

# Film Comic Reflecting Camera-Works

Masahiro Toyoura, Mamoru Kunihiro, and Xiaoyang Mao

University of Yamanashi, Takeda 4-3-11, Kofu, Yamanashi, Japan  
mtoyoura@yamanashi.ac.jp

**Abstract.** We propose a novel technique for automatically creating film comics reflecting the camera-works of an original movie. Camera-works are one of the most important effects contributing to the mise en scene of the movie. A skilled director can use the camera-works dexterously for drawing the attention of audiences, representing sentiments, and give a change of pace in the movie. When creating film comics, camera-works are detected from the original movie, and mapped to panels and layouts of special comic styles. The technique is called as *the grammar of manga*. Our new algorithm is presented for automatically tiling the stylized panels into comic pages based on the grammar of manga. The results of our subject study show that reflecting camera-works in film comics enables the stories being presented in a more readable, vivid and immersive way.

**Keywords:** Film comic, grammar of film language, manga representation, and camera-work.

## 1 Introduction

A film comic is a kind of comic book created by selecting images from a movie and arranging them into a book of comic style. By turning a movie into a comic, film comic enables a wide range of readers, from little babies to comic manias, to easily browse or enjoy movie stories yet through another form of art media. Most of popular cartoon animations have been formed into film comics. We can buy printed film comics at book stores or enjoy digital film comics online. The process for creating a film comic includes the tasks for selecting images from the movie, editing the images into the comic pages, and inserting the balloons for representing the lines of characters. Even now, these tasks are mainly done manually. It is usually very time-consuming to select a compact set of images well depicting the story from the huge number of images in original movie. The most difficult thing, however, is how to well convey the mise en scene of original movie in comic style, so as to make the resulting comic as immersive and enjoyable as the original movie. Recently several researches on automatic generation of film comic have been reported [1,2,3,4,5]. Although each of those researches succeeded in developing some effective methods or tools for the automation of a part or the whole of the creation process, none of them has addressed, in particular, the automatic mapping of mise en scene from movie to comic. In this paper, we challenge the issue through reflecting camera-works in comic. Typical camera-works and their effects are well-known as a part of *the grammar of film language* [6]. A skilled director can use the camera-works dexterously for drawing the attention of audiences, representing sentiments, and give a change of pace in the movie. On the other hand, camera-works are also well-known

as a part of *the grammar of manga* [7]. Modern Japanese manga, which is said to have been established by Osamu Tezuka, is actually featured with the introduction of the style of movies [8]. In manga, camera-works are represented as panels and layouts of special styles commonly accepted by creators and readers. We adopt those manga styles in representing the camera-works in film comic.

Our major contributions can be summarized as follows:

1. A new method for mapping the camera-works in a movie to the special panel styles and layouts of based on the grammar of manga.
2. A novel panel algorithm for automatically tiling the panels to form comic pages with camera-works representations.
3. A subject study to compare the film comics with and without camera-works.

Our experimental results show that reflecting camera-works in film comic can not only improve the readability of stories, but also enables viewers to better experience the mise en scene of original movies in a more immersive way.

The new mapping method and layout algorithm are implemented as a part of our fully automatic film comic generation system. Since the other parts of the system, including the selection of representative images, the detection of camera-works from movies and the balloon representation of lines of characters, are built upon existing technologies, we will mention them briefly in the related works and focus our discussion only on the major contributions, due to the length limit of manuscript. The remainder of the paper is organized as follows. Section 2 briefly reviews the related works. Section 3 describes the mapping of camera-works between the movie and comic. Section 4 presents the new algorithm for creating the page layouts of film comic. Section 5 discusses the results of the subject studies and Section 6 concludes the paper by showing several future research directions.

## 2 Related Works

A pilot work on film comic generation was done by W.I. Hwang et al. [1,3]. Although their system supports some stylized effects such as speed line and rotational trajectory which contribute to the mise en scene of movie, those effects need to be added by users manually. Preuß et al. [5] proposed an automatic system for converting movies to comics, but the method requires the screenplay of the movie to be given and hence limits its applications to a very special case. More recently, R. Hong et al. [2] presented fully automatic system which employs face detector, lip motion analysis, and motion analysis to realize the automatic script-face mapping and key-scene extraction. As another unique approach, Kunihiro et al. [4] used gaze information of viewers for positioning the balloons and trimming the images. They assume that the lines of characters can be obtained from the captions of the movie. Gaze information enables the system to identify the speaker even if the speaker does not have skin colored face, which occurs frequently in case of cartoon animation. The regions not attended by viewers would be trimmed to fit into a panel. Since film comics are usually created from cartoon animations, our system adopt the gaze-based approach for the trimming of images and the positioning of balloon.

