Skeleton Clustering by Multi-Robot Monitoring for Fall Risk Discovery

Deguchi, Yutaka
ISEE, Kyushu University

Takayama, Daisuke
ISEE, Kyushu University

Takano, Shigeru
Kyushu University

Scuturici, Vasile-Marian
CNRS, INSA de Lyon, France

他

出版情報：Journal of Intelligent Information Systems. 48 (1), pp.75-115, 2015-12-14. Springer-Verlag
バージョン：
権利関係：

Noname manuscript No.
(will be inserted by the editor)

Skeleton Clustering by Multi-Robot Monitoring for Fall Risk Discovery

Yutaka Deguchi · Daisuke Takayama
Shigeru Takano · Vasile-Marian Scuturici
Jean-Marc Petit · Einoshin Suzuki

Received: date / Accepted: date

Abstract This paper tackles the problem of discovering subtle fall risks using skeleton clustering by multi-robot monitoring. We aim to identify whether a gait has fall risks and obtain useful information in inspecting fall risks. We employ clustering of walking postures and propose a similarity of two datasets with respect to the clusters. When a gait has fall risks, the similarity between the gait which is being observed and a normal gait which was monitored in advance exhibits a low value.

In subtle fall risk discovery, unsafe skeletons, postures in which fall risks appear slightly as instabilities, are similar to safe skeletons and this fact causes the difficulty in clustering. To circumvent this difficulty, we propose two instability features, the horizontal deviation of the upper and lower bodies and the curvature of the back, which are sensitive to instabilities and a data preprocessing method which increases the ability to discriminate safe and unsafe skeletons.

To evaluate our method, we prepare seven kinds of gait datasets of four persons. To identify whether a gait has fall risks, the first and second experiments use normal gait datasets of the same person and another person, respectively. The third experiments consider that how many skeletons are necessary to identify whether a gait has fall risks and then we inspect the obtained clusters. In clustering more

Y. Deguchi · D. Takayama · S. Takano · E. Suzuki
ISEE, Kyushu University, 744 Motooka, Nishi, Fukuoka, Japan
E-mail: yutaka.kyushu@gmail.com

D. Takayama
E-mail: takayamad@gmail.com

S. Takano
E-mail: takano@inf.kyushu-u.ac.jp

E. Suzuki
E-mail: suzuki@inf.kyushu-u.ac.jp

V-M. Scuturici · J-M. Petit
Université de Lyon, CNRS, INSA-Lyon, Villeurbanne, France
E-mail: Marian.Scuturici@insa-lyon.fr

J-M. Petit
E-mail: jean-marc.petit@insa-lyon.fr

than 500 skeletons, the combination of the proposed features and our preprocessing method discriminates gaits with fall risks and without fall risks and gathers unsafe skeletons into a few clusters.

Keywords Skeleton Clustering · Human Monitoring · Mobile Robots · Service-oriented DSMS

1 Introduction

In these years, a rapid growth of the elderly population and dangers of accidental falls result in significant problems in the world. In 2050, older people (age ≥ 65) are forecasted to reach 997 million (11.6% of the world population) [United Nations, Department of Economic and Social Affairs, Among the various problems the elderly face in their daily lives, accidental falls are critical [Rubenstein(2006)], e.g., in the U.S., falls occur in 30-60% of older adults each year, and 10-20% of these falls result in injury, hospitalization, and/or death [Rubenstein(2006)]. Recently, many researchers proposed early detection methods for the accidental falls using static cameras or sensors [Gjoreski et al(2012)Gjoreski, Lustrek, and Gams, Rougier et al(2011)Rougier, Auvinet, Rousseau, Mignotte, and Meunier,Sixsmith and Johnson(2004)]. We tackle developing a monitoring system to discover fall risks because it is important to prevent injuries caused by the falls such as fractures and sprains. Since 17% of the causes of these falls are "Gait/balance disorders or weakness" [Rubenstein(2006)], our system monitors walking postures to discover fall risks related with instabilities of the postures.

We adopt a multi-robot system for monitoring a human because it can collect data of the human from a suitable position while the human walks. We have witnessed a rapid advancement of mobile robots and hence their applications in various domains such as industry, medicine, security, social infrastructure, and transportation [Dunbabin and Marques(2012),Leibrandt et al(2014)Leibrandt, Marcus, Kwok, and Yang, Neumann et al(2012)Neumann, Asadi, Lilienthal, Bartholmai, and Schiller,Palunko et al(2012)Palunko, Cruz, and Fierro, Pippin and Christensen(2014)]. Recently, many researchers developed mobile robots which work close to humans, e.g., navigation, assistance and monitoring [Ardiyanto and Miura(2014), Coradeschi et al(2013)Coradeschi, Cesta, Cortellessa, Coraci, Gonzalez, Karlsson, Furfari, Loutfi, Orlandini, Palumbo, Pec Fischinger et al(2013)Fischinger, Einramhof, Wohlkinger, Papoutsakis, Mayer, Panek, Koertner, Hofmann, Argyros, Vincze Martinson and Yalla(2014)]. Autonomous mobile robots possess a potential to drastically change our society by their ability to sense, move, and reason.

Among various kinds of reasoning performed by an autonomous mobile robot, we believe discovery is one of the most promising and challenging research issues. Naturally it is impossible for its designers and developers to foresee all kinds of situations that the robot encounters. For instance, a monitoring robot would benefit from its ability to analyze its observation and discover patterns to adapt to its target person [Deguchi et al(2014)Deguchi, Takayama, Takano, Scuturici, Petit, and Suzuki]. With this motivation in our mind, we have proposed several autonomous mobile robots which perform online clustering to discover useful patterns [Deguchi and Suzuki(2014), Kondo et al(2014)Kondo, Deguchi, and Suzuki,Suzuki et al(2012)Suzuki, Matsumoto, and Kouno, Suzuki et al(2014)Suzuki, Deguchi, Takayama, Takano, Scuturici, and Petit,Takayama et al(2014)Takayama, Deguchi, Taka which we call data mining robots.

In this paper, we develop a multi-robot monitoring system based on a service-oriented data stream management system (DSMS) [Gripay et al(2010)Gripay, Laforest, and Petit]

to discover fall risks inherent in a gait of a human. We define subtle fall risks as fall risks which are too small to be recognized at a glance and discover them to avoid falls. Our subtle fall risk discovery aims to (1) discriminate gaits with fall risks and gaits without fall risks and (2) obtain useful information in inspecting the fall risks. Our system collects posture data of a walking human using mobile robots and uses a clustering method to discover subtle fall risks.

This paper is organized in the following way. In the next section, we list related works. Section 3 shows human monitoring in our problem, a definition of the fall risk discovery problem and the motivation to tackle subtle fall risk discovery. In Section 4, we explain the detail of our multi-robot monitoring system based on a service-oriented DSMS. Section 5 presents our proposed methods for clustering and evaluation. In Section 6, we conduct three kinds of experiments to evaluate our proposed methods for clustering and discuss the results. Section 7 concludes this paper.

2 Related Works

Many researchers tackled the problems of analyzing and detecting the falls of the elderly people. From the 1980s to the 2000s, there were many researches about risk factors of accidental falls. Tinetti et al. investigated the individual chronic characteristics associated with the falls among elderly persons and tested the hypothesis that the risk of the falls increases as the number of chronic disabilities increases [Tinetti et al(1986)Tinetti, Williams, and Mayewski]. They revealed that falls among several elderly persons resulted from accumulated effects of multiple specific disabilities. Rubenstein surveyed studies about causes of falls, risk factors and therapeutic and preventive approaches [Rubenstein(2006)].

From the 2000s, many researchers developed methods for automatic early detection of accidental falls. Sixsmith and Johnson developed a prototype system which detects falls using thermal-imaging sensors and then raises alarms [Sixsmith and Johnson(2004)]. They employed a neural network to classify falls and other kinds of events using vertical-velocity of the target. Rougier et al. [Rougier et al(2011)Rougier, Auvinet, Rousseau, Mignotte, and Meunier] used a Kinect sensor to calculate the distance from the 3D centroid of the human body to the ground and the vertical velocity of the centroid. They determined the best thresholds from the training dataset to classify "sitting down actions", "crouching down actions" and "falls". Gjoreski et al. employed body-worn inertial and location sensors for fall detection [Gjoreski et al(2012)Gjoreski, Lustrek, and Gams]. They proposed a fall detection method based on the use of context information which consists of the user's body accelerations, activities and locations.

In recent years, mobile robots are used to monitor older adults and fall detection is applied as a part of monitoring. GiraffPlus[1] and Hobbit[2], which were developed in large-scale EU projects, achieved fall detection by mobile robots successfully. GiraffPlus is a complex system for early detection and adaptive support to changing individual's needs and the heart of the system is a unique telepresence robot [Coradeschi et al(2013)Coradeschi, Cesta, Cortellessa, Coraci, Gonzalez, Karlsson, Furfari, Loutfi, Orlandini, Palumb
The system generates alarms and informs remote caregivers about location of the

[1] http://www.giraffplus.eu/

[2] http://hobbit.acin.tuwien.ac.at/

faller when a fall occurs [Koshmak et al(2013)Koshmak, Linden, and Loutfi]. Hobbit consists of a socially assistive robot which supports older adults at home and detects a fall to call for help if necessary [Fischinger et al(2013)Fischinger, Einramhof, Wohlkinger, Papoutsakis, Mayer, Panek, In addition, the robot picks up objects to clear the floor for fall prevention.

We employ clustering of walking postures for subtle fall risk discovery because it is difficult to foresee all diverse risks. Posture clustering is studied by many researchers for gesture recognition, action recognition and so on. Wang used a pairwise clustering algorithm to extract salient postures which can characterize actions [Wang(2010)]. In the paper, non-Euclidean relational fuzzy c-means was adopted to cope with the fuzziness of the histogram of oriented gradients [Wang(2010)] which is extracted from two posture images. Diraco et al. presented an active vision system for an automatic fall detection and a posture recognition method for elderly healthcare [Diraco et al(2010)Diraco, Leone, and Siciliano]. As a part of the research, they proposed thresholding-based clustering of 3D postures using the distance from the ground to the human centroid as a feature for fall detection. Pal et al. tackled the problem of recognizing health care linked gestures of young individuals by using Kinect [Pal et al(2014)Pal, Saha, and Konar]. They applied fuzzy c-means and principal component analysis to feature vectors each of which consists of 171 normalized distances between each pair of joints of the skeleton.

3 Skeleton Clustering with Instability Features for Fall Risk Discovery

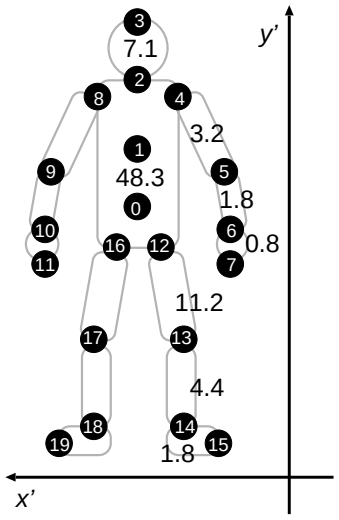
3.1 Monitoring Human by Mobile Robots

In this paper, we assume that human monitoring for fall risk discovery is conducted in a room of which the size is about 6.5m $\times$ 4.5m, so we use two mobile robots due to the space capacity. Our monitoring system involves two autonomous mobile robots $\mathbf{R}_1$ and $\mathbf{R}_2$, which monitor a target human $\mathbf{H}$, and a fixed PC, which manages the robots. The management PC communicates with each robot via a wireless network. We assume that $\mathbf{H}$ is in an indoor space, though our system could be used outdoors if we employ appropriate robots.

While the monitoring system is operating, $\mathbf{H}$ is either standing still or walking around. In this paper, we assume at a first step that there are no obstacles between $\mathbf{R}_k$ and $\mathbf{H}$ to focus postures of $\mathbf{H}$[3] because obstacles such as furnitures and other humans cause several problems for movement and monitoring of the robots. In this paper, we especially focus the functionality of the robot to discover fall risks.

To discover fall risks, our system collects postures of the target human and clusters them incrementally while the human walks around a room. After the walking for a few minutes, our system discovers useful information about the fall risks which the gait of the human has. In fall risk discovery, we aim to (1) discriminate gaits with fall risks and gaits without fall risks and (2) obtain useful information in inspecting fall risks. Since we conduct clustering of walking postures, the second

[3] It is surely better to conduct fall risk discovery at home. However, if fall risk discovery is conducted in another situation, it has an advantage over early fall detection in an aspect because prevention is better than cure. Fall risk discovery aims to find fall risks before the target human is likely to fall.



