



Enabling Informal Communication of Digital Stories

Philippe Debaty, Patrick Goddi, Rich Gossweiler, Rakhi Rajani,
Alex Vorbau, Josh Tyler
Consumer Applications and Systems Laboratory
HP Laboratories Palo Alto
HPL-2004-180
October 18, 2004*

E-mail: firstname.lastname@hp.com

digital media,
photo sharing,
multimodal,
camera phone,
storytelling

The convergence of communication and imaging capabilities in a single device, the camera phone, is changing the way people take, share and communicate around pictures. In this paper, we describe and discuss three complementary research prototypes - MemoryNet Viewer, Plog and StoryMail - that we built to explore how media can be used as part of everyday storytelling activities. Each system focuses on informal, casual and lightweight solutions for multimedia storytelling and conversation. We conducted small preliminary pilot studies that revealed interesting patterns of use of the media within social networks, which we plan on investigating further.

Enabling Informal Communication of Digital Stories

Philippe Debaty, Patrick Goddi, Rich Gossweiler, Rakhi Rajani, Alex Vorbau, Josh Tyler

Hewlett-Packard Labs

1501 Page Mill Road

Palo Alto, CA 94304

{firstname}.{lastname}@hp.com

ABSTRACT

The convergence of communication and imaging capabilities in a single device, the camera phone, is changing the way people take, share and communicate around pictures. In this paper, we describe and discuss three complementary research prototypes - MemoryNet Viewer, Plog and StoryMail - that we built to explore how media can be used as part of everyday storytelling activities. Each system focuses on informal, casual and lightweight solutions for multimedia storytelling and conversation. We conducted small preliminary pilot studies that revealed interesting patterns of use of the media within social networks, which we plan on investigating further.

Keywords

Digital Media, Photo Sharing, Multimodal, Camera Phone, Storytelling.

1. INTRODUCTION

Despite today's trend of combining communication and imaging capabilities into a single device (e.g. cell phone cameras), it is still difficult for people to use digital media as part of their everyday casual storytelling activities within their social networks. As observed in [3], there are many issues concerning the way people communicate and converse around their digital media.

In this paper we introduce and discuss three research prototypes focused on a casual form of multimedia storytelling that is designed to capitalize on the capabilities of current and future storytelling devices that integrate mobility, immediacy, and communication. These three complementary prototypes are:

- MemoryNet Viewer [1], is a peer-to-peer system for circulating, displaying, and annotating photos and videos within a social network.
- Plog [2], a system for easily capturing, sharing, and presenting digital photo-based and location tagged stories from camera phones.
- StoryMail, which enables camera phone users to record, transmit, and playback voicemails augmented with a sequence of digital pictures or video clips.

These systems emphasize informal, light-weight, and low effort mechanisms for delivering, presenting and conversing around personal digital media.

2. OUR APPROACH

By surveying current media sharing systems (see section 5), we identified five categories of functionality that are relevant to digital media-centric communication: capture, organization,

distribution, playback, and annotation. While our approach is to focus on informal, casual and lightweight solutions, in this section, we will position, motivate and distinguish our contributions within those categories.

2.1 Capture

The fact that camera phones integrate mobility, connectivity and photo taking capabilities on a single device, brings a new meaning to capture. People no longer take a picture for the sole purpose of saving a memory, they now can capture with the intent of immediately communicating an experience. Saving a memory with a traditional camera only takes one click of a button. We wanted to preserve this simplicity on camera phones, while supporting this new kind of capture. One of our prototypes, Plog, simplifies the phone interface to enable performing this in two clicks. The first click is to take a picture, while the second is to confirm that the picture should be communicated.

2.2 Organization

Prior art [3] has indicated that organizing pictures into stories or contextual clusters can enhance meaning, provide better recall and greater emotional value. We explored several concepts to make this process informal, lightweight and low effort. People tend to take pictures in clusters around an event. Plog uses capture time information to automatically determine these clusters, which define stories.

However, we found that people also have the desire to have a more reflective process about defining and editing a story after the capture time. Our StoryMail prototype focuses on making this reflective process as casual as possible by combining it with an oral narration of the story.

2.3 Distribution

The process of distributing stories from photographers or storytellers to the audience is difficult with current technologies. Many steps are involved in getting pictures from a digital camera to a computer to a website to your friend's house. In an effort to simplify this process, the Plog prototype adopts a syndication model. Pictures are automatically and immediately sent from the phone to a central server via GPRS. The pictures, organized in stories can then be fetched automatically by various clients from anywhere by simply using a Web browser.