Another research field closely related to film comic is video summary. While a film comic usually serves as an alternative art media of storytelling, video summary aims to enable users to quickly review videos and use it as the index to access to the detailed information when necessary [9]. Despite of such differences, they share many elemental technologies. As new styles of video summaries, storyboards [10], video tapestries [11,12] and comic-like layouts [13] are developed in recent years with which all important information are represented in a single image. Although they are effective for quickly capturing the summary of a video, film comic is more effective in getting readers immersed in a story by presenting a series of camera-works in a way temporally synchronized with the original video. Yoshitaka et al. have proposed to structuralize a movie based on the grammar of film language for making a video summary [14]. They also proposed to recognize the feelings of tension, liberty and loneliness from camera-works in another research [15].

Our system employs Porter's color histogram based approach [16] to segment a movie into shots and choose the image with the average color histogram of a shot as the representative image of the shot. The spatio-temporal image proposed in [17] is used for extracting camera-works from movies.

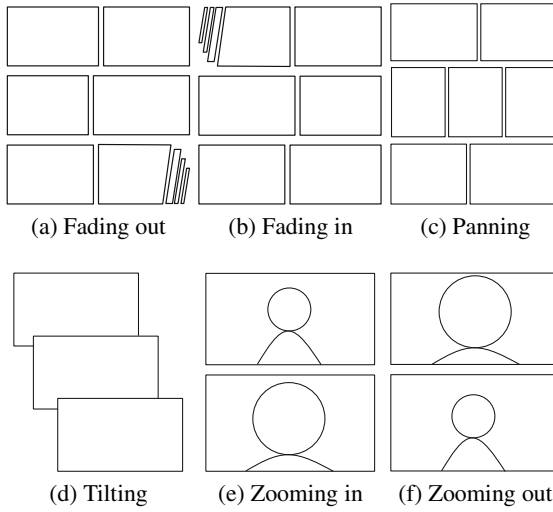
### 3 Mapping Camera-Works from Movie to Comic

In this section we explain typical camera-works and their corresponding comic representation. Typical camera-works include fading in/out, panning, tilting and zoom in/out.

*Fading out* is a camera-works that makes a scene disappear by gradually decreasing or increasing the brightness of the images. A shot ending with fading out gives a mild impression about the disappearing of the scene. Following the style of manga, we represent fading out with a stylized panel as shown in Figure 1(a). The fading out panel is positioned at the end of a page.

*Fading in* is a camera-work that brings a scene out by gradually increasing or decreasing the brightness of images. Same as fading out, the shot starting with fading in gives a mild impression on the appearing of the scene. We represent fading in with a panel as shown in Figure 1(b). A fading in panel should be positioned at the beginning of a page. A fading in is usually used right after fading out for achieving a smooth transition between two shots. Since a fading out panel is located at the end of a page, the fading in panel will be naturally positioned at the beginning of next page. Therefore, no special consideration is required in laying out the fading in panel.

*Panning* is a camera-work that changes the point of view along horizontal direction. It can be used to attract viewer's attention from right to left or left to right, enabling the viewer to catch the whole image of a wide object or impressing the viewer with the breadth of a landscape. It is also used for giving a sense of motion or fluidity to otherwise static shots. Inspired by some representations in manga, we represent panning with a stylized set of panels as shown in the middle row of Figure 1(c). Multiple panels corresponding to the images selected at successive positions during the panning are placed in the same row to convey the passage of time, continuousness, perspective or relationship between objects.



**Fig. 1.** Stylized panels and layouts for representing typical camera-works in film comic

*Tilting* is a camera-work that changes the point of view vertically up to down, or down to up. The former is also called as tilting down and the latter tilting up. The height or hugeness of objects can be emphasized with tilting. We represent the tilting with a page consisting of 3 panels corresponding to the images from the top, middle and bottom point of view, respectively (Figure 1(d)). Two adjacent panels are partially overlapped in such a way that the lower panel comes on top of the higher panel for tilting down, and vice versa for tiling up.

