Visual Analysis of Outdoor Surveillance Videos Using Principal Component Analysis

Kaoru Sugita

Fukuoka Institute of Technology, Japan

Received: November 30, 2025; Revised: January 21, 2026; Accepted: February 22, 2026; Published: March 06, 2026

Abstract

In this paper, we investigated the application of principal component analysis (PCA) to the visualization of outdoor videos for safety and security monitoring. We analyzed videos depicting daytime airplane takeoffs and landings, nighttime airplane landings, small birds flying, and small birds resting on an elevated bridge. The extracted frames were arranged on a two-dimensional plane according to their principal component scores. The results indicate that frame placement reflects inter-frame correlations and is strongly influenced by global visual factors such as illumination conditions and the size of moving subjects. A qualitative evaluation suggests that the visualization provides an intuitive overview of frame-level variations and reduces the workload required for scene exploration without continuous playback. However, the effectiveness decreases for scenes involving small or low-contrast subjects. These findings clarify the characteristics and limitations of PCA-based visualization for outdoor surveillance videos.

Keywords: Principal Component Analysis (PCA), Video Visualization, Outdoor Surveillance, Scene Analysis.

1 Introduction

In recent years, videos have increasingly been recorded in various public and private locations for safety and security purposes, such as monitoring facilities, transportation systems, and outdoor spaces. When an incident occurs, it is necessary to identify the problematic scene based on the information available at the time of the event. In such cases, reviewing the video can require a substantial amount of work, especially for long-duration footage. Therefore, it is necessary to reduce the workload of reviewing video content. With respect to finding specific scenes, many studies have focused on motion and action visualization (Botchen, 2008; Schoeffmann, Lux, Taschwer, & Boeszoermenyi, 2009; Schoeffmann K. T., 2009). Also, (Zhao, 2007) proposed a method for categorizing segmented scenes. Other researchers also have proposed various methods for categorizing video scenes. For example, (Liu, 2002) characterized video frames using a compact feature set created by repeatedly sorting and merging image features.

In our previous works, we developed tools to reduce the workload of reviewing videos with long playback times, including average image compositing software (Sugita, 2022), video segmentation (Sugita, 2024), representative frame extraction from segmented videos (Sugita, 2025), and visualization of subjects and their motion in video using principal component analysis (Sugita, 2025). These approaches make the general motion of a person in indoor environments visually discernible, providing a basis for determining whether playback is necessary to review a particular video or scene. In this paper, we investigate the application of PCA-based visualization to outdoor videos and examine how outdoor environmental factors influence the visualization results. By applying PCA to outdoor videos, we aim to provide a foundation for intuitive scene understanding and anomaly detection in safety and security monitoring.

The paper is organized as follows. Section 2 introduces a method for visualizing video content using principal component analysis. Section 3 presents visualization results, and Section 4 provides a qualitative evaluation. Finally, Section 5 concludes the paper.

2 Visualization Method for Video Content Using Principal Component Analysis

Reviewing video content through simple playback increases the review time proportionally to the video's duration (i.e., the number of frames). In this study, we focus on the differences between frames caused by camera movement and object motions, which are reflected as variations in subject appearance across frames.

As shown in Figure 1, we propose a method for visualizing subjects and their motion by arranging the frames that compose a video in a spatial layout based on their correlations. PCA is a technique that represents a large number of correlated variables using a smaller number of composite variables, called principal components. In PCA, the variable space is compressed along new axes (principal components) so that as much of the original information as possible is retained. In conventional PCA, a covariance matrix is calculated from normalized data, and principal components are derived from the corresponding eigenvalues and eigenvectors. Since the covariance matrix of normalized data is equivalent to the correlation matrix, this study employs PCA to reduce the dimensions of the correlation matrix to two dimensions. This enables the visualization of subject motion by arranging video frames in a two-dimensional space based on their inter-frame similarity. Specifically, PCA is applied to the correlation matrix obtained from video frames to calculate the principal component scores for each frame.

In this method, a correlation matrix R ($n \times n$) is used to represent the relationships among video frames.

A composite variable Y is defined as a linear combination of the columns of R with coefficients a derived from the eigenvectors of R, as expressed below:

$$Y = a_1R_1 + a_2R_2 + ... + a_nR_n$$

Here, R_i is the i-th column of the correlation matrix R, and the coefficient vector a corresponds to the eigenvector of R.

The first principal component Y_1 represents the direction that maximizes the variance of the correlation matrix R. It is obtained using the elements of the eigenvector associated with the largest eigenvalue, as follows:

$$Y_1 = a_{11}R_1 + a_{21}R_2 + ... + a_{n1}R_n$$

Similarly, the second principal component Y_2 is defined as the composite variable whose coefficients correspond to the elements of the eigenvector associated with the second largest eigenvalue of R:

$$Y_2 = a_{12}R_1 + a_{22}R_2 + ... + a_{n2}R_n$$

Each frame is then plotted in a two-dimensional space, where the first principal component Y1 (PC1) is assigned to the x-axis and the second principal component Y2 (PC2) to the y-axis.

This visualization enables observation of relationships between frames: frames with similar content or motion are positioned closer together, while those with significant differences are placed farther apart. In this way, subject motions can be effectively visualized, providing an intuitive means to review video content.