ID j	Name	weight w_j
0	HipCenter	6.90
1	Spine	6.90
2	ShoulderCenter	10.45
3	Head	3.55
4	ShoulderLeft	8.50
5	ElbowLeft	2.50
6	WristLeft	1.30
7	HandLeft	0.40
8	ShoulderRight	8.50
9	ElbowRight	2.50
10	WristRight	1.30
11	HandRight	0.40
12	HipLeft	12.50
13	KneeLeft	7.80
14	AnkleLeft	3.10
15	FootLeft	0.90
16	HipRight	12.50
17	KneeRight	7.80
18	AnkleRight	3.10
19	FootRight	0.90

Fig. 1 20 joints of Kinect. The left shows the position of each joint and the mass ratio [Clauser et al(1969)Clauser, McConville, and Young] of body parts. The right explains each joint and its weight ratio, where j represents the ID of the joint $\mathbf{p}_{j,t}$.

aim necessitates us to obtain a few clusters which (mostly) consist of unbalanced postures.

3.2 Fall Risk Discovery

Our monitoring system collects a skeleton[4] $\mathcal{S}(t, r(t))$ at time t which is observed by its vision system, Kinect for Windows[5] mounted on the robot $\mathbf{R}_{r(t)}$, as a posture of human $\mathbf{H}$. $\mathcal{S}(t, r(t))$ consists of 20 joints $(\mathbf{p}_{0,t}, \mathbf{p}_{1,t}, ..., \mathbf{p}_{19,t})$ as shown in Fig. 1. Kinect outputs $\mathbf{p}_{j,t}$ as a point $(\mathbf{p}_{j,t}.x, \mathbf{p}_{j,t}.y, \mathbf{p}_{j,t}.z)$ in the 3D space with the position of the sensor as its origin (shown in Fig. 2). Joint $\mathbf{p}_{j,t}$ also has tracking state $\mathbf{p}_{j,t}.s$ which indicates 1: Tracked or 2: Inferred when the robot $\mathbf{R}_{r(t)}$ monitors the human $\mathbf{H}$.

We define the fall risk discovery as a clustering problem of skeletons $\{\mathcal{S}(1, r(1)), ..., \mathcal{S}(T, r(T))\}$ $(r(t) = 1, 2)$ observed by the two robots $\mathbf{R}_1$ and $\mathbf{R}_2$, where T represents the number of the skeletons. Before clustering skeletons, we use shift and rotation transformation as shown in Equation (1) to convert $\mathbf{p}_{j,t}$ so that $\mathbf{p}_{0,t}$ becomes the point of origin, and $x'z'$ plane becomes parallel to the ground.

$$\begin{bmatrix} \mathbf{p}_{j,t}.x' \\ \mathbf{p}_{j,t}.y' \\ \mathbf{p}_{j,t}.z' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi_{r(t)} & -\sin\phi_{r(t)} \\ 0 & \sin\phi_{r(t)} & \cos\phi_{r(t)} \end{bmatrix} \begin{bmatrix} \mathbf{p}_{j,t}.x - \mathbf{p}_{0,t}.x \\ \mathbf{p}_{j,t}.y - \mathbf{p}_{0,t}.y \\ \mathbf{p}_{j,t}.z - \mathbf{p}_{0,t}.z \end{bmatrix} \quad (1)$$

[4] `http://research.microsoft.com/en-us/projects/vrkinect/`

[5] Kinect is a sensing device developed by Microsoft. Kinect consists of an RGB camera, an IR emitter and an IR depth sensor, a multi-array microphone and a 3-axis accelerometer. (`http://www.microsoft.com/en-us/kinectforwindows/`)

Here $\phi_{r(t)}$ is the elevation angle of the Kinect. As the clustering method, we use Balanced Iterative Reducing and Clustering using Hierarchies (BIRCH) [Zhang et al(1997)Zhang, Ramakrishnan, and Livny] as we did in [Deguchi(2014),Deguchi and Suzuki(2014),Takayama et al(2014)Takayama, Deguchi, Takano, Scuturici, Petit, a

BIRCH is a distance based incremental method to obtain clusters from data points, each of which corresponds to a skeleton in our problem. BIRCH constructs a data structure named clustering feature tree (CF-tree) which is similar to a B+-tree [Bayer and McCreight(1972)]. Each node of a CF-tree stores a clustering feature $\mathbf{CF}(X)$.

$$\mathbf{CF}(X) = \left(N, \sum_{i=1}^{N} \mathbf{x}_i, \sum_{i=1}^{N} \|\mathbf{x}_i\|^2 \right) = (N, \mathbf{LS}, SS) \tag{2}$$

$\mathbf{CF}(X)$ is a compressed representation of a set X which consists of N data points $\{\mathbf{x}_1, ...\mathbf{x}_N\}$ and $\mathbf{CF}(X)$ enables calculation of various extended distance measures [Zhang et al(1997)Zhang, Ramakrishnan, and Livny] including average inter-cluster distance $D(\mathbf{CF}_1, \mathbf{CF}_2)$.

$$\begin{aligned} D(\mathbf{CF}_1, \mathbf{CF}_2) &= \sqrt{\frac{\sum_{i=1}^{N_1} \sum_{j=N_1+1}^{N_1+N_2} \|\mathbf{x}_i - \mathbf{x}_j\|^2}{N_1 N_2}} \\ &= \sqrt{\frac{N_1 SS_2 + N_2 SS_1 - 2\mathbf{LS}_1 \cdot \mathbf{LS}_2}{N_1 N_2}} \end{aligned} \tag{3}$$

$\mathbf{CF}(X_1 \cup X_2)$ can be updated using information of $\mathbf{CF}(X_1)$ and $\mathbf{CF}(X_2)$ only.

$$\begin{aligned} \mathbf{CF}(X_1 \cup X_2) &= \left(N_1 + N_2, \sum_{i=1}^{N_1+N_2} \mathbf{x}_i, \sum_{i=1}^{N_1+N_2} \|\mathbf{x}_i\|^2 \right) \\ &= \left(N_1 + N_2, \sum_{i=1}^{N_1} \mathbf{x}_i + \sum_{j=N_1+1}^{N_1+N_2} \mathbf{x}_j, \sum_{i=1}^{N_1} \|\mathbf{x}_i\|^2 + \sum_{j=N_1+1}^{N_1+N_2} \|\mathbf{x}_j\|^2 \right) \\ &= \mathbf{CF}(X_1) + \mathbf{CF}(X_2) \end{aligned} \tag{4}$$

The CF-tree is updated incrementally when a new data point $\mathbf{x}_{\text{new}}$ is input. At first, the nearest leaf n_{near} from $\mathbf{x}_{\text{new}}$ is searched from the root of the CF-tree. For each internal node n_{inter}, the CF $\mathbf{CF}_{\text{input}}$ of $\mathbf{x}_{\text{new}}$ is added into the CF of n_{inter}.

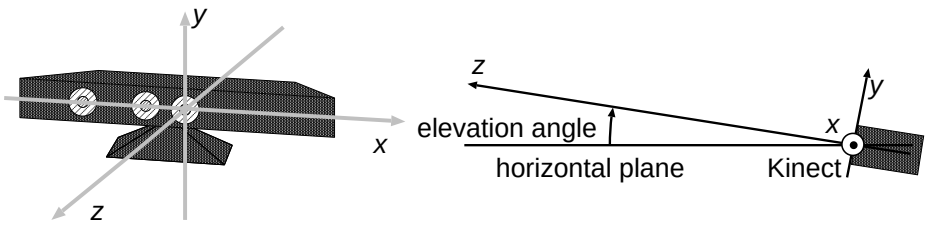


Fig. 2 The original 3D space of Kinect. The origin $(x, y, z) = (0, 0, 0)$ is located at the center of the IR depth sensor on the Kinect. The z axis is not parallel to the ground when the elevation angle of the Kinect is not equal to 0. In the right figure, the x axis is parallel to the depth direction of the figure.

The distance between $\mathbf{CF}_{\text{input}}$ and each child of n_{inter} is calculated and then n_{near} is searched from the subtree of the nearest child recursively. After n_{near} is searched, the diameter of a CF which is the merge of $\mathbf{CF}_{\text{input}}$ and the CF $\mathbf{CF}_{\text{near}}$ of n_{near} is calculated. $\mathbf{CF}_{\text{input}}$ is merged into $\mathbf{CF}_{\text{near}}$ when the diameter is less than the absorption threshold θ_{leaf} of BIRCH, otherwise $\mathbf{CF}_{\text{input}}$ is added into the CF-tree as a new leaf. To simplify BIRCH, we use the inter-cluster distance $D(\mathbf{CF}_1, \mathbf{CF}_2)$ for both internal nodes and the nearest leaf. Thus we employ $D(\mathbf{CF}_{\text{input}}, \mathbf{CF}_{\text{near}})$ instead of the diameter to judge whether $\mathbf{CF}_{\text{input}}$ is merged into $\mathbf{CF}_{\text{near}}$.

Each leaf of the CF-tree consists of similar data points, so the leaf is called a subcluster. BIRCH obtains a set of $K(\mathcal{E}, \theta_{\text{leaf}})$ clusters $C = \{c_1, ..., c_{K(\mathcal{E},\theta_{\text{leaf}})}\}$ by merging leaves if the distance of leaves is less than the absorption threshold θ_{leaf}. Here $\mathcal{E}$ represents a set of the input data. For the fall risk discovery, each data point of $\mathcal{E}$ is represented by features which are relevant to the degrees of the instabilities of a skeleton $\mathcal{S}(t, r(t))$ to obtain a set of clusters in terms of fall risks.

3.3 Subtle Fall Risk Discovery

In [Takayama et al(2014)Takayama, Deguchi, Takano, Scuturici, Petit, and Suzuki], we tackled the problem of discovering fall risks from walking postures of a human who occasionally took spontaneous unbalanced postures. In contrast to [Takayama et al(2014)Takayama, Deguchi, Takano, Sc our target human does not take spontaneous unbalanced postures. For effective early prediction of accidental falls, we should discover fall risks before the target human is in a dangerous situation in which he/she is likely to fall down.

Our target is an elderly person who does not have serious impairments on his/her bodies, but he/she may have fall risks. If we discover their fall risks, it can lead to mitigate the fear of falling in his/her future life. If he/she has fall risks, we define his/her gait as abnormal. In the abnormal gait, both safe and unsafe skeletons appear, where an unsafe skeleton is taken from a posture in which fall risks appear slightly as instabilities. On the other hand, a safe skeleton is taken from a stable posture.

Since unsafe skeletons are similar to safe skeletons, there are two difficult points. First, it is difficult to build a correct set of clusters which is accurate enough to evaluate the clustering results. Since we cannot find unsafe skeletons in abnormal gaits at a glance, it is difficult to label skeletons as safe or unsafe. Even if we look at the skeletons carefully, it is not easy to determine the boundary between the safe and the unsafe skeletons. Second, there is a dilemma between the sensitivity for the fall risks and robustness against the noise. In our problem, it is harder to reflect instabilities of postures than in [Takayama et al(2014)Takayama, Deguchi, Takano, Scuturici, Petit, and Suzuki]. In addition, several joints of the skeleton sometimes become inaccurate due to noise because Kinect was originally developed as a static sensor for video game consoles. Therefore, it is difficult to define features which are both sensitive for instabilities and robust against noise.

Normalized mutual information (NMI) is often used as an evaluation measure of clustering. However, it is not adequate due to the first difficult point because the measure needs the correct set of clusters. The correct set of clusters are significantly different among the labelers, most of whom cannot have confidence in their labels. On the other hand, it is easy to label gait without fall risks as normal and prepare

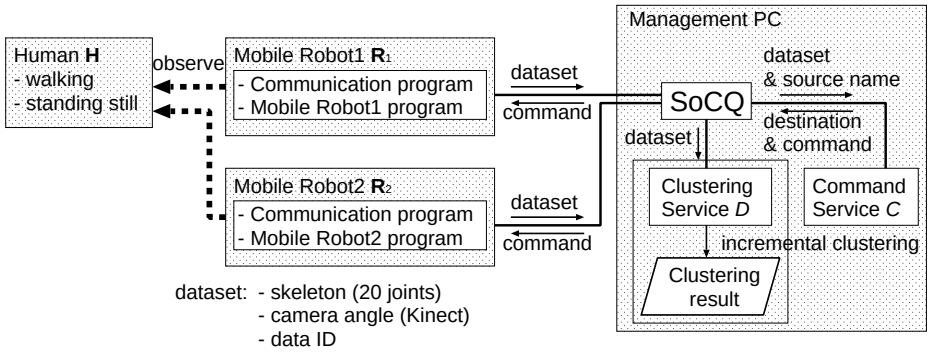


Fig. 3 System overview. In our monitoring system, SoCQ engine manages two mobile robots to monitor the target human.

a normal gait dataset in advance. Therefore, we define a similarity to a normal gait dataset and propose an evaluation method using the similarity in Section 5.3.

