International Journal on Artificial Intelligence Tools
Vol. 21, No. 2 (2012) 1240008 (13 pages)

DOI: 10.1142/S0218213012400088

World Scientific
www.worldscientific.com

CAMERA CONTROL FOR GENERATING COMICS FROM VIRTUAL MUSEUM VISITORS' EXPERIENCES

RUCK THAWONMAS* and KOHEI KATO
Intelligent Computer Entertainment Laboratory
Graduate School of Science and Engineering, Ritsumeikan University
1-1-1 Nojihigashi, Kusatsu, Shiga 525-8577, Japan
**ruck@ci.ritsumei.ac.jp*

We present a camerawork module for representing events in a virtual museum with a comic style. The module uses four basic camera shots, i.e., internal, external, apex, and group, to compose a shot sequence of two main events happening in a museum, i.e., move and stay. This module is fully implemented in a comic generation system we have been developing for summarizing user experiences in Second Life. We also discuss a user evaluation that compares comics whose camerawork is decided by the presented module with those decided by a baseline module that simply uses the same camera parameters as those of the Second Life's viewer. User evaluation results as well as visual comparisons confirm the effectiveness of the proposed camerawork module.

Keywords: Camerawork; virtual museum; camera shots.

1. Introduction

Comic is a promising representation style for summarizing various kinds of user experiences in an entertaining fashion. Applications of the comic-style representation include summarization of activities in a conference,[1] daily activities,[2] video sequences,[3,4] and game-play activities.[5–7] Comic-style summaries facilitate perceiving of main events, augmenting of personal memories, and promoting of communication among user communities. In this paper, we describe a part of our current project on summarizing user experiences in a museum located in metaverse such as Second Life (SL).[8,9] In particular, we focus on how to perform camerawork in such an environment. In addition, we propose four basic camera shots and three camera idioms, each of which specifies a sequence of those basic camera shots for a certain situation.

Our research group has developed a number of techniques[10–13] for our comic generation system[6] that aims at automatically generating comics from game log. In that system, we adopted the same approach as in Shamir *et al.* (2006)[5] where the game engine of a game of interest is used for rendering a comic based on information in the game log. Another approach adopted in Chan *et al.* (2009)[7] is that

of composing a comic from selected screen shots. Although the game engine must be accessible, our approach provides more room to play with camerawork and thus more varieties in comics.

Previously, in order to utilize the advantage of the game-engine approach, we proposed in Thawonmas *et al.* (2009)[12] and Thawonmas and Oda (2010)[13] a camerawork editor and an automatic camerawork module, respectively. The former allows the viewer to manually edit the camerawork of a generated comic while the latter automatically decides the camerawork of each frame based on rules derived from an analysis of online-game webcomics. However, camerawork of online-game experiences, where major activities are related to fights against monsters and acquisitions of items, is different from that of user experiences in virtual museums, where major activities are simply moving around and staying to view exhibits therein.

This work is related to automated camera control for generation of an animation or a movie. According to Friedman and Feldman,[14] there exist two approaches, i.e., idiom-based (or knowledge-based)[14–19] and constraint satisfaction (or optimization)[20–26] approaches. Recent works such as the one in Lino *et al.* (2010)[27] combine both of them. Those techniques are computationally expensive because they need to control the camera for a sequence of all consecutive shots which results in a video. However, part of research findings can still be applied to our work. In this paper, as our first step, we follow He *et al.*[17] in deriving basic shots and idioms beforehand. We then use them for deciding the camerawork of a selected comic frame and perform user evaluation to see their effectiveness. Optimization of camera parameters is, however, beyond the scope of this paper.

The contributions of this paper are as follows:

(1) Outline of our automatic comic generation system that is applicable to summarizing user experiences in virtual museums, galleries, or exhibitions;
(2) Description of the four basic camera shots, i.e., internal, external, apex, group, that are used to articulate the avatar's movement and its interactions with nearby exhibits in comic frames;
(3) Description of the three camera idioms, i.e., move, single view, multiple view, that specify which basic shots to use in a given situation, based on well-known cinematographer heuristics, and lead to exciting, lively and entertaining comics.