Our MemoryNet Viewer prototype explores a different angle with a peer to peer distribution model. Each user defines a friend list and pictures and annotations are automatically circulated within this social network.

2.4 Playback

Just like the organization process, playback can range from a very involved process requiring the whole attention of the

viewer, to a more peripheral or ambient model where media-centric stories are presented in a non intrusive way. Each of our three systems covers a different range of this spectrum. StoryMail presents a narrated slide-show and therefore requires the attention of the viewer. MemoryNet Viewer is at the other end of the spectrum by being designed to be viewed in an ambient, non-intrusive mode while still allowing for interactivity when desired. Plog is somewhere in the middle since it presents pictures as animated stories and alerts on new stories as they happen. Plog also focuses on rendering the stories in an engaging format, supporting for instance the creation of printed postcards.

Another aspect of Plog and MemoryNet Viewer is that stories are presented in the larger context of other stories, told from other people, encouraging the viewer to make implicit connections between memories and people.

2.5 Annotation

Annotating and adding information to media-centric stories is often too tedious for users, who can be either the storyteller himself or the viewers. To simplify this process, our approach is to make it an integral part of the act of communicating the story. Our MemoryNet viewer system records conversations that collocated users have when pictures being displayed in an ambient mode draw their attention. With StoryMail, people can annotate a story simply by recording a voice narration over a sequence of pictures.

2.6 System/Overall

Across all these categories of functionality, our focus is not on sharing pictures but on communicating via media-centric stories. A single picture can tell a story (MemoryNet viewer focus), but stories can come from multiple pictures with (StoryMail) or without (Plog) words juxtaposed.

Each of our systems explores a slightly different communication style. Whereas StoryMail allows for a narrative communication model, MemoryNet Viewer is designed to support and encourage a conversational style. Plog offers an implicit model, where users communicate simply by the intentional act of taking and sending a picture.

We also recognize that there are different moments when people have the desire to communicate stories. Plog focuses on enabling people to communicate right at the time of capture. With StoryMail, this communication moment happens sometime after the capture time, when the user feels that he has enough captured material to tell a self contained and vivid story. MemoryNet viewer focuses on the communication that happens when the pictures are being viewed.

Finally, whereas StoryMail and Plog focus on enabling media-centric communication only between remote users, MemoryNet Viewer also supports users being collocated and conversing together around a digital media artifact.

3. THE SYSTEMS

3.1 MemoryNet Viewer

The first system we developed, MemoryNet Viewer (MNV), is a peer-to-peer system for circulating and displaying photos, videos, and annotations within a social network. Users can add and remove people from a "friend list", which determines who

exchanges media with the user. The MNV displays friends' media in a continuous slideshow and was designed to operate as a peripheral device and viewed passively (see Figure 1).

Recognizing that people often consider annotation of media a tedious activity, the MNV encourages others to comment on media spontaneously -- when they happen to see something of interest. Users can "attach" text or audio comments to media which continue to circulate with the media object to others' displays. This form of passive viewing and opportunistic annotation, results in threads of conversation or stories around individual media objects that form over time.

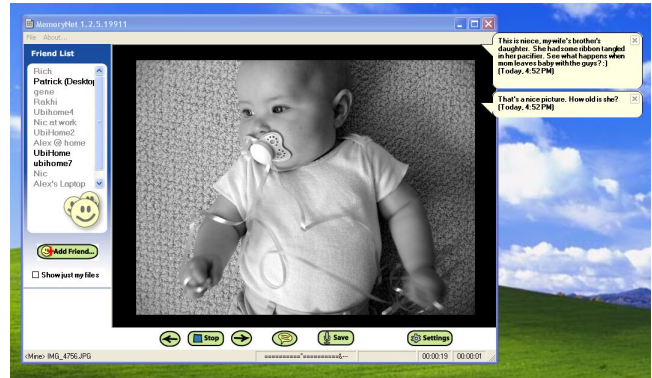


Figure 1. Screenshot of the MemoryNet Viewer Prototype

To avoid collocated users realizing in mid-story that they were not recording their comments we use a buffered audio system that always records and deletes the recording if the user does not opt to save the annotation.

The MemoryNet Viewer is a .NET Windows application and uses HTTP for inter-node communication. A lightweight, embedded web server answers requests for media files and the .NET Remoting mechanism delivers media metadata and internal messages. To reduce bandwidth usage, nodes cache media objects locally until the application is closed.