4 Proposed System for Human Monitoring

4.1 System Overview

Since unsafe skeletons are similar to safe skeletons, it is important to collect details of walking postures. A mobile robot has several advantages for monitoring over fixed cameras and sensors such as observation from a suitable position. In addition, a multi-robot system observes more information about the target by cooperative monitoring from different directions. However, there are troubles in developing, managing and controlling various components. We adopt a service-oriented data stream management system (DSMS), a Service-oriented Continuous Queries (SoCQ) [Gripay et al(2010)Gripay, Laforest, and Petit] engine, built on a middleware Ubiware [Scuturici et al(2012)Scuturici, Surdu, Gripay, and Petit]. SoCQ engine manages data sources effectively in a uniform manner and simplifies the application development by using declarative queries. Therefore, SoCQ enables us an agile development and an efficient management of the system.

The management PC receives observed data from each robot and issues commands to the robots for realizing a cooperative observation. In our system, vision systems to monitor the target human **H** are only those mounted on the robots. In using a low-cost sensor as the vision system of each robot, observed data are prone to noise while the robot is moving. Therefore, our robots are designed to move to suitable positions for monitoring at short intervals.

Fig. 3 shows a schematic overview of our monitoring system. From the viewpoint of DSMS, each robot provides a data stream about the postures of the human **H** and a service implementing a functionality which reflects a command (move left, move right or stay). These services are seen by SoCQ engine and can be called by the queries executed on the management PC. A clustering service (SoCQ compatible) based on BIRCH [Zhang et al(1997)Zhang, Ramakrishnan, and Livny] is also available on the management PC. SoCQ engine executes continuous queries integrating all these services and data streams.

4.2 Service-Oriented DSMS on the Management PC

The SoCQ engine is a service-oriented DSMS capable of managing dynamic objects, which are data/relations, streams and services, in a dynamic environment [Gripay et al(2010)Gripay, Laforest, and Peti
Each dynamic object is seen as a data source, may it come from a relation or a stream, or from a method/service execution. In our solution, most interactions with robots are implemented as continuous queries, which are run possibly forever and return data as soon as some conditions are satisfied. In other words, we see the robots and the surrounding space as a database-like environment. Each robot produces a skeleton stream and responds to commands issued by the SoCQ engine.

Fig. 4 shows the SoCQ description of the environment which consists of four parts: Receive data, Use BIRCH, Generate command and Send command. Each part includes a relation and a data stream with their associated continuous queries each of which is similar to tables in classical relational databases. The relation *skeletons* describes all the skeletons sent by each robot and the corresponding stream is *skeletonSupervision*. The relation $addUsingBIRCH$ enables clustering skeletons incrementally according to a service based on BIRCH [Zhang et al(1997)Zhang, Ramakrishnan, and Livny]. The relations $generateCommand$ and $receiveCommand$ represent commands for a mobile robot for realizing a cooperative monitoring. The streams of the SoCQ model is implemented as continuous queries: for instance, $supervision\_BIRCH$ provides the skeleton clustering for all the data sent by the robots, $supervision\_create$ and $supervision\_receive$ produce the commands and invoke the corresponding service, respectively.

For our service-oriented DSMS, each robot is integrated in the system as a data source which generates a data stream consisting of the skeleton $\mathcal{S}(t, r(t))$, the camera angle $\phi_{r(t)}$ and the data index at time t. When a command is sent to the robot, the robot is recognized as a data source that handles a service to respond to the command.

4.3 Mobile Robots

For monitoring a walking human, each mobile robot finds the human $\mathbf{H}$ and moves to a suitable position for monitoring $\mathbf{H}$ at short intervals. For realizing a cooperative monitoring, one of the two robots moves relatively long distance to observe skeletons of $\mathbf{H}$ from a different direction.

In our system, each robot acts according to a behavior strategy which is modeled as a 4-state finite automaton in Fig. 5. The three out of the four states, the searching, adjusting and monitoring statuses, are defined to monitor the target human $\mathbf{H}$. Only when the robot receives a repositioning command from SoCQ engine, it transits to the remaining state, the moving status, for realizing a cooperative monitoring. Each behavior of the robot consists of left and right turns and forward and backward moves.

In the searching status, the robot $\mathbf{R}_k$ alternately performs turning left/right and moving forward to find $\mathbf{H}$ and transits to the adjusting status when $\mathbf{R}_k$ finds $\mathbf{H}$. In the adjusting status, $\mathbf{R}_k$ uses left/right turns and forward/backward moves[6]

[6] A left/right turn precedes a forward/backward move so that the robot does not lose sight of the target human easily.

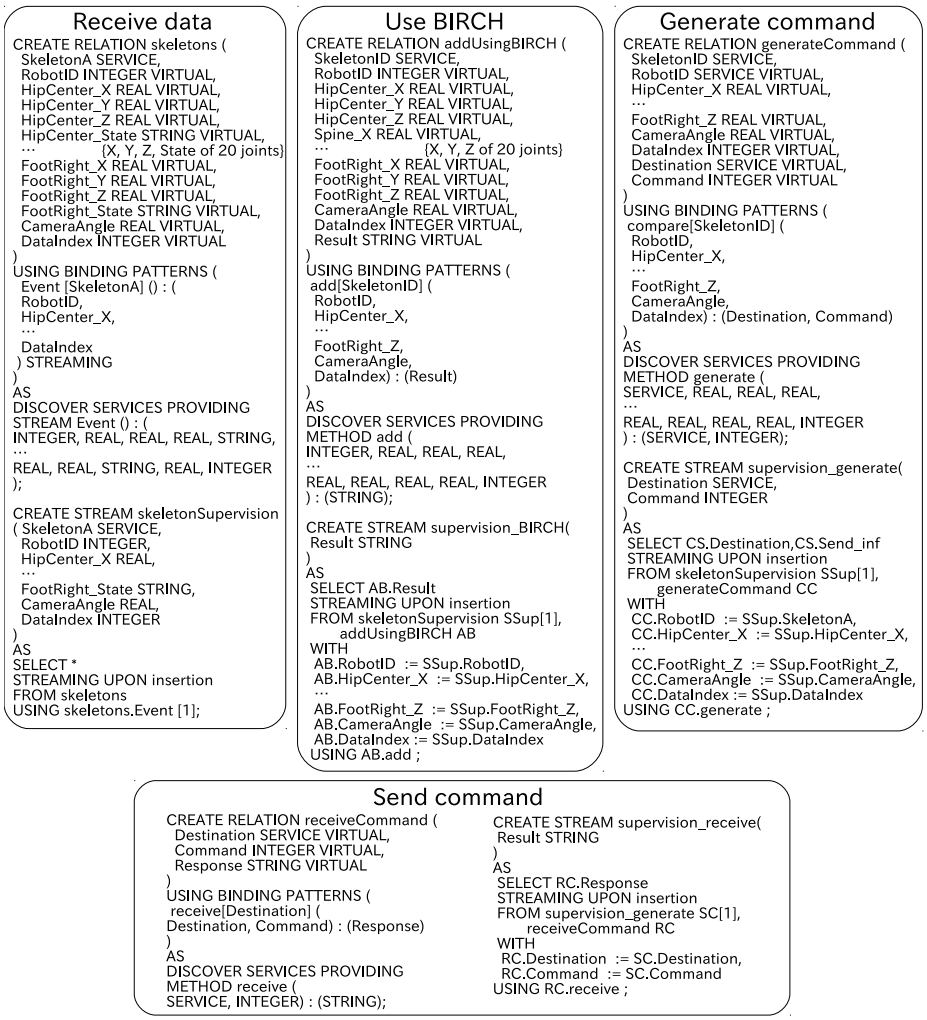


Fig. 4 SoCQ description of our monitoring system. There are four parts of descriptions and each part has a relation and a stream with associated continuous queries.

to move to a suitable position for monitoring. When $\mathbf{R}_k$ is on a suitable position, $\mathbf{R}_k$ transits to the monitoring status and stops for monitoring $\mathbf{H}$. If $\mathbf{R}_k$ receives a repositioning command from SoCQ engine, it transits to the moving status. In the moving status, $\mathbf{R}_k$ turns left/right so that $\mathbf{R}_k$ does not head to $\mathbf{H}$ or the other robot and then $\mathbf{R}_k$ moves forward for a while. After the forward move, $\mathbf{R}_k$ transits to the searching status to find $\mathbf{H}$ again.

Fig. 6 shows a situation that $\mathbf{R}_k$ is in a suitable position for monitoring $\mathbf{H}$. Here d_{ideal} is the ideal distance between $\mathbf{R}_k$ and $\mathbf{H}$. $\mathbf{R}_k$ has the monitored area where $\mathbf{R}_k$ observes skeletons of $\mathbf{H}$. If $\mathbf{H}$ is not in the area, skeletons are prone to noise and $\mathbf{R}_k$ is likely to lose sight of $\mathbf{H}$ immediately. d_{ideal} is determined by using the model in Fig. 7. The vision system, Kinect, is mounted on $\mathbf{R}_k$ of which the height is equal to h_{R}. Here h_{H}, ϕ_{R} and Φ represent the height of $\mathbf{H}$, the elevation angle

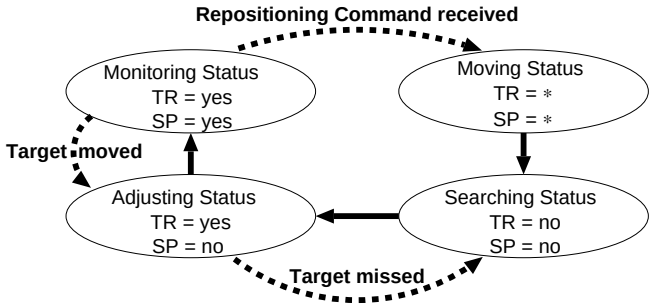


Fig. 5 State diagram of the behavior strategy of the robot. TR (target recognition) represents that whether the robot recognizes the target human and SP (suitable positioning) represents that whether the robot positions itself in a suitable position for monitoring. TR and SP are both binary variables. * is used as a wildcard value (yes or no).

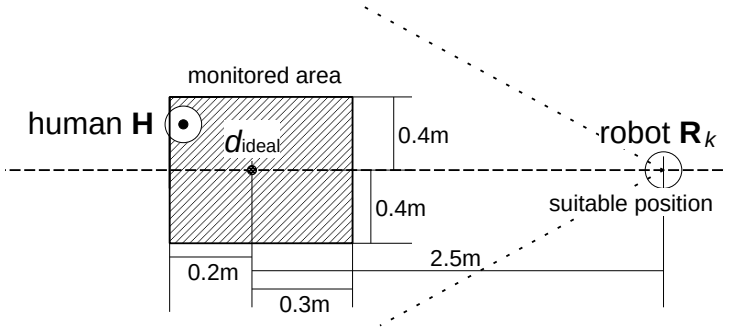


Fig. 6 Suitable position for monitoring. In the monitored area, observed skeleton does not be prone to noise.

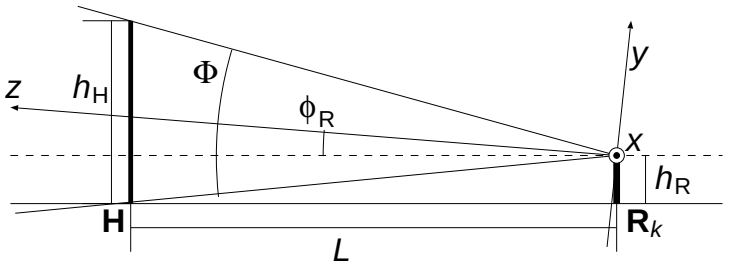


Fig. 7 Geometric model for setting the ideal distance for monitoring. The x axis is parallel to the depth direction of the figure.

of the Kinect and its field of view angle in the vertical direction, respectively. To monitor the whole body of $\mathbf{H}$, $\mathbf{R}_k$ needs to satisfy the following two inequalities.

$$L\tan(\Phi/2+\phi_{\text{R}})+h_{\text{R}}>h_{\text{H}} \tag{5}$$

$$L\tan(\Phi/2-\phi_{\text{R}})>h_{\text{R}} \tag{6}$$

In our monitoring system, we set an ideal distance $d_{\text{ideal}} = 2.5$m because we assume that h_{H} is equal to 1.7m and h_{R}, ϕ_{R} and Φ of the robot are equal to 0.46m,

Table 1 The detail of the robot's behavior. (var.) represents that the value is variable depending on the robot's observation.

