EXTRACTION AND CLASSIFICATION OF HUMAN GAIT FEATURES FROM ACCELERATION DATA

Takuma Akiduki\textsuperscript{1}, Kento Kawamura\textsuperscript{1}, Zhong Zhang\textsuperscript{1} and Hirotaka Takahashi\textsuperscript{2}

\textsuperscript{1}Department of Mechanical Engineering
Toyohashi University of Technology
1-1 Hibarigaoka, Tenpaku-cho, Toyohashi, Aichi 441-8580, Japan
\{akiduki; zhang\}@me.tut.ac.jp

\textsuperscript{2}Department of Information and Management Systems Engineering
Nagaoka University of Technology
1603-1 Kamitomioka, Nagaoka, Niigata 940-2188, Japan
hirotaka@kjs.nagaokaut.ac.jp

Received October 2017; revised February 2018

ABSTRACT. This paper discusses the methodology for extracting and classifying a style and characteristic component from a walking motion. The walking motion is measured using four wearable motion sensors for acquiring segmented body motion. To extract the style and characteristic component, we use the singular value decomposition of the measured data and evaluate the contribution of each sensor module for gait identification by using the degree of class separation. From these results, the characteristic component of human gait features can be extracted by using singular vectors of whole data of walking motion. In addition, the singular vectors of higher order modes can be used for identifying individuals by proper choice of the modes. Furthermore, using the degree of class separation, the important body segments for gait identification can be indicated by combinations of sensors with the high degree of separation.

Keywords: Human gait features, Singular value decomposition, Accelerometer

1. Introduction. Human activity recognition using wearable sensors such as accelerometers and gyroscopes is one of the key issues in ubiquitous and wearable computing. These technologies are widely used for understanding human activities such as car driving [1], sports [2], healthcare assessments [3, 4], or activities of daily living (ADLs) [5]. This paper discusses the problem of human motion analysis for accelerometer data, which is the time sequence data obtained from body-worn inertial sensors. The goals of human motion analysis generally include the classification or characterization of movements of any particular individual. The purpose of classification is to comprehend \textit{what} activity is being performed. On the other hand, the purpose of the characterization is to comprehend \textit{how} any activity is being performed. To achieve these goals, extraction methods of motion features from accelerometer data have been discussed by many researchers in the fields of human activity recognition (HAR). For example, the statistical features, such as the mean, variance or kurtosis are mostly used for recognizing variety of physical activities [6, 7]. On the other hand, few studies have focused on the features extraction method for evaluating how any activity is being performed to extract qualitative information from accelerometer data, such as the characteristics or quality of executing an activity. Khan \textit{et al.} [8] have introduced a symbolic representation of raw sensor data for

DOI: 10.24507/ijicic.14.04.1361
extracting hierarchical representations of rule structures that enable linkage between activities and their progressively abstract meanings. However, the activities targeted in [8] are complicated activities composed of several basic movements, such as surgery, cooking, and sports. Also, they have evaluated the combination or order of the basic movements for skill assessments. On the other hand, in our research, we focus on a certain specific activity involved in movement and coordination of the body parts such as arms, legs, and trunk. In addition, we evaluate the differences in the movement between users to find the point for evaluating the quality of activity using activity data, where the point means a body segment or sensor module.

Mishima et al. [9] have proposed an extraction method for similarities and differences in human body motion obtained full-body motion capture system (MoCap) using singular value decomposition (SVD). In addition, Akiduki et al. [10] have shown that the SVD method [9] can be applied to data of segmented body motion collected using body-worn accelerometers for extracting individual features. Moreover, Kamio et al. [11] have discussed the physical meaning of these extracted individual features obtained from SVD method. In [9, 10, 11], walking motion is targeted. The walking motion is one of the most fundamental movements in everyday life and does not require special training or proficiency. In addition, even with the same “walking” movement there is also a unique movement and habit for each person. For this reason, walking motion is also targeted in this paper. However, these results in [9, 10, 11] were obtained from data of one gait cycle per one subject. That is, we need to increase the number of gait cycle data to discriminate individuals by using extracted features.