User studies revealed an interesting social issue: users sometimes wanted to withdraw a regretted comment. With this in mind, we architected the system so that the control remains with the annotation author by storing the annotation on the author's computer.

3.2 Plog

Although MemoryNet Viewer proved very engaging and valuable, it also had some interesting shortcomings. Among them, there was no simple way to get images and video into the system. Moreover, images and video were independent objects, and could not be presented as multi-object sequences. The Plog system, in addition to exploring new models of casual capture, addresses these shortcomings of MemoryNet by providing a radically simple model for getting images into a system, and allowing them to be clustered together in stories using context data like time and location.

Technologically, Plog is very simple—the key innovation is in the user experience. Plog consists of four key pieces of technology:

- Simple "click and send" phone software that makes it much easier to take and share a photo—one click to take a photo,

one to send it (see top left of Figure 2—no more menus, forms, or manual text entry)

- The phone software also includes location information read from a GSM network (accurate to the size of a cell)
- A back-end server that clusters a user's photos into distinct story vignettes, based on time and location algorithms
- Story viewer software, run locally or over the web, that provides layout and animation for the stories, including a map of the photo locations, making them more entertaining and memorable (see Figure 2). The presentation is done with a scripting language that allows authors to create their own presentation styles. The software also enables the printing of picture stories in a postcard format.



Figure 2. Plog creates story vignettes from a simple “click and send” phone interface (top left) and animates them, with location information (the map at top right is zoomed in)

The Plog server syndicates photo content via RSS (Really Simple Syndication). We use RSS to communicate with the standalone Plog client story viewers, enabling us to decouple presentation from the data structure, as well as to integrate with other web services and software (exclusively internal projects so far).

3.3 StoryMail

Although audio annotations provide a more personal alternative to text annotations, with the MemoryNet viewer, we observed that most audio annotations were done during collocated storytelling and people were reluctant to tell audio stories to a screen when they were alone.

The StoryMail system acknowledges the value of voice annotations and integrates the more familiar model of voicemail as a method for combing images with voice. StoryMail also introduces an alternative model to Plog for organizing photos and video into stories through its interactive interface.

In essence, StoryMail is an extension of the traditional voice mail service model, which is familiar to millions of mobile and landline phone users. It enables camera phone users to record, send, receive and playback voicemails augmented with a set of digital pictures or video clips.

The client part of the StoryMail system provides an interface which allows the user to create a timed sequence of pictures

and/or video clips while narrating a story. The sequence is incrementally created simply by selecting previously taken picture and/or video clip at the desired time while telling the story. The system records both the voice track and the pictures sequence, resulting in a message, which can then be reviewed and distributed to friends and family. Figure 3 is a representation of such a message.

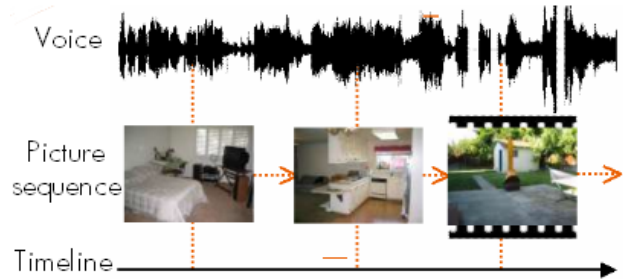


Figure 3. StoryMail message consisting of a narrated timed sequence of photos and video clips

We implemented two prototype systems, which are proofs of concepts for two complementary deployment models.

Our first prototype is based on a voice interface delivered by a server-side voice application, which we implemented using the HP Opencall Media Platform [4]. To invoke it, the user calls a number corresponding to his personal story mailbox. A voice menu lets him choose between recording, sending or playing back storymails. During the recording or playback phase, the voice part of the message is transmitted live over the dedicated voice channel of the phone while the digital media part is transmitted separately via the data channel. Our implementation uses VoIP protocols (SIP and RTP) to transmit the voice part, SMTP to transmit the data part. The client part consists of a software simulated cellphone running on a PC. It is represented on the left side of Figure 4.

In the second prototype, the audio track is recorded by the handset itself instead of being recorded by the server. Once the handset has recorded both the audio track and the corresponding sequence of pictures, it sends the whole message via email or MMS. We implemented this prototype on a camera phone running the pocket pc phone edition platform. It is shown on the right side of Figure 4.

Both prototypes use SMIL to represent the structure of a story mail.



Figure 4. Screenshots of our two StoryMail client prototypes.

