Frank Nack
CWI

Video Blogging: Content to the Max

Conrad Parker and Silvia Pfeiffer
Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia

The lure of video blogging combines the ubiquitous, grassroots, Web-based journaling of blogging with the richness of expression available in multimedia. Some claim that video blogging will be an important force in a future world of video journalism and a powerful technical adjunct to our existing televised news sources. Others point to the huge demands it imposes on networking resources, the lack of hard standards, and the poor usability of current video blogging systems as indicators that it's doomed to fail.

Like any nascent technology, video blogging has many unsolved problems. The field, however, is vibrant, the goals are fairly clear, and the challenges they pose to multimedia researchers are exciting indeed.

A cursory overview

From around 1997, blogging has become a quickly spreading passion among Internet literates.[1] A blog entry is essentially a regular text entry into a Web page to publish experiences, thoughts, and opinions. These usually accommodate diary entries but are also used for news, discussion, photo albums, and regularly updated lists of links to other interesting sites such as lists of publications, books, or music.

During world-shaking events such as the recent tsunami in Southeast Asia or the latest election of the US president, the best sources of information are often not the official TV news, newspapers, and government press releases but the personal diary entries of bloggers. Not only do the bloggers describe events from a personal and thus more immediate perspective, but they also provide validation and corrections to public media releases. In that way, waves of support can be started for people and communities who otherwise would not be heard.

The earliest Web logs were nothing more than regularly updated Web sites, like the grassroots technology news site slashdot.org, and readers had to visit the site on a daily basis to read the latest updates. Since then, much of this process has been automated. Although still viewable as HTML Web pages, such sites also publish machine-readable updates in syndication languages such as Rich Site Summary (RSS) or Atom (see the "More Information" sidebar). Rather than always having to remember to visit a site directly, a reader can configure his or her RSS browser software to automatically check for updates. This process is known as *subscription*, and, borrowing further from the language of print and broadcast media, the content source is a *channel*.

Although subscribing to a single channel provides only a small saving of effort, subscription becomes quite useful when you're interested in reading updates from a larger number of sources. You can configure such software to subscribe to multiple channels, in which case your RSS browser will regularly poll the various sites for updates. Because the updates' format is machine readable, the site can present results in summary form—for example, displaying headlines and authors, with individual items sorted by publication date. This process, known as *aggregation*, lets you create a

Editor's Note

The phenomenon of blogging—sites packed with links, quips, ideas, and arguments that once were the near-monopoly of established news outlets—has already changed the media world. Blogs can be as nuanced and well sourced as traditional journalism, but they have immediacy and they are personal.

The authors of this column investigate what's required to propel text blogging to the next level, namely video blogging. For that, they investigate first attempts in this direction and provide a view on existing technology that can help establish a publishing revolution more profound than anything since the printing press.

—*Frank Nack*

1070-986X/05/$20.00 © 2005 IEEE

Published by the IEEE Computer Society

completely customized news summary and have it automatically updated on an ongoing basis.

As blogging has exploded in popularity, with many thousands of individuals each offering their own channel for anyone to subscribe to, the aggregation task has also become more complex. This has given rise to stand-alone aggregators, namely Web sites that are subscribed to all the channels a particular community's members offer, republishing the aggregate as an easily readable Web page, or even offering it as its own, collectively larger, channel. Such sites form a common meeting point for members of a community, and they let the site moderator introduce new members who are then immediately visible to all.

Automated aggregation lets you view updates from widely distributed Web sites as an integrated whole, which is essentially what differentiates blogging from conventional Web site publishing.

Enter video blogging

If blogging, with its highly customizable subscriptions and automated aggregation, can help us keep abreast of the Web's vast quantities of text, can we apply the same principles to video publication so that we can keep up with both the latest newsreels and the highlights from our friends' and families' lives?

If so, then video blogging technology will allow anyone to publish their own channel. It will allow all sorts of communities to build automated aggregates of their members' footage—combining soapboxing, distributed discussion, and the social cohesion of blogging with the vibrant, immediate honesty of video. Ideally, we'd like video blogging to encompass everything we like about text blogging, with the technical quality we expect from broadcast media.