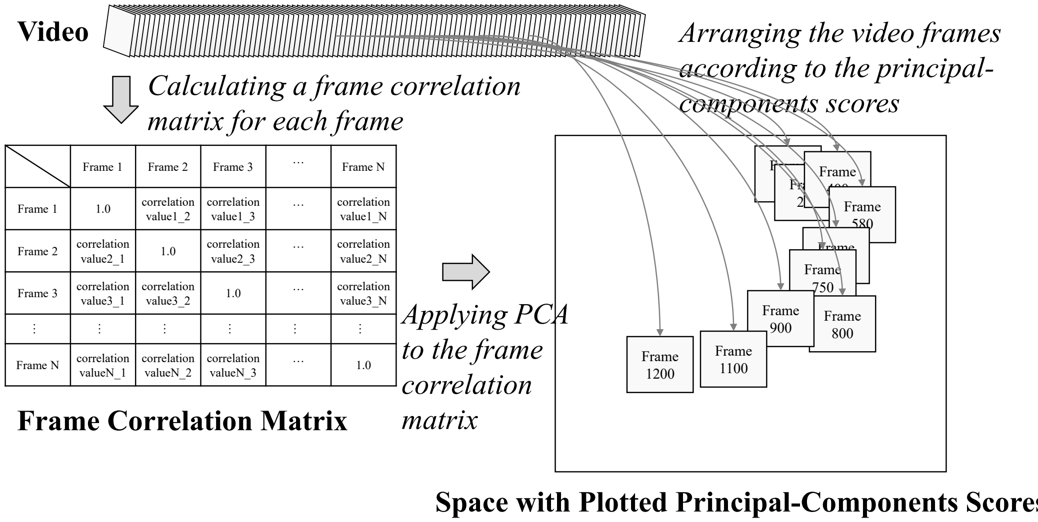

Figure 1. Visualization of the video content using principal component analysis

Figure 2. Outdoor videos

3 Visualization Results

In this visualization, 200 frames were extracted from each of the five outdoor videos shown in Figure 2. These videos were recorded at night and under four different sunlight conditions, with varying subject sizes. All videos were encoded in MP4 format with HD resolution (1280 × 720 pixels) at 30 frames per second, resulting in 1,800 frames per video (1 minute in duration). For each video, 200 frames were sampled at equal intervals, resized to 96 × 54 pixels, and projected onto a two-dimensional space based on their principal component scores. Scatter plots of all extracted frames are shown in Figures 3–7, where PC1 and PC2 are assigned to the x- and y-axes, respectively. The percentage shown on each axis indicates the proportion of variance explained by the corresponding principal component. In addition, frames corresponding to the maximum and minimum values of each principal component were selected as a minimal frame set and plotted in the same two-dimensional space, as shown in Figures 8–12.

Figure 3 presents the scatter plot of all frames from a video of an airplane taking off from a runway. In the original footage, the airplane is visible during takeoff, and variations in sunlight occur due to cloud movement. Frames located in the lower-left region correspond to moments when the airplane is visible, while frames in other regions mainly reflect changes in sunlight on the ground surface.

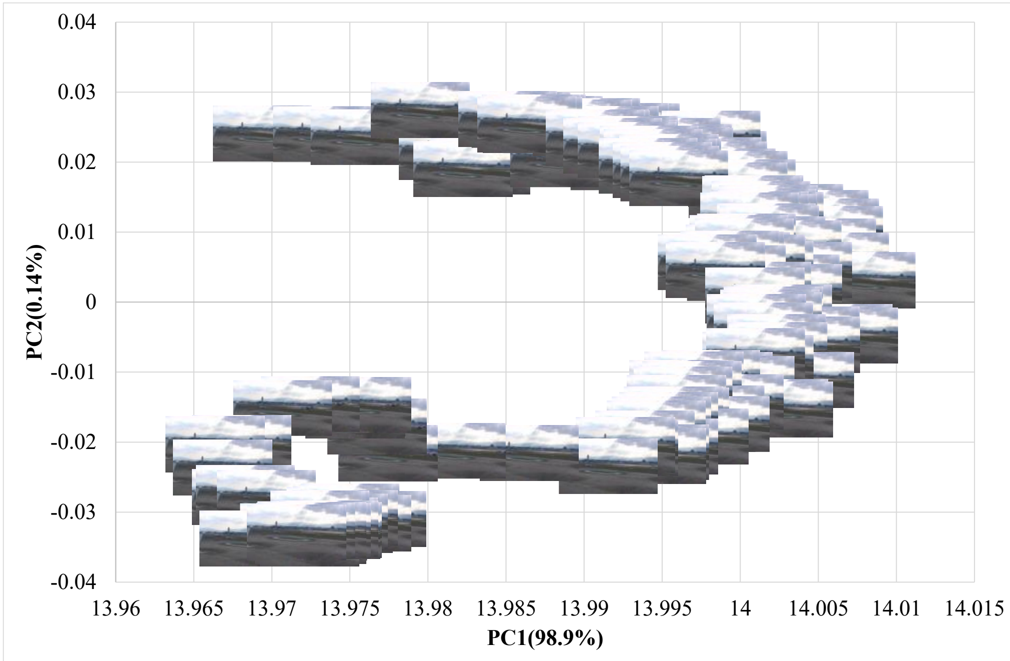

Figure 3. Visualization of Airplane takeoff during daytime

Figure 4 shows all frames from a daytime video of an airplane landing. The airplane is visible during landing, accompanied by gradual variations in sunlight caused by moving clouds. In the scatter plot, frames depicting the landing airplane appear in the left and lower regions, whereas the remaining frames primarily capture background illumination changes.

