Both actions require pressing the record button on the mobile phone, which may be more or less smoothly done depending on how you are holding the phone. There are examples of well-planned and smooth handling of these transitions, but more often the broadcasts begin with nothing in frame or with the camera being set up unsteadily, and end with the action of turning the camera towards the floor while the user is probably looking at the interface on the screen or searching for the record button on the phone.

In traditional filmmaking and videography, a common practice is to leave the camera running for a number of seconds at both the beginning and end to make room for editing out the actual start and ending, along with the button presses and unwanted camera movements they may contain. As editing is not an option in live broadcasting, these stages will, in effect, become transitions to and from the actual broadcast content.

Coordination Talk
The ways in which the broadcasters address other people in view are another salient feature in the corpus. In 41 videos out of 178, the broadcaster verbally comments or talks to people in the viewing context. In the following section, we will discuss these instances of conversation and we will argue that they are a way to coordinate and manage the appearance of the topic at hand. On-camera talk can be separated into two different categories: explicit and implicit directions.

Explicit Directions
There are altogether 36 videos in which the broadcaster explicitly gives verbal directions to the people who appear on camera. The most common phrase, "we are live," occurs on 15 occasions. It informs the people present that there may be someone not present who can see what they are doing. Although the broadcaster is not actually telling the people how to act, this is a hint that their behavior has to be accountable in a broader context. Reactions by the people in view of the camera towards being live on the internet vary from laughter and excited screaming to questions such as "are you serious?" and disregard. A common way of directing, occurring in 21 videos, is to request specific actions. Examples include the broadcaster asking people to introduce themselves, i.e., to "say hello" or "wave" to the camera while in the frame. But there are also occasions when the broadcaster asks dogs and babies to act in specific ways.

These types of directional comments are a way to make the content more visually interesting. In professional live video, such instructions are given ahead of going live. Thus, these comments indicate that these broadcasts were unplanned. The broadcaster may notice that the video is rather dull only when she is already on air, which causes her to ask for some form of action. We suggest that both these types of comments can be seen as a way to coordinate or influence the action in front of the camera.

Implicit Directions
There are five videos in which the broadcaster asks direct questions. Such talk can, of course, be understood as an interview and thus belong to the topic categories. However, we also find it interesting from a coordination perspective since the questions arguably serve dual functions, such as in the following example.

The video displays a table with plates and cutlery. The broadcaster turns the camera to show himself in third-person view, as well as the people sitting next to him. He says:

Broadcaster: Hey what's going on, Kyte family? I'm sitting here with... [turns the camera to show the person sitting next to him and puts his arm on his shoulders] Look, this is my nephew. [Nephew nods and smiles to the camera] He just graduated Law School. [Turns the camera to show both himself and the nephew in the video] I'm a proud uncle.

Nephew: Thanks.

Broadcaster: How does it feel? How does it feel? By the way it's attorney.

Nephew: [inaudible answer]

Broadcaster: Okay you will be an attorney after you pass the bar in two months.

Nephew: That's right, that's right.

Broadcaster: [turns the camera to frame only the interviewee] So how does it feel?

Nephew: It feels great.

Excerpt 1: Transcription of conversation in video

The interview format, both in terms of the conversation and the framing of the video, puts the interviewee at the center. By choosing to interview the people present, the broadcaster induces them to talk about a selected topic. This is a powerful way to coordinate the live appearance. The interview reveals who is sitting around the table, as well as why they are there. It is a format that implicitly coordinates the activities in front of the camera, but much more naturally than explicitly asking for actions.

In both cases discussed above, i.e., giving explicit and implicit direction, the broadcaster takes on the role of director. It is clearly visible that the person with the camera not only selects camera views and transitions, but also actively tries to manage the situation in front of the camera. Action, framing, and transitions all need to be controlled during the live broadcast, and what is not planned for in advance needs to be directed in real-time as the situation unfolds. This may explain why we see more verbal directions on-camera in amateur broadcasts.

DISCUSSION
The content analysis reveals an emerging and fascinating social medium. However, it is evident that mobile broadcasting is a medium whose users are still struggling to make use of its specific affordances.

From Testing to Actual Use
At this early stage, testing is by far the most frequent activity, surpassing all of the more qualified content categories combined, including broadcasts from social events, presentations, and performances. Given that this is an unestablished medium that has only been available since

2005, it is not surprising to find so many people just trying out the technology on these sites. In that sense, the large number of tests need not indicate a user experience problem. However, it is clear that the large total number of video clips at these services does not indicate that there are equally many users. Hence, the testers need to progress to actually providing live broadcasts of selected topics. We see some of the mentioned categories – e.g., demonstrations, tours, and performances – and their respective numbers of occurrences, as early indicators of how users are becoming familiarized with these services and what types of content we can come to expect in the future.

