Evaluation of a System to Analyze Long-Term Images from a Stationary Camera

Akira Ishii[✉], Tetsuya Abe, Hiroyuki Hakoda, Buntarou Shizuki,
and Jiro Tanaka

University of Tsukuba, Tsukuba, Japan
{ishii,abe,hakoda,shizuki,jiro}@iplab.cs.tsukuba.ac.jp

Abstract. Recording and analyzing images taken over a long time period (e.g., several months) from a stationary camera could reveal various information regarding the recorded target. However, it is difficult to view such images in their entirety, because the speed at which the images are replayed must be sufficiently slow for the user to comprehend them, and thus it is difficult to obtain valuable information from the images quickly. To address this problem, we have developed a heatmap-based analyzing system. In this paper, we present an experiment conducted using our analyzing system to evaluate the system and identify user processes for analyzing images provided by a stationary camera. Our findings should provide guidance in designing interfaces for the visual analytics of long-term images from stationary cameras.

Keywords: Data visualization · Big data management · Evaluating information · Information presentation · Heatmap · Surveillance system · Visual analytics · Lifelog

1 Introduction

Recording and analyzing images over a long time period (e.g., several months) from a stationary camera could reveal various information regarding the recorded target. For example, if department store staff members install a stationary camera to produce aerial images of a floor, then the recorded images can provide useful data for evaluating the layout of the floor. However, it is difficult to view such images in their entirety, because the speed at which the images are replayed must be sufficiently slow for the user to comprehend them, and thus it is difficult to obtain valuable information from the images quickly.

To address this problem, a considerable amount of research has explored the analysis of images from stationary cameras based on image recognition techniques, with the aim of revealing specific information [5,7,14]. On the other hand, we can focus on visual analytics [12,13] to make discoveries by observing unknown objects or phenomena that have not been established beforehand.

With this motivation, we have developed an analyzing system [8,9] with a heatmap-based interface, designed for performing visual analytics on long-term

Fig. 1. Omni-directional camera image.
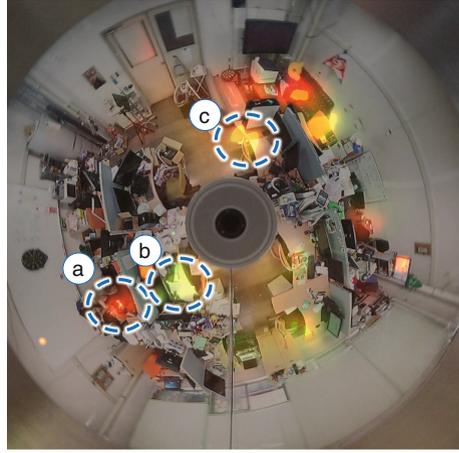
Fig. 2. Image consisting of two heatmaps: (a) a red colored heatmap representing changes of the images in a specific timeframe and (b) a green colored heatmap representing another timeframe. The image contains areas (c) overlaid with red and green colored heatmaps. (Color figure online)

images from a stationary camera (e.g., Fig. 1). This allows the user to analyze long-term images by displaying periods in which the images are changing. It also provides a heatmap that represents the changes to images within a specific timeframe. This heatmap serves as a summary of changes taking place within the timeframe. In addition, this system allows the user to compare two different timeframes by displaying two heatmaps (Fig. 2).

In this paper, we improve our heatmap-based analyzing system [9], and conduct an experiment with our system to evaluate it and identify user processes for analyzing images provided by a stationary camera. The result of our experiment shows that the participants can discover many facts regarding the recorded target quickly (an average of 24 discoveries in 30 min). We also observe that the discoveries can be classified according to five properties. Our findings in this paper are summarized as follows:

- Five properties that can classify the discoveries that the participants obtain using each function of our analyzing system.
- The revelation of the participants' analyzing processes that lead to these five properties.

Those findings should provide guidance in designing interfaces for the visual analytics of long-term images from stationary cameras.