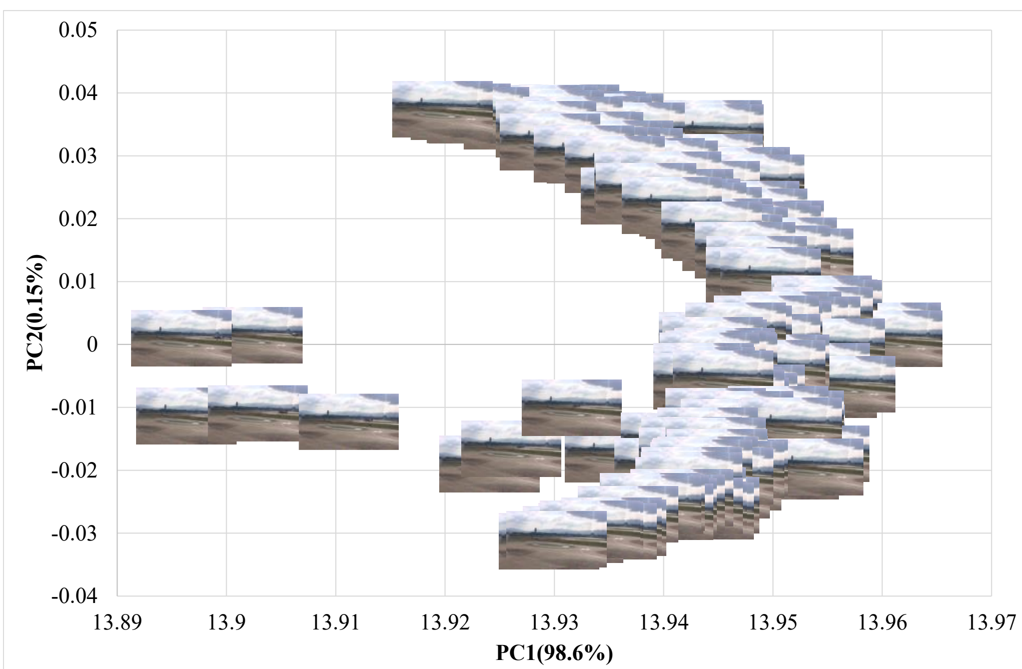

Figure 4. Visualization of Airplane landing during daytime

Figure 5 presents all frames from a nighttime landing video. In the original footage, the airplane moves from right to left across the scene. In the scatter plot, frames showing the landing process tend to be distributed toward the outer regions, while frames without the airplane are located closer to the center.

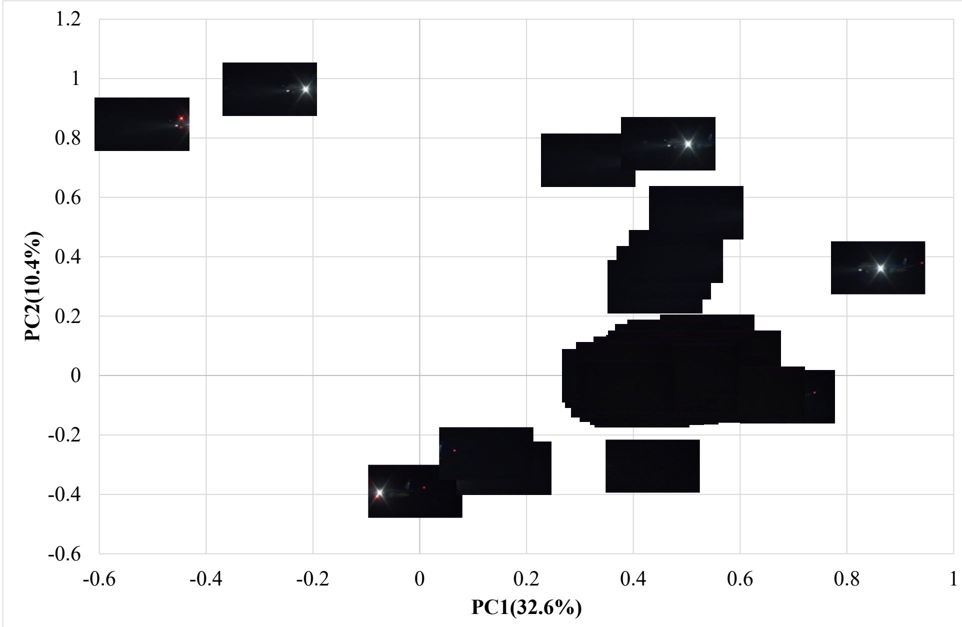

Figure 5. Visualization of Airplane landing during night

Figure 6 shows all frames from a video scene in which small birds appear and briefly fly, together with vehicles moving on an elevated bridge. In the scatter plot, frames depicting vehicles on the bridge appear in the lower-left region. However, frames containing flying birds are dispersed among other frames, and their positions in the principal component space are not clearly distinguishable.

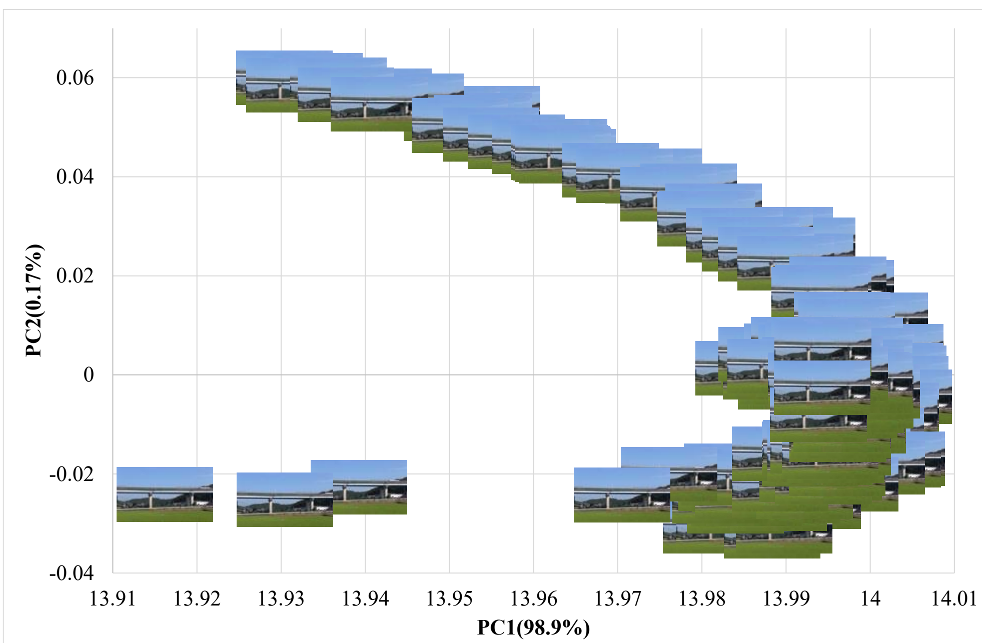

Figure 6. Visualization of small birds flying

Figure 7 shows all frames from a video of small birds resting on an elevated bridge. In the scatter plot, frames depicting resting birds are intermixed with other frames, making their distribution patterns unclear.