*Zooming in* is a camera-work that brings objects closer without physically moving a camera. It can attract viewer’s attention to the focusing objects, by weakening the impression of background and showing the closed up views of the objects. Since zooming in is unnatural changing of point of view for human, it also has the effect of heightening the tension. Following the style of manga, we present zooming in with a two panel page as shown in Figure 1(e). The two panels correspond to the frame before and after the zooming in, respectively.

*Zooming out* is a camera-work pushing objects further away without physically moving a camera. It gives an overall image of the scene and makes it easy to grab the relationship between objects. The feeling of liberation is also expected with zooming out, since the focus is released from the objects. Similar to zooming in, we use a two panel page as shown in Figure 1(f) to represent zooming out, with the two panels corresponding to the frames before and after the zooming out, respectively.

## 4 Panel Layout

In this section, we discuss how to arrange panels to create the pages of film comic, taking into consideration the stylized panels and layouts described in the previous section.

The panel layout algorithm should follow the rules below:

1. Assign one page for the shot of tilting, zooming in, and zooming out.
2. Assign the last panel of a page for the shot of fading out.
3. Assign one row for the shot of panning.
4. Tile the other images in such a way that each page contains  $m$  rows and  $n$  columns as possible. The remaining space is filled by expanding the widths of some panels.

As shown in Figure 2, let  $p$  denote the number of panning shots, and  $r_i (i = 1, 2, \dots)$  the number of other shots between two panning shots. First, we decide the number of pages in a way that each page contains  $m$  rows and  $n$  columns as possible. With  $n$  panels in one row, the number of total rows  $s$  is  $s = p + \sum_i [r_i/n]$ . Note that  $\lceil x \rceil$  is the ceiling of  $x$ . If each page consists of  $m$  rows, the number of pages  $t$  is  $t = \lceil s/m \rceil$ .

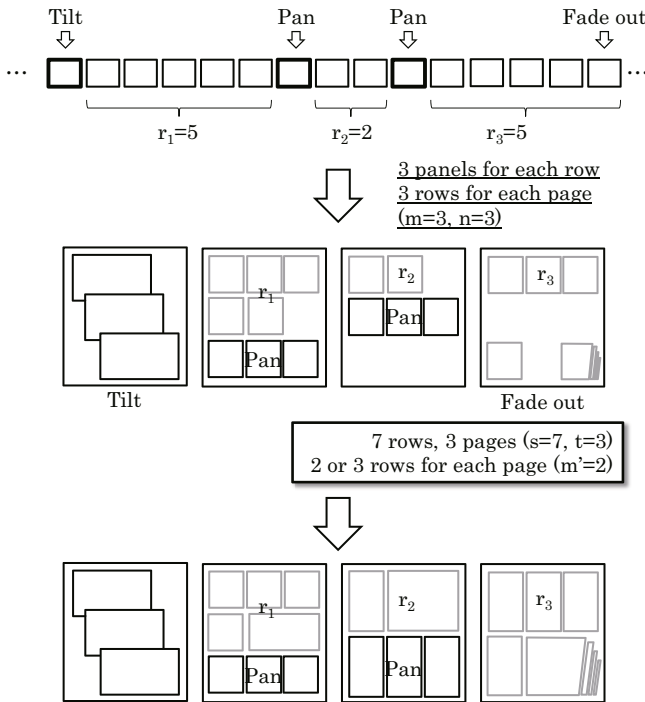


Fig. 2. Lay out panels in pages and rows

We cannot ensure that all pages have  $m$  rows exactly, except for the case that  $s$  is dividable by  $m$ . Let  $m'$  be the quotient of dividing the number of rows  $s$  by the number of pages  $t$ . We set the number of rows in each page to either  $m'$  or  $m' + 1$ , so as to keep the number of rows balanced among pages with a difference of 1 at most. In case  $m'$  is much less than  $n$ , a layout with narrow long panels will be created. When  $|m' - n| > m_{th}$ ,

we change  $n$  to  $n - 1$  and repeat this to decide new  $m'$  to improve the ratio of panels in row and column.  $m_{th}$  is a user given threshold to control the aspect ratio of panels.

As shown with the upper left image of Figure 3, a row right before a panning may not be fully filled with panels. The remaining space can be filled by expanding the width of a necessary number of panels. If those panels are not correctly chosen, however, we may get a layout with some panels across two rows. Such an example is shown in the upper right of Figure 3. To avoid generating such a layout, we should select the last image to expand for instance.

