Effect of shoe insoles on body sway in video-based person identification

Haruki Nakayama*, Takuya Kamitani*, Masashi Nishiyama*†, Yoshio Iwai*†,
*Graduate School of Engineering, Tottori University
†Cross-informatics Research Center, Tottori University

***Abstract*—Body sway has been attracting attention as an informative cue for identifying people. Body sway is a minute movement that occurs naturally when a person is standing, even when they consciously try to stop it. In the field of biomechanics, it has been analytically reported that shoes affect body sway. However, these analytical studies did not consider the effect of shoe insoles on body sway. Here, we investigated the identification performance of body sway recorded by a camera when insoles were placed in the shoes. We found that the identification performance significantly decreased, the low-frequency components of the frequency response of body sway increased, and the maximum magnitude of body sway increased when shoe insoles were worn.**

***Index Terms*—Body sway, Shoe insole, Identification, Video sequence**

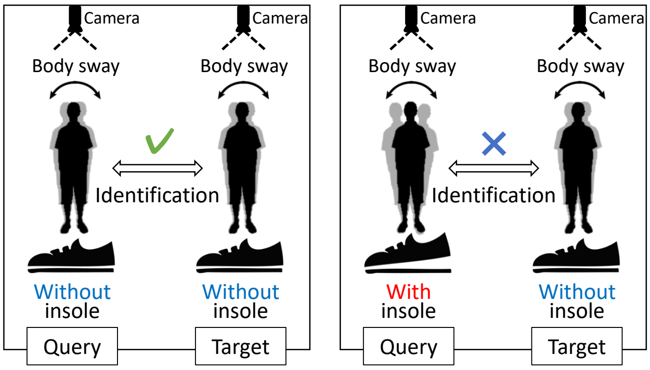


Fig. 1. Effect of shoe insoles on the performance of person identification using body sway. The performance depends significantly on whether or not shoe insoles are worn.

I. INTRODUCTION

To create a safe and secure society, techniques are needed to identify people using security cameras [1]–[3]. To accurately identify people, it is necessary to obtain informative cues that represent individuals from the video sequences obtained by a security camera. Body sway in standing individuals has been attracting attention in recent years as one such cue. Body sway is a minute movement that occurs naturally when a person is standing, even when they consciously try to stop it, as described in [4]. Kamitani et al. [5] showed that the body sway recorded by an overhead camera is an effective cue for person identification. However, they did not consider identification performance in real environments. We need to consider various factors to evaluate identification performance in real environments. In this study, we focus on shoes as a factor.

In the field of biomechanics, analytical studies have been performed using a footplate force sensor to evaluate the effect of shoes on body sway. One analytical study [6] focused on a standing person with and without shoes on. It reported that the magnitude of body sway is greater in people who wear shoes than in people who do not wear shoes. The analytical study [7] evaluated body sway in people wearing shoes with low heels and shoes with high heels. It reported that the velocity of body sway is greater in people who wear high heels than in people who wear low heels. However, these studies did not take into account the effect of insoles placed in the shoes. Therefore, shoe insoles could degrade the performance of person identification using body sway, as illustrated in Fig. 1.

In this study, we investigated the effect of shoe insoles on the performance of person identification by extracting the features of body sway from a video sequence. We used slippers and insoles to evaluate the effect of shoe insoles. In the following, we refer to the state of wearing slippers only as "without insoles" and the state of wearing slippers with insoles as "with insoles." The evaluation results reveal that, in the case of the without-insole target samples, the identification performance for the with-insole query samples decreased significantly when compared with the identification performance for the without-insole query samples. The power spectral density computed from the time-series signals of body sway showed that there was a significant difference in the low-frequency components. In addition, we found that the maximum magnitude of body sway increased when insoles were worn.

II. EVALUATION METHOD

A. *Shoe factors*

Body sway changes depending on which part of the shoe is varied. In particular, changes in the condition of the soles of the feet that support the body can have a significant impact on even the slight movement of body sway. We focus on the insole, which is a variation in the condition of the soles of the feet. Insoles are used to protect the ankles and to increase body height.

Assuming that the without-insole features are stored as the target samples, we evaluate the following hypothesis.