2. Comic Generation System

Figure 1 depicts an overview of our comic generation system.[12,13] The aim of this system is to summarize user experiences in a virtual museum in a comic fashion. The system includes the proposed camerawork module, described in Sec. 3, as the most right module.

At the state-identification module, the visit log is chronologically divided into multiple partitions (henceforth called scenes). Each belongs to either the **move**

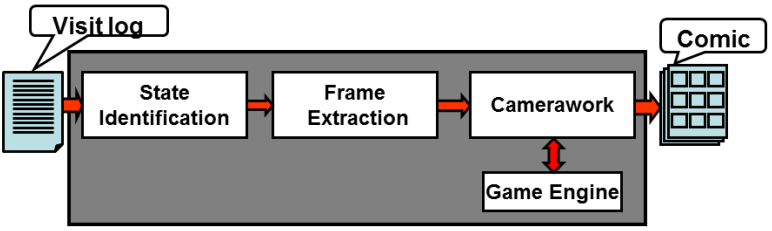


Fig. 1. Architecture of the comic generation system.

state or the **stay state**. The former state is assigned to a given scene in which the avatar continuously moves, and the second one is when the avatar stays in order to view a nearby exhibit.

At the frame-extraction module, given the desired total number of frames by the viewer, the system decides the number of frames to be extracted from each scene according to the scene's length. Thereby it is possible that no frames are extracted from a scene with relatively short length. If n frames are to be extracted from the ith scene that begins at t and ends at $t + length(i)$, the snapshot timing for the jth frame from this scene is at

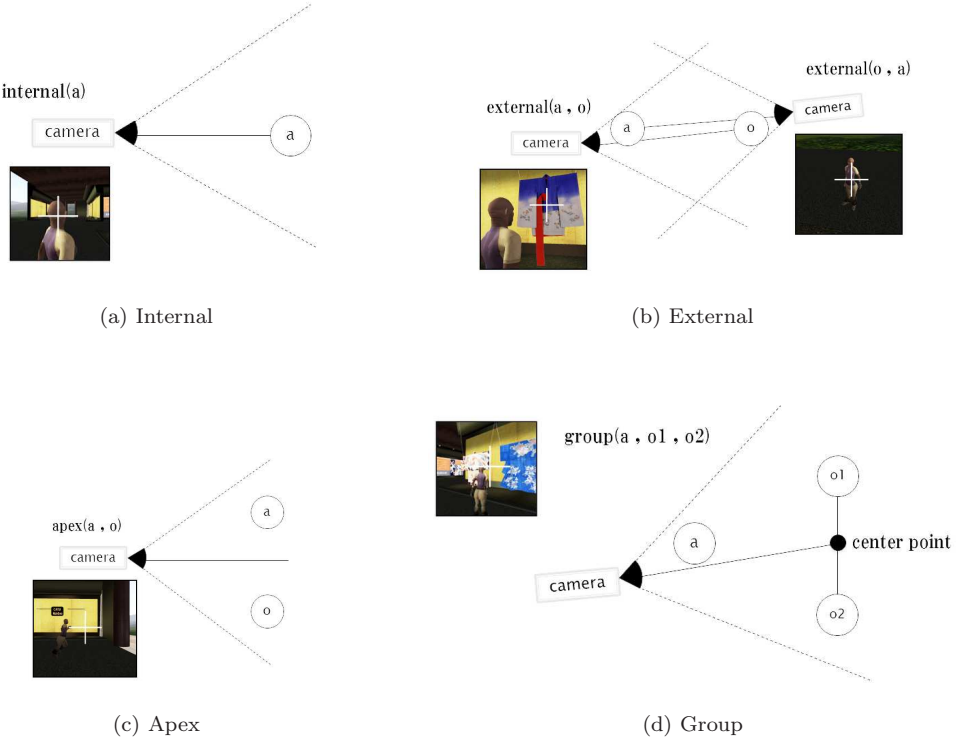
$$t + j(length(i)/(n+1))$$

for $j = 1$ to n, where $length(i)$ is the length of the ith scene.

