PAPER

Individuality-Preserving Gait Pattern Prediction Based on Gait Feature Transitions

Tsuyoshi HIGASHIGUCHI[†], ***Nonmember***, **Norimichi UKITA**[††a)], ***Senior Member***, **Masayuki KANBARA**[†], ***Member***, ***and*** **Norihiro HAGITA**[†], ***Fellow***

SUMMARY This paper proposes a method for predicting individuality-preserving gait patterns. Physical rehabilitation can be performed using visual and/or physical instructions by physiotherapists or exoskeletal robots. However, a template-based rehabilitation may produce discomfort and pain in a patient because of deviations from the natural gait of each patient. Our work addresses this problem by predicting an individuality-preserving gait pattern for each patient. In this prediction, the transition of the gait patterns is modeled by associating the sequence of a 3D skeleton in gait with its continuous-value gait features (e.g., walking speed or step width). In the space of the prediction model, the arrangement of the gait patterns are optimized so that (1) similar gait patterns are close to each other and (2) the gait feature changes smoothly between neighboring gait patterns. This model allows to predict individuality-preserving gait patterns of each patient even if his/her various gait patterns are not available for prediction. The effectiveness of the proposed method is demonstrated quantitatively. with two datasets.

key words: *gait, human skeleton, individuality-preserving prediction*

1. Introduction

In walking rehabilitation, it is important to take into account the gait pattern of the individual patient. Gait patterns have been analyzed with various factors such as aging [1], [2], fatigue [3], and medical conditions such as hemiplegia [4], cerebral palsy [5], and Parkinson's disease [6]. On the basis of these analyses, patients can be instructed in ways that take into account their specific symptoms. Recently, instructions are given not only by specialists (e.g., medical doctors and physical therapists) but also by robotic devices such as exoskeletal robots [7], [8] and multi-modal systems [9]. It is known that an unnatural gait pattern provided by the instruction may cause unusual muscle responses [10] or degrade the effect of rehabilitation [11], [12].

For measuring the current natural gait pattern of each patient, various off-the-shelf motion capture systems can be used. In general, however, a motion given in a rehabilitative instruction should be modified from the current gait pattern in order to improve his/her physical condition. For example, in order to increase the step length of a patient from s_{now} to $s_{now} + \Delta$, where s_{now} is the current step length, a gait cycle in $s_{now} + \Delta$ of this patient must be predicted because it is unknown. In this predicted motion, the individual gait patterns such as bow-legs should remain for comfortable rehabilitation optimized for each patient. To achieve this prediction, individuality-preserving gait prediction is the goal of this paper. In this work, individuality-preserving gait prediction for providing $s_{now} + \Delta$ to each patient is defined as follows: if individuality is preserved in a predicted gait when one of gait features (e.g., walking speed and step width) is changed, the 3D skeleton sequence of the predicted gait is similar to that of its source gait except for the gait feature changed for this prediction. In this definition, the individuality preservation does not mean person identification (e.g., gait recognition [13]) but means that the gait pattern of a target person is preserved. More specifically, we evaluate the individuality preservation in terms of two different criteria: the difference between a predicted 3D skeleton and its ground truth (see Sect. 4) and the deviations of gait features in the predicted 3D skeleton (see Sect. 5). If individuality is preserved, both of the difference in the 3D skeleton and the deviation of gait features must be small.

This paper addresses the following issues:

- Difficulty in individuality-preservation: The individuality-preserving gait prediction only from an observed current gait cycle is an ill-posed problem because it is difficult to specify what are individualities in this gait cycle. We define the individualities with a huge number of training gait patterns.
- Difficulty in prediction: The prediction of gait change is challenging than classification [14], [15] because of the complex dynamics of gait. In the literature on biomechanics, the focus is on the stochastic properties of the observed gait data (e.g., the effect of aging on gait [1], [2], [16]) rather than on the prediction of gait change.
- Difficulty due to a full-body structure (i.e., high-dimensional features): Previous studies focused on simple gait features rather than the entire geometric structure of the skeleton. For example, a change in the toe clearance was predicted based on acceleration in [17]. However, the motion of the full-, lower-, or upper-body skeleton is required for several applications such as an exoskeletal robot operation.
- Difficulty in gait-cycle prediction: A small number of future frames in a gait pattern can be predicted from the

Manuscript received March 5, 2018.
Manuscript revised June 12, 2018.
Manuscript publicized July 20, 2018.
[†]The authors are with Nara Institute of Science and Technology, Ikoma-shi, 630–0192 Japan.
[††]The author is with Toyota Technological Institute, Nagoya-shi, 468–8511 Japan.
a) E-mail: ukita@toyota-ti.ac.jp
DOI: 10.1587/transinf.2018EDP7082

past few frames when sophisticated machine learning is used [18]. While such short-term prediction is useful for gait tracking [19], [20], all frames in one gait cycle must be predicted to meet the goal of this study.

We resolve these issues by extending latent dynamics models that allow us to predict high-dimensional motions via the low-dimensional latent space. We extend this model so that transition between similar gait cycles are explicitly modeled for individuality-preserving gait cycle prediction.

2. Gait Data Representation and Gait Features

2.1 Gait Representation Using a Sequence of Skeletons

2.1.1 Gait Measurement by Kinect V2

A gait cycle is represented by the temporal sequence of a 3D skeleton measured by a motion capture system. Instead of using optical markers attached to a user's body, a Kinect V2 sensor was used in our experiments. Kinect V2 has a generic RGB camera and a set of IR camera and projector, and measures the depth of a 3D scene based on a Time-of-Flight technology [21]. The accuracy of Kinect V2 is better than that of Kinect V1 [22] and its effectiveness for detailed scientific gait analysis has been demonstrated [23], [24].