4. LESSONS LEARNED

While larger user studies of these technologies are either planned or underway, small pilot studies have revealed data that has allowed us to evaluate and further explore the issues detailed above. In evaluating the data we noticed three prominent themes.

4.1 New Images and Image Models

We observed that the combination of mobility, immediacy and connectivity supported new types of image capture and use. More photographs were being taken and of more every day items. We saw pictures of grocery lists, shoes, cars, road signs and others that were taken on a regular basis. It was also very interesting to note that picture taking changed because of the knowledge that the items would be shared immediately and that there was someone or some group of people who would see them soon. When users knew who their audience was going to be, they took photos of relevance to that person and we often saw humorous images that were sent to cause a reaction. The collections of image became less formal in their nature and more relevant to the people in our social networks. These uses were compelling to observe and we continue to investigate them.

4.2 Different Interactions with the Media across Time

Stories and interactions change in time and require different modalities based on both user goals and circumstances. With Plog, the story to be told was decided in the moment of capture and continued in sequences of photographs with very little text. Voice conversations sometimes followed the display of photographs and interestingly, with Plog, some viewers of the story vignettes created their own interpretations of what was going on. With the MemoryNet Viewer stories evolved over time through text annotations and groups of people were often found in front of a viewer, sharing either stories or the experience of viewing them. With StoryMail, a set of photographs was narrated and delivered as a package that was later viewed. The different models here presented different methods for communicating moments depending upon the situation and needs of both the sender and receiver. Other modalities such as the telephone sometimes supplemented the communication that was sparked by one of these applications.

4.3 Social networks dictated use of media

An additional observation relates to story distribution in a social network. It was common for multiple stories to be communicated using a particular set of pictures. For example one story in the dorm network, while a very different story in the family network. Users are very interested in maintaining control of their pictures and stories as they are made available. MemoryNet provided some controls but observations verify that this will be an important issue in the success of these types of systems.

The social intricacies of the use of images and the relevance and acceptability of storytelling with these applications are yet to be explored in depth. While we observed some positive uses and feedback we do not assume that this is what people want to do and do for prolonged periods of time. There were points at which the push of images became an intrusion and where the interaction models we proposed were not suited to the needs of

use. We have much anecdotal evidence that we are investigating further with larger studies that probe both the technical and social aspects of these technologies and their application.

5. RELATED WORK

Frohlich, in [3], categorizes "Photoware" into four quadrants based on whether the person being narrated to is present or remote and whether the sharing and listening is done at the same time or at different times. Each of these projects falls into the category of asynchronous and remote. Frohlich also makes the suggestion of including the ability to share photos instantly. While not truly instant, Plog and StoryMail reduce the potential gap significantly.

Current commercial sharing photowares like Ofoto [5] and TextAmerica.com [6]. Ofoto is a traditional web-based sharing web site and differs from these projects in that it does not support audio narration or videos in stories (MemoryNet, StoryMail) and that photo collections are announced via a URL in email or the like and not send directly to a public or private display (Plog). TextAmerica is a photo weblog site and shares similarities to Plog, but it does not auto-cluster photos into stories.

From our lab, Fleck describes in [7] a system for recording a narrated audio story over a series of photos for the purpose of voice recognition and keyword identification. StoryMail extends this work by transposing it to a telecom infrastructure and focusing on multimedia communication between cell phones.

Finally Picshare [8] and Picassa's Hello [9] are systems for having synchronous conversation around photos over the internet. MemoryNet and StoryMail both support audio narration of stories, but in asynchronous form.

6. CONCLUSION AND FUTURE WORK

We continue to develop systems and interfaces that probe casual storytelling patterns. Several pilots are either underway or planned. Future work involves better support for multiple overlapping social networks, and continued investigation into creating rich stories and interactions via very casual storytelling mechanisms.

7. REFERENCES

- [1] Reference removed for double blind reviewing.
- [2] Reference removed for double blind reviewing.
- [3] Frohlich, D., et. al., "Requirements for Photoware", Proceedings of. CSCW 2002, pp. 166-175.
- [4] OpenCall Media Platform (OCMP), <http://h20208.www2.hp.com/opencall/products/media/ocmp/index.jsp>
- [5] Ofoto, <http://www.ofoto.com>
- [6] TextAmerica, <http://www.textamerica.com>
- [7] Fleck, M., "Eavesdropping on Storytelling". HPL-2003-201.
- [8] Pilu, M., "PicShare: A Capture-Centric Remote Collaboration Tool", HPL-2000-14.
- [9] Picassa's Hello. <http://www.hello.com>