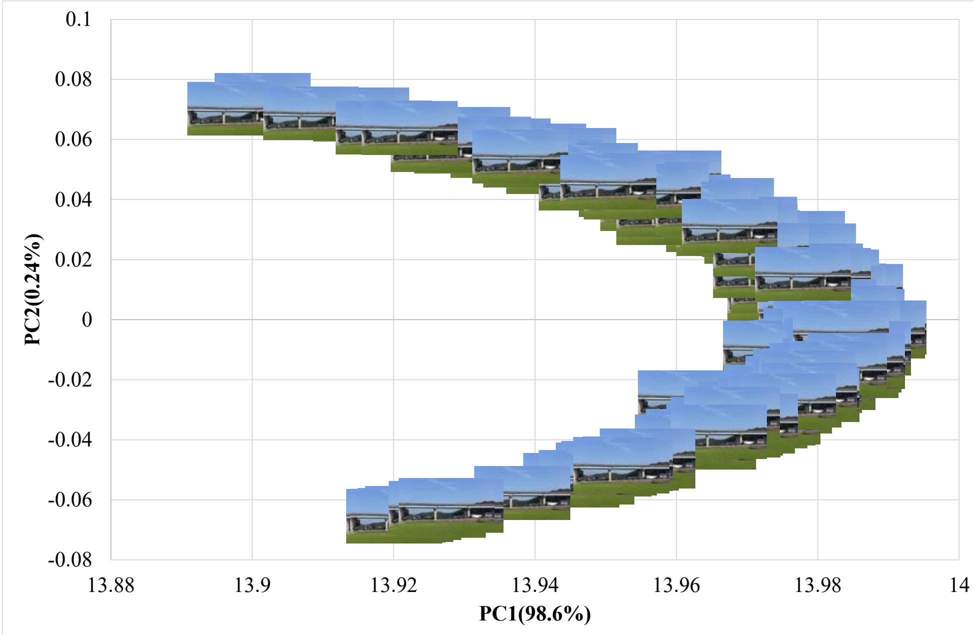

Figure 7. Visualization of small birds resting on an elevated bridge

Figure 8 shows the scatter plot of the minimal frame set corresponding to the maximum and minimum values of each principal component for a video of an airplane taking off. In this case, the frame corresponding to the minimum score of PC1 captures the moment when the airplane begins to take off. The remaining extreme frames mainly reflect variations in background illumination.

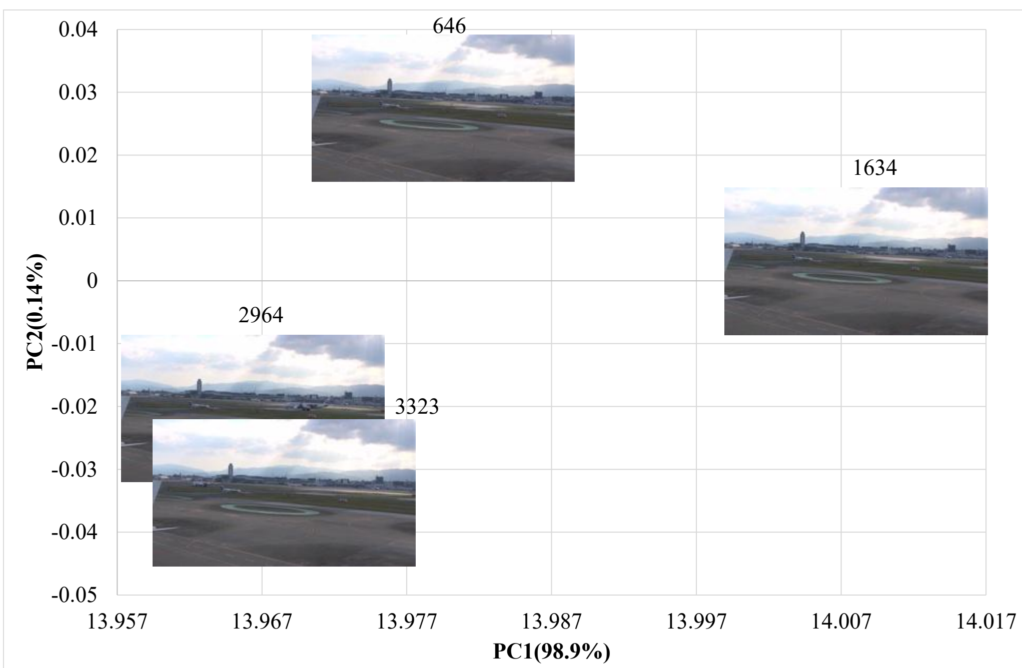

Figure 8. Visualization minimal set of Airplane takeoff during daytime

Figure 9 shows the minimal frame set for a daytime video of an airplane landing. The frame corresponding to the minimum score of PC1 captures the landing moment. In contrast, the other extreme frames primarily represent background elements rather than the airplane itself.

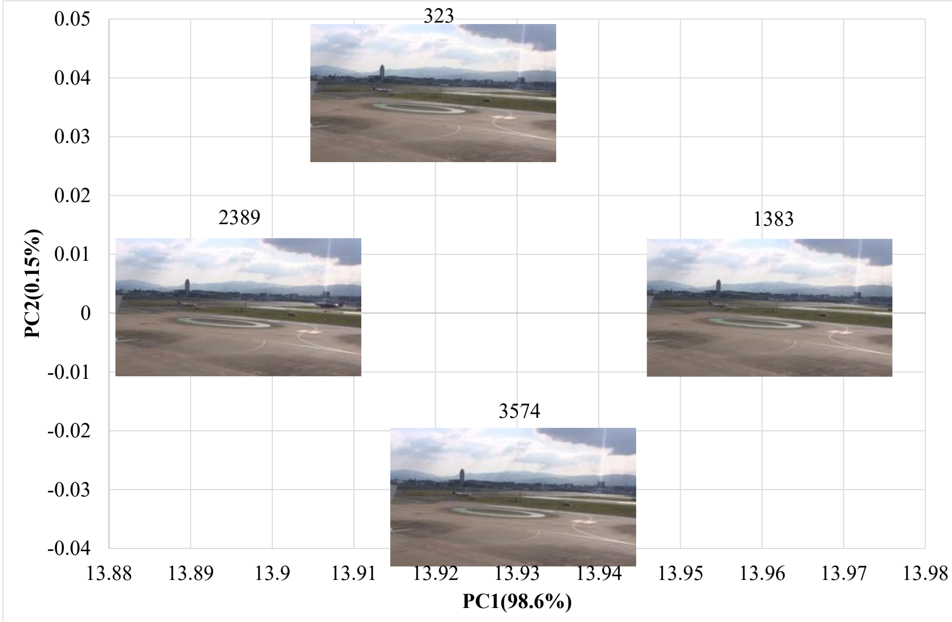

Figure 9. Visualization minimal set of Airplane landing during daytime

Figure 10 shows the minimal frame set for a nighttime landing video. Unlike the daytime example in Figure 9, this scene was recorded under low-light conditions. Here, the frame corresponding to the minimum score of PC1 does not depict the airplane, while the other extreme frames show the airplane at different positions within the scene.