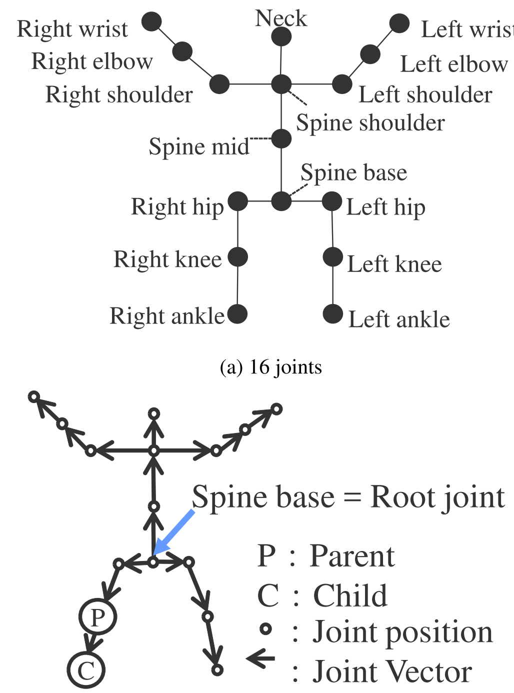
For easy and reliable measurement, a single Kinect V2 was used to record a subject from the front. When the sensor observation is made from the side of a subject, measurement accuracy is unreliable due to the narrow view angle of the sensor. Our experiments showed that a 3D skeleton can be measured accurately at a distance of around 3m between the subject and the sensor.

2.1.2 Normalization of Gait Data

The skeleton measured by Kinect V2 comprises 25 points including joints and endpoints. Since measurement accuracy of the endpoints is lower than that of the other joints [23], the skeleton used in our experiments comprised 16 joints, the neck, spine shoulder, spine mid, spine base, left/right wrists, elbows, shoulders, hips, knees, and ankles (Fig. 1 (a)), denoted by $\boldsymbol{B}$:

$$\boldsymbol{B} = \{Neck, SShoulder, SMid, SBase, LShoulder, RShoulder, LElbow, RElbow, LWrist, RWrist, LHip, RHip, LKnee, RKnee, LAnkle, RAnkle\} \quad (1)$$

The 3D coordinates of the raw measurement data (i.e., in the Kinect coordinate system) are spatially aligned so that (i) the z-axis coincides with the walking direction of a subject, (ii) the y-axis is equal to the vertical upward axis, and (iii) the origin at each frame coincides with the spine base. The walking direction is approximated by the 3D direction $\boldsymbol{D}_w$ from the spine base at the beginning frame to that at the ending frame. This approximate definition allows us to stably align the 3D coordinate system at each frame because it is difficult to estimate the instantaneous walking direction robustly against the measurement noise of Kinect V2.



(a) 16 joints

(b) Skeleton representation $\boldsymbol{S}_t$

Fig. 1 16 joints used for the skeleton representation of a human body.

For the aforementioned alignment, the 3D coordinates of i-th joint in the Kinect coordinate system (denoted by $\boldsymbol{K}_{t,i} = (K_{t,i,x}, K_{t,i,y}, K_{t,i,z})^T, 1)$ are transformed to those in the coordinate system (denoted by $\boldsymbol{P}_{t,i}$) at frame t as follows:

$$\boldsymbol{P}_{t,i} = \begin{pmatrix} \cos(\theta) & 0 & -\sin(\theta) & 0 \\ 0 & 1 & 0 & 0 \\ \sin(\theta) & 0 & \cos(\theta) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & -K_{t,i,x} \\ 0 & 1 & 0 & -K_{t,i,y} \\ 0 & 0 & 1 & -K_{t,i,z} \\ 0 & 0 & 0 & 1 \end{pmatrix} \boldsymbol{K}_{t,i} \quad (2)$$

where θ denotes the angle from $\boldsymbol{D}_w$ to the z-axis in the Kinect coordinate system. Note that we assume that Kinect V2 is spatially aligned so that the y-axis of its coordinate system is equal to the vertical upward axis.

One gait cycle is defined to be between two frames in which the z-coordinate of the right knee is at a temporally-local maximum.

Finally, the skeleton at t (denoted by $\boldsymbol{S}_t$) is represented by a set of normalized 3D vectors each of which is defined to fall between two neighboring body joints (Fig. 1 (b)). Then, normalized 3D vector $\boldsymbol{J}_{t,c}$ is expressed as follows:

$$\boldsymbol{J}_{t,c} = \frac{l_{mean}}{\|\boldsymbol{P}_{t,c} - \boldsymbol{P}_{t,p}\|} \left(\boldsymbol{P}_{t,c} - \boldsymbol{P}_{t,p}\right), \quad (3)$$

where $\boldsymbol{P}_{t,c}$ and $\boldsymbol{P}_{t,p}$ denote the 3D coordinates of the child and parent joints, respectively. l_{mean} denotes the mean of $\|\boldsymbol{P}_{t,c} - \boldsymbol{P}_{t,p}\|$ over all frames of all the gait data. Since one of the 16 joints is a root joint that has no parent, the skeleton is represented by $15 \times 3\text{D} = 45\text{D}$ vectors, as shown in

Fig. 1 (b), where each arrow indicates a 3D vector. The root joint in our experiments was the spine base.

$$\begin{aligned} &\boldsymbol{S}_t = \{\boldsymbol{J}_{t,c}\} \\ &c \in \boldsymbol{B} \setminus SBase \end{aligned} \tag{4}$$

2.2 Gait Features Influencing a Gait Pattern

Typical aging effects appear on gait features such as the walking speed, step width, and leaning forward angle [1], [25], [26]. The risk of falling is affected by several gait features including the walking speed, and lateral body sway [27]–[29]. In our experiments, the walking speed, step width, leaning forward angle, and lateral body sway were used as the gait features.

2.2.1 Walking Speed

Since the 3D coordinate system is defined so that the walking direction coincides with the z-axis, the walking speed f_v is computed only along z-axis. Specifically, the walking speed f_v is computed from the z-coordinates of the spine base at the beginning and ending frames (denoted by frames B and E, respectively) of a gait cycle as follows:

$$f_v = \frac{N^{(fps)}}{N^{(cycle)}}(P^{(K)}_{E,SBase,z} - P^{(K)}_{B,SBase,z}),$$

where $N^{(fps)}$ and $N^{(cycle)}$ denote the number of frames captured in a second and included in a gait cycle, respectively. $P^{(K)}_{t,j,z}$ denotes the z-coordinate of joint j at the t-th frame.