2   Related Work

Interfaces for analyzing images from stationary cameras have recently been explored. Romero et al. proposed Viz-A-Vis, which displays 3D heatmaps [10], and evaluated their system [11]. Their visualization system is different from ours, but we use their evaluating method as a reference. Viz-A-Vis provides 3D heatmaps that summarize the movement of people and objects within a certain timeframe. In contrast, our system provides 2D heatmaps. These 2D heatmaps let the user know where and when changes have frequently occurred during certain timeframes, and provide an easy comparison of two different timeframes. This allows the user to locate events of interest from the images. TotalRecall [3] focuses on transcribing and adding annotations on audio and video recorded at the same time for a hundred thousand hours. While the visualization of their system is similar to ours, ours focuses on comparing two different timeframes using different colored heatmaps. HouseFly [2] presents audio-visual data recorded in several rooms simultaneously using multiple cameras. Their system generates heatmaps and projects the heatmaps onto a 3D model of the recorded space. MotionFinder [1] generates a heatmap as a summary of the images recorded by a surveillance camera, which shows traces of movements across the scene. While that system is similar to ours in generating heatmaps, our research focuses on user discoveries that are obtained by observing heatmaps, and on the processes leading to such discoveries.

Image analyzing methods using crowdsourcing have also been recently explored. Zensors [4] detects objects in images from a stationary camera using crowdsourcing, and notifies users of changes in an image. While Zensors employs crowdsourcing to analyze the images for a specific purpose, our system allows users to analyze images by observing heatmaps by themselves for visual analytics.

Furthermore, automated image analyzing methods have been explored. VERT [6] is a technique to evaluate a summary of an automatically generated video by comparing it with one made by users. By contrast, we provide an analyzing system to users and evaluate the discoveries users obtain with our system.

3   Implementation

This section describes a specification of our system. Our system consists of a recording system and an analyzing system. The recording system obtains images from two stationary cameras that were mounted on the ceiling of the authors' laboratory rooms, preprocesses the images for the generation of a heatmap, and stores the images to NAS (network attached storage). The analyzing system generates heatmaps using the stored images, and presents the heatmaps to users.

3.1   Recording System

We run the recording systems that store the images from omni-directional cameras (Sharp Semiconductor LZ0P3551) mounted on the ceiling of our laboratory (two rooms with sizes of approximately $7.50\,\mathrm{m}\,\times\,7.75\,\mathrm{m}\,(58\,\mathrm{m}^2)$ and
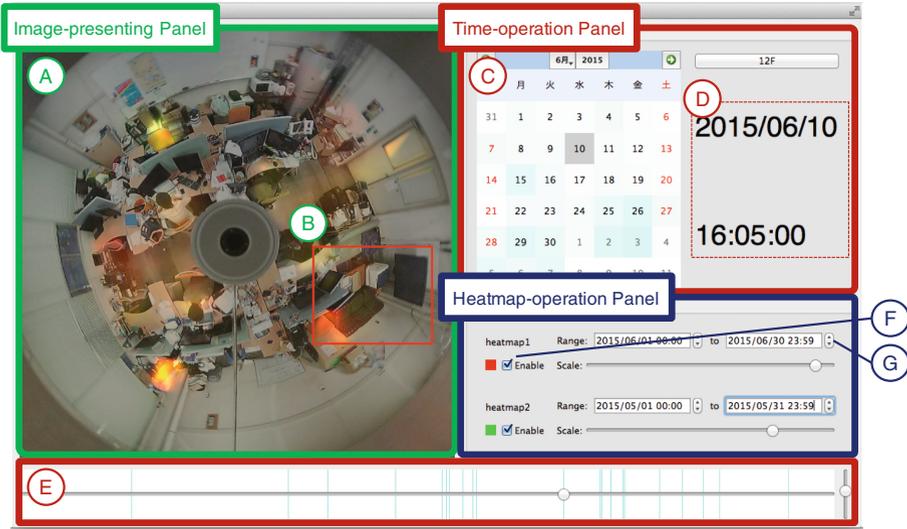
Fig. 3. Our analyzing system using heatmaps.

7.5 m × 15.0 m (113 m²), and heights of approximately 2.5 m and 2.7 m, respectively) with a 608 × 608 pixels spatial resolution at 1 fps (frames per second). This frame rate is a frequently employed in the video archives of surveillance systems, and results in the recording system producing 86,400 frames per day. The recorded images are stored to NAS (QNAP TS-859 and TS-859 Pro+). The recording systems run on two computers (MacBook Pro 13-inch Late 2011 and MacBook Pro Retina 15-inch Early 2013).