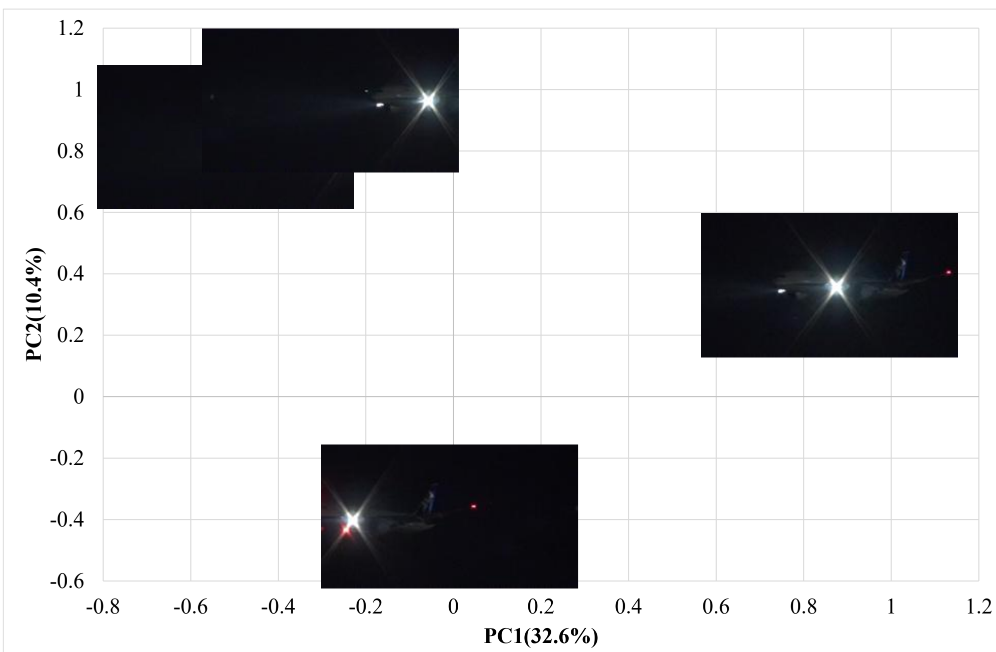

Figure 10. Visualization minimal set of Airplane landing during night

Figure 11 shows the minimal frame set for the same video scene in which small birds appear and briefly fly. The bird is not clearly visible in the selected extreme frames. Because the bird occupies only a small region of the frame, its influence on the overall pixel distribution is likely to be limited. The frame corresponding to the minimum score of PC1 shows vehicles moving on the bridge, and the remaining extreme frames do not exhibit clearly distinguishable differences in scene content.

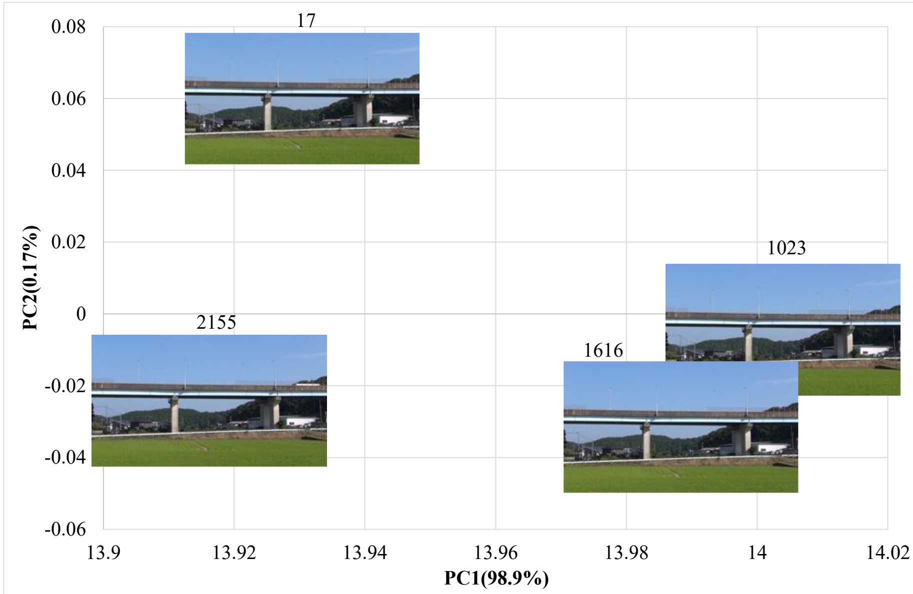

Figure 11. Visualization minimal set of small birds flying

Figure 12 shows the minimal frame set for a video scene in which a bird was perched on the bridge structure. Similar to Figure 11, the bird is not clearly visible in the selected extreme frames. The frame corresponding to the minimum score of PC1 shows vehicles on the bridge, while the remaining extreme frames show no clearly discernible differences in scene content. Even in the original video, the perched bird is difficult to identify without magnified observation.