Status	turn left/right		move forward	
	time [s]	velocity [deg/s]	time [s]	velocity [m/s]
monitoring	0.0	0	0.0	0.0
adjusting	(var.)	40	(var.)	$0.1+v_a$
searching	3.0	40	1.0	0.2
moving	$1.5+t_m$	40	0.6	0.4

10 degrees and 43 degrees, respectively. For our robot behavior, each turn or move is determined experimentally as shown in Table 1. Here v_a is the velocity so that $\mathbf{R}_k$ reaches a suitable position for monitoring in approximately one second and t_m is the time until $\mathbf{H}$ moves outside the visual field of $\mathbf{R}_k$.

4.4 Reposition Command for the Robot

Our two robots should monitor the target human $\mathbf{H}$ from different directions due to two reasons. The first reason is the advantage of multi-robot monitoring to obtain more information of $\mathbf{H}$. The second reason is the prevention of a bad situation for monitoring in which all skeletons are prone to noise. Since Kinect is suitable for monitoring from the front of $\mathbf{H}$[7], monitoring from the side of $\mathbf{H}$ corresponds to a bad situation and it is possibly subject to a mistrack or a sight loss of joints. In the monitoring from different directions, at least one of the robots avoids such a bad situation.

For realizing such a cooperative monitoring, a single robot $\mathbf{R}_{r(t)}$ moves if $\mathbf{R}_{r(t)}$ receives a repositioning command, which is generated by the generating command service in the management PC, from the SoCQ engine. When the both robots monitor $\mathbf{H}$ unstably, the repositioning command is not generated to prevent situations in which the both robots lose sight of $\mathbf{H}$. The repositioning command indicates left or right so that $\mathbf{R}_{r(t)}$ does not head $\mathbf{H}$ or the other robot.

The command is generated by the Algorithm 1 by comparing skeletons in two circular arrays $array[1]$ and $array[2]$ each of which is held by one of the robots. In the algorithm, there are three judgments: whether the monitorings of the both robots are stable, whether $\mathbf{R}_{r(t)}$ monitors $\mathbf{H}$ from a less suitable direction than the other robot $\mathbf{R}_{3-r(t)}$ and whether $\mathbf{R}_{r(t)}$ should move left or right. The algorithm consists of four parts: data storage (line 1 to line 5), the first judgment (line 6), the second judgment (line 9 to line 15) and the third judgment (line 16).

In data storage, the latest m sets of data are stored in each array. If there are old data, which are obtained more than τ before, in each array, the old data are removed. Here m is the size of each array and τ is a time threshold. For the first judgment, we check whether the both arrays are full. When the frequency of observing data is higher than the frequency of removing data in data storage, the both robots monitor $\mathbf{H}$ stably. For the second judgment, we use a classification method to identify whether $\mathbf{R}_{r(t)}$ monitors $\mathbf{H}$ from a less suitable direction. For the

[7] Because Kinect was initially developed as an input device for a video game console.

Algorithm 1 Generating the command which is sent to the robot $\mathbf{R}_{r(t)}$

Input: skeleton $\mathcal{S}(t, r(t))$, robot ID $r(t)$ $(= 1, 2)$ and camera angle $\phi_{r(t)}$
Output: command c_t

1: $\mathcal{S}'(t) \leftarrow ShiftAndRotate(\mathcal{S}(t, r(t)), \phi_{r(t)})$
2: $t_{\text{current}} \leftarrow GetTime()$
3: $\mathbf{e}(t) \leftarrow ConvertInstance(\mathcal{S}'(t))$
4: $Dir(t) \leftarrow InferDirection(\mathcal{S}'(t))$
5: $StoreData(array[r(t)], \mathbf{e}(t), Dir(t), t_{\text{current}}, \tau, m)$
6: $send\_flag \leftarrow BothCountMax(array[1], array[2], m)$
7: $c_t \leftarrow 0$
8: **if** $send\_flag =$ True **then**
9: $\quad worse\_count \leftarrow 0$
10: $\quad$ **for** $i = 1$ to ξ **do**
11: $\quad\quad$ **if** $JudgeWorse(array[r(t)], array[3 - r(t)], i)$ **then**
12: $\quad\quad\quad worse\_count \leftarrow worse\_count + 1$
13: $\quad\quad$ **end if**
14: $\quad$ **end for**
15: $\quad$ **if** $worse\_count > \nu$ **then**
16: $\quad\quad c_t \leftarrow SelectDirection(array[r(t)], array[3 - r(t)], \xi)$
17: $\quad\quad ClearArrays(array[r(t)], array[3 - r(t)])$
18: $\quad$ **end if**
19: **end if**

third judgment, we use heading directions of the latest m skeletons to determine whether $\mathbf{R}_{r(t)}$ should move left or right.

In lines 1-4 of the Algorithm 1, a feature vector $\mathbf{e}(t)$ and a heading direction $Dir(t)$ are extracted from a set of received data $\mathcal{S}(t, r(t))$, $r(t)$ and $\phi_{r(t)}$. $\mathbf{e}(t)$ and $Dir(t)$ are used in the second and third judgments, respectively, In $ShiftAndRotate$, we apply the shift and rotation transformation, which is the same as Equation (1), to the skeleton $\mathcal{S}(t, r(t))$. Concerning the transformed skeleton $\mathcal{S}'(t)$, the hip center joint is the point of origin and the xz plane is parallel to the ground. $ConvertInstance$ extracts $\mathbf{e}(t)$ from $\mathcal{S}'(t)$ and $InferDirection$ estimates $Dir(t)$ of $\mathcal{S}'(t)$. $StoreData$ stores $\mathbf{e}(t)$, $Dir(t)$ and the current time t_{current} in $array[r(t)]$ and removes old data. In line 6, $BothCountMax$ returns True when the both arrays are full. $JudgeWorse$ returns True when it judges that $\mathbf{R}_{r(t)}$ monitors $\mathbf{H}$ from a less suitable direction than the other robot by using $\mathbf{e}(t)$ included in the i-th latest element in each array. The second judgment is conducted in lines 10-14 by using $\mathbf{e}(t)$s included in the ξ most recent elements. When $worse\_count$ becomes larger than a threshold ν, the algorithm generates the repositioning command. In line 16, $SelectDirection$ sets one of two values meaning "move left" and "move right" to the command c_t[8] by comparing the average values of $Dir(t)$s included in the ξ most recent elements in each array. $ClearArrays$ clears all elements of the two arrays.

The heading direction $Dir(t)$ is estimated from the coordinates of the two shoulders $\mathbf{p}_{4,t}$ and $\mathbf{p}_{8,t}$ as in the following equation.

$$Dir(t) = \arctan \frac{\mathbf{p}_{8,t}.z' - \mathbf{p}_{4,t}.z'}{\mathbf{p}_{8,t}.x' - \mathbf{p}_{4,t}.x'} \quad \left(-\frac{\pi}{2} < Dir(t) < \frac{\pi}{2}\right) \tag{7}$$

[8] In line 7, the command $c_t = 0$ which has no effect to the behavior of $\mathbf{R}_{r(t)}$ is necessary for the generating command service. This is because a service which is managed by a SoCQ engine needs to respond to the engine.

In our monitoring system, we experimentally set $m = 15$, $\tau = 10s$, $\xi = 10$ and $\nu = 7$ for generating the repositioning command by using the Algorithm 1.

For the second judgment, we assume that the front of $\mathbf{H}$ is more suitable for monitoring of $\mathbf{H}$ than the side of $\mathbf{H}$ because skeletons taken from the side are more prone to noise. We define a front skeleton and a side skeleton as skeletons which are taken from the front and the side of $\mathbf{H}$, respectively. To identify whether a mobile robot $\mathbf{R}_{r(t)}$ monitors $\mathbf{H}$ from a less suitable direction than the other robot $\mathbf{R}_{3-r(t)}$, we employ a binary classification method, Adaboost [Freund and Schapire(1997)]. Let $T = \{e_0, e_1, ..., e_N\}$ be the training data, which consist of N examples. Each example $e_t = \{\mathbf{x}_t, y_t\}$ consists of an instance $\mathbf{x}_t$ and its class label $y_t (= 0, 1)$. A binary classification method learns a classifier from T and then the classifier is used to predict the class label of a new instance. In the classification, we classify two situations in which: ($y_t = 0$) $\mathbf{R}_{r(t)}$ and $\mathbf{R}_{3-r(t)}$ observe the front and the side skeletons, respectively and ($y_t = 1$) $\mathbf{R}_{r(t)}$ and $\mathbf{R}_{3-r(t)}$ observe the side and the front skeletons, respectively.

In our problem, elements in each instance $\mathbf{x}_t$ represent the difference between two skeletons in terms of suitability based on directions from the center of the whole human body to joints and tracking statuses of joints. Each element of an instance is extracted from same joints of two skeletons. An instance $\mathbf{x}_t$ is represented as follows.

$$\begin{aligned}\mathbf{x}_t = \big(&a(\mathbf{R}_{r(t)}, \mathbf{R}_{3-r(t)}, t, 1), ..., a(\mathbf{R}_{r(t)}, \mathbf{R}_{3-r(t)}, t, 19),\\ &b(\mathbf{R}_{r(t)}, \mathbf{R}_{3-r(t)}, t, 1), ..., b(\mathbf{R}_{r(t)}, \mathbf{R}_{3-r(t)}, t, 19)\big)\end{aligned} \tag{8}$$

Here the elements of $\mathbf{x}_t$ are represented as follows.

$$a(\mathbf{R}_{r(t)}, \mathbf{R}_{3-r(t)}, t, j) = \left|\eta(t_{\text{late}}(t, \mathbf{R}_{r(t)}), j)\right| - \left|\eta(t_{\text{late}}(t, \mathbf{R}_{3-r(t)}), j)\right| \tag{9}$$

$$\eta(t, j) = \arctan \frac{\mathbf{p}_{j,t}.x'}{\mathbf{p}_{j,t}.z'} \qquad \left(-\frac{\pi}{2} < \eta(t, j) < \frac{\pi}{2}\right) \tag{10}$$

$$b(\mathbf{R}_{r(t)}, \mathbf{R}_{3-r(t)}, t, j) = \mathbf{p}_{j,t_{\text{late}}(t,\mathbf{R}_{r(t)})}.s - \mathbf{p}_{j,t_{\text{late}}(t,\mathbf{R}_{3-r(t)})}.s \tag{11}$$

$a(\mathbf{R}_{r(t)}, \mathbf{R}_{3-r(t)}, t, j)$ and $b(\mathbf{R}_{r(t)}, \mathbf{R}_{3-r(t)}, t, j)$ are calculated from two skeletons which are observed by the two robots $\mathbf{R}_{r(t)}$ and $\mathbf{R}_{3-r(t)}$ at the latest times $t_{\text{late}}(t, \mathbf{R}_{r(t)})$ and $t_{\text{late}}(t, \mathbf{R}_{3-r(t)})$ before time t, respectively. $\eta(t, j)$ indicates the angle between the z' axis and the line passing through the hip center joint[9] and the j-th joint in the $x'z'$ plane as shown in Fig. 8. When the line is parallel to the z' axis, $\eta(t, j)$ is equal to 0 degrees. When $\mathbf{R}_{r(t)}$ observes a front skeleton and $\mathbf{R}_{3-r(t)}$ observes a side skeleton, $a(\mathbf{R}_{r(t)}, \mathbf{R}_{3-r(t)}, t, j)$ becomes large. In the same situation, $b(\mathbf{R}_{r(t)}, \mathbf{R}_{3-r(t)}, t, j)$ is likely to exhibit 0 or -1. This is because several joints of the side skeleton observed by $\mathbf{R}_{3-r(t)}$ are inferred, thus $\mathbf{p}_{j,t_{\text{late}}(t,\mathbf{R}_{r(t)})}.s$ and $\mathbf{p}_{j,t_{\text{late}}(t,\mathbf{R}_{3-r(t)})}.s$ are likely to be equal to 1 and 2, respectively. To create each instance for judging whether $\mathbf{R}_{r(t)}$ monitors $\mathbf{H}$ from a less suitable direction, $\mathbf{e}(t)$ in the Algorithm 1 is represented as follows.

$$\mathbf{e}(t) = (\eta(t, 1), ..., \eta(t, 19), \mathbf{p}_{1,t}.s, ..., \mathbf{p}_{19,t}.s) \tag{12}$$

To collect front skeletons and side skeletons for the second judgment, we conducted experiments as shown in Fig. 9. The two robots $\mathbf{R}_1$ and $\mathbf{R}_2$ did not move

[9] We assume that the hip center joint $\mathbf{p}_{0,t}$ is positioned on the center of the whole body.

