Proc. of IADIS International Conference Game and Entertainment Technologies 2011, Rome, Italy, pp. 43-50, Jul. 22-24, 2011.

FRAME SELECTION FOR AUTOMATIC COMIC GENERATION FROM MUSEUM PLAYLOG IN METAVERSE

Ruck Thawonmas
Intelligent Computer Entertainment Lab, Graduate School of Science & Engineering, Ritsumeikan University
1-1-1 Nojihigashi
Kusatsu, Shiga, 525-8577, Japan
ruck@ci.ritsumei.ac.jp
Tomonori Shuda
Intelligent Computer Entertainment Lab, Graduate School of Science & Engineering, Ritsumeikan University
1-1-1 Nojihigashi
Kusatsu, Shiga, 525-8577, Japan

ABSTRACT

The paper presents a system for generating comics from museum playlog in Metaverse. Metaverse is a 3D virtual world in which users can act freely, such as visiting museums or chatting with others, according to their own purposes. Compared with existing approaches for representing user experiences using snapshots or video clips, the comic approach can 1) allow users to grasp the whole story at a glance, 2) facilitate distinguishing of important frames, and 3) exploit varieties of comic writing techniques. In order to summarize user experience into comic's frames, detection of important experience, interesting exhibits in case of the museums, is an important task. In this paper, we propose a visiting-time based method for frame extraction. After describing each module of our system, we discuss a system evaluation where the effectiveness of the proposed frame extraction method is confirmed.

KEYWORDS

comic, metaverse, user experience, museum, playlog.

1. INTRODUCTION

Recently, virtual 3D space called metaverse has gained interests from educators and researchers as a promising educational and research platform. A representative of such is Second Life (SL) which is equipped with a function that enables users to build objects or architectures. Unlike online games, metaverse, in general, has no specific roles assigned to its users, and user experiences are arguably limitless including, for example, visiting to museums built by other users or chatting to other users. Other usages of metaverse include Machinima (Lowood 2006) that uses in-game avatars (user characters), objects, as well as architectures for filming and an experiment (Prendinger 2009) that aims at realization of an eco-friendly society.

A number of metaverse have functions that allow users to take snapshots of their experiences or record them into video clips. Such snapshots or video clips are used by the users to recall their memories or shown to other users via, for example, blogs or SNS sites. However, manually taking a snapshot each time imposes a burden to the user. And video clips require considerable time to grasp the story and distinguish between important and unimportant events therein. To circumvent these issues, we focus on the use of comics for representing user experiences.

In comic style, a given user experience can be summarized into a limited number of frames, provided that a proper summarization mechanism is used. Hence a whole story can be apprehended at one glance. In addition, a frame layout technique can be applied to emphasize important frames and underemphasize unimportant ones. In this paper, we focus on museum visiting, one of the major usages in metaverse, present a system for

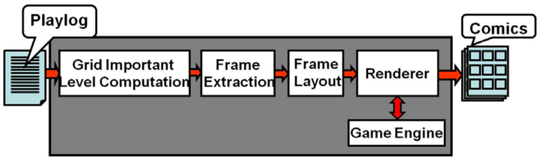

Figure 1. System architecture

Table 1. Example of partial playlog about user positions

loop_count=112 | tag=PlayerLoc | value={163.7, 152.3, 23.1 }

generating a comic based on the museum visiting experience of a user of interest, and propose a frame extraction method that uses the information on the user's visiting time to each visited exhibit.

2. SYSTEM

The authors previously proposed a comic generating system and related methods (Shuda and Thawonmas 2008; Thawonmas and Shuda 2008; Thawonmas and Oda 2010) for on-line games. The objectives therein are to summarize players' game experiences into comics so that the players can exchange their comics with others and recall their memories about the game. However, an online game, the research target in our previous system, is essentially different from metaverse, which is the research target of this paper.

In a typical online game, most of the players' actions, such as "attack a monster" or "open a treasure", are prepared and provided in advance by the game developer. Such actions are meaningful as important episodes and are thus worth memorizing. On the contrary, most of the actions in metaverse, such as "move" or "sit", have not much meaning. It is therefore difficult to extract important episodes from playlog. As far as the museum visit application is concerned, one might come up with a method for extracting comic frames that cover important exhibits based on weights assigned to them in advance. However, this method needs to reassign such weights each time the exhibition layout changes. In addition, it does not take into account the preference of a user of interest.