2.2.2 Step Width

The step width is equal to the distance between the left and right ankles in their respective stance phases. In accordance with the literature [30], in our experiments, the step width f_w is computed from the x-coordinates of the left and right ankles, each of whose y-coordinates is minimized:

$$f_w = ||P_{t'_l,LAnkle,x} - P_{t'_r,RAnkle,x}||,$$

where $t'_l = \underset{t}{\mathrm{argmin}}(P_{t,LAnkle,y})$, $t'_r = \underset{t}{\mathrm{argmin}}(P_{t,RAnkle,y})$. In other words, f_w is a distance from the z-axis (i.e., the walking direction $\boldsymbol{D}_w$). We define f_w by $\boldsymbol{D}_w$ rater than by the instantaneous walking direction because it is difficult to be computed robustly to measurement noise.

2.2.3 Leaning Forward Angle

The leaning forward angle, f_θ, is defined to be the angle between the vertical axis and the torso around the x-axis. The torso was defined so that its two endpoints are *SBase* and the midpoint of *LShoulder* and *RShoulder*. This midpoint was used instead of other candidates (i.e., *SMid*, *SShoulder*, and *Neck*) because 1) *Neck* can be bent and 2) it is difficult to correctly localize *SMid* and *SShoulder* because they are more inside the body rather than the right and left shoulders as validated in experiments; see Table 3.

$$f_\theta = \frac{1}{E-B+1}\sum_{t=B}^{E}\tan^{-1}\left(\frac{||P_{t,SS,y} + P_{t,SS,y}||}{||P_{t,SS,z} + P_{t,SS,z}||}\right),$$

where SS denotes *SShoulder*.

2.2.4 Lateral Body Sway

The lateral body sway represents the moving distance of the body centroid along the x-axis. The body centroid at t is determined by the weight-weighted centroid of 16 joints with the approximate weight ratios of the body parts (i.e., head = 8%, arms = 12%, torso = 46%, and legs = 34%). The lateral body sway, f_s, is the difference between the max and min x-coordinates of the body centroid in each gait cycle.

$$f_s = ||\max_t(P_{t,Centroid,x}) - \min_t(P_{t,Centroid,x})||.$$

3. Individuality-Preserving Gait Prediction

This section describes a methodology for individuality-preserving gait prediction. The learning process of our method, which is based on low-dimensional latent representation [18] introduced in Sect. 3.1, is proposed in Sect. 3.2. With a model acquired in the learning process, we can predict an individuality-preserving gait pattern, as described in Sect. 3.3.

3.1 Latent Representation of Gait Cycles Lying along the Transition of a Gait Feature

It is not easy to model gait patterns because the dimensions of the gait data are large; our 3D skeleton representation $\boldsymbol{S}_t$ is a 45D vector. When modeling high-dimensional data, dimensionality reduction is useful, and principal component analysis (PCA) is one of the most popular techniques. In the low-dimensional latent space, the distribution of data can be represented efficiently. This efficient representation allows us to model the dynamics of the gait data even with a small amount of training data. However, PCA is inappropriate for representing complex temporal data such as gait data.

Gait data are therefore modeled in a low-dimensional latent space acquired by Gaussian Process Dynamical Models (GPDM) [18]. Given a temporal sequence of D-dimensional observation data $\boldsymbol{Y} = (\boldsymbol{y}_1, \cdots, \boldsymbol{y}_T)$, where T denotes the number of training observation data, GPDM [18] acquires the following:

1. A mapping function from the latent space to the observation space:
   $$F_O(\boldsymbol{x}_t) = \boldsymbol{y}_t, \tag{5}$$
   where $\boldsymbol{x}_t$ denotes a d-dimensional latent variable ($d \ll D$) at time t.
2. A sequence of data in the latent space, which is denoted by $\boldsymbol{X} = (\boldsymbol{x}_1, \cdots, \boldsymbol{x}_T)$.
3. A mapping function from t to $t+1$ in the latent space:

$$F_D(\boldsymbol{x}_t) = \boldsymbol{x}_{t+1}.$$

The above two mapping functions are modeled as follows:

$$F_O(\boldsymbol{x}_t) = \sum_j \boldsymbol{b}_j \psi_j(\boldsymbol{x}_t) + \boldsymbol{n}_{y,t}, \quad (6)$$

$$F_D(\boldsymbol{x}_t) = \sum_i \boldsymbol{a}_i \phi_i(\boldsymbol{x}_t) + \boldsymbol{n}_{x,t}, \quad (7)$$

where ϕ_i and ψ_j are basis functions with weights $\boldsymbol{A} = [\boldsymbol{a}_1, \dots]$ and $\boldsymbol{B} = [\boldsymbol{b}_1, \dots]$, and $n_{x,t}$ and $n_{y,t}$ are noise. Under the assumption that the noise is zero-mean Gaussian, the following likelihood for $\boldsymbol{Y}$ can be obtained by marginalization of the basis functions:

$$p(\boldsymbol{Y}|\boldsymbol{X}, \theta) = \frac{1}{\sqrt{(2\pi)^{TD}||K_Y||^D}} \exp\left(-\frac{1}{2}\mathrm{tr}(\boldsymbol{K}_Y^{-1}\boldsymbol{Y}\boldsymbol{Y}^T)\right), \quad (8)$$

where $\boldsymbol{K}_Y$, in which $\boldsymbol{K}_{Y_{i,j}} = k_Y(\boldsymbol{x}_i, \boldsymbol{x}_j)$, is a kernel matrix with hyperparameters θ. Similarly, the likelihood for $\boldsymbol{X}$ can be obtained with hyperparameters θ_X. GPDM training is achieved by optimizing the hyperparameters θ and θ_X for maximizing the joint likelihood for $\boldsymbol{Y}$ and $\boldsymbol{X}$.

A mapping function from the observation space to the latent space is not explicitly provided by GPDM. From one-to-one correspondences between $\boldsymbol{Y}$ and $\boldsymbol{X}$, a mapping function from the observation space to the latent space can be obtained:

$$F_L(\boldsymbol{y}_t) = \boldsymbol{x}_t, \quad (9)$$

We represent F_L by Gaussian Process Regression, GPR.