3.2   Analyzing System

Our analyzing system generates heatmaps using the images stored on NAS, and presents the heatmaps to users. Figure 3 illustrates our analyzing system, which consists of Image-presenting Panel, Time-operation Panel, and Heatmap-operation Panel.

Image-Presenting Panel. A camera image view (A) displays a camera image at the date and time (D). Users can select a part (B) of the image (A) for further analysis.

Time-Operation Panel. The system applies blue color to the calendar (C) and the time slider (E), with the density depending on the amount of the movement in the area (B). This function allows users to find a range that they wish to analyze, and reduces the time used to analyze unnecessary images.

Heatmap-Operation Panel. Our system displays two different heatmaps with two different colors (red and green), each of which can be turned on/off

using the two checkboxes (F). Users can specify date and time ranges for the heatmaps using the date/time range pickers (G). Activated heatmaps are overlaid with the camera image view (A), as shown in Fig. 2.

Our analyzing system summarizes the movement of people and objects in a specified timeframe based on the number of changes of pixels in the camera images. The more movement there is in the image (A), the more densely the pixel is colored. The more movement there is in the area (B), the more densely the calendar (C) and the time slider (E) are colored. Therefore, our system allows users to recognize areas with little or much movement within a specified timeframe at a glance. Moreover, users can compare movement in two different timeframes by using two heatmaps.

4   Experiment

We conducted an experiment to examine which discoveries users obtained and how, using each function provided by our analyzing system.

4.1   Participants

Four participants (three males, one female) aged between 22 and 23 were recruited for the experiment. Note that the rooms recorded by the stationary cameras consisted of the laboratory of the participants. None of the participants had previously used our system, nor did they have prior knowledge regarding our system.

4.2   Apparatus and Experimental Environment

We employed a MacBook Pro 13-inch Mid 2010 (CPU: 2.4 GHz Core 2 Duo, RAM: 4 GB, OS: Mac OS X 10.9.5) as the computer for running our analyzing system. We recorded the whole experiment with a video camera, a voice recorder, and screen capture software (QuickTime Player 10.3).

4.3   Procedure

First, we informed the participants of the purpose and the procedure of the experiment. We then informed the participants that the reward for participation in the experiment included not only a basic reward, but also a bonus depending on the number of discoveries. After the basic explanation, we explained the use of our analyzing system to the participants. We then asked them to engage with the system until they felt that they completely understood how to use it, as a practice. We employed the images in which the participants were not recorded for the practice.

In the analyzing task, we asked the participants to use our analyzing system for 30 min, during which time they should attempt to make as many discoveries

as possible and inform the experimenters about each of these using think aloud protocol. In addition, we asked the participants to inform the experimenters of the facts that led them to each discovery (e.g., the color of the heatmap is dense in certain areas, as described in Sect. 5.3). In this analyzing task, we used images that were recorded over six months (July 1, 2014 to December 31, 2014; 4,416 hours; approximately 32 million images) using the stationary camera. Note that the participants were recorded in the images for this analyzing task.

After the analyzing task was completed, we asked the participants to answer a questionnaire related to our system. The experiment took approximately 60 min in total.
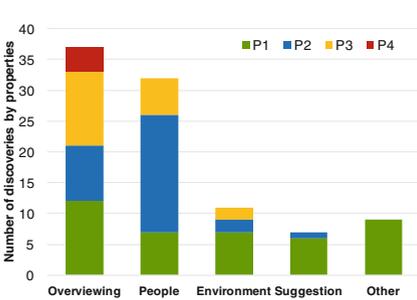

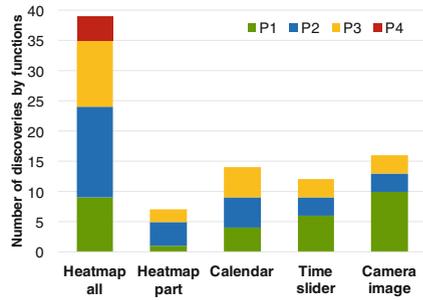**Fig. 4.** Number of discoveries by properties. (Color figure online)