3. Camerawork Module

The proposed camerawork module decides a shot sequence for each scene according to one of the three idioms: **move idiom**, **single-view idiom**, **multiple-view idiom**. For move scenes, the move idiom is used. For stay scenes, the camerawork module uses the single-view idiom when the number of exhibits near the avatar is one and the multiple-view idiom when there are two or more nearby exhibits. Four basic camera shots inspired by the work of He *et al.*,[17] i.e., internal, external, apex, and group, are used in these idioms. While basic shots in He *et al.* (1996)[17] were decided for avatars and interactions among them, our basic shots focus on the user avatar and its interactions between nearby exhibits or objects.

Figure 2 shows the camera placement and an example shot for internal, external, apex, and group, respectively, where a and o denote the avatar and an object of interest, respectively. The camera direction of each of these shots is towards the avatar's head, the second argument' center point, the center point between the avatar and the object, and the center point between the two objects, respectively. For each camera shot, three camera parameters, i.e., the camera angle, camera position, and zoom position, are decided in advance. As mentioned in Sec. 1, optimization of these parameters is beyond the scope of this paper.



(a) Internal (b) External

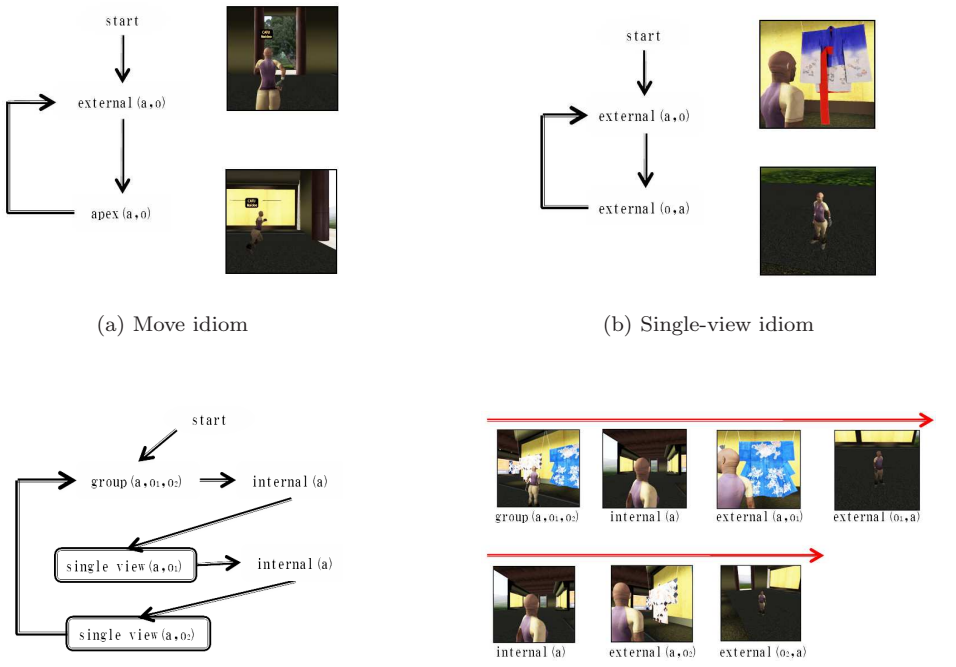
(c) Apex (d) Group

Fig. 2. (color online) Camera placement for the four basic camerashots proposed in this work.

Figure 3 shows the shot sequence and an example of each of the three camera idioms, where a and o denote the user avatar and an object of interest, respectively. We decided their shot sequences based on two cinematographer heuristics discussed in He *et al.* (1996)[17]: *avoid jump cuts* and *use establishing shots*. The former heuristic establishes a distinction between any two adjacent scenes. The latter heuristic allows the viewer to see all involving objects in the initial shot of a scene. The description of each idiom is given as follows:

Move idiom establishes a scene with external(a, o) so that the object o is seen over the shoulder of avatar a, where o is a virtual point located at a certain distance ahead of the avatar on the avatar's direction line. Then apex(a, o) is used, in which the camera focus is directed towards the point between the avatar and the virtual point, in order to express the avatar's movement.