In a study on a real-world museum (Sparacino 2003), the more interest a user has on a given exhibit, the higher probability is that he or she will spend longer time on it. In this paper, dividing the region of a target museum into $m \times n$ grids, we hence consider that exhibits a user of interest finds interesting reside on grids where the user spends relatively longer time than the other grids. Based on this assumption, we propose a formula for deciding grids' weights in (1) and show the system architecture in Fig. 1. In this figure, "playlog" indicates a file that contains information on the user traces, the user actions, and the object positions in the museum. An example of partial playlog about user positions is shown in Table 1, where "loop count" shows the number of samplings to acquire the user information. In this example, the user position when loop count = 112 was at the coordinate (163.7, 152.3, 23.1).

In the following, we describe each of the four modules shown in Fig. 1.

1. At the first module, the important level of each grid (i, j), $I_k(i, j)$ ($i = 1,2,\ldots,m; j = 1,2,\ldots,n$), is calculated according to (1) based on user trace k residing in the playlog.

$$I_k(i, j) = \frac{t_k(i, j)}{\bar{t}_k(\cdot)}, \tag{1}$$

   where $t_k(i, j)$ is the visit time at grid (i, j) of trace k and $\bar{t}_k(\cdot)$ is a normalization factor being the average visit time to each visited grid in trace k.

2. At the frame extraction module, a frame $Frame(f)$ is generated for each visited grid, where f indicates the current frame number. Based on our experience, we decide not to generate a new frame $Frame(f)$ if its corresponding grid is the same as the grid of $Frame(f-1)$ or that of $Frame(f-2)$. This is done to prevent over-generating frames when the user moves along grid borders, causing a movement pattern that repeatedly goes back and forth between a pair of grids. The frame snapshot timing at a given grid could be decided by a complex method that detects changes in colors according to HSV histograms (Calic 2007). However, we use the formula in (2), which requires less computational resources, for this task.

$$t(f) = a * t_{in}(f) + (1-a) * t_{out}(f) , \quad (2)$$

where a is a parameter whose range is $[0,1]$, and, in our system evaluation, it was assigned a value of 0.4.The term $t_{in}(f)$ is the loop count at which the user entered the grid of $Frame(f)$, and $t_{out}(f)$ is the loop count at which the user left this grid. Our reason to set the snapshot timing slightly nearer to $t_{out}(f)$ than to $t_{in}(f)$ is based on our observation that, after entering a new grid, most users spent time in the first half to locate an exhibit, and, therefore, the aforementioned shooting time could increase the chance in extracting a frame where the user is viewing the exhibit. Each frame $Frame(f)$ contains its important level $I_k(i,j)$, which is used for selecting frames, and the snapshot timing $t(f)$, which is used later for rendering the selected frames. In general, the number of generated frames exceeds the number of frames desired by the user. As a result, a mechanism for selecting frames is needed. We implement it by first selecting the first generated frame and the last one because we consider them important for storytelling and then sorting all of the remaining frames in decreasing order of important level and selecting the first $F-2$ frames, where F is the number of desired frames.

3. At the frame layout module, each selected frame is assigned a page and the position therein. This decision is based on the information on the number of desired frames, the page margins, and the number of frames per row and column.
4. At the renderer module, the system scans the playlog from the beginning and renders the snapshot image of the user experience with the snapshot timing of each selected frame. Each rendered image is then placed on the specific position in the predetermined comic page according to the information stored in the corresponding frame. Finally, all comic pages are outputted as an image file.

3. EVALUATION

After implementing the proposed system, we conducted a system evaluation. The objective of this user evaluation is to examine if the proposed system can generate a comic that properly summarizes the user experience about the exhibits in which a user of interest was interested during his or her visit.

3.1 System implementation

For the system evaluation, we targeted user experiences at a SL museum[1] designed and operated by members of Global COE (Center of Exellence) Progam "Digital Humanities Center for Japanese Arts and Culture" of Ritsumeikan Unviersity. This museum is aimed at a virtual exhibition of Kaga Okunizome Dyeing[2], kimono and bedding from the Ishikawa region in Japan during the latter part of the Edo period until the beginning of the Showa period. However, we would like to point out that our system is also applicable to other museums in SL as well as other metaverse.