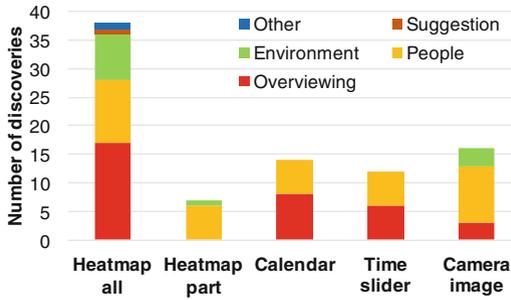**Fig. 5.** Number of discoveries by functions. (Color figure online)

Fig. 6. Number of discoveries by functions. (Color figure online)

5   Result

5.1   Discoveries and Classification

In the experiment, the participants (P1–P4) had an average of 24 discoveries (Total = 96, SD = 13.7). P1 had 41 discoveries, P2 had 31 discoveries, P3 had

20 discoveries, and P4 had four discoveries. Each participant obtained discoveries that were related to her/his colleagues or situations regarding the room. Furthermore, three participants obtained discoveries that related to themselves. We classified the discoveries by the following five properties, with reference to [11]. The result is presented in Fig. 4.

Discovery

    Overviewing. Discoveries obtained by paying attention to the entire image.

    People. Discoveries obtained by paying attention to the people in recorded images.

    Environment. Discoveries obtained by paying attention to the environment (e.g., objects in recorded images and changes in the appearance of the room).

Discovery Related to the System

    Suggestion. Opinions/discoveries that are related to the analyzing system (e.g., requests to extend functionality, proposals of a new function, and ideas to improve the system).

    Other. Other opinions/discoveries (e.g., suggestions for applications of our system).

We also classified the discoveries by the five functions of our analyzing system by considering which function the participants used when they obtained discoveries (Fig. 5). In this classification, one discovery is classified into multiple functions if participants used more than one functions for the discovery.

Heatmap/All. Discoveries obtained by paying attention to the whole heatmap, (A) in Fig. 3.

Heatmap/Part. Discoveries obtained by paying attention to a part of the heatmap, (B) in Fig. 3.

Calendar. Discoveries obtained by paying attention to the color of the calendar, (C) in Fig. 3.

Time Slider. Discoveries obtained by paying attention to the color of the time slider or comparing different images by operating the time slider, (E) in Fig. 3.

Camera Image. Discoveries obtained by paying attention to the camera image, (A) in Fig. 3.

As shown in Figs. 4 and 5, we found that the discoveries depended on the participant or the function used, based on the color pattern of each chart. P1 made more discoveries using Time slider and Camera image than other participants. P1 realized that movement occurred in a specific area from a heatmap, and then observed images in that area. P1 made many discoveries that were classified under Suggestion or Other, thus providing ideas for potential applications of our system. Figure 4 suggests that only P2 made more People discoveries than Overviewing ones. This is because P2 made more discoveries using Heatmap/part function than other participants, as shown in Fig. 5. The main

reason that P2 made more discoveries using Heatmap/part function was that P2 used two heatmaps to compare different timeframes several times. The distribution of the functions used by P3 was similar to that for P2. However, the distribution of the properties was different. This is because P3 tended to make discoveries by seeing things roughly (e.g., there were many people at a certain timeframe) using Heatmap/all function. As shown in Figs. 4 and 5, the number of discoveries by P4 was extremely low. Because all of the discoveries by P4 were classified into Overviewing and obtained using Heatmap/all function, we might fail to explain P4 sufficiently how to use our system or the intent of the experiment. In addition, from the videos that we recorded of the whole experiment using a video camera, we note that P4 hardly used any of the functions except for Heatmap/all.
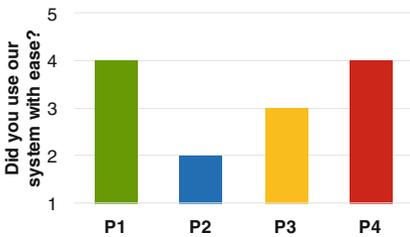

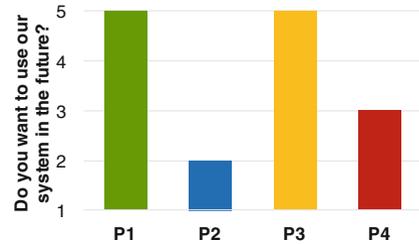**Fig. 7.** Did you use our system with ease?