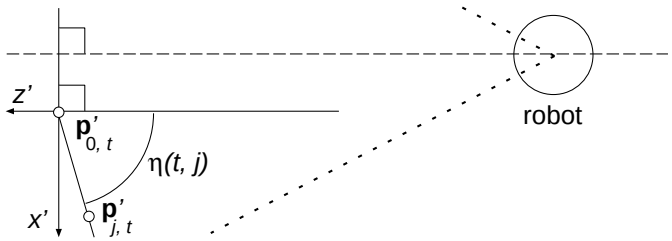


Fig. 8 Definition of the angle $\eta(t, j)$ between the z' axis and the line passing through the 0th joint $\mathbf{p}_{0,t}$ and the j-th joint $\mathbf{p}_{j,t}$ in the $x'z'$ plane. When the robot observes the side skeleton, the line is parallel to the z' axis, thus $\eta(t, j)$ is likely to be equal to 0.

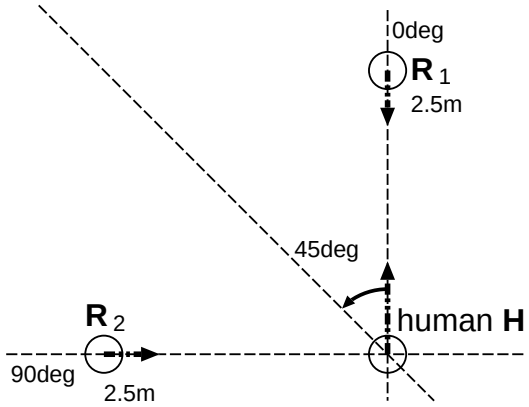


Fig. 9 Relative positions of the human and the robots in the experiment to collect front skeletons and side skeletons. The two robots $\mathbf{R}_1$ and $\mathbf{R}_2$ collect the front skeletons and the side skeletons, respectively.

and observed skeletons of the human **H**. **H** gradually changed the heading direction within a range of 0 to 45 degrees as shown in Fig. 9 while stepping at the position. Therefore $\mathbf{R}_1$ collected front skeletons and $\mathbf{R}_2$ collected side skeletons. In the experiment, we collected 363 front skeletons and 271 side skeletons.

Our classifier was learnt from 786984 examples which were generated from the collected skeletons and their right-and-left-reversed ones. In the training data, 393492 instances with labels $z_t = 0$ were generated by regarding $\mathbf{R}_1$ as $\mathbf{R}_{r(t)}$ and $\mathbf{R}_2$ as $\mathbf{R}_{3-r(t)}$. On the other hand, we regarded $\mathbf{R}_1$ as $\mathbf{R}_{3-r(t)}$ and $\mathbf{R}_2$ as $\mathbf{R}_{r(t)}$ to generate 393492 instances with labels $z_t = 1$. For evaluation of our classifier, we used 10-fold cross validation and then the accuracy of each fold is equal to 1.

When skeletons are observed from a human who walks in the room, the performance of the classifier may be lower due to noise but we believe that the classifier is accurate enough to be employed in our system. In situations in which it is difficult to judge whether $\mathbf{R}_{r(t)}$ monitors **H** from a less suitable direction, the both robots monitor **H** from directions which are similar to each other in terms of suitability. Therefore, the quality of the cooperative monitoring does not heavily depend on which robot moves for realizing the monitoring from different directions.

5 Proposed Method for Skeleton Clustering

5.1 Two kinds of Instability Features

To discover subtle fall risks, we represent skeletons by using features which are related with the instabilities of the human posture. For reflecting the degrees of the instabilities, we define two kinds of features: (δ_{HC}, α_{W}) and (δ_{UL}, α_{U}) as shown in Figs. 10 and 13.

5.1.1 Previous Instability Features

Initially, we considered the robustness against noise on several joints, thus we used a large number of joints or instability of the whole body for calculating features. We initially defined two instability features δ_{HC} and α_{W} which reflect large instabilities of postures. We call δ_{HC} and α_{W} the previous features.

$\delta_{\mathrm{HC}}(t)$ was defined as the horizontal deviation of the center $\mathbf{g}(t)$ of gravity of the whole body from the hip center joint $\mathbf{p}_{0,t}$, which we consider as the center of the whole body.

$$\delta_{\mathrm{HC}}(t) = f(\mathbf{g}(t), \mathbf{p}_{0,t}) \tag{13}$$

$$f(\mathbf{a}, \mathbf{b}) = \sqrt{(\mathbf{a}.x' - \mathbf{b}.x')^2 + (\mathbf{a}.z' - \mathbf{b}.z')^2} \tag{14}$$

$$\mathbf{g}(t) = \frac{\sum_{j=0}^{19} w_j \mathbf{p}_{j,t}}{\sum_{j=0}^{19} w_j} \tag{15}$$

Here w_j, of which the value is shown in the right of Fig. 1, represents the weight ratio of each joint [Takayama et al(2014)Takayama, Deguchi, Takano, Scuturici, Petit, and Suzuki]. To determine w_j, we used the mass ratios of body parts as shown in the left of Fig. 1 which were investigated experimentally in [Clauser et al(1969)Clauser, McConville, and Young]. We assumed that the mass of each part can be regarded as concentrated at the center of the part and the joints of the part are distributed so that each of their

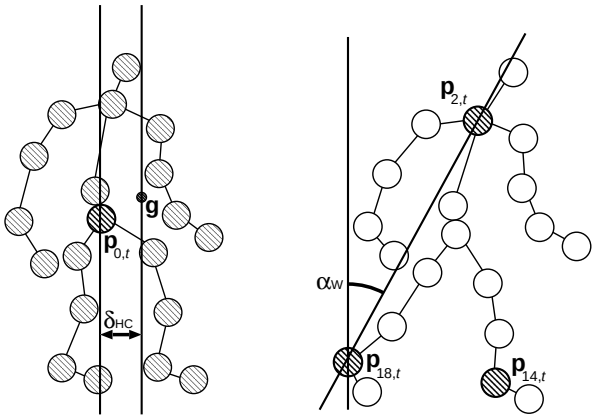


Fig. 10 Previous instability features $\delta_{\mathrm{HC}}(t)$ and $\alpha_{\mathrm{W}}(t)$ to reflect overall instabilities

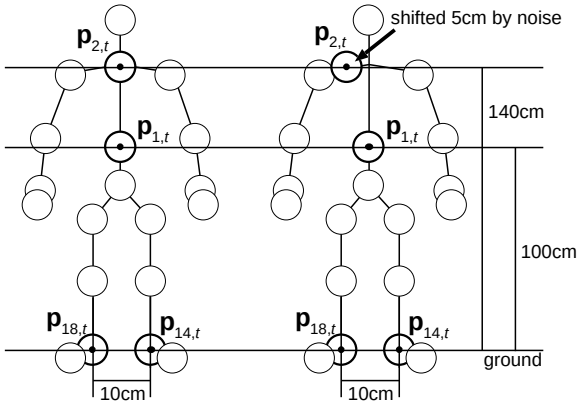


Fig. 11 Geometric model for explanation of the robustness of α_{W}

centers coincides with the center of the part. These assumptions allow us to consider that each joint owns the evenly distributed sum of the mass ratios of the body parts as its concentrated mass ratio.

$\alpha_{\mathrm{W}}(t)$ represents the degree of the inclination of the whole body. We defined $\alpha_{\mathrm{W}}(t)$ as the largest angle between the y' axis and the line passing through the shoulder center joint $\mathbf{p}_{2,t}$ and one of the ankle joints $\mathbf{p}_{14,t}$ and $\mathbf{p}_{18,t}$.

$$\alpha_{\mathrm{W}}(t) = \max_{j=14,18} \left[\arctan \frac{f(\mathbf{p}_{j,t}, \mathbf{p}_{2,t})}{|\mathbf{p}_{j,t}.y' - \mathbf{p}_{2,t}.y'|} \right] \quad \left(0 < \alpha_{\mathrm{W}}(t) < \frac{\pi}{2} \right) \tag{16}$$

δ_{HC} and α_{W} tend to be robust against noise. In calculating $\delta_{\mathrm{HC}}(t)$, we use $\mathbf{g}(t)$ which is the center of gravity of all joints. If a joint $\mathbf{p}_{n,t}$ is shifted to $\mathbf{p}'_{n,t} = \mathbf{p}_{n,t} + \mathbf{n}$ due to noise and the other joints are not shifted, the difference $\mathbf{d}$ from the accurate center of gravity is calculated as follows.

$$\mathbf{d} = \frac{\sum_{j=0}^{19} w_j \mathbf{p}'_{j,t}}{\sum_{j=0}^{19} w_j} - \frac{\sum_{j=0}^{19} w_j \mathbf{p}_{j,t}}{\sum_{j=0}^{19} w_j} = \frac{w_n \mathbf{n}}{\sum_{j=0}^{19} w_j} \tag{17}$$

Although the joints on the hands and the wrists are often shifted by relatively large distances due to noise, the shifts are mitigated by $w_7 / \sum_{j=0}^{19} w_j = 0.004$ and $w_6 / \sum_{j=0}^{19} w_j = 0.013$, respectively. Therefore $\mathbf{g}(t)$ is robust against $\mathbf{n}$ which is caused by noise. To explain the robustness of α_{W}, we compare α_{W} and the angle between the y' axis and the line passing through the spine joint $\mathbf{p}_{1,t}$ and the shoulder center joint $\mathbf{p}_{2,t}$ using the model shown in Fig. 11. Assume that the left and right skeletons in Fig. 11 are equivalent, but the shoulder center joint $\mathbf{p}_{2,t}$ in the right skeleton is shifted 5cm on the horizontal plane due to noise. Without noise, α_{W} and the compared angle are equal to 2.05 and 0.00, respectively. When $\mathbf{p}_{2,t}$ is shifted by noise, α_{W} and the compared angle exhibit a value between 0.00 and 4.09, and 7.13, respectively. Therefore, noise does not have relatively large influence on $\delta_{\mathrm{HC}}(t)$ and $\alpha_{\mathrm{W}}(t)$.

The previous features are effective in reflecting fall risks when the posture of the target human becomes largely unbalanced, but they are not suitable to represent subtle fall risks. $\delta_{\mathrm{HC}}(t)$ does not exhibit a large value because the human moves

his/her arms and legs unconsciously to keep balance. In using $\alpha_{\mathrm{W}}(t)$, values representing an unsafe skeleton and a safe skeleton are similar to each other because the small instability cannot be discriminated from the effects of the strides. Therefore the previous features have a low ability to reflect small instabilities and there is no difference between the normal and the abnormal datasets in the previous feature space as shown in the left of Fig. 12.

5.1.2 Proposed Instability Features

To discover subtle fall risks, we define δ_{UL} and α_{U} as shown in Fig. 13 for reflecting small instabilities of postures. We call δ_{UL} and α_{U} the proposed features. These features are more sensitive to instabilities than the previous features.

Since the lower body sustains the upper body, we consider that the horizontal deviation between their centers of gravity signifies instabilities. $\delta_{\mathrm{UL}}(t)$ is the deviation of the centers of gravity of the upper body $\mathbf{g}_{\mathrm{up}}(t)$ (joints 1-11) and the lower body $\mathbf{g}_{\mathrm{low}}(t)$ (joints 12-19).

$$\delta_{\mathrm{UL}}(t) = f\left(\mathbf{g}_{\mathrm{up}}(t), \mathbf{g}_{\mathrm{low}}(t)\right) \tag{18}$$

$$\mathbf{g}_{\mathrm{up}}(t) = \frac{\sum_{j=1}^{11} w_j \mathbf{p}_{j,t}}{\sum_{j=1}^{11} w_j}, \quad \mathbf{g}_{\mathrm{low}}(t) = \frac{\sum_{j=12}^{19} w_j \mathbf{p}_{j,t}}{\sum_{j=12}^{19} w_j} \tag{19}$$

In the analysis by Tinetti et al. [Tinetti et al(1986)Tinetti, Williams, and Mayewski], the 24 out of the 25 elderly who were recurrent fallers were unable or unwilling to extend their backs. When the back is curved, the center of gravity of the whole body is largely located forward and the walking becomes unstable. In this case, the probability that the human falls down becomes relatively high. For reflecting the curvature of the back, we define $\alpha_{\mathrm{U}}(t)$ as the angle between the upper body

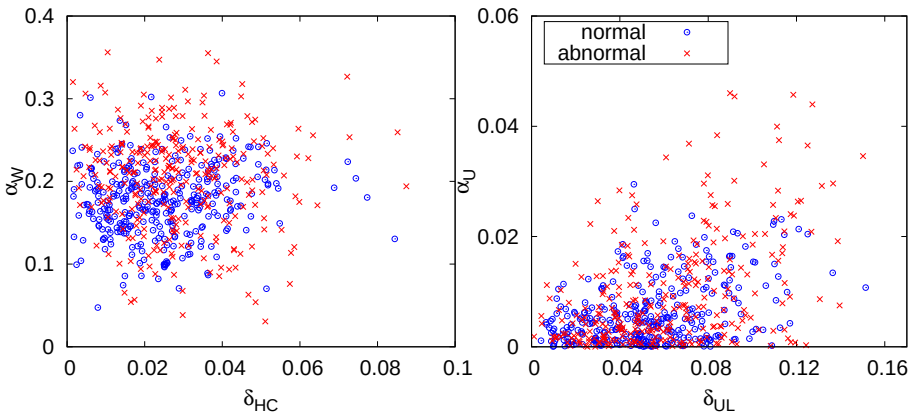


Fig. 12 Distributions of the skeletons in the normal and the abnormal gait datasets using two pairs of features. The left figure uses the previous instability features δ_{HC} and α_{W}. The right figure uses the proposed instability features δ_{UL} and α_{U}. The both figures use same datasets. δ_{UL} and α_{U} are effective to reflect fall risks. In these datasets, α_{U} is especially effective.

and the y' axis.

$$\alpha_{\mathrm{U}}(t) = \arctan \frac{f\left(\mathbf{p}_{1,t}, \mathbf{p}_{2,t}\right)}{|\mathbf{p}_{1,t}.y' - \mathbf{p}_{2,t}.y'|} \quad \left(0 < \alpha_{\mathrm{U}}(t) < \frac{\pi}{2}\right) \tag{20}$$

In the proposed feature space, it is easier to distinguish safe and unsafe skeletons than in the previous feature space. In the right of Fig. 12, there is a larger difference between a normal and an abnormal gait datasets than in the left. However, skeletons are widely distributed in the space and there is no obvious boundary between the safe and the unsafe skeletons. Even if most of safe and unsafe skeletons are separated in clusters, there are too many clusters to inspect the fall risks. To circumvent this problem, we propose a new feature transformation in Section 5.2.