In this paper, for evaluating differences in the movement between users, we use the singular value decomposition of the accelerometer data including about 20 gait cycle data per subject. Also, we evaluate a contribution of each sensor module for individual identification of gait cycle by using the degree of class separation to find the point for evaluating the quality of activity using activity data, where the point means a body segment or sensor module.

The paper is organized as follows. An overview of the experiment to collect the data on walking motion is given in Section 2. Then the extraction and evaluation method of individual gait features are presented in detail in Section 3. The experimental results and discussion are presented in Section 4 to evaluate the effectiveness of the method of evaluating the contribution of sensors for individual identification. Section 5 concludes this paper.

2. Overview of Experiment.

2.1. Capturing body movements. To acquire human activities, we have constructed a measuring system shown in Figure 1. This system includes wearable accelerometer modules (WAA-010, ATR-Promotions Inc.) shown in Figures 1(a), 1(b) and 1(c) for capturing segmented body motion of subjects. The dimensions of the sensor module are $39 \times 44 \times 8 \text{ mm}$ with a weight of 20 g. The four sensors are worn on the right lower leg (S1), left thigh (S2), lower back (S3) and left forearm (S4) on a subject, and the numbering of the segments is shown in Figures 1(a) and 1(b). These placements are referred to Bao and Intille [6]. Note that the earth-fixed reference coordinate system used is defined as right-handed Cartesian coordinate systems: $(G_X, G_Y, G_Z)$ as shown in Figure 1(a). By using this system, the motions of a subject can be collected as both acceleration and angular velocity along with three-axis of the local coordinate on the sensor module: $(S_X, S_Y, S_Z)$ as shown in Figure 1(c). Of these, all signals used in this paper are summarized in Table 1. And these signals are sampled at 100 Hz in each sensor module and transmitted.
Figure 1. Wearable accelerometer modules. (a) and (b) show sensor setting on a subject’s body, and (c) shows the accelerometer module and its coordinate system.

Table 1. Overview of measured three-axis accelerometer signals.

<table>
<thead>
<tr>
<th>Sensor Module#</th>
<th>Axis#</th>
<th>p in (1)</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>S1-{ax, ay, az}</td>
<td>1, 2, 3</td>
<td>Right lower leg</td>
</tr>
<tr>
<td>S2</td>
<td>S2-{ax, ay, az}</td>
<td>4, 5, 6</td>
<td>Left thigh</td>
</tr>
<tr>
<td>S3</td>
<td>S3-{ax, ay, az}</td>
<td>7, 8, 9</td>
<td>Lower back</td>
</tr>
<tr>
<td>S4</td>
<td>S4-{ax, ay, az}</td>
<td>10, 11, 12</td>
<td>Left forearm</td>
</tr>
</tbody>
</table>

Data collection. In this paper, data on walking motion is collected from 13 subjects (10 men and 3 women aged 24.3 ± 4.3). The subject wears the motion capture suits. At the same time, over the motion capture suits, the four sensors are placed on the position of S1, . . . , S4 shown in Figures 1(a) and 1(b). Authors instruct the subject to walk on the test course with 15 m straight flooring line according to a predefined protocol. In the protocol, each subject has an instruction to perform 5 walking with following conditions; Cond. N: walking with natural speed, Cond. S: walking with slow speed, and Cond. F: walking with fast speed on the course. The order of the instructions is Cond. N_{1st} → Cond. N_{2nd} → Cond. S → Cond. F → Cond. N_{3rd}. Before collecting data, we have explained the contents of the experiment. Moreover, we also have obtained informed consent from each subject to use obtained data for research purposes.

Feature Computation. To extract individual features of walking motion, we use the singular value decomposition of the measured data. Here, to expand the number of gait cycles, we introduce some parameters for the method of [9]. Also, to determine the
importance of the sensors in gait identification, we introduce a degree of separation of class.

