Weblog-style Video Annotation and Syndication

Daisuke Yamamoto
Graduate School of
Information Science, Nagoya Univ.
yamamoto@nagao.nuie.
nagoya-u.ac.jp

Shigeki Ohira
EcoTopia Science Institute,
Nagoya Univ.
ohira@nagao.nuie.
nagoya-u.ac.jp

Katashi Nagao
Center for
Information Media Studies,
Nagoya Univ.
nagao@nuie.nagoya-u.ac.jp

Abstract

In this paper, we present a method to extend the mechanisms employed for Weblogs in order to make video contents accessible to Weblog communities. We also present a mechanism that generates Weblog entries from the browsing history of the video contents. These entries help Weblog users to reach the related video contents. By extracting knowledge on the video contents from communication among members of Weblog communities and syndicating the knowledge in the form of RDF, we can use the mechanism for annotation-based applications, such as a semantic video retrieval and summarization system.

1 Introduction

We present a mechanism to distribute and advertise video contents and form communities centered on those contents effectively on the Internet. Conventional video contents distribution systems handle only the signal processing problem of stably distributing high-quality video streams, or the problem of defending the profit of contents holders, using methods such as encryption and accounting systems for multimedia contents. These systems do not consider advertisements or the spread of contents. To solve these shortcomings, we propose a video contents distribution system that suits Weblogs [9], which have become popular on the Internet in recent years. Weblog communities are formed by the mechanisms of Trackback [4] and RSS [1], and it is said that their influence is enormous. There is a wide variety of Weblog articles that introduce profitable information and commodities to others using photographs and links to related pages, and these articles are recognized as an effective means of advertising commodities. In addition, by applying the Weblog mechanism to video contents, this system can activate a community that centers on video contents, and by relating meta-information such as Trackback or comments to video contents, we can extract knowledge as video annotations.

2 Weblogizing Online Video Contents

We propose a mechanism that facilitates distribution of video contents and activates communication among members of their related communities by replacing Weblog entries with video contents (and their scenes). Some Weblog sites that treat video contents already exist, and the method of displaying animation on these Weblogs is to use an HTML Video plug-in. There are some problems associated with this, however, such as the fact that video files are simply too large to download many at a time, and we cannot browse video contents without watching the entire content to confirm what is on the video. Furthermore, we cannot treat the video contents scene-by-scene. To solve these problems, we separate the video contents from Weblog articles, and develop a mechanism for selecting arbitrary scenes of the video contents and linking them with appropriate Weblog articles.

2.1 Definition of Weblogs

Weblogs feature a mechanism to efficiently distribute frequently updated articles, called "entries," on the Web. Although there are various opinions about what constitutes an accurate definition of a Weblog, in this research we target Weblogs that comply with the following four requirements from a technical perspective.

1. Permalink [2]: Each entry can be specified by a unique and permanent URI.

2. Annotations: Users can add comments to each entry.

3. Trackback [4]: When a user writes entry A referring to entry B, the user can generate the reverse-link from entry B to entry A explicitly. As a result, we can form an interactive link between the entries.

4. XML Feed: The system can distribute updated Weblog information, such as RSS [1] and Atom [3].

Linking each entry via these mechanisms, we can activate a community that centers on Weblog entries.

These mechanisms have already been adopted and standardized in many of main Weblog services, such as MovableType (http://www.movabletype.org/) or Blogger (http://www.blogger.com/), and we can also use these mechanisms between different Weblog services.

2.2 Issues in Weblogizing Online Video Contents

To ensure that video contents can be included in a Weblog network, this system requires functions similar to existing Weblogs. In short, it is necessary to install the mechanisms of Permalink, Trackback, Annotation, and XML Feed. We call the application of the Weblog mechanisms to the contents "weblogizing," and we call the weblogized video contents "Videoblogs."

In contrast to an ordinary Weblog entry, the meaning and the situation of video contents might be different in each scene. Therefore, we have to consider the following conditions to weblogize the video contents. First, it should be possible to apply Trackback and Annotation, etc. to not only the entire contents, but also to arbitrary scenes. Second, when the user quotes a particular scene in the video, it becomes necessary to quote video scenes in a form by which the content can be understood. Finally, compatibility with existing Weblogs must be high.