- When with-insole features are used as query samples, the identification performance is worse than when the without-insole features are used as query samples.

B. Extraction of body sway features

Person identification is a technique that determines whether or not a person appearing in a query video sequence is stored in the target dictionary. It is important to extract features from the video sequence of body sway to obtain high identification performance. We extract spatio-temporal features by calculating the amount of body sway movement at each time step from the video sequence using an existing method [5]. Figure 2 shows the process of feature extraction. The details are as follows.
Step 1: We use a camera to record a color video sequence of body sway. The camera is placed above the body to observe the head, which is the part that moves the most in body sway. We assume that the person is standing in an upright position under the camera.
Step 2: We estimate a silhouette video sequence of the head region from the color video sequence. For the head region, the pixel value is assumed to be 1, otherwise, it is 0.
Step 3: We select the reference silhouette frame for each silhouette video sequence. We find a representative silhouette frame in the silhouette video sequence that is similar to other silhouette frames at other time steps. We use this reference silhouette frame.
Step 4: We create a difference-value video sequence to extract the movement of the head region at each time step. To do this, we calculate the difference in the pixel values of the silhouette frame at each time step and the reference silhouette frame.
Step 5: The difference-value video sequence is divided into local blocks to observe the spatial difference of the head shape. Specifically, the video sequence is divided into I blocks in the radial direction from the center of the head region. In Step 5 of Fig. 2, the number of local blocks is $I = 4$ as an example.
Step 6: We compute the amount of movement at each time to extract a time-series signal in the head region. We obtain the sum of the difference values in each local block and use this sum to represent the amount of movement at each time.
Step 7: We extract a feature for person identification. We calculate the power density (PSD) from the amount of movement in each local block. We then combine the PSDs of all blocks to determine the feature used for identification.

III. EXPERIMENTS

A. Dataset

We acquired without- and with-insole video sequences to evaluate the effect of shoe insoles on identification performance. In this experiment, we used slippers and insoles, as shown in Figure 3. The thickness of each insole was 3.5 cm. Twenty-two participants (age: 22.9 $\pm$1.0 years) participated in the experiment. The duration of each video sequence was set to 120 s. Each participant was videoed three times in the without- and with-insole states. As shown in Fig. 4(a), we asked the participants to stand with an upright posture under the camera. The arrangement of the experimental environment