Finding Relevant Topics

As soon as the users master the technology, they can get on with the broadcasting of live situations. However, finding such topics seems to be problematic as well. Many of the sampled videos are uneventful and border on tests in terms of production quality and camera use. The "home tours" that display the interior of homes are a salient example.

Interestingly, there is also something of a "screen paradox" in the material. The screens are typically TV sets and computer screens. Mobile broadcasting services provide users with a tool to display everyday situations occurring in their physical life, but instead there appears to be an extensive selection of screens in the videos, displaying parts of our digital and mediated life. We suggest that this might be due to difficulties in finding dynamic topics in everyday life, and that what happens on the screens might be considered as more interesting than our other life, thus paradoxically leading the broadcast back into the digital realm. This is, of course, a tentative interpretation of the content, and there might be a number of other explanations of this frequent occurrence.

The broadcasters' explicit talk during the broadcasts is a way to improve their quality while on the air. It is a way to add action to an everyday situation that runs the risk of becoming uninteresting while being broadcast publicly on the internet.

But the analysis also reveals that there are people who find ways to provide new forms of video topics. There are smaller groups of users streaming live from tourist sites, presentations, and social events. These broadcast performances differ from those studied by Shamma et al. [12], which were closely linked to desktop practices. Here we see webcasts from other locations and using less computer-dependent practices. There were several broadcasts of social groups and family, which underscores Reponen's [11] argument that it is useful in these situations. Broadcasts in the categories that could be said to have the strongest element of liveness, i.e., sudden events and performances, are rare but notable. These instances are arguably among the more advanced uses of the technology,

in the sense that they take advantage of the medium-specific properties of mobile broadcasts.

Live Directions as Amateur Activity

Previous research has shown that people do not edit mobile videos after recording them [6]. They record a video and show it unedited to their friends. This indicates that amateurs refrain from doing much post-production work. In a similar way, many of the videos contain real-time direction, which indicates that these shots are not so thoughtfully planned ahead of going live. They might find planning activities ahead of a broadcast, which is what professionals do, equally tedious. Thus, what we see in this amateur medium is real-time direction of various kinds, or a more clever use of interviewing techniques, which conceals the activity of coordination within the format per se.

Video Logs as a Transitional Category

As previously argued, mobile broadcasting applications differ from desktop webcam applications in that they allow users to capture content from mobile contexts and gain physical control of the camera. We argue that the video-log content category contains a particularly interesting diversity of ways of using of the mobile phone as a video broadcasting device, ranging from being essentially identical to a webcam to producing broadcasts on the move, actively referencing the passing background environment. The similarities to webcam use make the comparison between mobile camera phones and webcams a useful starting point for investigating what the mobility aspect contributes to this type of mobile broadcasting. The webcam can be said to be the predecessor of mobile live streaming, in terms of both use context and technology. Making your mobile phone work as a webcam (not emphasizing any of its mobile features) was actually the selling point of some of the earlier services in the field.

In a desktop setting, a common setup is to place the mobile phone with the camera facing oneself, while either testing or video chatting. This is the basic level of use of mobile streaming technology, as seen in the collected broadcasts in this study. Apart from the technology being a mobile phone, this type of use is very similar to webcam use, to the point where the content produced cannot be distinguished from the more established use of webcams in terms of framing, duration, or even image quality. This type of remediation – reconstructing and drawing upon the formats of earlier media and eventually refashioning them – is typical in any new medium [3].

As could perhaps be expected at this early stage, most of the broadcasts in the Vlog category fall into this early remediated format, transferring familiar use of live video to new technology but not actually taking advantage of any of the mobile properties of the camera phone. It is within this category that we can see this most explicitly, precisely because the form of the content is already set. The users may already have a viewer base and are merely changing

broadcast devices. But as we are looking for emerging media-specific usage of mobile webcasting, the few broadcasts that do display use of mobility are particularly interesting. E.g. the previously presented videologger reporting from downtown Toronto displays greater familiarity with mobile broadcasting, making well-planned framing decisions, timing camera movements to his own commentary, and maintaining awareness of the background setting as he moves through physical space.

Although the examples of mobile Vlogging are few in the corpus, they give us interesting indications of how existing categories can be further developed when users begin taking advantage of the specific properties of mobile broadcasting technology. As the format – here a third-person view of the broadcaster and a personal, diary-like commentary – remains the same, the possibilities that these properties afford become clear. More examples that point towards interesting new uses of the medium can be found in, e.g., the tour, presentation, and performance categories.

IMPLICATIONS FOR DESIGN
The content analysis of our corpus points to some possible areas of improvement of the design of current mobile broadcast services, and implications for the design of future services. Here, we present these in relation to the content categories.