Fig. 8. Do you want to use our system in the future?

Figure 6 shows the number of discoveries classified by functions. Heatmap/all was commonly used. Heatmap/part and Camera image were used mainly to make People discoveries. Calendar and Time slider were only used for Overviewing or People discoveries.

5.2   Qualitative Results

We asked participants to complete a questionnaire that consisted of two questions: "*did you use our system with ease?*" and "*do you want to use our system in the future?*" Each question included a five-point Likert scale styled form (1 = strongly disagree, 5 = strongly agree) and a comment form. The results are presented in Figs. 7 and 8. As shown in Figs. 7 and 8, the scores provided by P2 were lower than for the others. While P2 skillfully use all functions, he stated that "*the analyzing was fun for me, but I do not conceive application examples*" in the questionnaire. Furthermore, P2 provided many requests for extending the functionality, proposals for new functions, and ideas to improve our system. On the other hand, P1 and P3 awarded relatively high scores. From the videos that we recorded of the whole experiment and the comment form of the questionnaire, we conclude that the participants enjoyed looking back on and analyzing their research days.
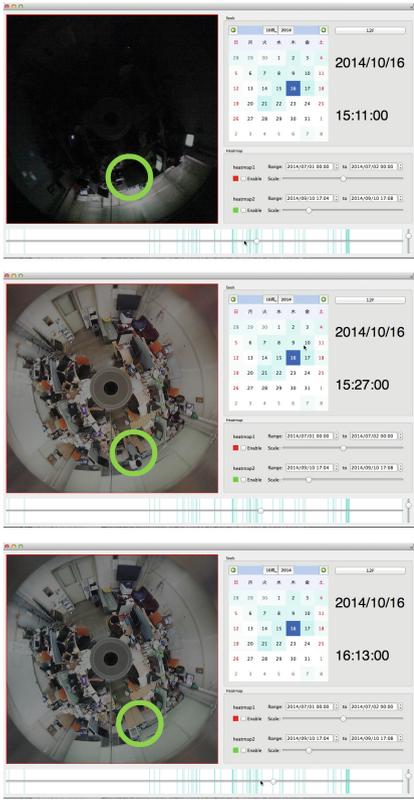
Fig. 9. P1 was aware of the time when a student placed something on the shelf in the laboratory.
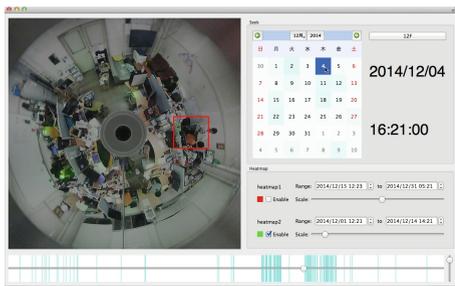

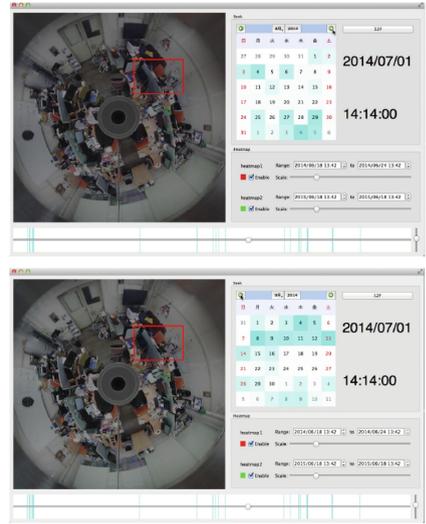**Fig. 10.** P2 was aware that he often stood up from his seat when he was in the laboratory.

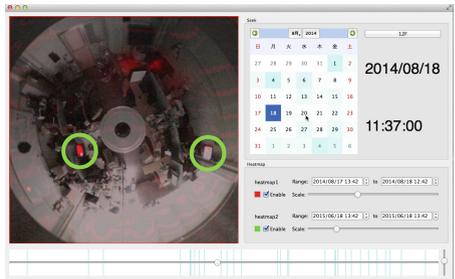**Fig. 11.** P3 was aware that he was frequently in the laboratory from late August to early September.