Figure 2 shows the museum building in which 19 exhibits and two posters are located. Our reasons for selecting this museum are that (1) the exhibition therein has a high cultural value because Kaga Okunizome dyeing is famous in Japan and (2) there is no copyright problem because the authors belong also to the aforementioned Global COE. For other representative museums in SL, please refer to the work by Urban et al. (Urban et al. 2007).

We implemented the system by adding the above four modules to the open-source SL client program (SL viewer program). For this work, we adopted a typical comic layout where the order to read is in the raster order, from top-left to bottom-right, and all frames have the same size. Figure 3 shows an example of a generated comic page of a user, say, user k. A partial trace information of this user is given in Table 2, where a row indicates the information on the visit grid at a given loop count, which is the timing that user k entered the grid. In addition, Figs. 4 and 5 show the time series of the important levels over the grids in visiting order and over the loop counts, respectively.

[1] http://slurl.com/secondlife/rits%20gcoe%20jdh/167/189/22
[2] http://shofu.pref.ishikawa.jp/shofu/okuni/english/index.html

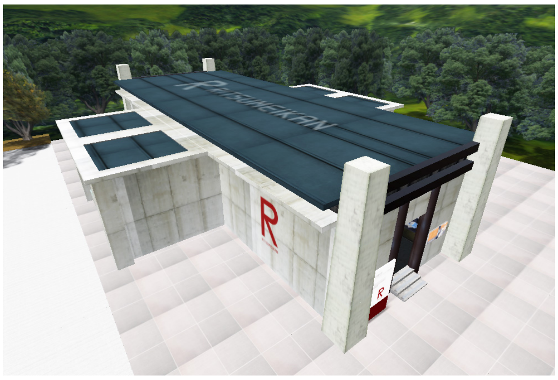

Figure 2. Target museum

Table 2. Example of user trace k

loop_count	grid position (i,j)	$I_k(i,j)$	f	Frame # in Fig.3
1	(5,1)	2.42	1	1st frame
360	(4,1)	1.38	2	2nd frame
448	(5,1)	2.42	-	-
838	(5,2)	0.46	3	-
935	(5,3)	4.61	4	3rd frame
1378	(6,3)	0.23	5	-
⋮	⋮	⋮	⋮	⋮

Figure 3. Example of a resulting comic page

3.2 Evaluation outline

There were 15 participants, who were undergraduate and graduate students in computer science, in the evaluation. Each of them was asked to visit the aforementioned museum at least for two minutes, spend time on exhibits they find interesting, and compare two comics, generated by the proposed system and a baseline system, in each of the following pairs:

(1) **One-page case**

Comic P_FEW consisting of 12 frames generated by the proposed system

vs

Comic B_FEW consisting of 12 frames generated by the baseline system

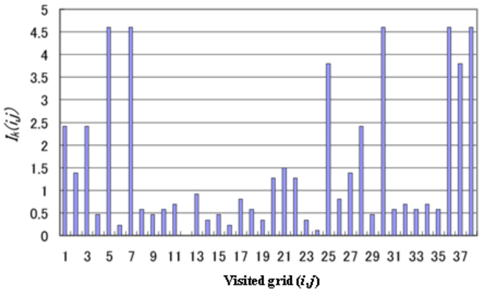

Figure 4. Important levels over the grids

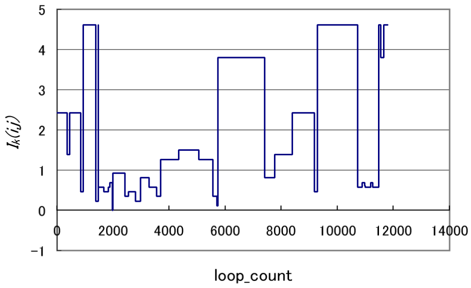

Figure 5. Important levels over the loop counts

(2) **Two-page case**

Comic P_MANY consisting of 24 frames generated by the proposed system

vs

Comic B_MANY consisting of 24 frames generated by the baseline system

The baseline system is the same as the proposed system except that a fixed-and-equal interval is used as the snapshot timing of each frame in a given comic, rather than the snapshot timing decided by (2). For each comic pair, one of the two questions that we asked the participants is "Which comic better represents your experience in terms of the time you spent to each of the visited exhibit?". The other one is "What are your reasons behind your selection?".