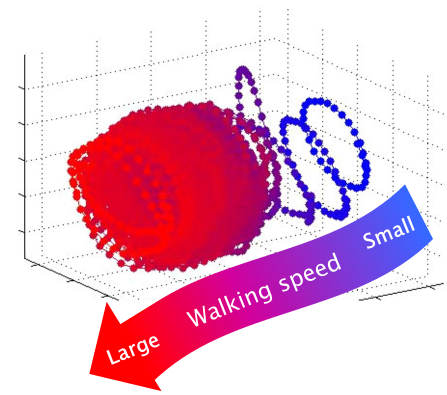
In our experiments, GPDM and GPR use the RBF kernel and are optimized with a full training set (i.e., with no sparse optimization) by a scaled conjugate gradient descent.

GPDM can be optimized not only from one sequence but also from multiple sequences of observation data (i.e., the gait cycles of many subjects, in our case).

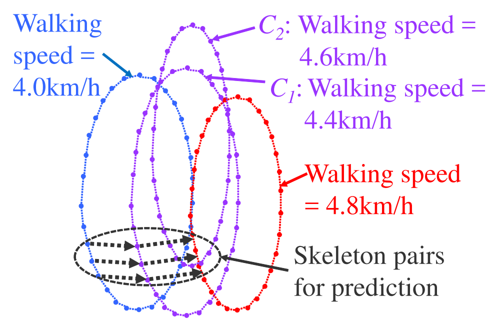
3.2 Learning Relationship between Gait Patterns and a Gait Feature Transition in the Latent Space

Figure 2 shows an example of gait cycles in the latent space. In this paper, GPDM is extended for representing the change in gait cycles in accordance with each gait feature. In the proposed method, the gait cycles should be arranged so that the gait features of neighboring cycles change smoothly as indicated by the color of the gait cycles shown in Fig. 2 (a). This arrangement makes it possible to derive gait prediction while preserving individuality, as shown in Fig. 2 (b) and discusses below in Sect. 3.2. GPDM arranges $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ close to each other in the latent space if their respective 3D skeletons, $\boldsymbol{y}_i$ and $\boldsymbol{y}_j$, are close to each other. With this property, gait cycles can be arranged along the transition of a gait feature in the latent space by embedding the gait feature into the observation data as follows. Each observation, $\boldsymbol{y}_t$, consists of a skeleton, $\boldsymbol{S}_t$, and a gait feature, $g \in \{f_v, f_w, f_\theta, f_s\}$:

$$\boldsymbol{y}_t = \left(\boldsymbol{S}_t^T, g\right)^T \quad (10)$$



(a) Gait cycles in 3D latent space



(b) Mapping between gait cycles

Fig. 2 Gait cycles lying along the transition of a gait feature. This example shows the distribution of gait cycles lying along the walking speed in the 3D latent space. Each point corresponds to a 3D skeleton and each ring-shaped sequence of points indicates a gait cycle. The walking speed becomes larger from gait cycles indicated by red to those indicated by blue.

All elements of this 46D vector $\boldsymbol{y}_t$ are linearly normalized to the range [0, 1] before being applied to GPDM in order to ensure computational stability. By smoothly arranging g of Eq. (10) in the latent space using GPDM, gait cycles represented by a sequence of $\boldsymbol{y}_t$ are smoothly changed according to the gait feature g as illustrated in Fig. 2 (a).

For individuality-preserving gait prediction, the mapping function, $F_{P,k}$, is learned in the latent space:

$$F_{P,k}(\boldsymbol{x}_t(g), \Delta) = \boldsymbol{x}_t(g + \Delta), \quad (11)$$

where $\boldsymbol{x}_t(g)$ corresponds to an observation $\boldsymbol{y}_t = \left(\boldsymbol{S}_t^T, g\right)^T$. $k \in$ {walking speed, step width, leaning forward angle, lateral body sway} denotes the ID of a gait feature. This mapping function is produced for each of the gait features and was implemented using Gaussian process regression in our experiments. For example, assume that k is the walking speed and $F_{P,k}(\boldsymbol{x}_t(g), \Delta)$ is learned with $\Delta = 0.1$ km/h. Given a 3D skeleton observed when $g = 1$ km/h, $F_{P,k}(\boldsymbol{x}_t(g), \Delta)$ predicts the 3D skeleton that will be observed when $g = 1.1$ km/h. Δ should be determined appropriately for each gait feature so that Δ is the minimum step for changing this gait feature. By iterative use of $F_{P,k}(\boldsymbol{x}_t(g), \Delta)$, we can predict a larger change also. For example, the following equation predicts the gait cycle of $(g + 2\Delta)$ from that of g:

$$\boldsymbol{x}_t(g + 2\Delta) = F_{P,k}(F_{P,k}(\boldsymbol{x}_t(g), \Delta), \Delta) \quad (12)$$

For ideal individuality-preserving gait prediction, $F_{P,k}(\boldsymbol{x}_t(g), \Delta)$ is learned using $\boldsymbol{x}_t(g)$ and $\boldsymbol{x}_t(g+\Delta)$ collected from the same subject. However, this condition is unrealistic because a gait cycle including $\boldsymbol{x}_t(g+\Delta)$ is unnatural when the one including $\boldsymbol{x}_t(g)$ is natural in the same subject. Our proposed method therefore assumes that only natural gait cycles are collected from each subject. To approximately meet the aforementioned condition, similar gait cycles of different people are treated as pairs of gait cycles, $\boldsymbol{x}_t(g)$ and $\boldsymbol{x}_t(g+\Delta)$. The generalization capacity of the Gaussian process allows us to predict a change in the gait pattern of different people having similar gait patterns as follows:

Step 1: GPDM gets the latent variables of all N gait cycles.

