Efficient Meta-information Annotation and View-dependent Representation System for 3D Objects on the Web

Yukiko Kawai, Shogo Tazawa
Faculty of Computer Science
Kyoto Sangyo University

Ryo Furukawa
Faculty of Information Sciences
Hiroshima City University

Hiroshi Kawasaki
Faculty of Engineering
Saitama University

Abstract

So far, there have been many techniques and applications developed for handling 3D contents efficiently on the Web. However, there is currently no application that successfully makes 3D contents pervasive on the Web. One of the main reasons for this can be considered be the fact that there are only a few pages on the Web where a 3D object is sufficiently annotated with meta-information; meta-information is fundamental for the Web. In this paper, (1) a semi-automatic annotation module which supports users in creating 3D contents with rich meta-information by collecting and filtering meta-data on the Web using 3D information; and (2) a view-dependent information representation module which shows information corresponding to the viewing direction of 3D objects are presented. In those modules, view-dependent information representation module is most important in our system, because such system is highly suitable to show meta-information to users, however, not proposed yet. We implemented the proposed system and conducted experiments which show that it is possible to realize semi-automatic annotation of meta-data on 3D objects and efficiently present view-dependent meta-information to users.

1 Introduction

Recently, 3D content is widely used in the fields of movies, games, electronic museums and so on. On the other hand, in terms of 3D content on the Web, a lot of systems have been proposed and developed from a simple 3D object up to a large-scale virtual environment [1]. However, Web 3D systems have not been used pervasively yet. The main reasons for this can be considered to be as follows: (1) it is not easy to create a 3D object; and (2) there is not enough valuable information added to 3D objects. For the 1st issue, we have developed a system which can construct a 3D model from a real object easily by using an ordinary projector and a camera [2]. By using the system, users can capture 3D objects without difficulty and easily publish them on the Web. For the 2nd issue, a lot of research into visualization techniques to represent 3D contents have been proposed and developed[3, 4]. However, in those systems, effective representations of web pages are the main issue and an annotation method for 3D objects to enrich the meta-information has not been discussed. Based on this fact, we propose an efficient annotation and representation method for 3D objects on the Web. The proposed system consists of the following methods; a data collection and filtering method for the Web that is specialized for 3D objects, a view-dependent meta-data representation method, and a personalized adaptation method based on user preferences and browsing history. The uniqueness of our system is in the view-dependent meta-information representation method, because such system is highly suitable to show meta-information to users efficiently, however, not proposed yet.

2 Related work

Research on 3D contents for the Web has a long history and can be categorized into two types: (1) the research about virtual environments which focuses on multi-user interaction and collaboration [1, 5], and (2) the research about highly-interactive Web contents for unspecified users [4, 6]. In the former type, research regarding virtual reality for multi-users to collaborate on the Internet have been conducted [5]. On the other hand, although the latter type is also intensively researched, such as 3D data compression methods for transferring via the Internet and search engines for 3D objects [6], 3D contents on the Web are not popular yet. One of the reason is that the added value on the 3D contents is insignificant. Since one of the most important factors of the Web is meta-information, and annotation of meta-information on a 3D object is still a difficult task, the solution is strongly required. Moreover, an efficient representation of information about a 3D object based on user preferences and view-dependency is also important. Croquet [4] can represent meta-information of 3D objects

978-1-4244-2175-6/08/$25.00 ©2008 IEEE

on the Web and has an annotation system of meta-data on 3D objects, however, it is still a laborious task to annotate enough meta-data on 3D objects. In addition, as far as we know, there is no research conducted to realize view-dependent meta-information representation method, which is hightly suitable for interactive Web 3D systems.

3 System outline and details

The system consists of the following three functions as shown in Fig. 1.

1. Semi-automatic annotation method of meta-data on 3D objects: This function is processed on the server system; collecting information from the Web using 3D information, analyzing them and automatically annotating those meta-data on 3D objects.
2. Representation of meta-data dependent on view-direction: This function works on the client system using plug-in system on a web browser.
3. Personalized adaptation of meta-data based on user preferences and history: This function works on the server system using a *cookie* on the client PC and changes meta-information dynamically.