Single-view idiom establishes a scene with external(a, o) and then switches to external(o, a) in order to prevent shooting similar consecutive shots, where o is the nearest exhibit within a given distance from the avatar or the aforementioned virtual point if there are no exhibits nearby.



(a) Move idiom

(b) Single-view idiom

(c) Multiple-view idiom

(d) Multiple-view shot sequence

Fig. 3. (color online) Shot sequence and example shots of the three camera idioms proposed in this work.

Multiple-view idiom establishes a scene with group(a, o_1, o_2) in which the nearest exhibit o_1 and the second nearest exhibit o_2 are both seen over the shoulder of the avatar. Then, in order to make the scene more exciting and lively, this idiom uses a combination of internal shots as well as single-view idioms.

For a scene of interest, the camerawork of each of its frames is decided according to the basic shot, one by one from top to bottom, in the corresponding idiom until the last frame of this scene. If there are still remaining frames when the bottom basic shot has been reached, the camerawork of the first remaining frame will be decided according to the top basic shot again. This process is repeated until the camerawork of all frames is decided.

4. Results and Discussions

We fully implemented the proposed camerawork module in the aforementioned comic generation system developed based on the open-source 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.

Fig. 4. (color online) The virtual museum in Second Life used in this work.

In this work, we targeted user experiences at a SL museum[a] designed and operated by members of the Global COE (Center of Excellence) Progam Digital Humanities Center for Japanese Arts and Culture of Ritsumeikan University. This museum exhibits Kaga Okunizome Dyeing, 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 note that our system is also applicable to other virtual museums, galleries, and exhibitions in SL as well as other metaverse.

Figure 4 shows the museum building in which 19 exhibits and two posters are located. Our reasons for using 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 also belong to the aforementioned Global COE. For other representative museums in SL, please refer to the work by Urban *et al.*[28]

4.1. *Evaluation outline*

We compared comics whose camerawork was decided by the proposed camerawork module and by a baseline camerawork module. The baseline module used the same camera parameters as those of the SL viewer program. Each of them was fully implemented in our comic-generation system. In our evaluation, a comic page consists of 4×3 frames, and all frames have the same size.

We first requested 20 participants, who are undergraduate or graduate students in our department, to watch each of the three video clips, showing typical visits of the following three visitor types:[29]

busy visitor who has a quick tour and skips most exhibits,[b]

[a]http://slurl.com/secondlife/rits%20gcoe%20jdh/167/189/22
[b]http://www.youtube.com/watch_popup?v=ijxwUrHejG8&vq=medium

selective visitor who spends longer time at exhibits that they are interested in and shorter time at or simply skips other exhibits,[c] and
greedy visitor who deliberately views most exhibits.[d]

In our evaluation, their comics have one, two, and three pages, respectively.

After watching each video clip, all participant were asked to compare two comics that were generated from the corresponding visit log used for the video clip, without being told which camerawork was used for each of these comics. In particular, each participant was asked to answer the following five questions for each comic in the typical five-level Likert scale:

Q1. Is the comic exciting?
Q2. Is the comic lively?
Q3. Does the comic well represent the clip's content?
Q4. Is the comic intelligible?
Q5. Is the comic entertaining?

They were also asked to write a short reason behind their rating for each question.

4.2. *Results and Discussions*

Tables 1, 2, and 3 show the average score of each camerawork for the busy, selective, and greedy types, respectively. The proposed camerawork outperforms, with statistical significance (t-test),the baseline camerawork in Q1 (exciting) and Q5 (entertaining) for all visitor types. With respect to Q2 (lively), the proposed camerawork outperforms the baseline camerawork for all visitor types and with statistical significance for the selective and greedy types. Regarding Q3 and Q4, the baseline camerawork outperforms the proposed camerawork with statistical significance for the busy type, but the differences between them for the selective and greedy types are not statistically significance. These two criteria open questions for future work.

Table 1. Average score of each camerawork for the busy type.