Step 2: Let $\boldsymbol{X}_i = (\boldsymbol{x}_{i,1}, \cdots, \boldsymbol{x}_{i,T_i})$ be a sequence of T_i latent variables corresponding to the i-th gait cycle. Assume that gait feature k of $\boldsymbol{X}_i$ has the value g. Gait cycles whose gait feature k falls between $g+\Delta-\delta$ and $g+\Delta+\delta$ are extracted from all the gait cycles. The threshold $\delta = 0.1\Delta$, in our experiments. Figure 2 (b) shows an example where the gait feature is the walking speed, $g = 4.0$ km/h, and $\Delta = 0.4$ km/h. In this example, only gait cycle C_1 is extracted because the walking speed of C_2 is beyond the threshold.

Step 3: Let $\boldsymbol{X}'_j$ be the j-th gait cycle in a set of gait cycles extracted in step 2. The root mean square error (RMSE) between $\boldsymbol{X}_i$ and $\boldsymbol{X}'_j$ is computed. RMSE between $\boldsymbol{X}_i$ and $\boldsymbol{X}'_j$ is computed as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{T_i}\sum_{p=1}^{T_i} \|\boldsymbol{x}_{i,p} - \boldsymbol{x}_{j,q'}\|^2} \tag{13}$$

$$q' = \operatorname*{argmin}_{q}^{T_j} \|\boldsymbol{x}_{i,p} - \boldsymbol{x}_{j,q}\|^2$$

Let $\hat{\boldsymbol{X}}'$ be the one whose RMSE is the smallest over all the extracted gait cycles. The pair of $\boldsymbol{X}_i$ and $\hat{\boldsymbol{X}}'$ is used for learning the mapping function $F_{P,k}$ in the next step. In the example shown in Fig. 2 (b), each dotted arrow depicts a pair. The primary objective of this process is to find the closest pair framewise for learning smooth mapping $F_{P,k}$. Therefore, temporal continuity of a gait cycle is neglected in Eq. (13), while it is useful for gait-wise matching taking into account temporal continuity such as dynamic time warping. Note that some gait cycles having a higher value of g (more specifically, values near the highest g) may not have the paired sample, $\hat{\boldsymbol{X}}'$, and are not used in the following steps.

Step 4: After steps 2 and 3 have been performed for $\forall i \in \{1, \cdots, N\}$, the mapping function $F_{P,k}$ is learned from all the pairs of $\boldsymbol{X}_i$ and $\hat{\boldsymbol{X}}'$.

Step 5: Steps 2, 3, and 4 are performed for all gait features.

3.3 Gait Prediction through the Latent Space

Given a gait cycle, $\boldsymbol{Y} = (\boldsymbol{y}_1, \cdots, \boldsymbol{y}_T)$, where $\boldsymbol{y}_t$ is expressed by Eq. (10), $\boldsymbol{y}_t$ is projected to the latent space by $\boldsymbol{x}_t = F_L(\boldsymbol{y}_t)$ (Eq. (5)). From $\boldsymbol{x}_t$, individuality-preserving gait prediction for gait feature k is achieved by $\boldsymbol{x}'_t = F_{P,k}(\boldsymbol{x}_t, \Delta)$ (Eq. (11)). $\boldsymbol{x}'_t$ is then projected back to the observation space by $\boldsymbol{y}'_t = F_O(\boldsymbol{x}'_t)$ (Eq. (9)). A 3D skeleton $\boldsymbol{S}_t$ that consists of 15 $\boldsymbol{J}_{t,c}$ (Eq. (4)) is extracted from $\boldsymbol{y}'_t$ to acquire the 3D coordinates of all joints $\boldsymbol{P}'_{t,i}$ (Eq. (3)).

4. Evaluation with Ground Truth

The proposed method was evaluated quantitatively by comparing predicted skeletons with their respective ground truth. This comparison was conducted using the walking speed as a gait feature. Each subject walked on a treadmill at 10 different speeds ranging from 2.0 km/h to 5.6 km/h at intervals of 0.4 km/h. Kinect V2 measured the subject as illustrated in Fig. 3. The treadmill allowed the subject to walk at the predefined speed. One gait cycle was measured at each speed, after the subject had walked until the gait became stable and natural. We recruited 24 subjects with ages ranging between 22 and 42 years of age. In total, $10 \times 24 = 240$ gait cycles were measured. With these data, we conducted cross validation, where the gait cycles of 5 and 19 subjects were used for testing and learning, respectively. Δ was 0.4 km/h in accordance with the speed interval of the learning samples.

From the predicted data $\boldsymbol{y}'_t$ at each t-th frame, the 3D coordinates of all joints $\boldsymbol{P}'_{t,i}$ ($i \in \boldsymbol{B}$ where $\boldsymbol{B}$ denotes a set of all joints defined in Eq. (1)) were acquired as described in Sect. 3.3. Then the 3D error distance between each joint of the predicted skeleton, $\boldsymbol{P}'_{t,i}$, and its ground truth, $\boldsymbol{P}^{(gt)}_{t,i}$, is computed one by one in order to evaluate the RMSE of $\boldsymbol{P}'_{t,i}$.

$$\mathrm{RMSE} = \sqrt{\frac{1}{T|\boldsymbol{B}|}\sum_{t=1}^{T}\sum_{i\in\boldsymbol{B}} \|\boldsymbol{P}^{(gt)}_{t,i} - \boldsymbol{P}'_{t,i}\|^2}, \tag{14}$$

where T denotes the number of frames in this gait.

Table 1 shows the mean RMSEs (mm) over all subjects. In Table 1, the row and column indicate the walking speeds of the source and predicted gait cycles, respectively. While our model can predict only $s_{now} + \Delta$ km/h gait cycle from s_{now} km/h gait cycle where s_{now} is the walking speed of the source gait, i-times iteration of our prediction model allows us to predict $s_{now} + i\Delta$ km/h gait cycle, as shown in Eq. (12). It is called the iterative prediction process. Single use of our prediction model is called one-time prediction process. A value in the row of s_s km/h and the column of s_d km/h denotes the mean RMSE (mm) between the ground truth of

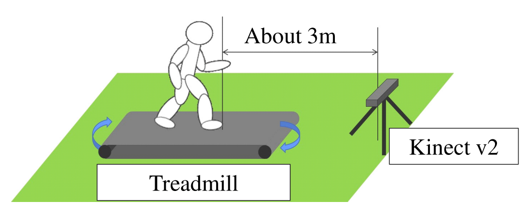


Fig. 3 Gait measurement with a treadmill using Kinect V2.