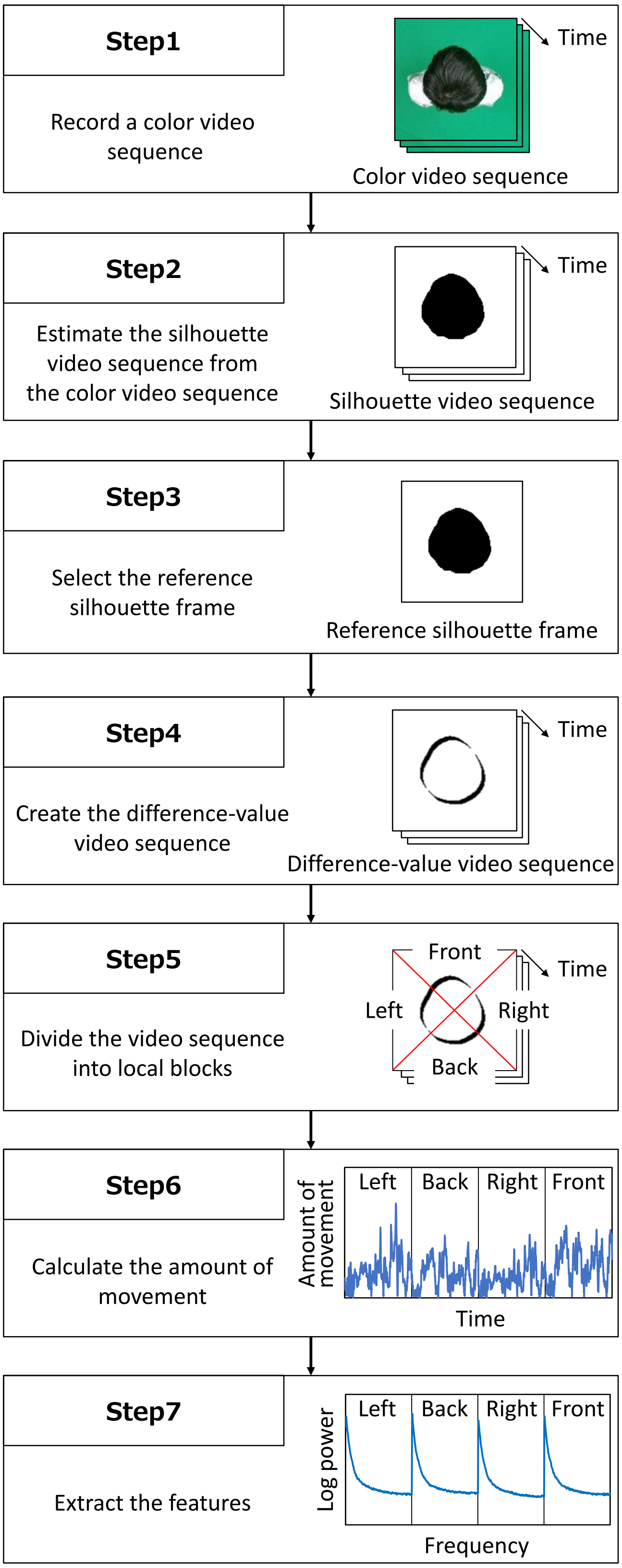


Fig. 2. Process of feature extraction for person identification using the video sequences of body sway

Fig. 3. Slippers and insoles used in the experiment

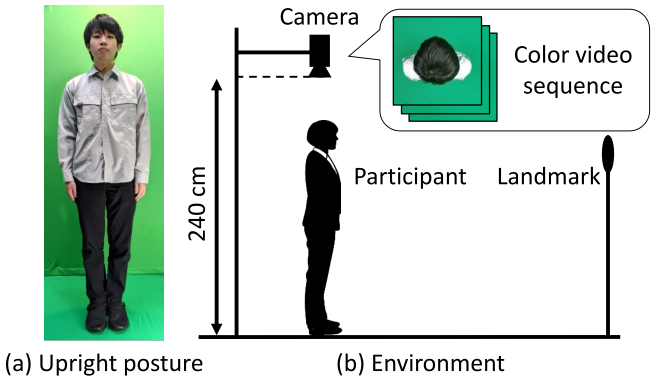


Fig. 4. Experimental setting

is illustrated in Figure 4(b). The camera was positioned at a height of 240 cm. This camera has a resolution of 1920 × 1080 pixels and a frame rate of 30 frames per second. The video sequence was cropped to 200 × 200 pixels in size around the person.

B. Performance of person identification

First, we describe the experimental conditions for the identification method presented in Section II-B. Based on the assumptions described in Section II-A, we evaluated the identification performance when the without-insole data were used as the target sample and the without- and with-insole data were used as the query samples. In each of the trials, different samples from the three collected samples were used for the target sample and query sample. The measure of accuracy is the percentage of rank-1 answers that were correct. For identification, we use the nearest neighbors method based on Euclidean distance.

Figure 5 shows the identification performance results for the without- and with-insole query samples. We confirmed that the identification performance for with-insole queries is about 18 points lower than it is for without-insole queries. We used the Student t-test to determine if there was a significant difference in the identification performance.

The significance level was set to $p < .01$. The results show that there was a significant difference in the identification performance results of the without- and with-insole samples.

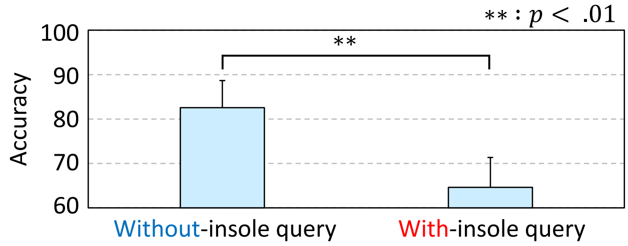


Fig. 5. Identification performance results for without- and with-insole query samples. The significance level was set to $p < 0.01$ and is indicated by **.