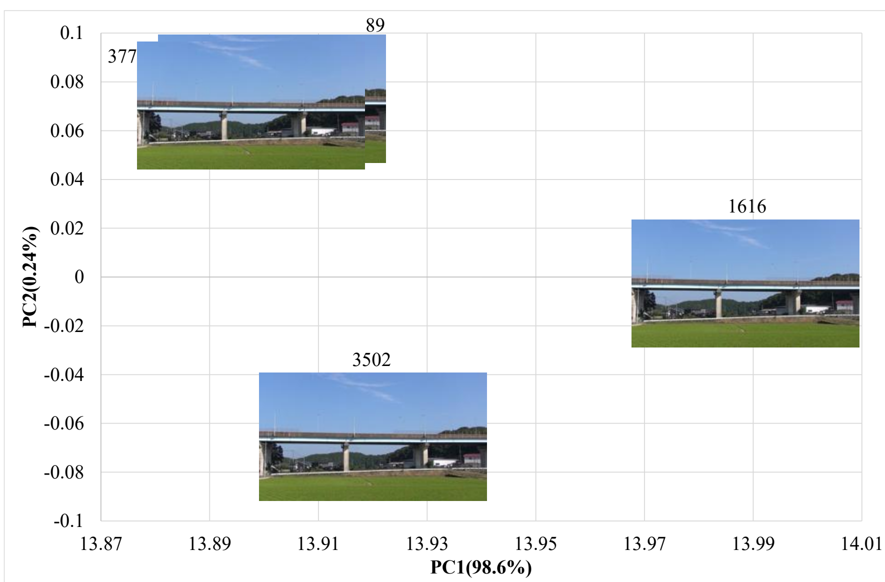

Figure 12. Visualization minimal set of small birds resting on an elevated bridge

Overall, the scatter plots show that frames containing large or high-contrast objects, such as an airplane with its lights on during a nighttime landing, are clearly separated in the principal component space. In contrast, smaller or low-contrast objects, such as daytime airplanes or birds, have little impact on the overall pixel distribution and are therefore difficult to discern in the PCA-based scatter plots. However, daytime airplanes can still be identified with careful observation. Additionally, variations in background illumination, such as sunlight or moving clouds, gradually influence frame positions in the scatter plots, subtly reflecting changes in the scene over time. These observations highlight both the strengths and limitations of using PCA for visualizing dynamic outdoor scenes.

4 Qualitative Evaluation

A qualitative evaluation was conducted with nine students to assess several aspects of video scene analysis, as shown in Figures 13–17. Participants were asked to evaluate (1) the workload involved in finding specific scenes (Figure 13), (2) the visibility of subjects and their motion under video playback without skipping (Figure 14), (3) the visibility of subjects and their motion when navigating the playback position using the seek bar (Figure 15), and (4) the visibility of subjects in PCA-based scatter plots, including both the full 200-frame set (Figure 16) and the minimal frame set containing only the maximum and minimum principal component values (Figure 17). Responses were collected using a questionnaire to capture overall tendencies on a five-point scale.

Figure 13 presents the results on the workload of scene finding. As expected, identifying specific scenes during standard video playback tended to be evaluated as demanding. The workload for seek bar–based navigation was comparable to that of the PCA-based scatter plot using 200 frames. This may be because the playback time was short (one minute), and intermediate frames could be checked without much difficulty. In contrast, the PCA-based scatter plot using only the minimum and maximum frames tended to impose a higher workload. Since only four representative frames were available, it was difficult to follow the intermediate progression of the subject.

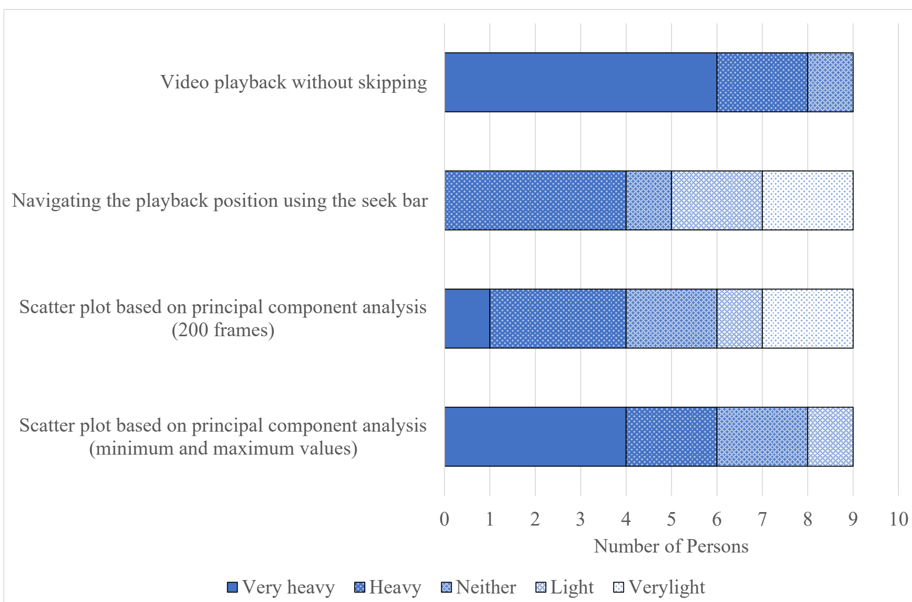

Figure 13. Questionnaire results on the workload of scene finding

Figure 14 shows the results regarding the visibility of subject motion during video playback without skipping. For daytime aircraft takeoff and landing and nighttime aircraft landing, the visibility was generally evaluated as high. For videos of flying small birds, the evaluations were more varied. In the case of small birds perched on an elevated bridge, the visibility tended to be relatively low. These results suggest that even during standard video playback, some subjects are inherently difficult to observe. This is especially true when the subject is small or when the contrast with the background is limited. Although daytime aircraft do not have the strong contrast seen in nighttime scenes, their relatively large size makes them easier to track than small birds.

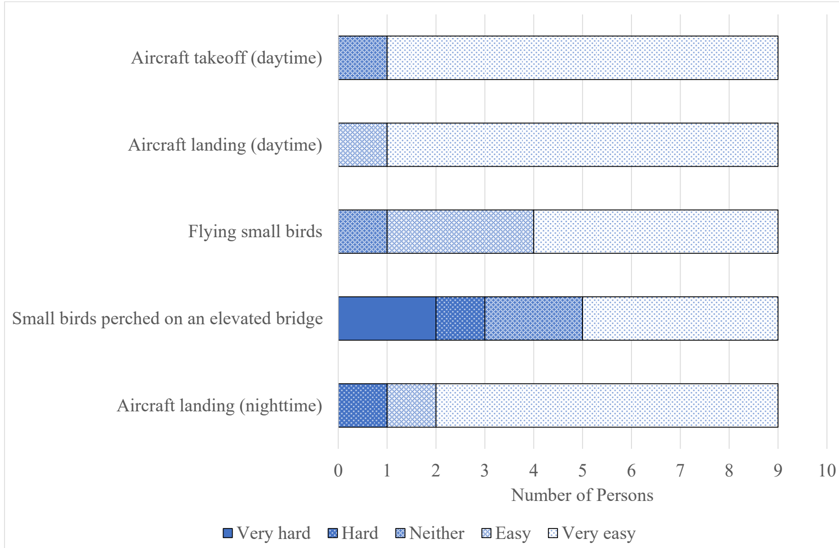

Figure 14. Questionnaire Results on the Visibility of Subject Motion in Video Playback Without Skipping

Figure 15 presents the results on the visibility of subject motion when navigating the playback position using the seek bar. The overall tendency was similar to that observed during standard video playback. Large subjects such as aircraft were generally easy to track. Small or low-contrast subjects, particularly perched birds, remained difficult to observe. These findings indicate that seek-bar navigation does not substantially change the relative difficulty caused by subject size. Even when the playback position can be adjusted freely, visibility limitations due to small subject size cannot be fully resolved.