The layout of each page was empirically set to 4 frames × 3 rows, leading to 12 frames per page. In addition, we partitioned the museum space into 8×8 grids such that each grid can accommodate one or two exhibits. The camerawork for rendering comic frames was set to the default camerawork of the SL viewer. Figure 3 shown earlier is the first page of the resulting comic P_MANY from one of the participants. Figures 6 and 7, respectively, show the resulting comics P_FEW and B_FEW from the same participant.

3.3 Results and discussions

Table 3 shows the evaluation result. It can be seen that, for the one-page case, the comics generated by the proposed system are better than those generated by the baseline system in representing the user experiences in the targeted museum. Compiling the participants' reasons behind their comic selections, we summarize our analysis results as follows:

(1) Due to the use of the fixed-and-equal snapshot timing, the baseline system could not extract frames that cover exhibits to which participants spent relatively long time. On the contrary, the proposed system could. For example, the kimono in the fourth frame (the top-right one) of Fig. 6, having a relatively high important level $I(Frame(4)) = 1.26$, was not shown in Fig. 7. In addition, Fig. 6 does not display the pair of exhibits in the third frame of Fig. 7 because the participant spent relative short time on the corresponding grid. For the two-page case, however, the proposed system renders also this pair of exhibits as shown in the sixth frame ($I(Frame(6)) = 0.92$) of Fig. 3.

(2) As shown in Fig. 8, for the two-page case, there existed many resemble frames in comics generated by the baseline system when a participant spent relatively long time to a particular grid. Such phenomena occurred less frequently in the proposed system.

However, from Table 3, the two systems have similar results for the two-page case. This is because the proposed system must select most of the generated frames in order to meet the number of frames, i.e., 24. As a result, the proposed frame selection mechanism, based on the important level, contributes less in this case. However, as far as summarization of user experiences is concerned, the result in the one-page case is more important than that of the two-page case in evaluation of the proposed system.

Figure 6. Example of a resulting comic page (P_FEW) by the proposed system

Figure 7. Example of a resulting comic page (B_FEW) by the baseline system

Figure 8. Example of a resulting comic page (B_MANY) by the baseline system that shows resemble frames

Table 3. Evaluation results

Comics	Votes	Comics	Votes
P_FEW	11	P_MANY	8
B_FEW	4	B_MANY	7
Total	15	Total	15

4. RELATED WORK

In addition to the authors' previous work (Shuda & Thawonmas 2008; Thawonmas & Shuda 2008; Thawonmas and Oda 2010) on comic generation for online games, other work includes comic generation for online games using a combination of screenshots (Chan et al. 2009), rather than using the targeted game's engine for re-rendering frames as done in our approach, and for first person shooters (Shamir et al. 2006). Comics were also used for summarization of experiences in a conference (Sumi et al. 2002), daily experiences (Cho et al. 2007), and videos or movies (Calic et al. 2007; Hwang et al. 2006; Tobita 2010).

Collections of images were used for storytelling by generating slides from images taken at multiple locations on a given map (Fujita & Arikawa 2008). A scripting system (Zhang et al. 2007.) was proposed for automating cinematics and cut-screens that facilitate video-game production processes such as the control of transitions between images, and the annotation of texts and sounds to image backgrounds. For segmentation of videos, a method (Xu et al. 2009) exists that uses histograms of human motions, but this method aims at fast movements such as sports or dances, not slow movements typically seen in SL avatars while they are visiting museums.

5. CONCLUSIONS AND FUTURE WORK

We proposed and implemented a system for generating comics from user experiences during their museum visits. The system enables extraction of frames that contain interesting exhibits, from the view point of a user of interest. This was achieved by partitioning the museum space into multiple grids and determining their important levels based on the user's visiting time on each of them. The conducted system evaluation confirmed the effectiveness, in properly representing the aforementioned user experiences, of the proposed system, especially when the number of frames was limited.