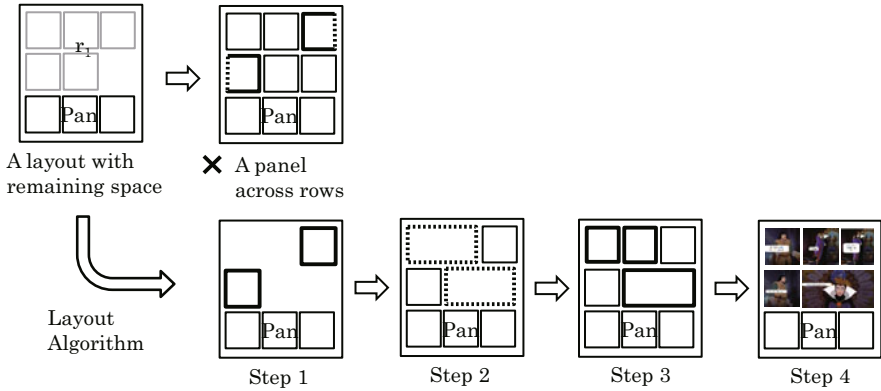


Fig. 3. Layout with/without a panel across rows

For any  $n$  and  $r_i$ , we can get a correct layout with the following algorithm.

**Layout algorithm for filling the remaining space (Figure 3):**

- Step 1.** If  $n$  is odd, assign a panel of width 1 at either the leftmost or the rightmost of each row.
- Step 2.** Divide the remaining space into the slots of width 2.
- Step 3.** Let  $v = r \bmod n$ , the remainder of dividing  $r$  by  $n$ . Randomly choose  $v$  slots and assign 1 panel of width 2 to each of them. The remaining slots are assigned with 2 panels of width 1.
- Step 4.** Fill all shots with the  $r$  images sequentially. The images are resized according to the width of the panels.

The above algorithm has been proven to be able to generate correct layouts always without panels across two rows. We include the formal mathematical proof in Appendix A.

## 5 Experimental Results

We conducted a subject study to investigate whether introducing camera-works improves the readability of stories and contributes to the representation of mise en scene

of original movies. The subjects were 20 university students. They were asked to watch both the comic without stylized representation for camera-works, called as A, and the comic with the stylized panels and layouts reflecting the camera-works, called as B. B includes panels representing fading in/out, panning, tilting and zooming in/out. The movies used for the test were “Snow White”, “Pinocchio” and “Shrek 3”. The lines of characters were extracted from captions, and the gaze information of pre-viewers was used for positioning the balloons and trimming the images. The examples of generated comic pages are shown in Figures 4 and 5<sup>1</sup>.

Our questionnaire includes the following questions together with a free feedback description. We presented A in first to a half of subjects, and B in first to the other half to eliminate the order effect.

### Questions:

1. Is the story easy to understand?
2. Is the comic presented in a vivid style?
3. Are the scenes immersive?
4. Can you feel sentiments of characters?
5. Can you easily focus on the subjects in each scene?
6. Did you enjoy reading?

The results are shown in Table 1. The evaluation results are summed up in the nominal scale with 3 values for adopting  $\chi^2$ -test. When  $\chi^2(2)$  is larger than 9.2, 6.0 and 4.6,  $p$  is smaller than 0.01, 0.05 and 0.10, respectively.

**Table 1.** Results of the subject study. \*:  $p \leq 0.05$ , \*\*:  $p \leq 0.01$ , **bold values** are the largest ones.

Q	Snow White				Pinocchio				Shrek 3			
	$\chi^2(2)$	-	0	+	$\chi^2(2)$	-	0	+	$\chi^2(2)$	-	0	+
1	1.6	4	<b>8</b>	<b>8</b>	9.1*	3	4	<b>13</b>	9.1*	3	4	<b>13</b>
2	24.1**	2	1	<b>17</b>	19.9**	1	3	<b>16</b>	19.9**	1	3	<b>16</b>
3	19.6**	2	2	<b>16</b>	12.4**	2	4	<b>14</b>	12.4**	2	4	<b>14</b>
4	9.7**	2	5	<b>13</b>	9.1*	3	4	<b>13</b>	9.1*	3	4	<b>13</b>
5	7.6*	2	6	<b>12</b>	4.3	4	5	<b>11</b>	4.3	4	5	<b>11</b>
6	15.7**	3	2	<b>15</b>	13.3**	1	5	<b>14</b>	13.3**	1	5	<b>14</b>