3.1 Semi-automatic annotation of meta-data on 3D objects

As shown in the upper right of the Fig.1, this function helps users to annotate meta-data on measured 3D objects by using collected and filtered information from the Web using 3D information. A 3D content creator can build a database for 3D objects efficiently by just submitting a name of the object (for example *digital camera*) to the system, and then, selecting several terms which are recommended by the system (for example, *lens or flash*).

Details of the recommendation process are as follows. First, a creator submits an objects name to the server system, then the system simply re-submits the term to a search engine (*e.g.* google or yahoo) with several additional keywords such as "*specification*". Next, by using the top ranked web pages, the server calculates the term appearance frequency by the equation $t_f = log$(occurrence number $+$ 1)/log(total number of word occurrence), and then ranks terms in order of t_f.

Then, the top-ranked terms are presented to the content creator and the selected terms are re-submitted to a search engine with additional spatial terms; in our implementation, we divide an entire space into six, such as front, back, top, bottom, left and right. Thus a database for a 3D object has the table of six directions and the system stores the search results in each table of the database.

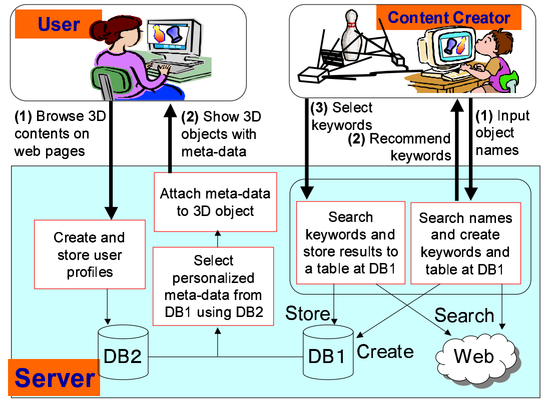


Figure 1. System outline.

3.2 Representation of meta-data dependent on view-direction

To show a meta-data to a user effectively, different data representation dependent on user's view-direction may be a promising solution. This is a simple implementation of the fact that, even if an object is the same, meta-data is usually different dependent on view direction. For example, the information of Mt.Fuji seen from the foot is, *e.g.* "how to climb Mt. Fuji" or "Best season to visit", and it is totally different from the information of Mt. Fuji seen as a bird's-eye view, *e.g.* "height" or "vegetation".

Another advantage of our view-dependent representation is that it makes the annotation process simple. Usually an annotation is applied on either each polygon or a set of polygons [7]. With this case, if a number of polygons increases, an algorithm and a data structure for annotation become complicated. In recent years, a number of polygons used has increased rapidly because 3D scanners are commonly used, therefore, an annotation for 3D objects becomes a difficult task and an efficient method is required. In our approach, since an annotation is not applied on each polygons, but only on a direction of a 3D object, an increase of polygons does not affect the annotation process at all. If precise annotations are required, it can be easily realized by simply dividing a 3D object into an appropriate size.

For an actual implementation, a direction angle of a view point from the origin of coordinate which is defined for each 3D objects is calculated, and then, meta-information designated on the angle are presented. As a result, unlike the usual 3D contents on the Web, it is possible to represent view-dependent information which makes the best use of characteristics of a free-viewpoint of 3D objects.

3.3 Personalized adaptation of meta-data based on user preferences and history

As shown in Fig.1, we also implemented a personalized adaptation system to increase the efficiency of the system. With the system, when a user decides a view direction of a 3D object, most adapted meta-information to individual users are effectively presented.

Such personalized adaptive meta-information (*i.e.* keywords) are selected by the following steps. Let $V_{ave} = (\tilde{v}_1, \tilde{v}_2, \tilde{v}_3, \tilde{v}_4, \tilde{v}_5)$ be the five dimensional vector whose components are the five greatest values of t_f for feature words which are extracted from all the web pages where the user has browsed. Let $Vp_i = (v_{1_i}, v_{2_i}, v_{3_i}, v_{4_i}, v_{5_i})$ be the vector whose components are t_f of the same words of V_{ave} in a given web page p_i. Then, a similarity of the two vectors is calculated as an angle between them using an inner product and actual equation is as follows:

$$sim(V_{ave}, V_{p_i}) = \frac{\tilde{v}_1 v_{1_i} + \cdots + \tilde{v}_5 v_{5_i}}{\sqrt{\tilde{v}_1^2 + \cdots + \tilde{v}_5^2}\sqrt{v_{1_i}^2 + \cdots + v_{5_i}^2}} \quad (1)$$

The web page p_i which has the largest value of $sim(V_{ave}, V_{p_i})$ is recommended.