Current video blogs

Current video blogs are essentially text blogs with externally linked videos for each entry. Though the fragments of video content form a cohesive diary, they're always introduced and navigated to via text. Although this is an understandable consequence of existing blogging technology, it's a far cry from the user experience we expect from broadcast media.

Typical examples of current video blogs are Charlene's video blog at http://crule.typepad.com/, where video is included through hyperlinks, and Jay Dedman's video blog at http://momentshowing.typepad.com/momentshowing/, which has inline video files. Quite a famous video blog is Peter Jackson's about the shooting of King Kong at http://www.kongisking.net/.

A video blog that makes use of more sophisticated features of Apple's QuickTime framework

More Information

Video bloggers experiment with bleeding-edge technologies to create next-generation blogs that stay in the world of video without requiring the world of text to glue them together. Some of the less commonly known standards and technologies mentioned in this article are described here in more detail.

RSS

RSS is a variably used abbreviation for Rich Site Summary, RDF site summary, or Really Simple Syndication. RSS is an XML-based metadata description and syndication format. For several versions of RSS, their details, and history, see http://www.rss-specifications.com/rss-specifications.htm.

Atom

Atom defines a feed format for representing, and a protocol for editing, Web resources such as Web logs, online journals, and similar content. The format feed representation is analogous to RSS but more comprehensive, and the editing protocol is novel. Atom isn't standardized yet: A Working Group was formed at the Internet Engineering Task Force (IETF) in August 2004, and several Internet drafts have been published (http://www.ietf.org/html.charters/atompub-charter.html).

BitTorrent

BitTorrent is a protocol designed for transferring files in a peer-to-peer network. Portions of files are exchanged directly between peers in a BitTorrent network, and a central server, called the tracker, coordinates the action of all peers. As BitTorrent works with parts of files, not just complete files, it's better suited to exchange videos and clips of videos than other common peer-to-peer protocols (http://bittorrent.com/).

CMML

The Continuous Media Markup Language is a simple language for authoring annotation tracks for time-continuous data. It's similar in spirit to HTML and provides timed metadata and outgoing hyperlinks. It can be used to create both free-text annotations and structured metadata for video clips, and to attach representative images and hyperlinks to such clips. CMML is currently being standardized through the IETF (http://www.annodex.net/specifications.html).

Annodex

Annodex is a multimedia exchange format combining a number of technologies including CMML and Ogg (an encapsulation format for multitrack multimedia, specified by the IETF in RFC 3533), allowing both incoming and outgoing hyperlinks. Any point of interest in Annodex media can be linked to from elsewhere, such that the content is delivered immediately from the desired offset. Annodex is currently being standardized through the IETF, and Internet drafts are being developed at http://www.annodex.net/specifications.html.

Figure 1. Screenshot of Adrian Miles' VOG 2.0.

is Adrian Miles' VOG 2.0 site at http://hypertext.rmit.edu.au/vog/, as Figure 1 shows. Miles teaches the theory and practice of hypermedia and interactive video at RMIT University, Australia, and uses his blog to demonstrate some of the ideas.[2] His video blog differs because he includes timed hyperlinks to other Web resources inside his videos, and has postproduced speech tracks with timed transcriptions of his speech inside the QuickTime video files. It's an exceptional video blog, requiring skills and tools not usually available to the more common blogger.

Although those pioneering the creation of video blogs are finding ample room for expression, and they obviously enjoy pushing the limits of current technology, users cite a few problems with current video blogs (see, for example, the Slashdot discussion on 9 November 2004 at http://slashdot.org/article.pl?sid=04/11/09/1446224&tid=149&tid=1):

- there's no way to add comments in video form;
- video items can't be easily found via search engines;
- video items can't be aggregated easily;
- interesting clips can't be viewed on their own; instead the video must be played back in its entirety; and
- as with most multimedia on the Web, sufficiently high bandwidth for reasonable quality video isn't widely available.

Essentially, people struggle with two classes of problems: first, the need for navigation and information access in video blogs; second, the distribution of blogs.