Table 1 Mean RMSEs (mm) of the 3D coordinates of skeleton joints in predicted gait cycles. Gait cycles were predicted for different walking speeds. The velocity indicated in the row and column denotes the walking speeds of the source and predicted gait cycles, respectively. Each RMSE was computed by subtraction of the predicted 3D skeleton from its corresponding ground truth.

	Predicted walking speed (km/h)								
	2.4	2.8	3.2	3.6	4.0	4.4	4.8	5.2	5.6
2.0	28	31	33	33	36	38	43	44	49
2.4		27	31	31	33	35	42	43	48
2.8			30	31	29	34	41	42	48
3.2				29	28	31	33	35	41
3.6					29	29	33	34	40
4.0						33	36	36	42
4.4							31	32	37
4.8								29	34
5.2									33

Table 2 Comparative experiments. The mean RMSEs (mm) of the 3D coordinates of skeleton joints in predicted gait cycles are computed by three different methods.

	Predicted walking speed (km/h)								
	2.4	2.8	3.2	3.6	4.0	4.4	4.8	5.2	5.6
Ours	28	27	30	29	29	33	31	29	33
GPDM	35	41	40	38	34	44	45	52	56
Mean	50	49	64	50	44	41	42	84	44

the s_d km/h gait cycle and the s_d km/h gait cycle predicted from the s_s km/h gait cycle.

The following observations are made from the experimental results:

- Iterative prediction processes increase the error. For example, from the 2.0 km/h gait cycle, the mean RMSE when predicting the 5.6 km/h gait cycle (i.e., 49 mm) was larger than when predicting the 2.4 km/h gait cycle (i.e., 28 mm).
- The mean RMSEs in one-time prediction process were a bit larger at faster speeds. For example, the mean RMSE of the prediction from the 5.2 km/h gait cycle to the 5.6 km/h gait cycle (i.e., 33 mm) is slightly greater than that from 2.0 km/h to 2.4 km/h (i.e., 28 mm). This is because the range of joint motion is larger at faster walking speeds.

The effectiveness of the proposed method was demonstrated by comparing its RMSEs with those of different methods. Table 2 shows mean RMSEs in one-time prediction process results (e.g., from a 2.0 km/h gait to a 2.4 km/h gait) with the mean RMSEs of our proposed method in the top row. To confirm the effect of our extension proposed in Sect. 3.2, the latent space was modeled by GPDM [18] with no gait features in the observation data (i.e., $\boldsymbol{y}_i = \boldsymbol{S}_i^T$). The results of only GPDM are shown in the second row of Table 2. The bottom row shows the mean RMSEs between the means of all observed gait cycles at s_d km/h, which was regarded as a base result for individuality-ignoring gait prediction, and each observed gait cycle at s_d km/h.

At all walking speeds, the proposed method was superior. The effect of individuality-preservation can be seen from the fact that both approaches outperformed the base result. The superiority of the proposed method over the one using only GPDM [18] is achieved by employing a gait feature in the observation data (i.e., $\boldsymbol{y}_i = (\boldsymbol{S}_i^T, g)$). Since g allows the gait cycles to be arranged in the latent space so that their gait features change smoothly between neighboring gait cycles (as illustrated in Fig. 2), it is easy to model the individuality-preserving mapping function $f_{P,k}$.

Table 3 Mean RMSE (mm) between the predicted 3D coordinates and their ground-truth values in each joint. For this evaluation, all results shown in Table 1 are used.

Neck	SShoulder	SMid	Sbase	LShoulder	RShoulder
10	31	48	60	11	15
LElbow	RElbow	LWrist	RWrsit	Lhip	RHip
22	30	40	39	42	48
LKnee	Rknee	LAnkle	RAnkle	–	–
23	28	65	59	–	–

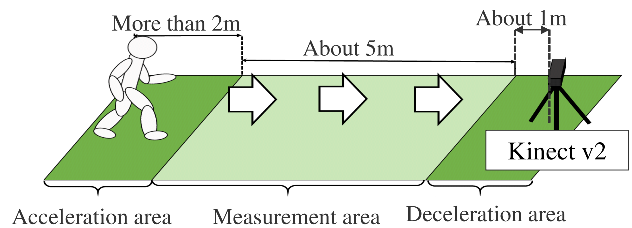


Fig. 4 Gait measurement on a floor using Kinect V2.

For detailed analysis, the mean RMSE of each joint over all subjects is shown in Table 3. Here, we focus on the RMSEs of lower-body joints because these joints are dominant for gait analysis. In the literature [31], it is shown that 50mm error is high enough for several rehabilitation tasks. While 50mm error is not small enough, we regard this criterion as a soft threshold in screening for inappropriate prediction methods for gait rehabilitation tasks. In Table 3, we can see that a maximum error up to 50mm is guaranteed in all joints except noisy ankles.

5. Evaluation with Realistic Condition

Gait data were measured from 206 subjects with a mean age of 78.5 in an elderly care home. Kinect V2 measured the gait pattern of each subject when the subject walked in a straight line for 7m. The experimental environment is illustrated in Fig. 4. While the official specification guarantees that Kinect V2 reliably tracks a body skeleton within 4.5m, a subject was observed within the range of 1m to 6m in our experiments as shown in this figure. This is because the gait cycle of a few subjects could not be captured within 4.5m, while most gait cycles were captured within 4.5m. Each subject was measured five times, and one gait cycle was extracted from each trial. In total, $5 \times 206 = 1030$ gait cycles were collected and used for experiments.

In this dataset, the ground truth of a predicted gait is not available. For validating the proposed method, the effect of individuality preservation is quantitatively evaluated by

Table 4 Difference/distribution in the leaning forward angle (degrees). Its standard deviation (SD) over all measured gait cycles is shown in the leftmost column. In other columns, the mean difference (degree) between source and predicted gaits is shown.