If the user actually browses the page p_i, V_{ave} will be updated by averaging all the components of V_{ave} and V_{p_i}. Therefore, after preferred web pages are browsed several times, feature words corresponding to user preferences will be top-ranked with order of t_f, and the web page where such feature words appear frequently will be recommended.

4 Experiment results

4.1 Semi-automatic annotation of meta-data on 3D objects

The object used in this experiment is a digital camera. We added the term "specifications" to the keywords to search for the camera from the Web. In our implementation, we used Yahoo!API as the web search engine.

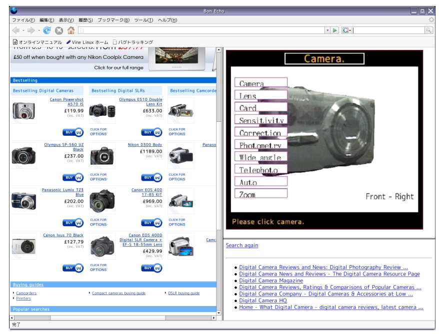
Then, we checked all the top 50 recommended keywords to see whether they were feasible as keywords for the digital camera or not. Such feasibility testing was subjectively conducted by four individuals. As the results, feasibility was 66% for this case and we can confirm that the system can effectively assist the user's task of annotating 3D objects with meta-data.

4.2 Representation of meta-data dependent on view-direction

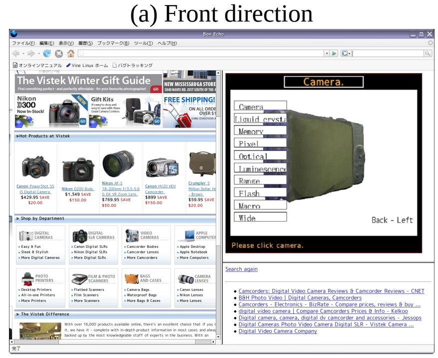
4.2.1 3D content of a scanned digital camera

Figure 2 shows a snapshot of our GUI system implemented in Adobe Flash. In this system, 3D objects are shown in the upper right frame and the keywords recommended by the system according to user preference are also super-imposed in the same frame. In the bottom right frame, collected and stored meta-information related to keywords are presented and the left frame shows the hyperlinked web page of meta-information that is selected by a user.

As shown in the Fig. 2 (a) and (b), keywords for the 3D object are different according to user's viewpoints. Therefore, a user can rotate the 3D object to change the keywords until they find the preferred keyword. Once the preferred keyword is found, the user can check the related meta-information by just clicking the keywords with a mouse pointer; at the same time, such browsing histories are stored and personalized adaptation is conducted.



(a) Front direction



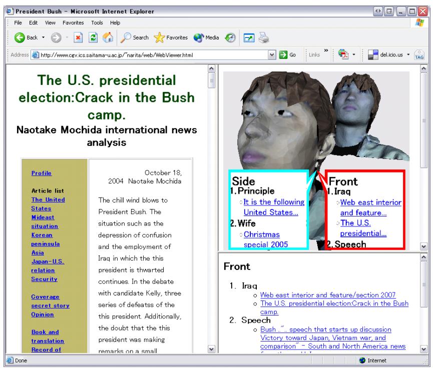
(b) Left direction

Figure 2. GUI interface with 3D and meta data in the top right area. View-dependent meta-information is presented.