The current workaround to the second problem involves specifying the location of the video attached to a blog entry in an RSS "enclosure," which marks it for later downloading. The actual video is then distributed via a common peer-to-peer network, usually BitTorrent, to reduce congestion. You must then configure your software to subscribe to such channels well in advance, often downloading the video attachments overnight. Although this facilitates efficient use of current network resources, it's hardly the "on-demand" experience that Web surfers are used to for text. Moreover, it limits your ability to randomly flick between channels.

With more high-bandwidth networks being rolled out to homes, the bandwidth problem will eventually solve itself. Until such time, low resolution and highly compressed video, distributed via peer-to-peer networks, provides a fairly acceptable compromise, letting enthusiasts express themselves and experiment with production techniques despite existing limits on networking resources.

The problem of navigation in and access to video blogs is more complex, because varying processes such as syndication, search, commentary, hyperlinking, and aggregation must be considered. Let's consider these issues in detail.

Making use of video blogs

With the current crop of video blogs, only the purely textual introductions to video items can be aggregated and searched on. To view such items, we have to navigate the textual introductions, which is somewhat more cumbersome than the purely video environment of news updates we're accustomed to on television.

It would indeed be a more powerful means of video blogging if we could stay completely in the video domain. We'd then be able to interact with video blogs through hyperlinks inside blog entries, add comments that are video clips themselves, search for video clips directly, and aggregate video blogs of different video blog channels by creating a video summary. How could that be done?

Video blog syndication

Syndication in text blogging involves the publishing of machine-readable items containing metadata (such as the author's name and date of publication) and content reformulated in a subset of HTML. To similarly syndicate video, video will need a standard metadata markup language. Potential starting points might be an existing blogging language like RSS or Atom, or a multimedia metadata syntax like the Continuous Media Markup Language (CMML)[3] or MPEG-21.[4] Because overlaps exist between these approaches, a combination might be the best solution. Such syntax would need to be used in conjunction with open-standard video file formats to ensure that syndicated content can be viewed freely on a wide range of systems. Such standard formats include Apple's QuickTime (http://www.apple.com/quicktime/), Ogg Theora (http://www.theora.org/), or MPEG-4 (http://www.digitalpreservation.gov/formats/fdd/fdd000037.shtml).

Video blog search

Given the still poor capabilities of most video formats concerning metadata handling, it's not surprising that multimedia search on the Web, and thus among video blogs, remains rather hit-and-miss. The aim is to make video blogs as easily searchable by Web search engines as normal Web pages.

Web search engines increasingly support scanning of RSS and Atom feeds, which allows more tightly coupled searches of actual blog entries. In fact, the original blogging company Blogger (http://blogger.com/) was acquired by Google to bring about such developments.

Although a multimedia syndication language would be just as amenable to scanning and indexing by search engines, it would unfortunately be able to index only the metadata referring to a video blog entry, not the actual video content itself. Ideally, the search engine should index the video content itself, by scanning for embedded transcriptions and timed metadata or even by performing automated analysis of the video content directly.

Video blog comments

That readers can add comments to blog entries is very popular, allowing friendly advice and spontaneous discussions to take place in remote corners of the Web. It's easy to imagine how lively these discussions would be in video format—if any viewer could easily provide feedback for others to watch and themselves respond to.

The technology for such video forums is already being explored in various areas of the telecommunications industry, but usually in the context of developing vendor-specific applications for limited numbers of users. In such environments, users are simply able to "point-and-shoot" to reply to commentary posted by others. Applying these concepts to the distributed forums of the Internet poses new challenges, but has the potential to create extremely vibrant communications between people in all places and time zones, wherever interesting topics of discussion arise.

Video blog hyperlinks

Text blog entries and their comments often concern real-world events as well as digital material, such as Web sites, images, and even other people's blog entries. Bloggers can do this by simply hyperlinking to the URL of the Web site under discussion or to the *permalink* of a blog entry. A permalink is a URL to a blog entry in the blog's archives, which will remain valid after the entry is no longer listed on the blog's front page. For video blogs to allow this behavior requires that the video material itself contain outgoing hyperlinks in a standard manner, such as that of Annodex.