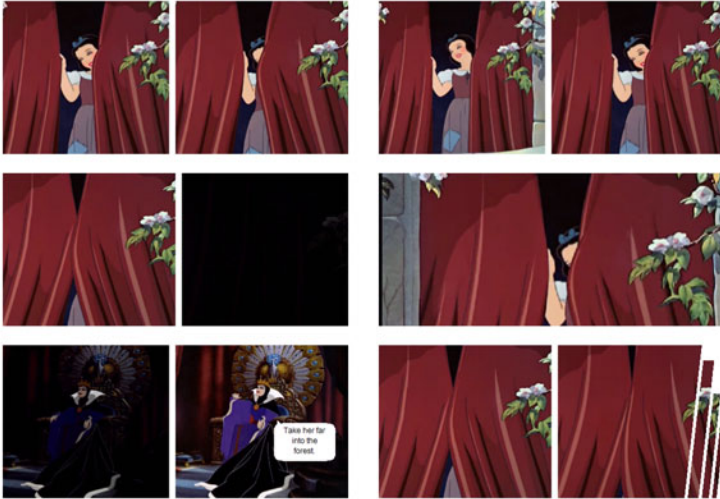
For Q1 of “Pinocchio” and “Shrek 3”, more subjects preferred B, the comic pages with stylized panels and layouts reflecting camera-works. We can conclude that reflecting camera-works improves the readability of stories. However, B was not preferred for Q1 of “Snow White”. In the free feedback description, a subject answered that there are too many panels being emphasized in the comic pages reflecting camera-works. Another subject answered that he got tired since it was too long. There are many stylized panels and layouts in case of “Snow White” in fact, since the original movie had many camera-works. It seems that the excessive presentment had prevented subjects from understanding the story of “Snow White” well. To narrow down to the panels that really

<sup>1</sup> The intellectual property of “Snow White” shown in Figure 4 and Figure 5 has expired.



(a) Page including pan shot  
(Without camera-work representation)

(b) Page including pan shot  
(With camera-work representation)

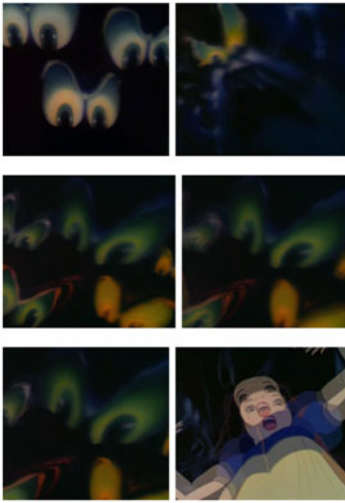


(c) Page including fade out shot  
(Without camera-work representation)

(d) Page including fade out shot  
(With camera-work representation)

**Fig. 4.** Generated comic pages with/without stylized panels and layouts. (a) and (b) are the pages including panning shots, and (c) and (d) including fading out shots. (a), (c) are comic pages without reflecting camera-works, while (b), (d) are those with the stylized panels and layouts reflecting camera-works. Read from top left to bottom right.

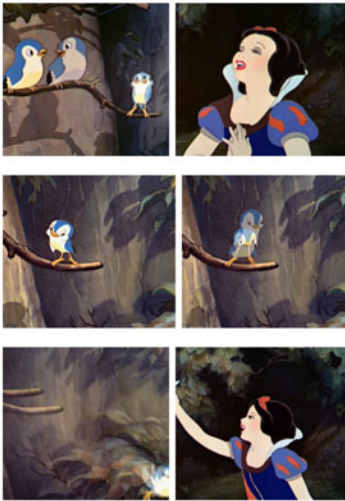




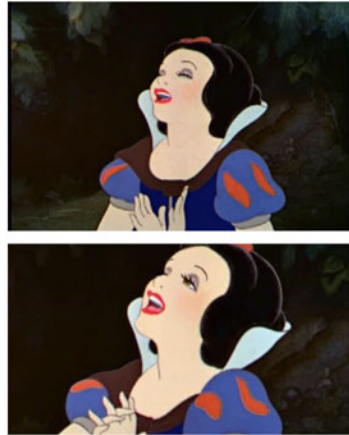
(a) Page including tilt shot  
(Without camera-work representation)



(b) Page including tilt shot  
(With camera-work representation)



(c) Page including zoom-in shot  
(Without camera-work representation)



(d) Page including zoom-in shot  
(With camera-work representation)

**Fig. 5.** Generated comic pages with/without stylized panels and layouts in manga. (a) and (b) are the pages including tilting shots, and (c) and (d) including zooming in shots. (a), (c) are comic pages without reflecting camera-works, while (b), (d) are those with the stylized panels and layouts reflecting camera-works. Read from top left to bottom right.

need to be emphasized, more advanced video understanding technologies is required. We are going to address the problem as a future work.