Fig. 12. P3 was aware that the screens of these two monitors were frequently changing.
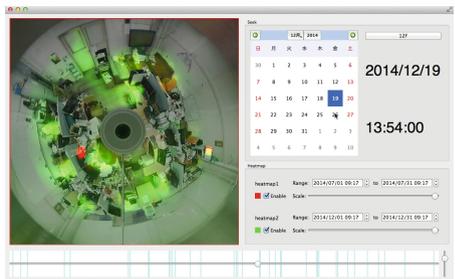

Fig. 13. P4 was aware that some students were in the laboratory at the end of the year.

5.3   Analyzing Processes

In this section, we describe some of the analyzing processes of the participants. To examine each participant's analyzing processes, we analyzed the screen captures. We found that all of the participants first browsed the images recorded in July, and then browsed the images recorded in August and later in the sequence. After that, each participant acted differently.

P1 discovered that at around 16:00 on October 16th a person placed an object on the shelf in the laboratory, as shown in Fig. 9. We classified this discovery as *People–Camera image.* P1 browsed the images recorded from July to December, and discovered that an object was placed on the shelf in the laboratory at a particular time (the green circles in Fig. 9). In order to reveal the time, P1 first used Calendar function, and then revealed the date. Next, P1 used Time slider function to find that there was no object present at 15:11, and found that the person began placing the object at 15:27 and that the process of placing the object was completed at 16:13. Thus, P1 arrived at the conclusion stated above.

P2 discovered that he often left his seat while he was in his laboratory, as shown in Fig. 10. We classified this discovery as *People–Time slider.* P2 first selected the area of his seat within the image, as shown in Fig. 10. Then, P2 used Calendar function, and explored each timeframe in which he was in his laboratory. As a result, P2 noticed that the blue part of Time slider was not continuous, but rather discrete, and concluded as above. Note that this discovery reveals that our tool is useful for self-behavioral analysis, because this discovery by P2 was related to himself.

P3 discovered that he was in his laboratory more frequently between late August and early September than other timeframes, as shown in Fig. 11. We classified this discovery as *People–Calendar and Heatmap/part.* P3 first selected the area of his seat within the image, as well as P2. Then, P3 used Calendar function, browsed the images recorded from July to December, and noticed that the color of the calendar was dense between late August and early September. Therefore, P3 concluded as above. In addition, P3 noticed that the color of the heatmap was dense in certain areas (the green circles in Fig. 12), and discovered that there were computer monitors in those areas. We classified this discovery as *Environment–Heatmap/all.* After P3 inspected the calendar, he noticed that the color of the heatmap was dense in the areas where the generating timeframe of the heatmap was one day, and concluded as above. Moreover, P3 discoveried that one of the computer monitors displayed a screen saver, while the other displayed a clock.

P4 discovered that there were many students present at the end of the year, as shown in Fig. 13. We classified this discovery as *Overviewing–Heatmap/all.* P4 browsed the images recorded in December, and examined the calendar. P4 set the generating timeframe of the red heatmap to July, and that of the green heatmap to December. At this time, P4 compared the two heatmaps, and she noticed that the density of the green colored heatmap was higher. Therefore, P4 concluded as above.

6   Discussions

In our experiment, we employed images that recorded the participants, in order to ensure consistency in the experimental condition of the participants. As a result, there were many discoveries made that related to the participants themselves. To be precise, 11 of the 96 discoveries (approximately 11.4 %) consisted of such discoveries. Two participants (P3 and P4) stated in the questionnaire that "*I looked back on my life pattern, and my motivation to go to the laboratory increased.*" Therefore, we surmise that our system is useful for analyzing the users themselves.

There was a bias present in the functions of our system that the participants in the experiment chose to use (there was one participant who did not use all of the functions). Therefore, we propose that we should limit the available functions depending on the purpose of the analysis. For example, if a user wants to perform an analysis regarding Environment, then only Heatmap/all and Camera image functions should be provided, considering Fig. 6. In addition, we plan to explore to possibility of reusing the analyzing processes that we found in our research, to provide a wizard that is specialized for each analyzing purpose. As a result, the wizard would enable users to perform an analysis without having a deep knowledge of our system.