Embedded hyperlinks allow multimedia to become a fully integrated part of the Web. Content that can be made to reference other content can be used for criticism and endorsement, as well as to let authors create structured Webs of video. For such entries to fully reference each other, the video format and retrieval method must support both outgoing and incoming hyperlinks. Outgoing hyperlinks enable a video to link "out" to a Web page or another video; as the viewer watches some commentary, the links can be offered for viewing the material under discussion. This is crudely possible in Apple's QuickTime and RealNetworks' RealMedia (https://common.helixcommunity.org/nonav/2003/HCS_SDK_r5/htmfiles/rmff.htm), and available in an HTML-compatible manner in the open standard, Annodex.

Incoming hyperlinks are a retrieval method that lets the portion of a video under discussion be directly referenced by a specifically constructed URL pointed to from another video, Web page, or email document. A specification for a URL syntax to describe temporal offsets is available from the Internet Engineering Task Force

Blogging Resources

Here are additional resources from the emerging video blogging community:

- First-ever video blogger conference, VLoggerCon 2005, New York City, held 22 January 2005: http://vloggercon.blogspot.com/
- Information on how to get started on video blogging: http://videoblogging.info/
- Video blogging Wiki: http://www.me-tv.org/
- Portal to different video blogs: http://www.vidblogs.com/

(IETF)[5] for Annodex. MPEG-21[4] aims to produce an even more detailed method of addressing visual subregions and structured elements.

Aggregation of video blogs

Automated aggregation lets updates from widely distributed Web sites be viewed as an integrated whole. Aggregation of video blogs allows the automated creation of continuous feeds from distributed sources, related to a particular topic. A centralized aggregation of video feeds may be similar in form to a television show, in that it provides a continuous set of related content.

To easily produce seamless and interesting aggregations of collected content requires techniques of automated video editing and video summarization. This process is entirely customizable if we remember that one of blogging's fundamental principles is separating content from presentation. An aggregator of text blogs produces a coherent summarization of updates from many Web sites without visually appearing as a crudely patched-together amalgamation of those sites. It retrieves only the textual content of each entry, and applies its own layout and visual styles.

A video aggregator, retrieving only raw footage and metadata related to each entry, is free to apply its own consistent editing styles, including transitions, titling overlays, and temporal edits for brevity. Researchers have extensively studied such techniques in the field of video summarization, and video blogging provides an avenue for exploration outside of the usual domains of professionally produced content. It also poses challenges related to acquiring and combining material from diverse sources to account for variations in production quality and encoding methods.

Visual future

Developing the standards and technologies for video blogging requires a combination of approaches from various areas including media representation, information retrieval, multimedia content analysis, and video summarization. Like the development of the Web and text blogging before, video blogging will only come about through open development and collaboration between engineers and researchers from diverse fields. Most strikingly, it will be fueled by the passion and enthusiasm of those creating content—those who go to the trouble of recording their lives and opinions within the fledgling medium, shaping it as a lively and useful resource for generations of Internet users to come. We should do our best to provide technology that helps make this possible. **MM**

References

1. B.A. Nardi, D.J. Schiano, and M. Gumbrecht, "Blogging as Social Activity, or, Would You Let 900 Million People Read Your Diary?" *Proc. ACM Conf. Computer-Supported Cooperative Work*, ACM Press, 2004, pp. 222-231.
2. A. Miles, *Softvideography*, online essay; http://hypertext.rmit.edu.au/essays/softvideography.pdf.
3. S. Pfeiffer, C. Parker, and A. Pang, "The Continuous Media Markup Language (CMML), Version 3.0" (work in progress); http://www.annodex.net/TR/draft-pfeiffer-cmml-02.txt.
4. I. Burnett et al., "MPEG-21: Goals and Achievements," *IEEE MultiMedia*, vol. 10, no. 4, 2003, pp. 60-70.
5. S. Pfeiffer, C. Parker, and A. Pang, "Specifying Time Intervals in URI Queries and Fragments of Time-Based Web Resources" (work in progress); http://www.annodex.net/TR/draft-pfeiffer-temporal-fragments-03.txt.

Readers may contact Conrad Parker and Silvia Pfeiffer at Conrad.Parker@csiro.au and Silvia.Pfeiffer@csiro.au.

Readers may contact Media Impact editor Frank Nack at frank.nack@cwi.nl.