For Q2, Q3 and Q4 of all movies, B was preferred by more subjects. Since camera-works usually can help giving a vivid impression or rhythmic pace to a movie, we can expect that a comic with camera-works reflected can also achieve the same effects. A subject answered that he could not concentrate on reading A compared with B, since A had uniform panels. The result of Q4 showed that effects of camera-works can also facilitate viewers to better feel the sentiments of characters with our proposed method. For Q5, B was better evaluated for “Snow White” only. In “Pinocchio” and “Shrek 3”, there were the shots including two and more people talking together, which might be a reason for preventing viewers from focusing on the subject. For Q6, B was preferred by most users for all movies. We can conclude that a film comic generated with our proposed method enable viewers to view the story in a more enjoyable and immersive way.

Subjects also answered that echoic and mimetic words would improve the quality of the comic and inserting additional panels without balloons may make the comic more rhythmic. Those feedbacks are important clues to the further improvement of the technology in our future work.

## 6 Conclusions and Future Works

We proposed a new method for automatically generating film comics with stylized panels and layouts reflecting the camera-works of the original movie. We have confirmed that the comic generated with the proposed technique can better convey the mise en scene of the original movie.

One of our future works is to develop a technique for narrowing down to the most important panels that really need to be emphasized in the comics. A subject of the evaluation experiment pointed out that the comics were over accentuated when too many camera-works were used in a movie. We will seek more powerful image understanding and motion analysis technologies for estimating the importance of camera-works. We would like to provide an interactive editing tool enabling users to easily reflect their personal styles, while continuing exploring the potential of automatic approach in the mapping of mise en scene between movie and comic.

## References

1. Chun, B.-K., Ryu, D.-S., Hwang, W.-I., Cho, H.-G.: An Automated Procedure for Word Balloon Placement in Cinema Comics. In: Bebis, G., Boyle, R., Parvin, B., Koracin, D., Remagnino, P., Nefian, A., Meenakshisundaram, G., Pascucci, V., Zara, J., Molineros, J., Theisel, H., Malzbender, T. (eds.) ISVC 2006, Part II. LNCS, vol. 4292, pp. 576–585. Springer, Heidelberg (2006)
2. Hong, R., Yuan, X.T., Xu, M., Wang, M., Yan, S., Chua, T.S.: Movie2comics: a feast of multimedia artwork. In: Proceedings of the International Conference on Multimedia, pp. 611–614 (2010)

3. Hwang, W.I., Lee, P.J., Chun, B.K., Ryu, D.S., Cho, H.G.: Cinema comics: Cartoon generation from video stream. In: International Conference on Computer Graphics Theory and Applications, pp. 299–304 (2006)
4. Kunihiro, M., Mao, X.: The automatic generation of the film comics from the animation with eyes information. In: Symposium on Visual Computing, Graphics and CAD (2008) Article 13 (in Japanese)
5. Preuß, J., Loviscach, J.: From movie to comics, informed by the screenplay. In: ACM SIGGRAPH (Poster) (2007)
6. Arijon, D.: Grammar of the Film Language. Silman-James Press (1991)
7. Tsukamoto, H.: Manga Bible (2007)
8. Ohtsuka, E.: Guide for Movie Style Manga. Ascii Books (2010) (in Japanese)
9. Boreczky, J.S., Girgensohn, A., Golovchinsky, G., Uchihashi, S.: An interactive comic book presentation for exploring video. In: ACM Conference on Computer-Human Interaction (CHI), pp. 185–192 (2000)
10. Goldman, D.B., Curless, B., Salesin, D., Seitz, S.M.: Schematic storyboarding for video visualization and editing. ACM Transactions on Graphics (Proceedings of ACM SIGGRAPH) 25(3), 862–871 (2006)
11. Barnes, C., Goldman, D.B., Shechtman, E., Finkelstein, A.: Video tapestries with continuous temporal zoom. ACM Transactions on Graphics (Proceedings of ACM SIGGRAPH) 25(3) (2010)
12. Correa, C.D., Ma, K.L.: Dynamic video narratives. ACM Transactions on Graphics (Proceedings of ACM SIGGRAPH) 29(3) (2010)
13. Myodo, E., Takagi, K., Yoneyama, A.: Generation of efficient comic-like layout for video summaries. Forum on Information Technology (FIT) 3, 291–292 (2008) (in Japanese)
14. Yoshitaka, A., Deguchi, Y.: Video summarization based on film grammar. In: IEEE Workshop on Multimedia Signal Processing, pp. 333–336 (2005)
15. Yoshitaka, A., Matsui, R., Hirashima, T.: Extracting movie rendition based on camera work detection and classification. Transactions of Information Processing Society of Japan 47(6), 1696–1707 (2006) (in Japanese)
16. Porter, S.V.: Video Segmentation and Indexing using Motion Estimation. PhD thesis, University of Bristol (2004)
17. Ngo, C.W., Pong, T.C., Zhang, H.J.: Motion analysis and segmentation through spatio-temporal slices processing. IEEE Transactions on Image Processing 12, 341–355 (2003)