SD	Mean difference (degrees)		
	Walking speed	Lateral sway	Step width
8.2	2.9	1.1	1.2

comparing gait features between source and predicted gaits. If the proposed method is able to preserve individuality in a predicted gait when a gait feature is changed, other gait features in the predicted gait should be similar to those in the source gait; this is a definition of individuality preservation. The change in the gait feature was evaluated by comparing it with the distribution of the feature over all measured gait cycles. In our experiments, the leaning forward angle was used as a gait feature compared between the source and predicted gaits, while each of other gait features (i.e., walking speed, lateral sway, and step width) is changed for gait prediction. This is because, in our observed gait cycles, the distribution of the leaning forward angle was relatively smaller within each subject, which is good for validating the effect of individuality preservation.

The prediction models were learned for five gait features; walking speed, step width, leaning forward angle, and lateral body sway. In total, five prediction models were learned for each test data. For all prediction models, the latent space had three dimensions. Δ was set to 1.16 km/h, 13 mm, 19 mm, and 4.3 degrees for walking speed, lateral body sway, step width, and leaning forward angle, respectively. For evaluation, onefold cross validation was executed.

The mean difference of the leaning forward angle between the source and predicted gait cycles is shown in Table 4. For comparison, the standard deviation of the leaning forward angle over the gait cycles of all subjects, which corresponds to the sample distribution, was also computed. As expected, the mean difference is smaller than the standard deviation.

6. Concluding Remarks

This paper proposed a method for individuality-preserving gait prediction. A gait cycle is predicted for the variation in five gait features related to aging and a risk of falling. The complex change in the gait cycle is predicted while preserving individuality by latent space modeling extended with the gait feature. The effectiveness of the proposed method was demonstrated quantitatively compared with prediction methods that ignore individuality.

Future work includes detailed quantitative evaluation against ground truth. Quantitative evaluation against ground truth was conducted only for walking speed in this paper, but it would be interesting to extend this evaluation to other gait features. It is, however, difficult to obtain natural varying gait patterns from a single individual, while such gait patterns may be collected by observation over a long term (e.g., 10 years or more). Instead of the long-term observation, a gait dataset derived from a huge number of subjects may provide a dataset of similar gait patterns that can be treated as mimicking those from a single subject.

This study was supported by JSPS KAKENHI Grant Number 15H01583.

References

[1] H.B. Menz, S.R. Lord, and R.C. Fitzpatrick, "Age-related differences in walking stability," Age and Aging, vol.32, no.2, pp.137–142, 2003.
[2] M.C. Bisi and R. Stagni, "Complexity of human gait pattern at different ages assessed using multiscale entropy: From development to decline," Gait & Posture, vol.47, pp.37–42, 2016.
[3] M. Arif, Y. Ohtaki, R. Nagatomi, and H. Inooka, "Analysis of the effect of fatigue on walking gait using acceleration sensor placed on the waist," International Journal of Engineering Intelligent Systems for Electrical Engineering and Communications, vol.18, no.2, pp.85–95, 2010.
[4] S. Morita, H. Yamamoto, and K. Furuya, "Gait analysis of hemiplegic patients by measurement of ground reaction force," Scand. J. Rehabil. Med., vol.27, no.1, pp.37–42, 1995.
[5] X. Wang and Y. Wang, "Gait analysis of children with spastic hemiplegic cerebral palsy," Neural Regeneration Research, vol.7, no.20, pp.1578–1584, 2012.
[6] P.C.R. Santos, L.T.B. Gobbi, D. Orcioli-Silva, L. Simieli, J.H. van Dieën, and F.A. Barbieri, "Effects of leg muscle fatigue on gait in patients with Parkinson's disease and controls with high and low levels of daily physical activity," Gait & Posture, vol.47, pp.86–91, 2016.
[7] S.K. Banala, S.H. Kim, S.K. Agrawal, and J.P. Scholz, "Robot assisted gait training with active leg exoskeleton (ALEX)," IEEE Trans. Neural. Syst. Rehabil. Eng., vol.17, no.1, pp.2–8, 2009.
[8] K. Shamaei, M. Cenciarini, A.A. Adams, K.N. Gregorczyk, J.M. Schiffman, and A.M. Dollar, "Design and evaluation of a quasi-passive knee exoskeleton for investigation of motor adaptation in lower extremity joints," IEEE Trans. Biomed. Eng., vol.61, no.6, pp.1809–1821, 2014.
[9] N. Ukita, D. Kaulen, and C. Röcker, "A user-centered design approach to physical motion coaching systems for pervasive health," Smart Health, Lecture Notes in Computer Science, vol.8700, pp.189–208, Springer International Publishing, Cham, 2015.
[10] Y. Masugi, T. Kitamura, K. Kamibayashi, T. Ogawa, T. Ogata, N. Kawashima, and K. Nakazawa, "Velocity-dependent suppression of the soleus H-reflex during robot-assisted passive stepping," Neuroscience Letters, vol.584, pp.337–341, 2015.
[11] K. Kamibayashi, T. Nakajima, M. Takahashi, and K. Nakazawa, "Changes in input-output relations in the corticospinal pathway to the lower limb muscles during robot-assisted passive stepping," 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp.4140–4144, 2011.
[12] J. Tajino, A. Ito, M. Nagai, X. Zhang, S. Yamaguchi, H. Iijima, T. Aoyama, and H. Kuroki, "Discordance in recovery between altered locomotion and muscle atrophy induced by simulated microgravity in rats," Journal of Motor Behavior, vol.47, no.5, pp.397–406, 2015.
[13] Y. Makihara, T. Tanoue, D. Muramatsu, Y. Yagi, S. Mori, Y. Utsumi, M. Iwamura, and K. Kise, "Individuality-preserving silhouette extraction for gait recognition," IPSJ Trans. Computer Vision and Applications, vol.7, pp.74–78, 2015.
[14] R.K. Begg, M. Palaniswami, and B. Owen, "Support vector machines for automated gait classification," IEEE Trans. Biomed. Eng., vol.52, no.5, pp.828–838, 2005.
[15] T. Higashiguchi, T. Shimoyama, N. Ukita, M. Kanbara, and N. Hagita, "Classification of gait anomaly due to lesion using full-body gait motions," IEICE Trans. Inf. & Syst., vol.E100-D, no.4, pp.874–881, April 2017.