In order to weblogize video contents, we have to equip our system with the four requirements and the three conditions listed above. If a unique ID is assigned to the contents, requirement 1 can be easily achieved. In this section, we propose the implementation of requirements 2, 3, and 4.

2.3 Implementation of Annotation for Video Contents

The first necessary mechanism for a Videoblog is one with which users can easily make annotations to the video contents. For this we employ iVAS (intelligent Video Annotation Server) [13], which is from our previous research.

With iVAS the user makes annotations to the online video contents while watching it. We proposed three types of annotation in iVAS, using the electronic bulletin-board interface as shown in Figure 1. That system includes: text annotation, which is the mechanism whereby users can make comments about arbitrary scenes in the video content; impression annotation, which is a mechanism where users can associate video content with subjective impressions on the content by clicking subjective buttons, such as a gniceh button and a gbooh button; and evaluation annotation, with which users can evaluate each text annotation by clicking the O-X buttons.

We can communicate with other users about video contents by using these mechanisms, and we have developed a video retrieval system that employs the annotations of video contents.
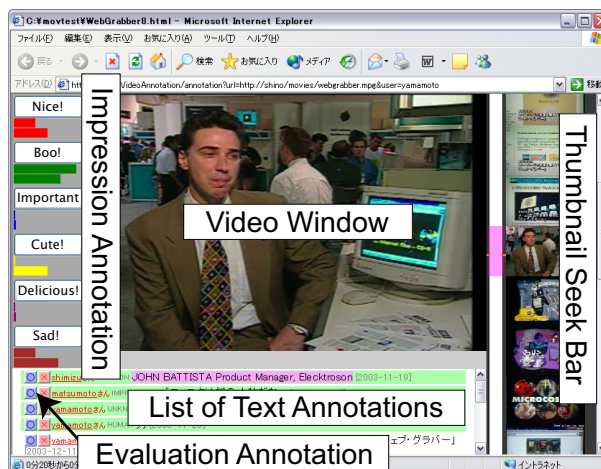

Figure 1. Example of Videoblog Annotations

2.4 Implementation of Trackback

Trackback is the second mechanism we apply for a Videoblog. Trackback is a structure that automatically links reference heads and the reference origin of entries. Therefore, unlike mutual links in Websites, we can construct highly detailed link entries. Furthermore, since Trackback links by human subjectivity, it provides stronger links with consideration given to an article's content. The form of Trackback is standardized [4], and it transmits the name of the site, the title of the entry, a summary of the entry, and URI information on the entry to the target entry's Trackback URI. This is called "Trackback ping." A Trackback URI is the usually in following form, and the user specifies the target entry by inputting this URI.

```
http://[server address]/tb/[entryID]
```

The structure of Videoblog's trackback is the same as that for a Weblog, and is organized by sending a Trackback ping. To transmit Trackback to an arbitrary scene of the video contents, we have to transmit the contents ID, the start time and the end time of the scene. Because this is strictly standardized excluding Trackback URI, in order to bury this information under a Trackback URI, we enhance it into the following form.

```
http://[ivas server address]/tb/
    [content ID]/[begin time]-[end time]
```

According to this form, a Videoblog can accept a Trackback without changing the existing Weblog.

2.5 Video Annotation Syndication based on RDF

We apply the mechanism of XML Feed to the video contents as the third mechanism for Videoblogs. We syndicate not only the updated information of contents, but also information such as annotations and trackback in the form of RDF. In concrete terms, by specifying the video contents' URI and the annotation ID to rdf:about attributes, we can express them in the following forms.
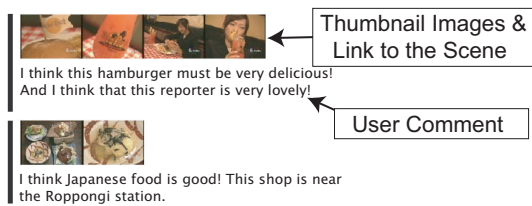
Figure 2. Quotation of Video Contents with Thumbnail Images

```
<annotation rdf:about=
  "http://[ivas server address]/content/
  [content ID]#[annotation ID]">
    ...Annotation XML Data...
</annotation>
```

As a result, we can collect annotations distributed to two or more Videoblog servers, and we can use these annotations for various applications. Although general XML Feed syndicates only the latest updated information, in this system, by syndicating meta information on all contents, we can promote its use for video annotations. Consequently, we can introduce the concept of the Semantic Web into the multimedia contents.