5.1.3 Normalization of Instability Features

To use the features for clustering, it is important to normalize them because δ_{HC} and δ_{UL} are deviations whereas α_{W} and α_{U} are angles. To normalize features online, we use a normal gait dataset $\mathcal{D}_{\mathrm{post}}$ which is prepared in advance to calculate several statistics such as the mean and the variance. When we inspected the distributions of X_{post} which is a set of the feature values extracted from $\mathcal{D}_{\mathrm{post}}$, we found that α_{U} follows an exponential distribution while the remaining three follow normal distributions[10]. α_{U} rarely exhibits a high value because a human takes a balance to prevent falls unconsciously. For each instability feature x, we judge whether x follows a normal distribution or an exponential distribution and convert x to x_{N} which follows a normal distribution [Deguchi(2014)] as shown in the Algorithm 2.

The Algorithm 2 consists of judgment whether X_{post} follows a normal distribution or an exponential distribution (line 1 to line 6), recalculation of the mean and the variance (line 8 to line 12) and conversion of the input feature x (line

[10] We consider a normal distribution and an exponential distribution only because these distributions are not necessary to set parameters such as a value k for degrees of freedom in a chi-squared distribution. It is difficult to determine such parameters because instabilities of a posture is caused by plural fall risks.

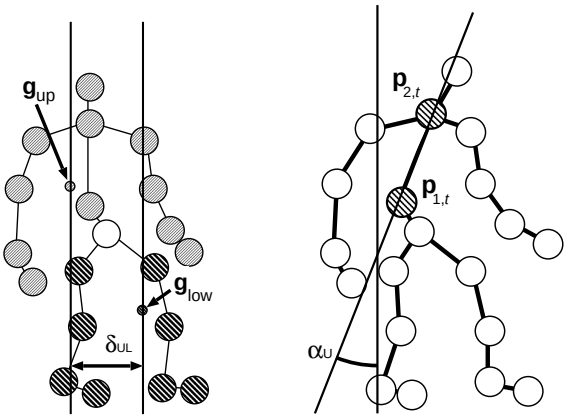


Fig. 13 New instability features $\delta_{\mathrm{UL}}(t)$ and $\alpha_{\mathrm{U}}(t)$ to reflect subtle fall risks

Algorithm 2 Approximate conversion from an exponential distribution to a normal distribution

Input: feature values $X_{\mathrm{post}} = \{x_1, x_2, ..., x_n\}$, input feature x
Output: approximated feature x_{N}
1: $\mu \leftarrow \sum_{i=1}^{n} x_i/n$
2: $\sigma^2 \leftarrow \sum_{i=1}^{n}(x_i - \mu)^2/n$
3: $flag_{\mathrm{Exp}} \leftarrow$ False
4: $sim_{\mathrm{N}} \leftarrow Correlation(X_{\mathrm{post}}, N(\mu, \sigma^2))$
5: $sim_{\mathrm{Exp}} \leftarrow Correlation(X_{\mathrm{post}}, Exp(1/\mu))$
6: **if** $sim_{\mathrm{Exp}} > sim_{\mathrm{N}}$ **then**
7: $\quad flag_{\mathrm{Exp}} \leftarrow$ True
8: $\quad$ **for** $i = 1$ to n **do**
9: $\quad\quad x_i \leftarrow \sqrt{x_i}$
10: $\quad$ **end for**
11: $\quad \mu \leftarrow \sum_{i=1}^{n} x_i/n$
12: $\quad \sigma^2 \leftarrow \sum_{i=1}^{n}(x_i - \mu)^2/n$
13: **end if**
14: **if** $flag_{\mathrm{Exp}} =$ True **then**
15: $\quad x_{\mathrm{N}} \leftarrow \sqrt{x}$
16: **else**
17: $\quad x_{\mathrm{N}} \leftarrow x$
18: **end if**

15). In lines 1 and 2, the mean μ and the variance σ^2 of X_{post} are calculated. In lines 4 and 5, *Correlation* calculates the correlation coefficients sim_{N} and sim_{Exp} between X_{post} and a normal distribution $N(\mu, \sigma^2)$ and between X_{post} and an exponential distribution $Exp(1/\mu)$, respectively. When sim_{N} is smaller than sim_{Exp}, we consider that X_{post} follows $Exp(1/\mu)$. Line 9 converts x_i so that X_{post} follows a normal distribution approximately. In lines 11 and 12, μ and σ^2 are recalculated, which are used for the feature transformation defined in Section 5.2. Lines 1 to 13 are conducted in advance because it needs X_{post} only. If X_{post} follows $Exp(1/\mu)$, line 15 converts the input feature x in the same way as in line 9.

5.2 Focus-on Transformation for an Instability Feature

In using our instability features, skeletons are widely distributed in the space of the proposed features δ_{UL} and α_{U} against our intuition and this fact deteriorates clustering performance. In this section, we define a feature transformation which enables a clustering method to keep the number of clusters small while yielding several clusters which mostly consist of unsafe skeletons. To realize such a transformation, we focus the intensity of attention to skeletons. When we manually separate safe and unsafe skeletons, we would pay more attention to skeletons which are not categorized obviously. We assume that the intensity of our attention can be approximated as a normal distribution $N(x_a, \sigma_a)$. For estimating $N(x_a, \sigma_a)$, we use only a set of feature values X_{post} extracted from a normal gait dataset $\mathcal{D}_{\mathrm{post}}$, which is prepared in advance, to transform a feature value online. By using $N(x_a, \sigma_a)$, we emphasize distances among ambiguous skeletons and not emphasize distances among obvious skeletons.

We propose the focus-on transformation which transforms an instability feature x to a new feature χ. The focus-on transformation consists of two steps: estimating

the intensity of the attention $N(x_a, \sigma_a)$ and transforming x to χ. In the first step, we estimate $N(x_a, \sigma_a)$ by using the mean μ and the standard deviation σ of X_{post}. x_a is the feature value corresponding to a skeleton to which we pay the most attention and σ_a represents the degree of spread of $N(x_a, \sigma_a)$. Most of the skeletons in $\mathcal{D}_{\text{post}}$ seem obviously safe, so we assume that the probability $P(X_{\text{post}} \geq x_a)$ is negligible. Therefore, x_a is determined as follows.

$$0.01 = \frac{1}{2} \int_{x_a}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(t-\mu)^2}{2\sigma^2}\right] \tag{21}$$

Here the values 0.01 and $1/2$ are chosen to represent a rare case and reflect the one-side long tail distribution of X_{post}, respectively. σ_a cannot be determined uniquely due to the difference of the individuals. In this paper, we use $\sigma_a = 2\sigma$ to cover a wide range because it is impossible to foresee to which skeleton a person pay more attention to distinguish the safe and the unsafe skeletons. The first step, which calculates x_a and σ_a, is conducted before our system starts to monitor the target human because $\mathcal{D}_{\text{post}}$ is prepared in advance.

In the second step, the transformation converts $x_{\text{B}} - x_{\text{A}}$ to $\chi_{\text{B}} - \chi_{\text{A}}$ using $N(x_a, \sigma_a)$ so that $\chi_{\text{B}} - \chi_{\text{A}}$ is equal to the probability $P(x_{\text{A}} \leq x_a \leq x_{\text{B}})$. Intuitively, when the area between x_{A} and x_{B} is close to x_a, $\chi_{\text{B}} - \chi_{\text{A}}$ is emphasized significantly in the transformed space because $P(x_{\text{A}} \leq x_a \leq x_{\text{B}})$ is high. The relation between $\chi_{\text{B}} - \chi_{\text{A}}$ and $x_{\text{B}} - x_{\text{A}}$ is shown in the following equation.

$$\begin{aligned}
\chi_{\text{B}} - \chi_{\text{A}} &= \int_{x_{\text{A}}}^{x_{\text{B}}} \frac{1}{\sqrt{2\pi\sigma_a^2}} \exp\left[-\frac{(t-x_a)^2}{2\sigma_a^2}\right] dt \\
&= \int_{-\infty}^{x_{\text{B}}} \frac{1}{\sqrt{2\pi\sigma_a^2}} \exp\left[-\frac{(t-x_a)^2}{2\sigma_a^2}\right] dt \\
&\quad - \int_{-\infty}^{x_{\text{A}}} \frac{1}{\sqrt{2\pi\sigma_a^2}} \exp\left[-\frac{(t-x_a)^2}{2\sigma_a^2}\right] dt
\end{aligned} \tag{22}$$

For example, we assume that a situation in which x_{B} is equal to $x_{\text{A}} + \sigma_a/5$. If x_{A} is equal to x_a, $\chi_{\text{B}} - \chi_{\text{A}}$ is equal to 0.0793. On the other hand, if x_{A} is equal to $x_a + \sigma_a$, $\chi_{\text{B}} - \chi_{\text{A}}$ is equal to 0.0436. We define the focus-on transformation from x to χ as follows.

$$\chi = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma_a^2}} \exp\left[-\frac{(t-x_a)^2}{2\sigma_a^2}\right] dt \tag{23}$$

5.3 Evaluation Method Using Similarities between the Datasets

5.3.1 Similarity between two datasets

As we explained in Section 3.3, an evaluation measure which needs a correct set of clusters, e.g., NMI, is inappropriate for our problem because it is difficult to label each skeleton accurately. To circumvent this problem, we define a similarity between two datasets and use the similarity in our evaluation method. The similarity focuses on the occurrence frequency of skeletons with respect to the obtained clusters. When a dataset is taken from an abnormal gait, more skeletons belong

to clusters which consist of skeletons with high instabilities. Thus, the frequencies of two normal gait datasets should be similar to each other, but the frequencies of a normal and an abnormal gait datasets should be dissimilar.

To calculate the degree of the similarity $\Delta_{\mathcal{X},\Gamma}(\mathcal{D}_1, \mathcal{D}_2)$ between two datasets $\mathcal{D}_1 = \{\mathbf{d}_1, ..., \mathbf{d}_{M_1}\}$ and $\mathcal{D}_2 = \{\mathbf{d}_{M_1+1}, ..., \mathbf{d}_{M_1+M_2}\}$ with respect to the obtained clusters, we first create a merged dataset $\mathcal{D} = \{\mathbf{d}_1, ..., \mathbf{d}_{M_1+M_2}\}$. A set of clusters $\mathcal{C} = \{c_1, ..., c_{K(\mathcal{D},\mathcal{X},\Gamma)}\}$ is obtained by clustering $\mathcal{D}$ based on BIRCH using a preprocessing method Γ with a set of features $\mathcal{X}$. Here $K(\mathcal{D}, \mathcal{X}, \Gamma)$ is the number of the obtained clusters. We construct a histogram H_j expressing the occurrence frequency of skeletons in a dataset $\mathcal{D}_j$ with respect to $\mathcal{C}$ as follows.

$$H_j(i) = \frac{N_{c_i,\mathcal{D}_j}}{M_j} \tag{24}$$

Here $N_{c_i,\mathcal{D}_j}$ is the number of skeletons from $\mathcal{D}_j$ included in c_i. In H_j, we regard each cluster c_i as a bin $H_j(i)$ of H_j. We define $\Delta_{\mathcal{X},\Gamma}(\mathcal{D}_1, \mathcal{D}_2)$ as the Bhattacharyya coefficient [Bhattacharyya(1946)] between H_1 and H_2, which is widely used for representing the similarity between two histograms.