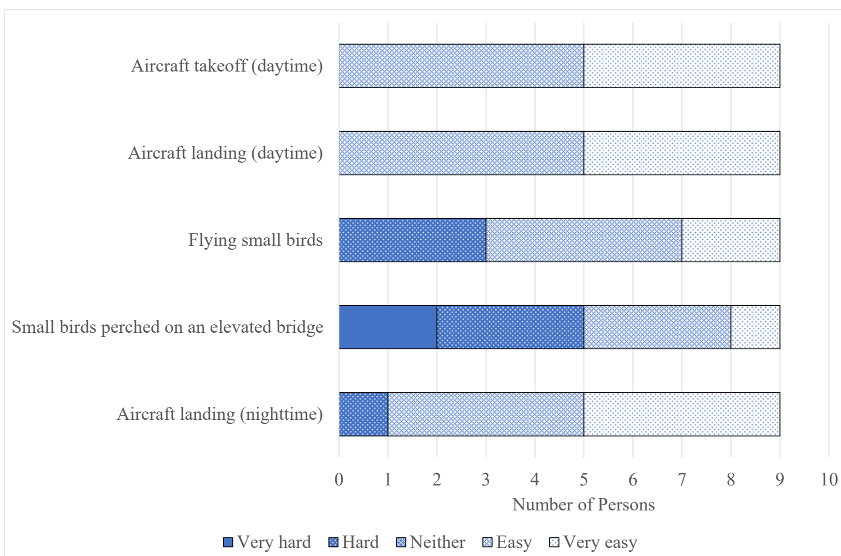

Figure 15. Questionnaire Results on the Visibility of Subject Motion in Seek Bar–Based Playback Navigation

Figure 16 presents the results for the PCA-based scatter plot using 200 frames. The differences in frame content could be recognized in the scatter plot. However, the visibility of subjects and their motion within each frame tended to decrease compared with standard video playback. This reduction is attributed to the fact that many frames were displayed in a limited space. As a result, each frame was shown at a reduced size. Fine details and subtle motion became difficult to observe, particularly for small or low-contrast subjects.

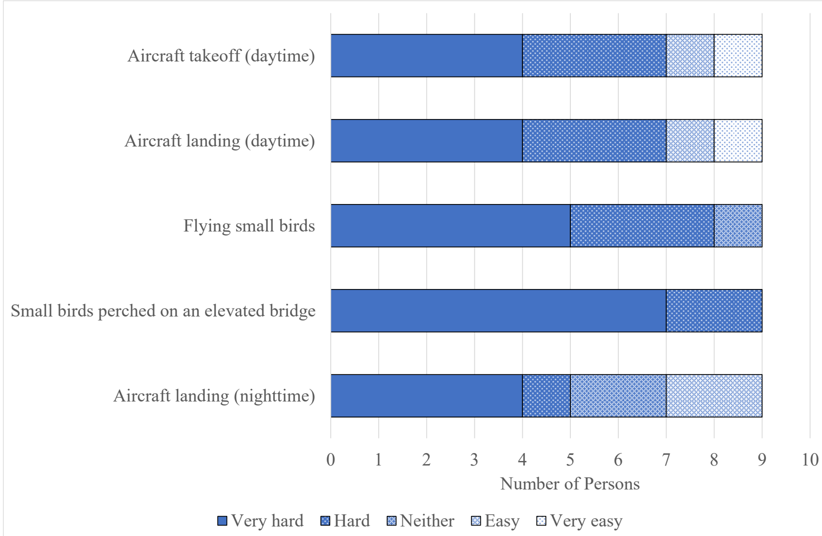

Figure 16. Questionnaire Results on the Visibility of Subject Motion in a Scatter Plot Based on Principal Component Analysis (200 Frames)

Figure 17 presents the results for the PCA-based scatter plot containing only the frames corresponding to the minimum and maximum values of PC1 and PC2. By reducing the number of frames and enlarging each individual frame, the subjects became easier to see compared with the 200-frame scatter plot. On the other hand, because only a few representative frames were displayed, it was sometimes difficult to understand the intermediate motion of the subject. These results indicate a trade-off between frame size and temporal continuity in the proposed visualization method.

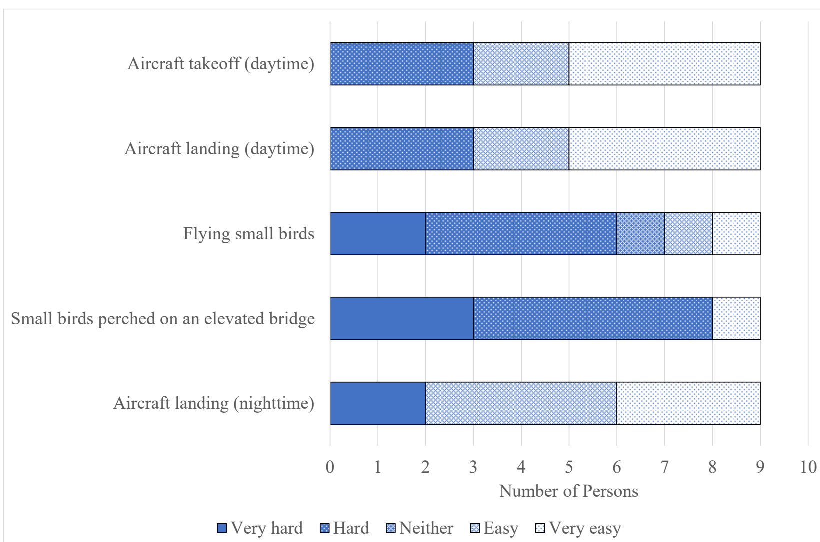

Figure 17. Questionnaire Results on the Visibility of Subject Motion in a Scatter Plot Based on Principal Component Analysis (Minimum and Maximum Values)