In the following, note that $R^n$ denotes real $n$-dimensional space, and $x \in R^n$ denotes an $n$-dimensional column vector, respectively. If $A = (a_{ij}) \in R^{n \times m}$ is a matrix of $n$ rows and $m$ columns, then $A^\top$ is the transpose of $A$.

3.1. Preparation for data matrix. Consider a sequence of segmented motion data, which is cyclic and consequently has a gait period. Then $p$th time series of the $\gamma$th gait cycle in the multivariate time series data with discrete time steps is as follows:

$$x^\gamma_p = (x^\gamma_p(1), x^\gamma_p(2), \ldots, x^\gamma_p(N))^\top \in R^N,$$

where $p = 1, 2, \ldots, S$ and $\gamma = 1, 2, \ldots, L$.

Here, $N$ is a length of the time series, and $S$ is the number of variables in the multivariate time series, that is, $S = (\text{the number of sensors}) \times (\text{the number of axes on the sensor})$. And the correspondence between index $p$ and axis number of the sensor is shown in Table 1. The total number of gait cycles for all subjects is $L = \sum_{\alpha=1}^{M} L_{\alpha}$, where $L_{\alpha}$ is a number of the gait cycles in subject $\alpha$, and $M$ is the number of subjects. Moreover, a set of the time series for the $\gamma$th gait cycle data is also as follows:

$$X^\gamma = (x^\gamma_1, x^\gamma_2, \ldots, x^\gamma_S) \in R^{NS}.$$  

In the following, we call time-series data of a $\gamma$th gait cycle: $X^\gamma$ as $\gamma$th gait frame. For comparing the motions with each gait frame, (2) is rewritten as follows:

$$\ast a^\gamma = \left(\{x^\gamma_1\}^\top, \{x^\gamma_2\}^\top, \ldots, \{x^\gamma_S\}^\top\right)^\top \in R^{NS}.$$  

Note that $\ast a^\gamma$ is a column vector, which represents a human gait pattern in the $\gamma$th gait cycle for subject $\alpha$. Then, the following data matrix for comparing motions is defined:

$$\ast D = (\ast a^1, \ast a^2, \ldots, \ast a^L) \in R^{NS \times L}.$$  

Finally, a matrix $D$ is defined as a result of standardization for each column of the matrix $\ast D$ where $\ast a^\gamma$ is centered to have mean 0 and scaled to have standard deviation 1.

$$D = (a^1, a^2, \ldots, a^L) \in R^{NS \times L}.$$  

Also, the matrix $D$ represents a set of human gait patterns for all subjects.

3.2. Singular value decomposition. In this paper, we suppose that the matrix $D$ contains information on both similarities and differences for each subject. The similarities are a common component of matrix $D$, and the differences can be defined a set of components obtained by subtracting the common component from the matrix $D$. These components can be extracted by using singular value decomposition (SVD) [9]. The SVD of data matrix $D$ is given by

$$D = UV^\top \Sigma,$$

where $U$ and $V$ are unitary matrices, and $\Sigma$ is a diagonal matrix. The diagonal elements of $\Sigma$ are called singular values $\sigma_i$ ($i = 1, 2, \ldots, L$), which are non-negative real numbers and $\sigma_i \geq \sigma_j$ ($i \leq j$). Each column vector of $U$ is a left singular vector $u_i \in R^{NS}$. Each column vector of $V$ is a right singular vector $v_i \in R^L$. The $i$th element of each $\Sigma$, $U$ and $V$ is called the $i$th mode by Mishima et al. [9].

Here, the matrix $D$ can be represented as follows using $u_i$ and $v_i$, which are linear independent respectively.

$$D = \sigma_1 u_1 v_1^\top + \sigma_2 u_2 v_2^\top + \cdots + \sigma_L u_L v_L^\top.$$
where \( u_{ni} \) is the \( n \)th element of \( u_i \), and \( v_{\gamma i} \) is the \( \gamma \)th element of \( v_i \). At (6), \( u_i \) represents a motion feature for the \( i \)th mode, \( v_{\gamma i} \) indicates the contribution ratio of the \( \gamma \)th gait cycle to the \( i \)th mode, and \( \sigma_i \) represents the contribution ratio of the \( i \)th mode to the matrix \( D \). From \( \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_L \) in (6), the lower order mode has a larger component in the matrix \( D \). Moreover, the 1st mode is a typical component that is common among all gait frames. Then the singular value of the 1st mode \( \sigma_1 \) is the largest among all modes, and \( v_{\gamma 1} \) have to be almost constant for all subjects. Here, \( u_1 \) represents the motion characteristic common to body motion, so in this paper, \( u_1 \) is called a style component of a body motion.