Improving on Users' Production Strategies
It is apparent that many users struggle with the ways in which live video should be handled. Even if they have a clear idea of what to broadcast, it is not obvious how they should go about it. Here we suggest that the services themselves should provide guidelines on how to carry out the production, both on the web interfaces and on the mobile devices. We argue that assisting more inexperienced users in acquiring basic camera operation techniques, as well as visual storytelling techniques, is an integral part of bringing these tools, previously reserved for professionals, to amateur users.

Support in the Web Interface
The desktop web interfaces could include a selection of "best practices" or "editor's picks" based on interesting production strategies, in addition to the most recent and most viewed clips which are already available.

Support in the Mobile Devices
There are also opportunities to assist the users on the mobile phones per se. Such support could consist of either simple demo guidelines in video format, or a more ambitious version that draws upon further phone capabilities than just video playback.

A more advanced version would include interactive tutoring in the camera that would allow the user to try out standard camera techniques, and automatically recognize users' skills through image recognition, or through other sensor technology in the device. The tutoring could include the use

of transitions, such as panning, tilting, and tracking shots. It could also allow them to explore some standard variations of framing and their typical applications – e.g., close-ups, medium-shots, and landscapes. The applications would present standard filming situations and then track the strategies chosen to give feedback on how well the user completes the task of producing, e.g., a steady tracking shot or a well-framed portrait in an interview situation.

In this we recognize that the service providers should be careful not to impose fixed formats that would inhibit creativity and interesting new ways of using live video as a medium. Being overly didactic is certainly a risk with services intended for non-professionals, who may not necessarily want to follow conventional formats.

Management of Beginnings and Endings
The handling of broadcast initiation and termination is a clearly visible problem in the videos. In most cases the videos start and end out of focus and wobbly. This problem is probably inherent in the services as they are available now. We suggest that more controlled broadcast starts could be aided by a countdown similar to an automatic time release on a still image camera, separating in time the pressing of the record button and the framing of the first image. The endings could become smoother if the termination could be done by any key on the device.

CONCLUSION
It is obvious from the number of clips found on the investigated websites that this medium triggers people's interest. They are curious about the concept and try it out. Some of these users also take live video to new places, both in terms of topics and ways of using it. But not everyone seems to take this step. It was somewhat to be expected that the use of mobile live video would be remediated in available formats, such as the Vlog. This could be expected with any new medium involving some degree of unfamiliar technology. But it is clear that mobile webcasting has not yet fulfilled its potential, foreseen by researchers, to become the latest in a long line of successful social media, and to support group interaction and empower citizens. There remains a challenge for the designers of these services to develop the concept in order to support people's appropriation and thereby democratize a medium which up to now has been entirely in the hands of well-trained professional TV-producers. Just providing a way to stream video from mobile phones does not seem to be enough.

ACKNOWLEDGEMENTS

REFERENCES

1. Bergstrand F, Landgren J, Information sharing using live video in emergency response work, In proc. Of SCRAM 2009, Göteborg, Sweden

2. Blythe, M., Cairns P. Critical Methods and User Generated Content: the iPhone on YouTube. In proc ACM CHI2009 Boston, US, pp 1467-1476

3. Bolter J D., Grusin R., Remediation: understanding new media, MIT Press, 2003

4. Bordwell, D. & Thompson, K. 2001. Film Art: An Introduction, 6th ed. The McGraw-Hill Companies, Inc. New York, USA.

5. Engström, A., Esbjörnsson, M. and Juhlin, O. (2008). Mobile Collaborative Live Video Mixing. In Proc MobileHCI 2008. ACM Press, 157-166

6. Kirk, D., Sellen, A., Harper, R. & Wood, K. 2007. Understanding Videowork. In proc. ACM CHI'07, San Jose, USA.

7. Krippendorff K., Content analysis-An introduction to its methodology, Sage, London 2004

8. Mascelli, J.V (1965). The five Cs of cinematography, Silman James Press, USA.

9. Multisilta J, Milrad M, Sharing Experiences with social mobile media, In proc. Mobile HCI 2009, Bonn Germany.

10. O'Hara, K., Black A., et al Everyday practices with mobile video telephony, In proc. ACM CHI 2006, Montreál Canada

11. Reponen, E. 2008. Live @ Dublin - Mobile Phone Live Video Group Communication Experiment. In proc. Euro ITV'08, Salzburg, Australia.

12. Shamma, D. A., Churchill, F. E., Bobb, N. et al. 2009. Spinning Online: A Case Study of Internet Broadcasting by DJs. C&T'09, Penn., USA.

13. Tazaki, A. (2006). InstantShareCam: Turning Users From Passive Media Consumers to Active Media Producers. Presented at the Workshop Investigating new user experience challenges in iTV: mobility & sociability, held at CHI'2006.

14. qik.com/browse/videos/all?subcat=recent. Visited 2009 09 17 at 22.35 CET

15. Quinn Patton, M., Qualitative evaluation and research methods, 1990, Sage, London UK