(25)

When $\mathcal{D}_2$ is a normal gait dataset and $\Delta_{\mathcal{X},\Gamma}(\mathcal{D}_1, \mathcal{D}_2)$ is low, there are several clusters which mostly consist of skeletons from $\mathcal{D}_1$ and these skeletons can be regarded as unsafe. In such a case, $\mathcal{D}_1$ is considered as abnormal and the clusters are useful in inspecting the fall risks.

5.3.2 Evaluation Measure

We evaluate the combination of a preprocessing method Γ and a set of features $\mathcal{X}$ from two aspects: to recognize two normal gaits as belonging to the same category and to discriminate a normal and an abnormal gaits. Thus, we prepare two normal gait datasets $\mathcal{D}_{\text{stable}}$ and $\mathcal{D}_{\text{normal}}$ and an abnormal gait dataset $\mathcal{D}_{\text{abnormal}}$. If the similarity between $\mathcal{D}_{\text{stable}}$ and $\mathcal{D}_{\text{normal}}$ is high and the similarity between $\mathcal{D}_{\text{stable}}$ and $\mathcal{D}_{\text{abnormal}}$ is low, the combination shows high performance.

We propose an evaluation measure $E(\mathcal{X}, \Gamma)$ as follows.

$$E(\mathcal{X}, \Gamma) = \Delta_{\mathcal{X},\Gamma}(\mathcal{D}_{\text{normal}}, \mathcal{D}_{\text{stable}}) - \Delta_{\mathcal{X},\Gamma}(\mathcal{D}_{\text{abnormal}}, \mathcal{D}_{\text{stable}}) \tag{26}$$

$E(\mathcal{X}, \Gamma)$ measures the difference between the similarities $\Delta_{\mathcal{X},\Gamma}(\mathcal{D}_{\text{normal}}, \mathcal{D}_{\text{stable}})$ and $\Delta_{\mathcal{X},\Gamma}(\mathcal{D}_{\text{abnormal}}, \mathcal{D}_{\text{stable}})$. In cases in which $E(\mathcal{X}, \Gamma)$ exhibits negative values or values close to zero, we can confidently regard the clustering method using Γ with $\mathcal{X}$ as ineffective. This is because we cannot identify whether a dataset $\mathcal{D}$ contains a gait with fall risks by using $\Delta_{\mathcal{X},\Gamma}(\mathcal{D}, \mathcal{D}_{\text{stable}})$. In a case in which $E(\mathcal{X}, \Gamma)$ exhibits a high value, the clustering method obtains a set of clusters in which a normal and an abnormal gait datasets have different occurrence frequencies of skeletons. Thus, we can regard the method as effective to discover fall risks.

Table 2 Datasets used in the experiments. Abnormality on the body part(s) refers to the body part(s) with tools for simulating the physical situation of the elderly. #skeletons represents the number of skeletons.

Dataset	#skeletons	Person	Abnormality on the body part(s)
A0	799	**A**	none
A1	658	**A**	visual deficit α
A2	764	**A**	visual deficit β
A3	701	**A**	mobility limitation 1
A4	638	**A**	mobility limitation 1 & weakness 1
A5	680	**A**	mobility limitation 2
A6	609	**A**	mobility limitation 2 & weakness 2
B0	593	**B**	none
B1	833	**B**	visual deficit α
B2	691	**B**	visual deficit β
B3	789	**B**	mobility limitation 1
B4	785	**B**	mobility limitation 1 & weakness 1
B5	746	**B**	mobility limitation 2
B6	913	**B**	mobility limitation 2 & weakness 2
C0	979	**C**	none
C1	1048	**C**	visual deficit α
C2	1071	**C**	visual deficit β
C3	1288	**C**	mobility limitation 1
C4	1296	**C**	mobility limitation 1 & weakness 1
C5	1151	**C**	mobility limitation 2
C6	1144	**C**	mobility limitation 2 & weakness 2
D0	1319	**D**	none
D1	858	**D**	visual deficit α
D2	991	**D**	visual deficit β
D3	517	**D**	mobility limitation 1
D4	811	**D**	mobility limitation 1 & weakness 1
D5	426	**D**	mobility limitation 2
D6	1186	**D**	mobility limitation 2 & weakness 2

6 Experiments

6.1 Conditions

In the experiments for evaluating our methods, we used seven kinds of datasets of four persons **A**, **B**, **C** and **D** as shown in Table 2. Each dataset consists of skeletons in a five-minute walk. To collect datasets, we conducted experiments in the room as shown in Fig. 14.

The robot platform is TurtleBot2 with Kobuki[11] (Kobuki). Our Kobuki had a notebook PC Panasonic CF-SX3. The width, the depth and the height of Kobuki are about 37cm, 37cm and 46cm, respectively, and a Kinect was mounted on the top of the robot. The service-oriented DSMS was run on a notebook PC Panasonic CF-SX2 and the PC and the robots were connected to a wireless LAN router NEC Aterm WG1800HP (we used IEEE802.11a). Specifications of the PCs are shown in Table 3. A person walked in the environment and our system observed skeletons using the two mobile robots which are shown in Fig. 15.

To compare all datasets, we manually removed non-human skeletons and broken skeletons which largely deviate from walking postures of a human from the

[11] `http://kobuki.yujinrobot.com/home-en/about/reference-platforms/turtlebot-2/`

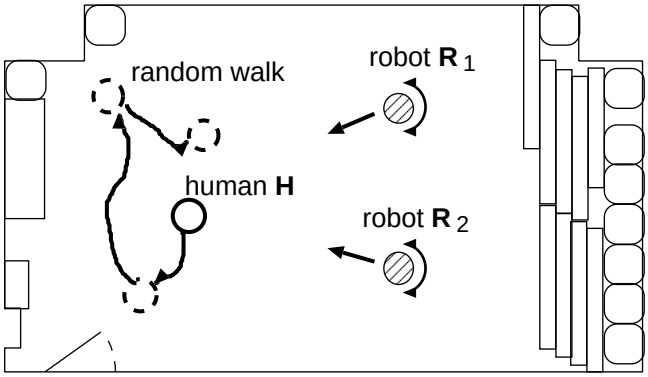


Fig. 14 Environment of the experiments (width × height ≃ 6.5m × 4.5m)

Table 3 Specifications of the PCs

	Mobile robot	Service-oriented DSMS
PC	CF-SX3	CF-SX2
CPU	Core i7-4500U	Core i7-3540M
Operation frequency	1.8GHz	3.0GHz
(using Intel Turbo Boost Technology)	max 3.0GHz	max 3.7GHz
Main memory	4GB RAM	4GB RAM

Fig. 15 TurtleBot2 with Kobuki

datasets. These skeletons were mistakenly taken from the wall or the floor or contain serious errors such as having the twice longer right arm than the left or both feet being 10cm above the ground. We consider that it is easy to remove these non-human and broken skeletons automatically by using a classification method because they largely differ from normal skeletons. **A0**, **B0**, **C0** and **D0** are normal gait datasets which consist of safe skeletons. In other datasets, the target human wore one or more tools on his/her eyes, one leg or both legs to become in a poor

Fig. 16 Tools for simulating the physical situation of the elderly. The left: the goggle simulates a visual deficit. The center: the band limits the movement of the knee joint. The right: the 1kg load simulates the decrease of the physical strength.

physical situation for walking[12]. We call the influence of such tools abnormalities. Since the target human often took unsafe postures due to the abnormalities, each of these datasets consists of safe and unsafe skeletons and is regarded as abnormal. The tools (Sanwa 104-998)[13] for realizing these abnormalities are shown in Fig. 16. They are used to simulate the physical situation of the elderly. The goggle was used with two kinds of filters: clouded or narrowed. The person with "visual deficit α" used the goggle with the narrowed filter and with "visual deficit β", the person walked using the goggle with the both filters. The band which limits the movement of the knee joint and the weight which places a 1kg load on the leg simulate the decrease in physical strength. The person with "mobility limitation 1" and "mobility limitation 2" wore the one or two bands on his/her one knee and both knees, respectively. In addition to the mobility limitations, the person with "weakness 1" and "weakness 2" wore the one or two weights on his/her one ankle and both ankles, respectively.

Fig. 17 shows a snapshot of the experiments for the abnormal gait dataset **A4**. We compared our method against a simplified method using the z-score normalization, which is often used for normalization of features, instead of our focus-on transformation. The z-score normalization assumes that a feature value v follows a normal distribution and converts v into v' to follow the standard normal distribution as shown in Equation (27).

$$v' = \frac{v - \mu_v}{\sigma_v} \tag{27}$$

Here μ_v and σ_v are the mean and the standard deviation of vs which were extracted from a normal gait dataset $\mathcal{D}_{\text{post}}$. For each method we tested three sets of the instability features: the previous two only, the proposed two only and all four.

The value for the absorption threshold θ_{leaf} of BIRCH yields various numbers of clusters; for a specific combination of the dataset $\mathcal{D}$, the instability feature set $\mathcal{X}$ and the preprocessing method $\mathit{\Gamma}$ of clustering. Thus we define a normalized threshold θ'_{leaf} as follows.

$$\theta'_{\text{leaf}} = \frac{\theta_{\text{leaf}}}{\max_{\mathcal{D}} \mathit{\Theta}_{\text{leaf}}(\mathcal{X}, \mathit{\Gamma}|\mathcal{D})} \tag{28}$$

[12] Although it is the best way to prepare datasets taken from healthy elderly people, it is difficult for us to prepare such datasets. This is because we cannot guarantee their safety completely.

[13] The tools are introduced to various domains, e.g., education, social welfare, home builder, airline company and vocational training. From 2011, the tools were used in several news programs. Therefore, we consider that the tools simulate the physical situation of the elderly adequately.

Fig. 17 Snapshot of the experiments (**A4**)

Here $\Theta_{\mathrm{leaf}}(\mathcal{X}, \Gamma|\mathcal{D})$ represents the minimum value of θ_{leaf} yielding a single cluster from $\mathcal{D}$ by using a method using Γ with $\mathcal{X}$. To simplify the experiments, we assume that the clustering method always obtains a single cluster when θ_{leaf} is greater than $\Theta_{\mathrm{leaf}}(\mathcal{X}, \Gamma|\mathcal{D})$. A larger value for θ_{leaf} gathers less similar skeletons into a cluster, thus the value tends to yield smaller number of clusters. Clustering using $\theta'_{\mathrm{leaf}} = 1$ gathers all skeletons in each dataset into a single cluster and using $\theta'_{\mathrm{leaf}} = 0$ does not gather any skeletons.

To evaluate a combination of Γ and $\mathcal{X}$, our method uses the similarity between two normal gait datasets $\mathcal{D}_{\mathrm{stable}}$ and $\mathcal{D}_{\mathrm{normal}}$ and the similarity between a normal gait dataset $\mathcal{D}_{\mathrm{stable}}$ and an abnormal gait dataset $\mathcal{D}_{\mathrm{abnormal}}$. $\mathcal{D}_{\mathrm{stable}}$ and $\mathcal{D}_{\mathrm{normal}}$ were selected or created from the normal gait datasets of the four person **A**, **B**, **C** and **D**. On the other hand, $\mathcal{D}_{\mathrm{abnormal}}$ was selected from the abnormal gait datasets of the persons. We conducted two kinds of experiments: (1) the datasets which were used for $\mathcal{D}_{\mathrm{stable}}$, $\mathcal{D}_{\mathrm{normal}}$ and $\mathcal{D}_{\mathrm{abnormal}}$ were taken from a single person and (2) $\mathcal{D}_{\mathrm{normal}}$ and $\mathcal{D}_{\mathrm{abnormal}}$ were taken from a single person while $\mathcal{D}_{\mathrm{stable}}$ was taken from another person. In the experiments (1), we consider a case that the gait of the target human is monitored periodically (e.g., every month) for a long term and we can prepare a normal gait dataset of the target for $\mathcal{D}_{\mathrm{stable}}$. In the experiments (2), we consider a more practical case that a normal gait dataset of the target is unavailable, but a normal gait dataset of another person is available for $\mathcal{D}_{\mathrm{stable}}$.

In addition to the experiments, we evaluated our clustering method from practical viewpoints: (3.1) how many skeletons our method needs to judge whether a gait is normal or abnormal and (3.2) whether several clusters mostly consist of unsafe postures which are useful in inspecting fall risks. The experiments (3.1) exhibit that how long the target elderly person has to walk for fall risk discovery.