3   Content Syndication based on Videoblog

Here we propose a supporting system with which the user can easily quote arbitrary scenes of the video contents and then produce an article concerning those contents. As a result, the user can advertise contents in his or her own Weblog, and can activate the circulation of those contents. In addition, by associating Videoblogs with Weblogs, we can acquire comments in Weblog entries as text annotation of the video contents. In this research, the form of the quotation in the video contents is composed of a link to the scene and a comment on the scene, along with a thumbnail image. An example can be seen in Fig 2.

3.1   Semi-Automatic Generation of Video Articles from a Videoblog

We assume that users watch video contents by using iVAS. Because annotations are made in real time in iVAS while the user watches the video contents, we cannot necessarily acquire sufficient annotations. However, if the user likes the contents, he or she will probably have the desire to write an article about it sooner or later. If the user writes the article after viewing the contents, it is important that scenes are marked and annotated, perhaps with an impression annotation. We consider iVAS the best tool for that.

The system then extracts the target scene of the video article from the annotation produced by a user when he or she used iVAS in real time. The system recommends scenes that might be of interest to the user, and offers a template on which to write a Weblog entry for the video contents. Moreover, the system automatically associates a user's Weblog
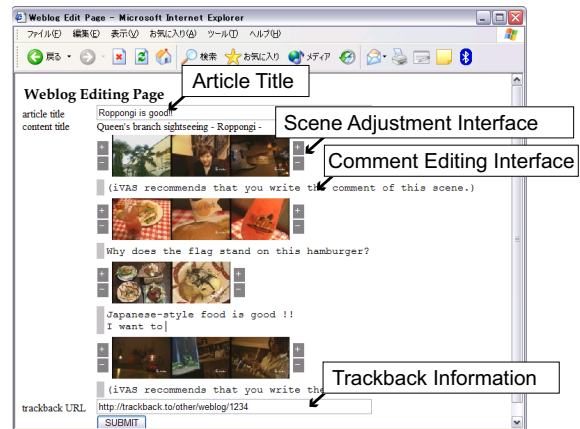

Figure 3. Weblog Editing Page for Video Article

entry with scenes of interest to the user by linking trackbacks. The user then describes the entry for these scenes based on this recommendation result.

The mechanism for recommending target scenes of the video article is roughly as follows. First, we introduce an interest level that indicates how much interest the user has in the scene: the more users give annotations to the scene, the higher the interest level rises. Concretely, the interest level is high for scenes that have had text annotations or many impression annotations made about them. Furthermore, the system recommends and ranks scenes that generate high interest.

3.2   Updating Video Articles and Annotations

Users can correct video articles generated semi-automatically just as they would for a conventional Weblog entry. This correction result is reflected in iVAS as text annotations, and is used to improve the accuracy of applications such as video retrieval systems. We can correct the following two items. The first is correcting the range of the targeted annotation scene: we can correct the start and the end points of the scene. The second is correcting text. Figure 3 shows an example of the interface. The system records the corrected result in the annotation database as an annotation of the superscription attribute. As a result, we can acquire highly reliable annotations corrected manually.

4   Applications based on Video Annotation

By considering the information that this system obtains as video annotations, we can extract knowledge from video contents and employ it for various services.

To date, there have been several studies focusing on video annotation. For instance, Informedia [12] automatically makes annotations based on the machine processing of both voice- and image-recognition technology. technology. Because analytical techniques and analytical accuracy differ
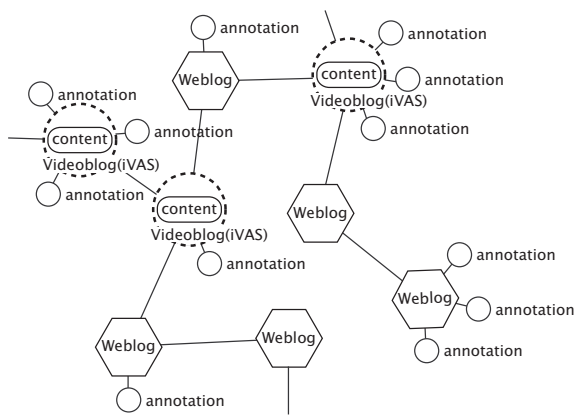
Figure 4. Weblog Network

depending on the type of contents, it is difficult to apply this mechanism to general contents. By manually supplementing and correcting annotations, however, some researchers have been able to achieve detailed annotations [6, 11, 10]. Unfortunately, producing detailed annotations consumes a high human cost.