[16] R. Begg and J. Kamruzzaman, "A machine learning approach for automated recognition of movement patterns using basic, kinetic and kinematic gait data," Journal of Biomechanics, vol.38, no.3, pp.401–408, 2005.
[17] D.T.H. Lai, A. Shilton, E. Charry, R. Begg, and M. Palaniswami, "A machine learning approach to k-step look-ahead prediction of gait variables from acceleration data," 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp.284–387, 2009.
[18] J.M. Wang, D.J. Fleet, and A. Hertzmann, "Gaussian process dynamical models for human motion," IEEE Trans. Pattern Anal. Mach. Intell., vol.30, no.2, pp.283–298, 2008.
[19] N. Ukita and T. Kanade, "Gaussian process motion graph models for smooth transitions among multiple actions," Computer Vision and Image Understanding, vol.116, no.4, pp.500–509, 2012.
[20] N. Ukita, "Simultaneous particle tracking in multi-action motion models with synthesized paths," Image Vision Comput., vol.31, no.6-7, pp.448–459, 2013.
[21] M. Hansard, S. Lee, O. Choi, and R. Horaud, Time-of-Flight Cameras: Principles, Methods and Applications, SpringerBriefs in Computer Science, Springer, Oct. 2012.
[22] S. Zennaro, M. Munaro, S. Milani, P. Zanuttigh, A. Bernardi, S. Ghidoni, and E. Menegatti, "Performance evaluation of the 1st and 2nd generation Kinect for multimedia applications," International Conference on Multimedia and Expo, 2015.
[23] X. Xu, R.W. McGorry, L.-S. Chou, J.-H. Lin, and C.-C. Chang, "Accuracy of the Microsoft Kinect™ for measuring gait parameters during treadmill walking," Gait & Posture, vol.42, no.2, pp.145–151, 2015.
[24] B. Galna, G. Barry, D. Jackson, D. Mhiripiri, P. Olivier, and L. Rochester, "Accuracy of the Microsoft Kinect sensor for measuring movement in people with Parkinson's disease," Gait & Posture, vol.39, no.4, pp.1062–1068, 2014.
[25] M. Schimpl, C. Moore, C. Lederer, A. Neuhaus, J. Sambrook, J. Danesh, W. Ouwehand, and M. Daumer, "Association between walking speed and age in healthy, free-living individuals using mobile accelerometry — A cross-sectional study," PLoS ONE, vol.6, no.8, e23299, 2011.
[26] M. Lövdén, S. Schaefer, A.E. Pohlmeyer, and U. Lindenberge, "Walking variability and working-memory load in aging: A dual-process account relating cognitive control to motor control performance," The Journals of Gerontology, Series B, Psychological Sciences and Social Sciences, vol.63, no.3, pp.P121–P128, 2008.
[27] J.M. Hausdorff, M.E. Nelson, D. Kaliton, J.E. Layne, M.J. Bernstein, A. Nuernberger, and M.A.F. Singh, "Etiology and modification of gait instability in older adults: A randomized controlled trial of exercise," Journal of Applied Physiology, vol.90, no.6, pp.2117–2129, 2001.
[28] B. Mariani, S. Rochat, C.J. Büla, and K. Aminian, "Heel and toe clearance estimation for gait analysis using wireless inertial sensors," IEEE Trans. Biomed. Eng., vol.59, no.11, pp.3162–3168, 2012.
[29] D.D. Espy, F. Yang, T. Bhatt, and Y.-C. Pai, "Independent influence of gait speed and step length on stability and fall risk," Gait & Posture, vol.32, no.3, pp.378–382, 2010.
[30] M.S. Orendurff, A.D. Segal, G.K. Klute, J.S. Berge, E.S. Rohr, and N.J. Kadel, "The effect of walking speed on center of mass displacement," J. Rehabil. Res. Dev., vol.41, no.6A, pp.829–834, 2004.
[31] K. Otte, B. Kayser, S. Mansow-Model, J. Verrel, F. Paul, A.U. Brandt, and T. Schmitz-Hübsch, "Accuracy and reliability of the Kinect version 2 for clinical measurement of motor function," PLoS One, vol.11, no.11, e0166532, 2016.

Tsuyoshi Higashiguchi got the M.E. degree from Nara Institute of Science and Technology. His research theme was evaluation of human physical motion such as gait.

Norimichi Ukita is a professor at the graduate school of engineering, Toyota Technological Institute, Japan (TTI-J). He received the Ph.D degree in Informatics from Kyoto University, Japan, in 2001. After working for five years as an assistant professor at NAIST, he became an associate professor in 2007 and moved to TTI-J in 2016. He was a research scientist of PRESTO, JST during 2002–2006. He was a visiting research scientist at Carnegie Mellon University during 2007–2009. His main research interests are multi-object tracking and human pose and activity estimation.

Masayuki Kanbara was received a Ph.D. degree in engineering from Nara Institute of Science and Technology (NAIST) in 2002. He was an assistant professor at the Information Science Department at NAIST since 2002. He was a visiting researcher of University of California, Santa Barbara in 2008–2009. He has been an associate professor at NAIST since 2010. Fields of research are focused on augmented reality, computer vision and Human-Robot Interaction.

Norihiro Hagita received the Ph.D. degree in electrical engineering from Keio University in 1986. In 1978, he joined Nippon Telegraph and Telephone Public Corporation (Now NTT). He is currently Board Director of ATR and ATR Fellow, director of the Social Media Research Laboratory Group and the Intelligent Robotics and Communication Laboratories. He is the chairman of ATR Creative. He is also a visiting professor of Nara Institute of Science and Technology, Osaka University, and Kobe University. His major interests are cloud networked robotics, human-robot interaction, ambient intelligence, pattern recognition and learning, and datamining technology. He has served as a chairman of technical committee in Network Robot Forum in Japan.