Question	Proposed Camerawork	Baseline Camerawork	p value
Q1	3.75	2.05	0.00
Q2	3.15	2.45	0.10
Q3	2.85	3.7	0.00
Q4	2.5	3.6	0.00
Q5	3.2	2.4	0.01

[c]http://www.youtube.com/watch_popup?v=H8eUJ6JZGXo&vq=medium
[d]http://www.youtube.com/watch_popup?v=DtVFf5gAUZI&vq=me dium

Table 2. Average score of each camerawork for the selective type.

Question	Proposed Camerawork	Baseline Camerawork	p value
Q1	3.55	2.1	0.00
Q2	3.25	2.4	0.05
Q3	3.2	3.05	0.63
Q4	3.45	3.3	0.67
Q5	3.7	2.5	0.00

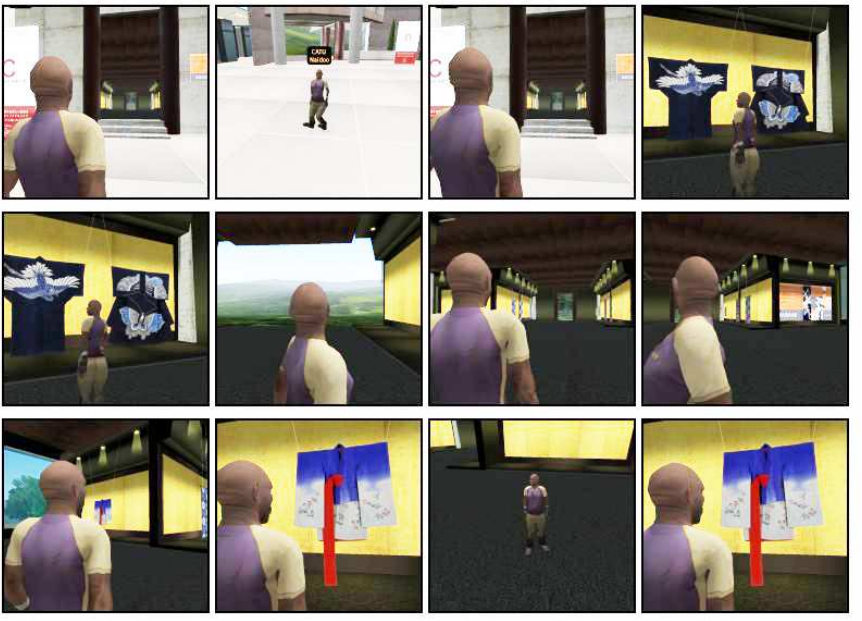
Table 3. Average score of each camerawork for the greedy type.

Question	Proposed Camerawork	Baseline Camerawork	p value
Q1	4.05	2	0.00
Q2	3.9	2.45	0.00
Q3	3.15	3.65	0.19
Q4	3.4	3.2	0.58
Q5	3.85	2.4	0.00

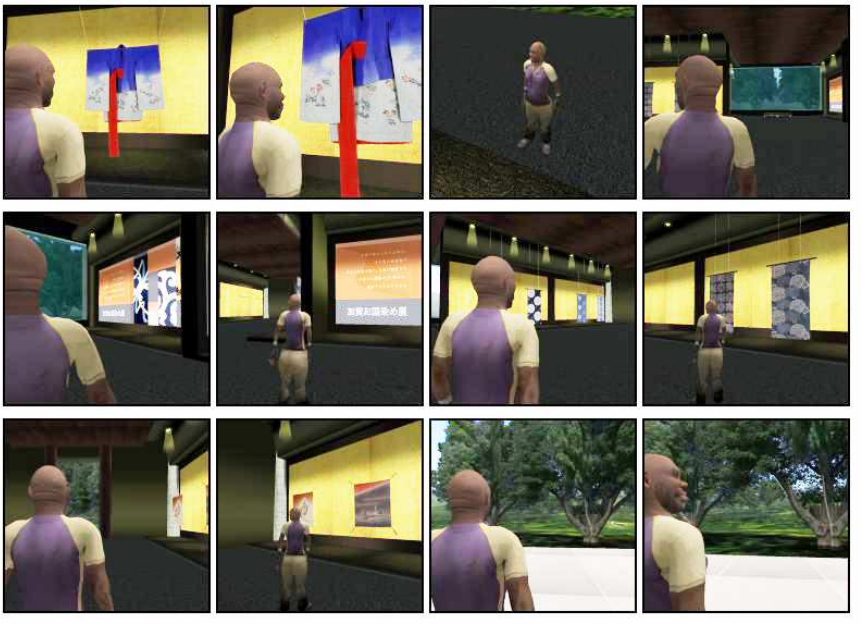
Figures 5 and 6 show the selective-type-visitor comic with the proposed camerawork and that with the baseline camerawork, respectively. For the former comic, the idiom and shot sequence of pages 1 and 2 are shown in Tables 4 and 5, respectively. Note that consecutive view (or move) idioms are seen in these tables because frames were not generated from those move (or stay) scenes, each lying between a pair of stay (or move) scenes, with short lengths. Comparing Figs. 5 and 6, one can see that similar consecutive frames exist in the comic with the baseline camerawork while this is not seen in the comic with the proposed camerawork. The same trend is seen for comics summarizing museum-visit experiences of the other two visitor types.