## Appendix A. The Validity of Scheduling Algorithm Described in Section 4

Let  $n$  be the number of panels in a row,  $r_i$  the number of images. The number of rows  $k$  required for arranging  $r_i$  images is calculated as  $k = \lceil r_i/n \rceil$ , which is an integer and satisfies the following equation.

$$(k - 1)n < r_i \leq kn, \quad k \geq 1. \quad (1)$$

In addition, let  $l$  and  $b$  denote the quotient and remainder of dividing  $kn$  by  $r_i$ , respectively. We have

$$kn = lr_i + b, \quad 0 \leq b < r_i. \quad (2)$$

To prove that a way of scheduling panels without panels across two rows exists, we first give Lemma 1 and Lemma 2.

**Lemma 1.** If  $k \geq 2, l = 1$ .

*Proof.* By Eq. (1),  $r_i \leq kn$ , therefore,  $l \geq 1$ . If  $l \geq 2, kn \geq 2r_i$  is satisfied by Eq. (2). By Eq. (1),

$$kn \geq 2r_i > 2(k-1)n. \quad (3)$$

Hence,

$$k > 2k - 2. \quad (4)$$

The above inequation is not true when  $k \geq 2$ . Thus,  $l < 2$  is proved by contradiction. Since  $l$  is an integer by definition,  $l = 1$ .  $\square$

**Lemma 2.**  $b \leq n - 1$ .

*Proof.* Assume  $b \geq n$ . By Eq. (1),

$$kn = r_i + b \geq r_i + n. \quad (5)$$

Then,

$$(k-1)n \geq r_i. \quad (6)$$

It conflicts with Eq. (1). Hence,  $b < n$ . Since  $b$  is an integer by definition,  $b \leq n - 1$ .  $\square$

Based on Lemma 1 and Lemma 2, we can derive Proposition as follows.

**Proposition.** For any  $n$  and any  $r_i$ , a way of scheduling panels without panels across two rows exists.

*Proof.* When  $k = 1$ , no panel that acrosses two rows exists, since panels are set in one row. When  $n = 1$ , no panel that acrosses two rows exists, since one panel is set in each row. When  $k \geq 2, n \geq 2$ , the remining space of  $b$  panels can be space is filled by doubling the width of  $b$  images selected from  $r_i$  images.  $r_i - b$  images are remining panels of width 1. As  $r_i - b$  equals to  $kn - 2b$  by Eq. (2) and Lemma 2,  $b$  panels of width 2 and  $kn - 2b$  panels of width 1 should be set in  $kn$ . Now, by Lemma 2,

$$kn - 2b \geq kn - 2(n-1). \quad (7)$$

Since  $n \geq 2$ ,

$$kn - 2(n-1) = (k-2)n + 2 \geq 2k - 2. \quad (8)$$

Furthermore, since  $k \geq 2$ ,

$$2k - 2 \geq k.$$

Hence,  $kn - 2b$ , or the number of panels of width 1, is larger than  $k$ , the number of rows.

$b$  panels of width 2 can be arranged in the following ways. When  $n$  is odd, one panel of width 1 is set at either the leftmost or the rightmost of each row. The remining space can be divided into the slots of width 2, which can be filled with panels of width 2 or pairs of panels of width 1. When  $n$  is even, the whole space can be divided into the slots of width 2, which can be filled with panel of width 2 or pairs of panels of width 1.  $\square$