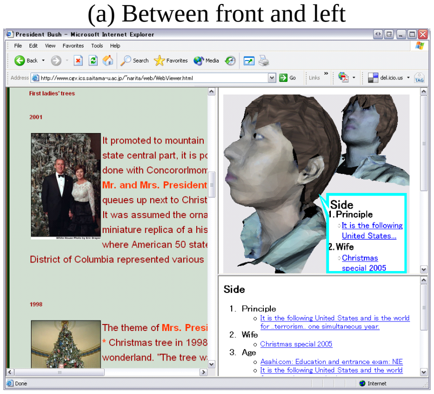
4.2.2 3D content of measured data of a human face

Next, we conducted an experiment using a human face. Here, we show another implementation of our GUI using Windows ActiveX. As shown in Fig. 3, when the user clicks the 3D object with a mouse pointer, a pop-up window with annotated keywords appears. As shown in Fig. 3 (a), both keywords for front and left appear at the same time for in-between directions. By clicking the keywords in the pop-up window, either a hyperlinked web page is shown on the left frame, or the view position jumps to another position dependent on the types of meta-information.

In this experiment, we supposed that the measured data was the face of President George W. Bush and we used "President George W. Bush" as the object's name. As shown in the Fig. 3 (a) and (b), since the spatial keywords such as "Front" or "Side" are frequently used as the semantic meaning for human, recommended keywords are quite different from the case of the digital camera; in this case, we can still confirm the effectiveness of the view-dependent data representation technique.



(a) Between front and left



(b) Left direction

Figure 3. Result of view dependent representation of meta-data.

4.3 Personalized adaptation of meta-data based on user preferences and history

Meta-data is selected and recommended to the a user based on the browsing histories of users as described in Sec. 3.3. In this experiment, browsing histories of keywords related to each view direction are used to perform a personalized adaptation. After browsing web pages ten times, the result of the precision ratio (the number of appropriate recommendation/the total number of recommendation) was 60% and the recall ratio (the number of appropriate recommendation/the total number of appropriate data) was 60%.

5 Conclusions

In this paper, we proposed a system which efficiently reduces users interaction for annotating meta-data on 3D objects by using Web and 3D information. We also proposed an efficient meta-data representation method which consists of a view-dependent data representation technique and a personalized adaptation technique. The experiments show that it is possible to reduce interactive operations necessary for annotating meta-data on 3D contents, and also confirm that it is possible to represent meta-information dependent on view direction.

6 Acknowledgment

This work was supported in part by SCOPE(Ministry of Internal Affairs and Communications, Japan).

References

[1] LINDEN RESEARCH, INC., "Second Life," .

[2] Ryo Furukawa and Hiroshi Kawasaki, "Uncalibrated multiple image stereo system with arbitrarily movable camera and projector for wide range scanning," in *IEEE Conf. 3DIM*, 2005, pp. 302–309.

[3] Akira Kotani, Satoshi Oyama, and Katsumi Tanaka, "A Spatial Interface for Aspect Browsing of Multiple Web Contents," *The Database Society of Japan (DBSJ) Letters*, vol. 4, no. 1, pp. 141–144, 2005.

[4] David A. Smith, Andreas Raab, David P. Reed, and Alan C. Kay, "Croquet: A Menagerie of New User Interfaces," in *Conference on Creating, Connecting and Collaborating through Computing(C5)*, 2004, pp. 4–11.

[5] C Cruz-Neira, J Leigh, C Barnes, S M Cohen, S Das, R Engelmann, R Hudson, M Papka, T Roy, L Siegel, C A Vasilakis, T A DeFanti, , and D J Sandin, "Scientists in wonderland: A report on visualization applications inthe cave virtual reality environment," in *Proceedings IEEE Symposium on Research Frontiers in Virtual Reality*, Oct. 1993, pp. 59–66.

[6] Thomas A. Funkhouser, Patrick Min, Michael M. Kazhdan, Joyce Chen, J. Alex Halderman, David P. Dobkin, and David Pokrass Jacobs, "A search engine for 3d models," *ACM Trans. Graph.*, vol. 22, no. 1, pp. 83–105, 2003.

[7] Yasuhide Okamoto, Jun Takamatsu, Masataka Kagesawa, Ken Okada, and Katsushi Ikeuchi, "3D database system of mercede church: The use of 3D models as an interface to information," in *11th International Conference on Virtual Systems and Multimedia (VSMM2005)*, 2005.