5. Conclusions and Future Work

This paper described our camerawork module that decides the camerawork of each comic frame representing user experiences in a virtual museum. First, we decided the four basic camera shots, internal, external, apex, and group, such that the avatar movement and the avatar interactions between nearby exhibits are articulated in comic frames. Then we derived the three camera idioms, move, single view, multiple view, based on two well-known cinematographer heuristics *avoid jump cuts* and *use establishing shots*. Conducted user evaluation confirmed the effectiveness of the proposed camerawork module in that the resulting comics were more exciting, lively and entertaining than those with the baseline camerawork.

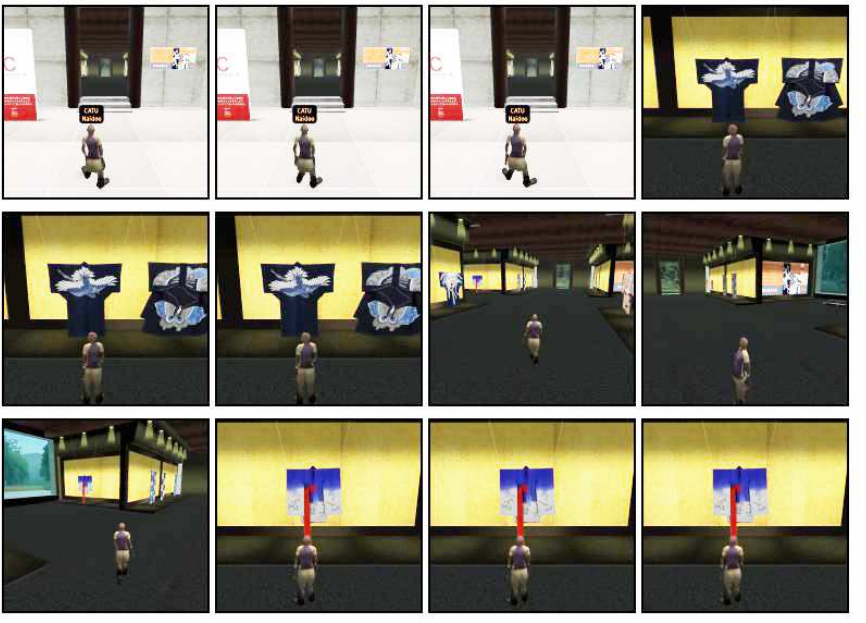


(a) Page 1



(b) Page 2

Fig. 5. (color online) Comic for the selective-type visitor with the proposed camerawork.



(a) Page 1

(b) Page 2

Fig. 6. (color online) Comic for the selective-type visitor with the baseline camerawork.