While, the high-order terms after the second term in (6), the component in the matrix \( D \) is smaller than in the first mode. And \( v_{\gamma i} \) \( (i \geq 2) \) have to change with each gait frame or each subject. So, in this paper, \( u_2 \) or more is called a characteristic component of the body motion.

### 3.3. Class variance as a degree of separation of class.

As the evaluation index of the sensors for gait identification, we introduce a degree of separation between each class. The degree of separation is defined as follows:

\[
J_{\sigma,\text{med}} = \frac{\text{median}\left[\sigma_{2,\alpha}^2\right]}{\text{median}\left[\sigma_{w,\alpha}^2\right]}, \quad \text{for } \alpha = 1, \ldots, M, \tag{7}
\]

where

\[
\sigma_{w,\alpha}^2 = \frac{1}{|X_\alpha|} \sum_{x \in X_\alpha} (x - m_\alpha)^\top (x - m_\alpha), \quad \sigma_{b,\alpha}^2 = L_\alpha (m_\alpha - m)^\top (m_\alpha - m).
\]

Here, \( \sigma_{w,\alpha}^2 \) and \( \sigma_{b,\alpha}^2 \) are within-class variance and between-class variance for subject \( \alpha \), respectively. And \( X_\alpha \subset \mathbb{R}^k \) is a gait frame of a subject in the \( k \)-dimensional feature vector, and \( X_\alpha \) represents a set of gait frames of the subject \( \alpha \). The size of \( X_\alpha \) is \( L_\alpha \). Moreover, \( m_\alpha \in \mathbb{R}^k \) represents a mean vector of the set of \( X_\alpha \), and \( m \) is mean vector for all gait frame of all subjects. Sensor data with a higher degree of separation in (7) contribute to gait identification.

### 4. Results and Discussions.

In this section, to evaluate the effectiveness of the method of evaluating the contribution of sensors for individual identification in Section 3, we have computed the degree of class separation with the procedure shown in Figure 2. The degree of class separation was computed while changing the combination of sensors, such as S1, S1, and S2 since sensors with high scores in common in all combinations contribute to gait identification. In this paper, to simplify the discussion, we used the data of walking motion with condition \( N_{1\text{st}}, N_{2\text{nd}}, \) and \( N_{3\text{rd}} \) in the following section. The reason for this is that the conditions S and F change the walking speed, and it is necessary to consider the difference in movement depending on the walking speed in addition to the individual difference factor. Before constructing matrix \( D \) in (4), time series for about 20 gait cycles per each subject (Max. 27 cycles, Min. 18 cycles) were clipped from the whole walking data of each subject. Since the length of one gait cycle was different for each subject, the clipped time series have even lengths between by processing through interpolation algorithm. In this paper, the cubic spline algorithm was used. As the results, the length of clipped time series was \( N = 134 \). In the accelerometer data, \( S = (4 \text{ segments}) \times (3 \text{ axis}) \)
Figure 2. Scheme showing a procedure for calculating the degree of class separation by changing the combination of sensors. Note that $S^{(*)}$ is determined by the combination of sensors. For example, when $\{S1, S2\}$, $S^{\{S1,S2\}} = 6$ since using the time series of S1-\{ax, ay, az\} and S2-\{ax, ay, az\}.

Figure 3. An example of time-series of walking by subject ‘as’ with the condition N\textsubscript{1st}.

when using all time series in Table 1. Also, the number of subjects was $M = 13$, and the number of gait frames of all subject was $L = 281$. Figure 3 shows an example of time-series for walking data of subject ‘as’ with N\textsubscript{1st}, which includes 6 gait frame data for feature computation in the following.