C. Comparison of features

We compared the without- and with-insole features to examine the reason for the decrease in identification performance. The feature extraction method is described in Section II-B. For each participant, we calculated the difference between the without- and with-insole features at each frequency. As shown in step Step 5 of Fig. 2, four blocks were used. For each of the four blocks, we calculated the absolute value of the difference between the features. Figure 6(a) shows the average values of the without-insole features, Figure 6(b) shows the average values of the with-insole features, and Figure 6(c) shows the absolute difference between the without- and with-insole features. We found that the low-frequencies components of the left and right blocks are generally different.

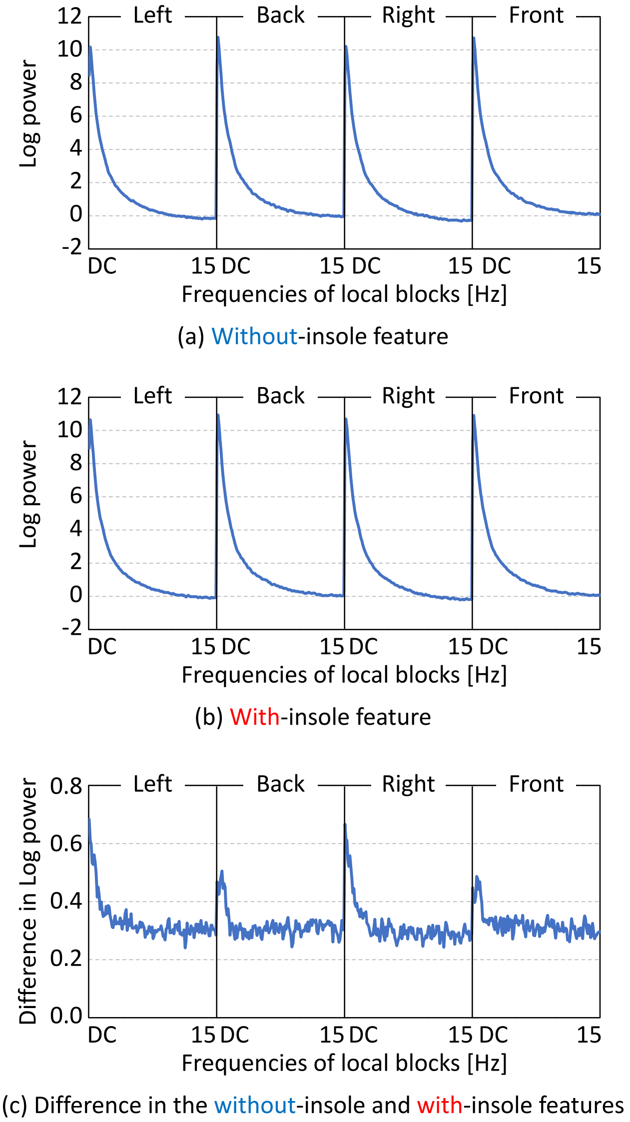
IV. EFFECT OF SHOE INSOLES ON BODY SWAY

A. Purpose

We also investigated how shoe insoles change the body sway. In one analytical study [6], the effect of shoes on body sway was analyzed in two directions: the anterior–posterior (AP) and medial–lateral (ML) directions. The observation in the AP direction corresponds to observing the person from the front, whereas the observation in the ML direction corresponds to observing the person from the side. In [6], it was reported that wearing shoes can significantly increase the maximum magnitude of body sway in the AP direction. However, the analytical study used a footplate force sensor to observe the magnitude of body sway, whereas we use a camera instead.

B. Investigation of the maximum magnitude of body sway

To extract the maximum magnitude of body sway from a video sequence, we first perform feature extraction using steps Step 1 to Step 6 in Fig. 2, and the amount of movement is then calculated. The maximum value of the amount of movement from the start to the end of the video is selected as the maximum magnitude of the body sway. The maximum magnitude of body sway is analyzed in both the AP and ML directions, in accordance with the analytical study. As shown in Fig. 7(a), the magnitude of body sway in the AP direction is the sum of the amount of movement in the front and back blocks. As shown in Fig. 7(b), the magnitude of body sway in the ML direction is the sum of the amount of movement in the left and right blocks.