On the other hand, we focus on the point that a lot of users watch the video contents, and automatically extract annotations from natural intelligent operations such as bulletin-board communications and Weblog descriptions among others. Because a variety of annotations is acquired from many users, the annotation cost is low. Moreover, we can acquire comments and users' impressions, which are difficult to obtain through image and voice recognition. There are some related studies on electronic bulletin board-style video annotation, such as one of our previous studies [13] and the research of Brgeron [5]. These investigations, however, discuss only stand-alone video contents. By linking the video contents and those of Weblogs by trackbacks, our proposed system can form a novel Weblog network (Fig 4). By using the semantic relation between an article on a reference origin and a video content on a reference head, we can extract many varieties of annotations. For example, the content of an article referring to specific scenes of video contents can be indirectly treated as a comment annotation to the video. Moreover, two or more contents referred from similar Weblog articles feature some similarity, just like the similarity shared between articles.

Using these annotations, we can leverage semantic retrieval and summarization [10, 13] based on video annotations. Also, analytical research on Weblog communities [8] and trend analysis [7] are applicable to video contents by employing the features of Weblogs.

5 Conclusion

The proposed mechanism can tightly connect video contents, Weblogs, and users. The mechanisms explained in this work bring about the advantage that connections among video contents, Videoblogs, and ordinary Weblogs facilitate the distribution of contents in Weblog communities. In addition, we can use the system for applications based on annotation by extracting the video annotations from the contents. As one of our future works, we are planning to open this system to the public and conduct experiments in a real environment.

References

[1] RDF site summary (RSS) 1.0. http://web.resource.org/rss/1.0/spec, 2001.

[2] E. Aimeur, G. Brassard, and S. Paquet. Using personal knowledge publishing to facilitate sharing across communities. In *Proceedings of the Twelfth International World Wide Web Conference (WWW2003)*, 2003.

[3] AtomEnabled Alliance. *AtomEnabled*. http://www.atomenabled.org/, 2004.

[4] Benjamin and M. Trott. *mttrackback - Track-Back Technical Specifications*. movabletype.org, http://www.movabletype.org/docs/mttrackback.html, 2002.

[5] D. Brgeron, A. Gupta, J. Grudin, and E. Sanocki. Annotations for streaming video on the web: System design and usage studies. In *Proceeding of the Eighth International World Wide Web Conference*, pp. 61–75, 1999.

[6] M. Davis. An iconic visual language for video annotation. In *Proceedings of the IEEE Symposium on Visual Language*, pp. 196–202, 1993.

[7] N. S. Glance, M. Hurst, and T. Tomokiyo. Blogpulse: Automated trend discovery for weblogs. In *Proceedings of the Thirteenth International World Wide Web Conference(WWW2004)*, 2004.

[8] R. Kumar, J. Novak, P. Raghavan, and A. Tomkins. On the bursty evolution of blogspace. In *Proceedings of the Twelfth International World Wide Web Conference(WWW2003)*, 2003.

[9] C. Lindahl and E. Blount. Weblogs: simplifying Web publishing. *IEEE Computer*, 36(11):114–116, 2003.

[10] K. Nagao, S. Ohira, and M. Yoneoka. Annotation-based multimedia summarization and translation. In *Proceedings of the Nineteenth International Conference on Computational Linguistics (COLING-02)*, pp. 702–708, 2002.

[11] J. R. Smith and B. Lugeon. A visual annotation tool for multimedia content description. In *Proceedings of the SPIE Photonics East, Internet Multimedia Management Systems*, 2000.

[12] H. D. Wactlar, T. Kanade, M. A. Smith, and S. M. Stevens. Intelligent access to digital video: Informedia project. *IEEE Computer*, 29(5):140–151, 1996.

[13] D. Yamamoto and K. Nagao. iVAS: Web-based video annotation system and its applications. In *Proceedings of the 3rd International Semantic Web Conference, Demonstration Session*, 2004.