4.1. Extracting for style and characteristics component of motion. Figure 4 shows the results of SVD of matrix $D$ with upper 13 modes out of $L = 281$ modes. The singular value at the 1st mode reached a peak of $\sigma_1 = 578.7$ in Figure 4(a). The singular
value at the 2nd mode drops suddenly, and after that decreases moderately. On the other hand, Figure 4(b) shows the right singular vector \( v_i \) for each gait frame with a grayscale image. The right singular vector at the 1st mode \( v_{1,1} \) remains constant at approximately 0.06 in Figure 4(b) for all gait frames of all subjects. These results also indicate that the 1st mode affects equally in all gait frames of all subjects, that is, the 1st mode represents the *style* component of walking motion common to all subjects. On the other hand, the higher than the 2nd modes are helpful in identifying individuals, that is, the higher than the 2nd modes represent *characteristic* component for each subject. These results are essentially consistent with the results discussed in [9]. Thus, we see that from the experimental results of Figure 4, SVD is an effective method for extracting both the style and the characteristic component of motion.

Thus, we see that from the experimental results of Figure 4, SVD is an effective method for extracting both the *style* and the *characteristic* component of motion.

4.2. Clustering for gait features. To visualize the relationships between the subjects and each mode, we reduce the dimension of matrix \( D \). Then the matrix \( D \) is approximated by using the left singular vectors as follows.

\[
D_k = U_k^T D \in \mathcal{R}^{k \times L}, \quad \text{where } U_k = (u_1, \ldots, u_k) \in \mathcal{R}^{NS \times k}. \tag{8}
\]

Here \( U_k \) is a set of \( k \) left singular vectors. And the columns and row elements of \( D_k \) are corresponding to each gait frame and each mode, respectively. Figure 5 shows the results of (8), where each marker is corresponding to elements of matrix \( D_{k|k=3} = (d_{ij}) \in \mathcal{R}^{3 \times 281} \). And PC1, PC2, and PC3 shown in Figure 5 are corresponding to the 1st, 2nd, and 3rd principal component respectively from principal component analysis. In each two-dimensional space, there are \( L = 281 \) points, and each point represents a gait frame. Then the relationship of placement between each point represents similarity between each gait frame. We can observe the dots in Figure 5 are concentrated in the narrow range of the PC1 axis and distributed over a wide range of the PC2 and PC3 axes. Then, on the PC2-PC3 plane shown in Figure 5, the dots are divided into roughly four groups corresponding to the subjects: \( \alpha = \{5\}, \{13\}, \{11,12\}, \text{ and } \{1,2,3,4,6,7,8,9,10\} \). These results suggest that the walking motion of each subject can be discriminated by using higher order modes of \( k = 2 \) or more in the \( D_k \) space. Also, these results suggest that
PC1 indicates a role for the style component, and high-order modes are corresponding to characteristic components of walking motion.

4.3. Contribution of sensors for gait classification. To evaluate sensors that contribute to individual identification of walking, we compute the degree of class separation of (7) for PC2-PC3 space, i.e., $\chi = (d_{2j}, d_{3j})^T$ ($j = 1, \ldots, L$) in (7). Table 2 shows the results of (7), where each column represents a combination of sensors and its degree of class separation. In Table 2, each number from S1 to S4 within $\{\ast\}$ means the sensor module numbers shown in Table 1 and its combinations. For example, in the case of $\{\text{S1}\}$, the matrix $D$ is constructed using only the tree axis acceleration data of S1. That is, the matrix $D \in \mathbb{R}^{(N \times 1 \times 3) \times L}$ in (4) is constructed with $p = 1, 2, 3$ in (1). Likewise, in the case of $\{\text{S2, S3}\}$, the matrix $D$ is constructed using the acceleration data of the total of six axes of S2 and S3. That is, the matrix $D \in \mathbb{R}^{(N \times 2 \times 3) \times L}$ in (4) is constructed with $p = 4, 5, \ldots, 8, 9$ in (1). Also, Figure 6 shows the results of $\sigma_{w,\alpha}$ and $\sigma_{b,\alpha}$ in (7) when only one sensor is used.