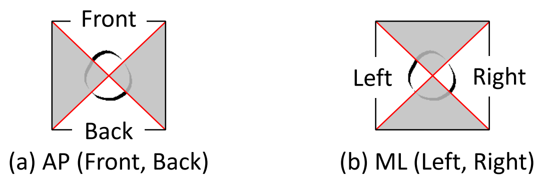


(a) Without-insole feature

(b) With-insole feature

(c) Difference in the without-insole and with-insole features

Fig. 6. Comparison of the average without- and with-insole features



(a) AP (Front, Back)

(b) ML (Left, Right)

Fig. 7. Local blocks for the AP and ML directions

C. Changes in the maximum magnitude of body sway

Table I compares the maximum magnitude of the body sway in the AP and ML directions in the without- and with-insole states. To analyze the results, we performed the Wilcoxon signed-rank test with a significance level of $p < .01$. The results show that there was a significant difference in the maximum magnitude of body sway in both directions. By contrast, in [6], a significant difference was found only in the AP direction. We consider that this difference is due to the fact that the analytical study focused on body sway without and with shoes, whereas we focused on body sway without and with insoles. Furthermore, this difference in the maximum magnitude of the body sway appeared in the low-frequency components, as described in Section III-C.

TABLE I
COMPARISON OF THE AVERAGE MAXIMUM MAGNITUDE OF BODY SWAY IN THE WITHOUT- AND WITH-INSOLE STATES. THE SIGNIFICANCE LEVEL WAS SET TO $p < 0.01$ AND IS INDICATED BY **.

	Without insoles (pixels)	With insoles (pixels)	Statistical significance
AP	1988.2 ± 568.0	2225.1 ± 637.7	**
ML	1439.0 ± 400.2	1858.5 ± 517.2	**

V. CONCLUSIONS

We investigated the effect of shoe insoles on the performance of person identification by extracting the feature of body sway from video sequences. The evaluation results revealed that the identification performance decreased significantly when with-insole query samples were used instead of without-insole query samples. The power spectral density computed from the time-series signals of body sway showed that there was a significant difference in the components of the low-frequency components. In addition, we found that the maximum magnitude of body sway is increased when insoles are worn.

In future work, we would like to develop a method that is robust to the presence of shoe insoles and increase identification performance. We also plan to expand the scope of evaluation to include more factors in addition to shoe insoles.

REFERENCES

[1] A. Bedagkar-Gala and S. Shah, "A survey of approaches and trends in person re-identification," *Image and vision computing*, vol. 32, no. 4, pp. 270–286, 2014.

[2] L. Zheng, L. Shen, L. Tian, S. Wang, J. Wang, and Q. Tian, "Scalable person re-identification: A benchmark," *In Proceedings of the IEEE international conference on computer vision*, pp. 1116–1124, 2015.

[3] L. Zheng, H. Zhang, S. Sun, M. Chandraker, Y. Yang, and Q. Tian, "Person re-identification in the wild," *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1367–1376, 2017.

[4] A. Bottaro, M. Casadio, P. G. Morasso, and V. Sanguineti, "Body sway during quiet standing: Is it the residual chattering of an intermittent stabilization process?" *Human Movement Science*, vol. 24, no. 4, pp. 588–615, 2005.

[5] T. Kamitani, H. Yoshimura, M. Nishiyama, and Y. Iwai, "Temporal and spatial analysis of local body sway movements for the identification of people," *IEICE Transactions on Information and Systems*, vol. 102, no. 1, pp. 165–174, 2019.

[6] B. G. Susana, V. C. Rafael, L. A. Collange, N. F. M. Hugo, Pasini, and S. O. Claudia, "Interference of high-heeled shoes in static balance among young women," *Human Movement Science*, vol. 31, no. 5, pp. 1247–1252, 2012.

[7] A. Mika, L. Oleksy, R. Kielnar, and M. Swierczek, "The influence of high- and low-heeled shoes on balance in young women," *Acta of Bioengineering and Biomechanics*, vol. 18, no. 3, pp. 97–103, 2016.