As our future work, we plan to increase the expressivity of comics. One possible research theme is to strengthen the comic layout mechanism. This would increase the variety of comic frames, such as large frames, small frames, slanted frames, etc. Another theme is to improve the camerawork so that a more proper set of camera parameters, i.e., the camera angle, camera position, and zoom position, is applied to a given frame. The mechanisms proposed in our previous work for online games (Thawonmas and Shuda 2008; Thawonmas and Ko 2010) will be extended, respectively, for these two research themes.

ACKNOWLEDGEMENT

The second author is now with Sony Corporation. This work is supported in part by the MEXT Global COE Program - Digital Humanities Center for Japanese Arts and Cultures, Ritsumeikan University.

REFERENCES

Bianchi, A. and Zancanaro M., 1999. Tracking users' movements in an artistic physical space, Community of the Future: Proc. of i3 Annual Conference, pp. 103-106.

Calic, J. et al., 2007. Efficient Layout of Comic-like Video Summaries, IEEE Transactions on Circuits and Systems for Video Technology, Vol. 17, No. 7, pp. 931-936.

Chan, C.-J. et al., 2009. Automatic Storytelling in Comics: A Case Study on World of Warcraft, CHI Extended Abstracts 2009, pp. 3589-3594.

Cho, S.B. et al., 2007. Generating Cartoon-Style Summary of Daily Life with Multimedia Mobile Devices, Proc. of IEA/AIE 2007, pp. 135-144.

Fujita, H. and Arikawa, M., 2008. Animation of Mapped Photo Collections for Storytelling, IEICE TRANS. INF. & SYST., Special Section/Issue on Human Communication III, Vol. E91-D, No. 6, pp. 1681-1692.

Hwang, W.-I., et al. 2006. Cinema comics: Cartoon generation from video stream. Proc. of GRAPP 2006, pp. 299-304.

Lowood, H., 2006. High-performance play: The making of machinima, Journal of Media Practice, Vol. 7, No. 1, pp. 25-42.

Prendinger, H., 2009. The Global Lab: Towards a Virtual Mobility Platform for an Eco -Friendly Society, Transactions of the Virtual Reality Society of Japan, Vol. 14, No. 2, pp. 163-170.

Shamir, A. et al., 2006. Generating Comics from 3D Interactive Computer Graphics, IEEE Computer Graphics and Applications, Vol. 26, No. 3, pp. 53-61.

Shuda, T. and Thawonmas, R., 2008. Frame Selection for Automatic Comic Generation from Game Log, Lecture Notes in Computer Science, Vol. 5309 (ICEC 2008: Entertainment Computing), pp. 179-184

Sparacino, F., 2003. Sto(ry)chastics: A Bayesian Network Architecture for User Modeling and Computational Storytelling for Interactive Spaces, Lecture Notes in Computer Science, Vol. 2864 (UbiComp 2003: Ubiquitous Computing), pp.54-72.

Sumi, Y. et al., 2002. ComicDiary: Representing individual experiences in a comic style, Ubicomp 2002, pp. 16-32.

Thawonmas, R. and Oda, K. 2010. Rule-Based Camerawork Controller for Automatic Comic Generation from Game Log, Lecture Notes in Computer Science, Vol. 6243 (ICEC 2010: Entertainment Computing), pp. 326-333.

Thawonmas, R. and Shuda, T., 2008. Comic Layout for Automatic Comic Generation from Game Log, Proc. of 1st IFIP Entertainment Computing Symposium, Vol. 279, pp. 105-115.

Tobita, H., 2010. DigestManga: interactive movie summarizing through comic visualization, CHI. Extended Abstracts 2010, pp. 3751-3756.

Urban, R. et al., 2007. A Second Life for Your Museum: 3D Multi-User Virtual Environments and Museums, Proc. of Museums and the Web 2007, San Francisco, April 2007.

Xu, J. et al., 2009. Temporal segmentation of 3-D video by histogram-based feature vectors, IEEE Transactions on Circuits and Systems for Video Technology archive, Vol. 19, No. 6, pp. 870-881.

Zhang, W. et al., 2007. Story scripting for automating cinematics and cut-scenes in video games, Proc. of the 2007 conference on Future Play, pp. 152-159.