Overall, the questionnaire results indicate several tendencies regarding workload and efficiency. Identifying specific scenes during standard video playback tended to require a relatively high workload. In contrast, the PCA-based scatter plots reduced the workload associated with scene finding. In particular, they allowed users to grasp differences among frames without continuously playing the video. In addition, unlike conventional video playback, the proposed method eliminates the need to operate the playback position using a seek bar. These characteristics suggest that the scatter plot–based visualization provides efficiency advantages for scene exploration. However, when only the minimum and maximum principal component frames were displayed, the limited number of frames made it difficult to understand intermediate changes in the subject. From the viewpoint of visibility, several tendencies were also observed. Large subjects such as aircraft were generally easy to track in the original videos as well as when using seek-bar navigation. Small or low-contrast subjects, such as perched birds, tended to be more difficult to observe. In the PCA-based scatter plots with 200 densely plotted frames, differences in frame content could be identified. However, the reduced frame size led to deterioration in the visibility of subjects and their motion. When only the minimum and maximum principal component frames were displayed and each frame was enlarged, visibility within each frame improved. On the other hand, temporal continuity became less clear. These results indicate that frame size and frame density are critical factors affecting the ease of observing subject motion, and they highlight the trade-off between visibility and temporal continuity in the proposed visualization method. Therefore, in outdoor surveillance videos, the proposed method may be less effective for scenes involving small or distant subjects, or low-contrast conditions caused by illumination.

5 Conclusions

In this paper, we investigated the application of a PCA-based visualization method to outdoor surveillance videos. Video frames were arranged on a two-dimensional plane according to their principal component scores, enabling an overview of frame-level variations without continuous playback. The method was applied to videos of daytime airplane takeoffs and landings, nighttime airplane landings, small birds flying, and small birds resting on an elevated bridge.

The results indicate that frame placement in the visualization is largely influenced by dominant visual factors in each scene. In particular, variations in illumination and the size of moving subjects strongly affect the principal component scores, while small subjects tend to have limited impact on the spatial distribution of frames. These findings suggest that the method reflects inter-frame correlations and is sensitive to global visual variations rather than subtle motion of small objects.

The qualitative evaluation also suggested that the visualization method can reduce the workload required for scene exploration in outdoor surveillance videos. By presenting frames in a two-dimensional layout, the method provides an intuitive overview and eliminates the need to operate the playback position using a seek bar. However, the results also revealed a trade-off between frame density and visibility. When many frames were displayed in a limited space, the reduced frame size degraded the visibility of small or low-contrast subjects. Therefore, the method may be less effective for scenes involving small or distant subjects or low-contrast conditions caused by illumination.

In future work, we plan to investigate the influence of subject size and illumination conditions on the visualization results. We also aim to improve the processing efficiency and apply the method to a wider variety of outdoor surveillance videos. In addition, increasing the number of participants in the questionnaire survey will help strengthen the reliability of the qualitative evaluation.

References

[1] Botchen, R. P. (2008). Action-based multifield video visualization. In IEEE Transactions on Visualization and Computer Graphics, 14(4) (pp. 885-899). IEEE.
[2] Liu, Y. &. (2002). Video frame categorization using sort-merge feature selection. In Workshop on Motion and Video Computing Proceedings (pp. 72-77). IEEE.
[3] Schoeffmann, K. T. (2009). Video browsing using motion visualization. In 2009 IEEE International Conference on Multimedia and Expo (pp. 1835-1836). IEEE.
[4] Schoeffmann, K., Lux, M., Taschwer, M., & Boeszoermenyi, L. (2009). Visualization of video motion in context of video browsing. In 2009 IEEE International Conference on Multimedia and Expo (pp. 658-661). IEEE.
[5] Sugita, K. (2022). Implementation of an Average Image Composite Software for Viewer Visualizing Behavior

During Learning Content. In Proceedings of International Conference on P2P, Parallel, Grid, Cloud and Internet Computing. 3PGCIC 2022. Lecture Notes in Networks and Systems, 571 (pp. 346-352). Springer.
[6] Sugita, K. (2024). Development of Scene Segmentation to Improve Work Efficiency of Learner Monitoring. In Research Briefs on Information and Communication Technology Evolution, 10 (pp. 124–133).
[7] Sugita, K. (2025). Finding Representative Frames from Surveillance Video for Visualizing Viewer Behavior. In Proceedings of International Conference on P2P, Parallel, Grid, Cloud and Internet Computing. 3PGCIC 2024. Lecture Notes on Data Engineering and Communications Technologies, 232 (pp. 312-317). Springer, Cham.
[8] Sugita, K. (2025). Visualization of Subjects and Their Motion in Video Using Principal Component Analysis. In International Journal of Mobile Computing and Multimedia Communications (IJMCMC), 16(1) (pp. 1-35).
[9] Zhao, Y. W. (2007). Scene segmentation and categorization using NCuts. In 2007 IEEE Conference on Computer Vision and Pattern Recognition (pp. 1-7). IEEE.

Author’s Biography

Kaoru Sugita received the B. Eng., M. Eng. and Doctor of Software and Information Science degrees from Saitama Institute of Technology, Toyo University and Iwate Prefectural University in 1996, 1998 and 2003, respectively. From April 1998 to November 2002, he was working for ATA Company. From April 2003 to March 2004, he was a Post Doctor Researcher at Iwate Prefectural University. From September 2003 to March 2004, he was working with NetBridge Company. He joined Fukuoka Institute of Technology in 2004 and works currently as a professor in Department of Information and Communication Engineering, Fukuoka Institute of Technology, Japan. His research interests include multimedia communication systems. He is member of VRSJ, JSKE, IPSJ and IEICE.