Table 4. Idiom and shot description for page 1 of Fig. 5.

single-view: external(a, o)	single-view: external(o, a)	single-view: external(a, o)	multiple-view: group(a, o_1, o_2)
multiple-view: group(a, o_1, o_2)	multiple-view: internal(a)	move: external(a, o)	single-view: external(a, o)
move: external(a, o)	single-view: external(a, o)	single-view: external(o, a)	single-view: external(a, o)

Table 5. Idiom and shot description for page 2 of Fig. 5.

single-view: external(a, o)	single-view: external(a, o)	single-view: external(o, a)	move: external(a, o)
move: external(a, o)	multiple-view: group(a, o_1, o_2)	move: external(a, o)	multiple-view: group(a, o_1, o_2)
move: external(a, o)	multiple-view: group(a, o_1, o_2)	move: external(a, o)	single-view: external(a, o)

However, in terms of content representation and intelligibility, the proposed camerawork and the baseline one did not differ with statistically significance. For the former criterion, in order to increase the viewer's perception about spatial behaviors of the avatar, we plan to modify the frame extraction module so that more frames in move scenes are extracted. For the latter, we plan to revise our idioms so that other cinematographer heuristics such as *don't cross the line of interest* hold. Other future work includes optimization of camera parameters for each camera shot.

Acknowledgments

This work is supported in part by the MEXT Global COE Program — Digital Humanities Center for Japanese Arts and Cultures, Ritsumeikan University, and by Grant-in-Aid for Scientific Research (C), No. 23500192, the Japan Society for Promotion of Science.

References

1. Sumi, Y., Sakamoto, R., Nakao, K., and Mase, K., ComicDiary: Representing Individual Experiences in a Comic Style. In: Borriello, G. and Holmquist, L. E. (eds.), *UbiComp 2002*. LNCS, Vol. 2498, pp. 16–32. Springer, Heidelberg (2002).
2. Cho, S. B., Kim, K. J., and Hwang, K. S., Generating Cartoon-Style Summary of Daily Life with Multimedia Mobile Devices. In: Okuno, H. G., Ali, M. (eds.) *IEA/AIE 2007*. LNCS (LNAI), Vol. 4570, pp. 135–144. Springer, Heidelberg (2007).
3. Calic, J., Gibson, D. P., and Campbell, N. W., Efficient Layout of Comic-like Video Summaries. *IEEE Transactions on Circuits and Systems for Video Technology*, 17(7) (2007) 931–936.