7   Future Work

In our experiment, we used images in which the participants were recorded. Because this is a particular situation, we will conduct a further experiment using images in which the participants have not been recorded, and reveal which participants obtain discoveries in such a situation.

All of the participants that were recruited for the experiment had a computer science background. Therefore, we will recruit participants that have different backgrounds to conduct a further experiment examining which discoveries they will make using our system and how.

In the experiment, we used images recorded over only a six-month period. However, because we also have images recorded over a period of more than 20 months and continue to record images, we plan to conduct a further experiment using the longer-term images. In addition, we plan to apply our system with images recorded in different locations (e.g., a hallway or a large shared room).

8   Conclusion

In this paper, we have improved our system for recording and analyzing images using a stationary camera. In addition, we have conducted an experiment to evaluate our analyzing system, and examined what participants discover using each function of the system. The result of the experiment was that the participants made an average of 24 discoveries, which we classified by five properties and by

five functions of the system. Furthermore, we revealed the analyzing processes of the participants. We believe that those findings provide guidance in designing interfaces for the visual analytics of long-term images from stationary cameras.

References

1. Buono, P.: Analyzing video produced by a stationary surveillance camera. In: Proceedings of the International Conference on Distributed Multimedia Systems, DMS 2011, pp. 140–145 (2011)
2. DeCamp, P., Shaw, G., Kubat, R., Roy, D.: An immersive system for browsing and visualizing surveillance video. In: Proceedings of the International Conference on Multimedia, MM 2010, pp. 371–380. ACM, New York, NY, USA (2010)
3. Kubat, R., DeCamp, P., Roy, B.: TotalRecall: visualization and semi-automatic annotation of very large audio-visual corpora. In: Proceedings of the 9th International Conference on Multimodal Interfaces, ICMI 2007, pp. 208–215. ACM, New York, NY, USA (2007)
4. Laput, G., Lasecki, W.S., Wiese, J., Xiao, R., Bigham, J.P., Harrison, C.: Zensors: Adaptive, rapidly deployable, human-intelligent sensor feeds. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, CHI 2015, pp. 1935–1944. ACM, New York, NY, USA (2015)
5. Leykin, A.: Visual human tracking and group activity analysis: a video mining system for retail marketing. Ph.D. thesis, Department of Computer Science and Cognitive Science, Indiana University (2007)
6. Li, Y., Merialdo, B.: VERT: automatic evaluation of video summaries. In: Proceedings of the International Conference on Multimedia, MM 2010, pp. 851–854. ACM, New York, NY, USA (2010)
7. Corporation, N.E.C.: Fieldanalyst. http://www.nec-solutioninnovators.co.jp/sl/fieldanalyst/. Accessed 1 Feb 2016. (in Japanese)
8. Nogami, R., Shizuki, B., Hosobe, H., Tanaka, J.: An analysis support interface using frame difference for a video from a stationary camera. In: Proceedings of the Interaction 2011. Information Processing Society of Japan (2011). (in Japanese)
9. Nogami, R., Shizuki, B., Hosobe, H., Tanaka, J.: An exploratory analysis tool for a long-term video from a stationary camera. In: Proceedings of the 5th IEEE International Symposium on Monitoring and Surveillance Research, ISMSR 2012, vol. 2, pp. 32–37 (2012)
10. Romero, M., Summet, J., Stasko, J., Abowd, G.: Viz-A-Vis: toward visualizing video through computer vision. IEEE Vis. Comput. Graph. **14**(6), 1261–1268 (2008)
11. Romero, M., Vialard, A., Peponis, J., Stasko, J., Abowd, G.: Evaluating video visualizations of human behavior. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 2011, pp. 1441–1450. ACM, New York, NY, USA (2011)
12. Thomas, J.J., Cook, K., et al.: A visual analytics agenda. IEEE Comput. Graph. Appl. **26**(1), 10–13 (2006)
13. Wong, P.C., Thomas, J.: Visual analytics. IEEE Comput. Graph. Appl. **24**(5), 20–21 (2004)
14. Xing, Y., Wang, Z., Qiang, W.: Face tracking based advertisement effect evaluation. In: The 2nd International Congress on Image and Signal Processing, pp. 1–4 (2009)