<table>
<thead>
<tr>
<th>Layout</th>
<th>$J_{\sigma,\text{med.}}$</th>
<th>Layout</th>
<th>$J_{\sigma,\text{med.}}$</th>
<th>Layout</th>
<th>$J_{\sigma,\text{med.}}$</th>
<th>Layout</th>
<th>$J_{\sigma,\text{med.}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>${\text{S1}}$</td>
<td>3.979</td>
<td>${\text{S3, S4}}$</td>
<td>2.521</td>
<td>${\text{S1, S2, S3}}$</td>
<td>2.278</td>
<td>${\text{S1, S2, S3, S4}}$</td>
<td><strong>2.222</strong></td>
</tr>
<tr>
<td>${\text{S2}}$</td>
<td>2.460</td>
<td>${\text{S2, S4}}$</td>
<td>0.801</td>
<td>${\text{S1, S2, S4}}$</td>
<td>1.837</td>
<td>${\text{S1, S2, S4}}$</td>
<td>3.330</td>
</tr>
<tr>
<td>${\text{S3}}$</td>
<td><strong>4.223</strong></td>
<td>${\text{S2, S3}}$</td>
<td>2.262</td>
<td>${\text{S1, S3, S4}}$</td>
<td><strong>4.024</strong></td>
<td>${\text{S1, S3}}$</td>
<td>2.491</td>
</tr>
<tr>
<td>${\text{S4}}$</td>
<td>3.788</td>
<td>${\text{S1, S4}}$</td>
<td>2.419</td>
<td>${\text{S2, S3, S4}}$</td>
<td>2.276</td>
<td>${\text{S1, S2}}$</td>
<td></td>
</tr>
</tbody>
</table>

From Table 2, S3 is the highest degree of separation in the case of one sensor. In addition, the combinations including the sensor S3 shows a high degree of separation. In other words, these results show that the contribution degree of sensor S3 is the highest in gait identification.

On the other hand, Figure 7 shows scatter plots of PC2-PC3 space corresponding to the sensor layout pattern. From Table 2, we can identify important body segment for gait identification from combinations of sensors with the high degree of separation. And these results suggest that the physical meaning of individual features can be considered...
EXTRACTION AND CLASSIFICATION OF HUMAN GAIT FEATURES

Figure 6. An example of within-class variance $\sigma^2_{w,\alpha}$ and between-class variance $\sigma^2_{b,\alpha}$ for subject $\alpha$.

Figure 7. Scatter plot of the matrix $D_3$ corresponding to the combination of the sensors.

based on the correspondence between the combination of sensors and the distribution of data shown in Figure 7.

5. Conclusions. In this paper, we discussed the methodology for extracting and classifying of gait features from walking motion. The motions were measured using four wearable accelerometer modules for acquiring segmented body motion. To extract human
gait features, we used the singular value decomposition of the measured data and evaluated the contribution of each sensor module for individual identification of gait cycle by using the degree of class separation. From these results, the characteristic component of human gait features can be extracted by using singular vectors of whole data of walking motion. In addition, the singular vectors of higher order modes can be used for identifying individuals by proper choice of the modes. For example, we showed that subjects used in this paper can be divided into roughly four groups using the singular vectors of the 2nd and 3rd modes. Furthermore, using the degree of class separation, the important body segments for gait identification can be indicated by combinations of sensors with the high degree of separation. By using these methods, we can identify the point of focus to find differences between subjects in walking motion, where the point means the body segment or the position of sensor modules. As further work, we need to investigate the physical meanings of each mode for understanding characteristics of a subject from the motion data.

Acknowledgments. This work was supported in part by MEXT/JSPS KAKENHI a Grant-in-Aid for Young Scientists (No. 16K06156; T. Akiduki and No. 17K05437; H. Takahashi).

REFERENCES