4. Tobita, H., DigestManga: interactive movie summarizing through comic visualization, In: *CHI Extended Abstracts 2010*, pp. 3751–3756 (2010).
5. Shamir, A., Rubinstein, M., and Levinboim, T., Generating Comics from 3D Interactive Computer Graphics. *IEEE Computer Graphics and Applications*, 26(3) (2006) 53–61.
6. Shuda, T. and Thawonmas, R., Automatic Comic Generation Using Online-Game Play Log. *Journal of Game Amusement Society*, 3(1) (2009) 40–45 (in Japanese).
7. Chan, C.-J., Thawonmas, R., and Chen, K.-T., Automatic Storytelling in Comics: A Case Study onWorld ofWarcraft. In: *CHI Extended Abstracts 2009*, pp. 3589–3594 (2009).
8. Thawonmas, R. and Fukumoto, A., Frame Extraction Based on Displacement Amount for Automatic Comic Generation from Metaverse Museum Visit Log. In: *Intelligent Interactive Multimedia Systems and Services, Smart Innovation, Systems and Technologies, 11* (*IIMSS-11*), pp. 153–162, DOI:10.1007/978-3-642-22158-3_16 (2011).
9. Thawonmas, R. and Shuda, T., Frame Selection for Automatic Comic Generation from Museum Playlog in Metaverse. In: *Proc. of IADIS International Conference Game and Entertainment Technologies 2011*, Rome, Italy, pp. 43–50, July 22–24 (2011).
10. Thawonmas, R. and Shuda, T., Comic Layout for Automatic Comic Generation from Game Log. In: *Proc. IFIP International Federation for Information Processing* (*ECS-2008*), Vol. 279, pp. 105–115 (2008).
11. Shuda, T. and Thawonmas, R., Frame Selection for Automatic Comic Generation from Game Log. In: Stevens, S. M. and Saldamarco, S. J. (eds.), *ICEC 2008*. LNCS, Vol. 5309, pp. 179–184. Springer, Heidelberg (2008).
12. Thawonmas, R., Oda, K., and Shuda, T., Camerawork Editor for Automatic Comic Generation from Game Log. In: *CD-ROM Proc. of Nicograph International 2009*, Kanazawa, Japan, June 19–20 (2009).
13. Thawonmas, R. and Oda, K., Rule-Based Camerawork Controller for Automatic Comic Generation from Game Log. In: Yang, H. S., Malaka, R., Hoshino, J., and Han, J. H. (eds.), *ICEC 2010*, LNCS, Vol. 6243, pp. 326–333, Springer, Heidelberg (2010).
14. Friedman, D. and Feldman, Y., Automated Cinematic Reasoning about Camera Behavior. *Expert Systems and Applications*, 30(4) (2006) 694–704.
15. Karp, P. and Feiner, S., Automated Presentation Planning of Animation Using Task Decomposition with Heuristic Reasoning. In: *Proc. of Graphics Interface '93*, pp. 118–127 (1993).
16. Christianson, D., Anderson, S., He, L., Salesin, D., Weld, D., and Cohen, M., Declarative Camera Control for Automatic Cinematography. In: *Proc. of the Thirteenth National Conference on Artificial Intelligence*, pp. 148–155 (1996).
17. He, L., Cohen, M. F., and Salesin, D. H., The Virtual Cinematographer: A Paradigm for Automatic Real-Time Camera Control and Directing. *Computer Graphics*, 30 (1996) 217–224.
18. Butz, A., Animation with CATHI. In: *Proc. of Ninth Innovative Applications of Artificial Intelligence* (*IAAI-97*), pp. 957–962 (1997).
19. Tomlinson, B., Blumberg, B., and Nain, D., Expressive Autonomous Cinematography for Interactive Virtual Environments. In: *Proc. of Fourth International Conference Autonomous Agents*, pp. 317–324 (2000).
20. Drucker, S. M. and Zeltzer, D., Intelligent Camera Control in a Virtual Environment. In: *Proc. of Graphics Interface '94*, pp. 190–199 (1994).
21. Drucker, S. M. and Zeltzer, D., Camdroid: A System for Intelligent Camera Control. In: *Proc. of SIGGRAPH Symposium on Interactive 3D Graphics* (1995).

22. Bares, W. H. and Lester, J. C., Intelligent Multi-Shot 3D Visualization Interfaces. *Knowledge-Based Systems*, 12(8) (1999) 403–412.
23. Halper, N. and Olivier, P., CAMPLAN: A Camera Planning Agent. In: *Proc. of AAAI 2000 Spring Symposium on Smart Graphics*, pp. 92–100 (2000).
24. Elson, D. K. and Riedl, M. O., A Lightweight Intelligent Virtual Cinematography System for Machinima Production. In: *Proc. of AIIDE 2007*, pp. 8–13 (2006).
25. Burelli, P. and Jhala, A., Dynamic Artificial Potential Fields for Autonomous Camera Control in 3D Environments. In: *Proc. of Artificial Intelligence and Interactive Digital Entertainment* (*AIIDE '09*), (2009).
26. Burelli, P. and Yannakakis, G. N., Combining Local and Global Optimisation for Virtual Camera Control. In: *Proc. of IEEE Conference on Computational Intelligence and Games*, pp. 403–410 (2010).
27. Lino, C., Christie, M., Lamarche, F., Schofield, G., and Olivier, P., A Real-time Cinematography System for Interactive 3D Environments. In: *Proc. of 2010 ACM SIGGRAPH Eurographics Symposium on Computer Animation*, pp. 139–148 (2010).
28. Urban, R. *et al.*, A Second Life for Your Museum: 3D Multi-User Virtual Environments and Museums. In: *Proc. of Museums and the Web 2007* (2007).
29. Sparacino, F., Sto(ry)chastics: A Bayesian Network Architecture for User Modeling and Computational Storytelling for Interactive Spaces. In: Dey, A. K., Schmidt, A., and McCarthy, J. F. (eds.), *UbiComp 2003*. LNCS, Vol. 2864, pp. 54–72 (2003).