The experiments (3.2) inform us that how useful information is included in the obtained clusters to inspect the fall risks.

6.2 Using Datasets of One Person

In the experiments (1), we employed a half of the normal gait dataset as $\mathcal{D}_{\text{stable}}$ and regarded the remaining half as $\mathcal{D}_{\text{normal}}$ so that we used the datasets taken from a single person for $\mathcal{D}_{\text{stable}}$, $\mathcal{D}_{\text{normal}}$ and $\mathcal{D}_{\text{abnormal}}$. For example, to obtain the evaluation measure $E(\mathcal{X}, \Gamma)$ for **A1**, the normal gait dataset $\mathbf{A0} = \{\mathcal{S}(0, r(0)), ..., \mathcal{S}(T, r(T))\}$ was divided into the first half $\mathbf{A0}_{\text{1st}}$ and the remaining half $\mathbf{A0}_{\text{2nd}}$ and then we calculated $E(\mathcal{X}, \Gamma)$ as follows.

$$\mathbf{A0}_{\text{1st}} = \{\mathcal{S}(0, r(0)), ..., \mathcal{S}(\lfloor T/2 \rfloor, r(\lfloor T/2 \rfloor))\} \tag{29}$$

$$\mathbf{A0}_{\text{2nd}} = \{\mathcal{S}(\lfloor T/2 \rfloor + 1, r(\lfloor T/2 \rfloor + 1)), ..., \mathcal{S}(T, r(T))\} \tag{30}$$

$$E(\mathcal{X}, \Gamma) = \Delta_{\mathcal{X},\Gamma}(\mathbf{A0}_{\text{1st}}, \mathbf{A0}_{\text{2nd}}) - \frac{\Delta_{\mathcal{X},\Gamma}(\mathbf{A1}, \mathbf{A0}_{\text{1st}}) + \Delta_{\mathcal{X},\Gamma}(\mathbf{A1}, \mathbf{A0}_{\text{2nd}})}{2} \tag{31}$$

The results are shown in Figs. 18-21. In each figure, there are six kinds of plots each of which uses a different dataset as $\mathcal{D}_{\text{abnormal}}$. We show the results in the figures for each person because there is the difference among representations of fall risks due to the difference of the individuals. The vertical axis represents the measure $E(\mathcal{X}, \Gamma)$ of our evaluation method for a combination of a preprocessing method Γ and a set of features $\mathcal{X}$. A high value for $E(\mathcal{X}, \Gamma)$ indicates that the method obtains clusters in which $\mathcal{D}_{\text{normal}}$ and $\mathcal{D}_{\text{abnormal}}$ have different occurrence frequencies of skeletons. The horizontal axis represents the normalized threshold θ'_{leaf}[14] of BIRCH. A high value for θ'_{leaf} indicates that the method obtains a few clusters which are useful in inspecting the fall risks in the skeletons. When both $E(\mathcal{X}, \Gamma)$ and θ'_{leaf} exhibit high values, the method has a high ability to identify abnormal gaits and obtain useful information in inspecting the fall risks. In each right plot in Figs. 18-21, the target person had larger abnormalities on his/her walking postures than in the left plot in the same row. Therefore the right three plots in each figure tend to show higher performance than the left three plots.

In the datasets taken from the persons **A** and **D** in which fall risks appear as instabilities of postures, our method using the focus-on transformation with the proposed features outperforms the other methods. For example in **A4**&**A0**, $E(\mathcal{X}, \Gamma)$ of our method exhibits 0.021 when θ'_{leaf} is equal to 0.55. In contrast to our method, the other methods exhibit no more than 0.002 for $E(\mathcal{X}, \Gamma)$ in $\theta'_{\text{leaf}} \geq 0.55$. Although the combination of the z-score normalization and the proposed features has a peak at $E(\mathcal{X}, \Gamma) = 0.11$ in $\theta'_{\text{leaf}} = 0.13$, there are 81 clusters. Therefore, the combination of the focus-on transformation with the proposed features is effective to discover fall risks which appear in walking postures as instabilities.

In the datasets **C2**&**C0**, **C4**&**C0** and **C6**&**C0**, fall risks appear as slower walking speed and shorter strides, but not instabilities of postures. Consequently, the previous features are as effective as the proposed features. When the target human walks more slowly, the center of gravity of his/her body does not move

[14] We employ the normalized threshold θ'_{leaf} as the horizontal axis instead of the number of clusters because we confirm the difficulty to select an optimal value for θ'_{leaf}.

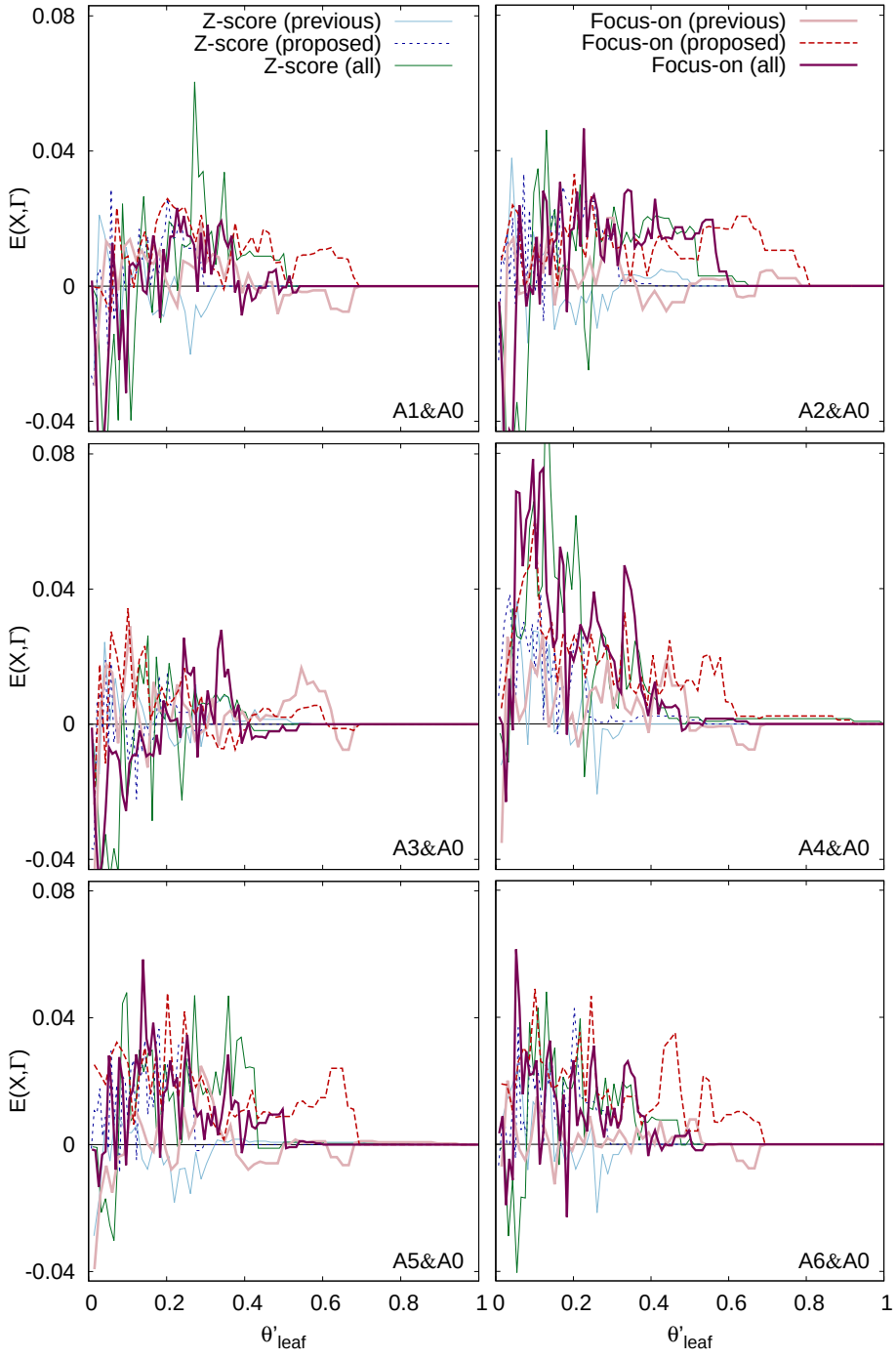


Fig. 18 The evaluation measure $E(\mathcal{X}, \Gamma)$ in terms of the threshold θ'_{leaf}: the results of the experiments (1) (on the dataset of the person **A**). The combination of the focus-on transformation and the proposed features shows high performance except **A3**&**A0** because the fall risks appeared in skeletons as instabilities.

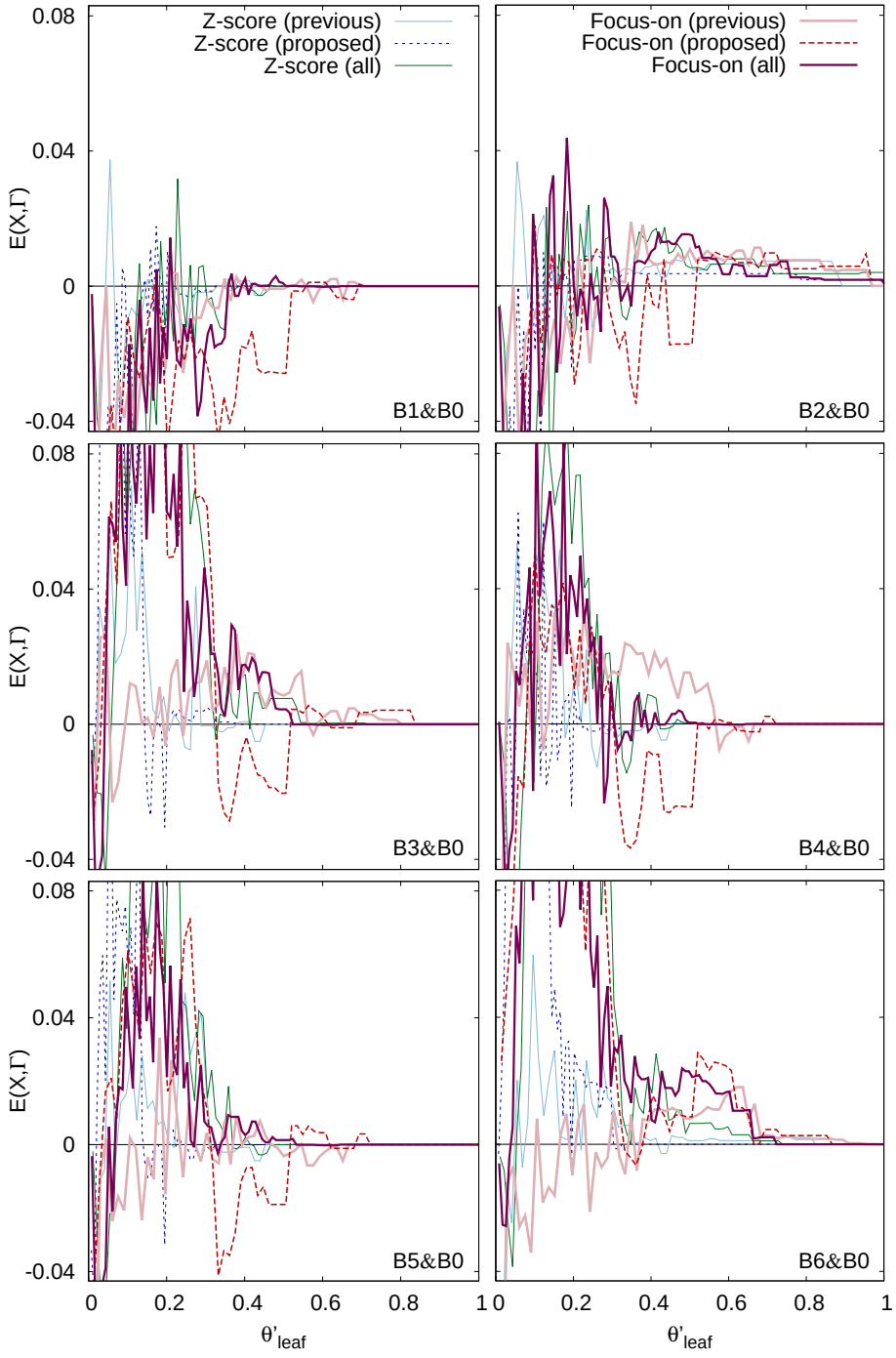


Fig. 19 The evaluation measure $E(\mathcal{X}, \Gamma)$ in terms of the threshold θ'_{leaf}: the results of the experiments (1) (on the dataset of the person **B**). The datasets, especially **B0**, are too noisy for the proposed features, thus the method using the focus-on transformation with the proposed features exhibits negative values for $E(\mathcal{X}, \Gamma)$.

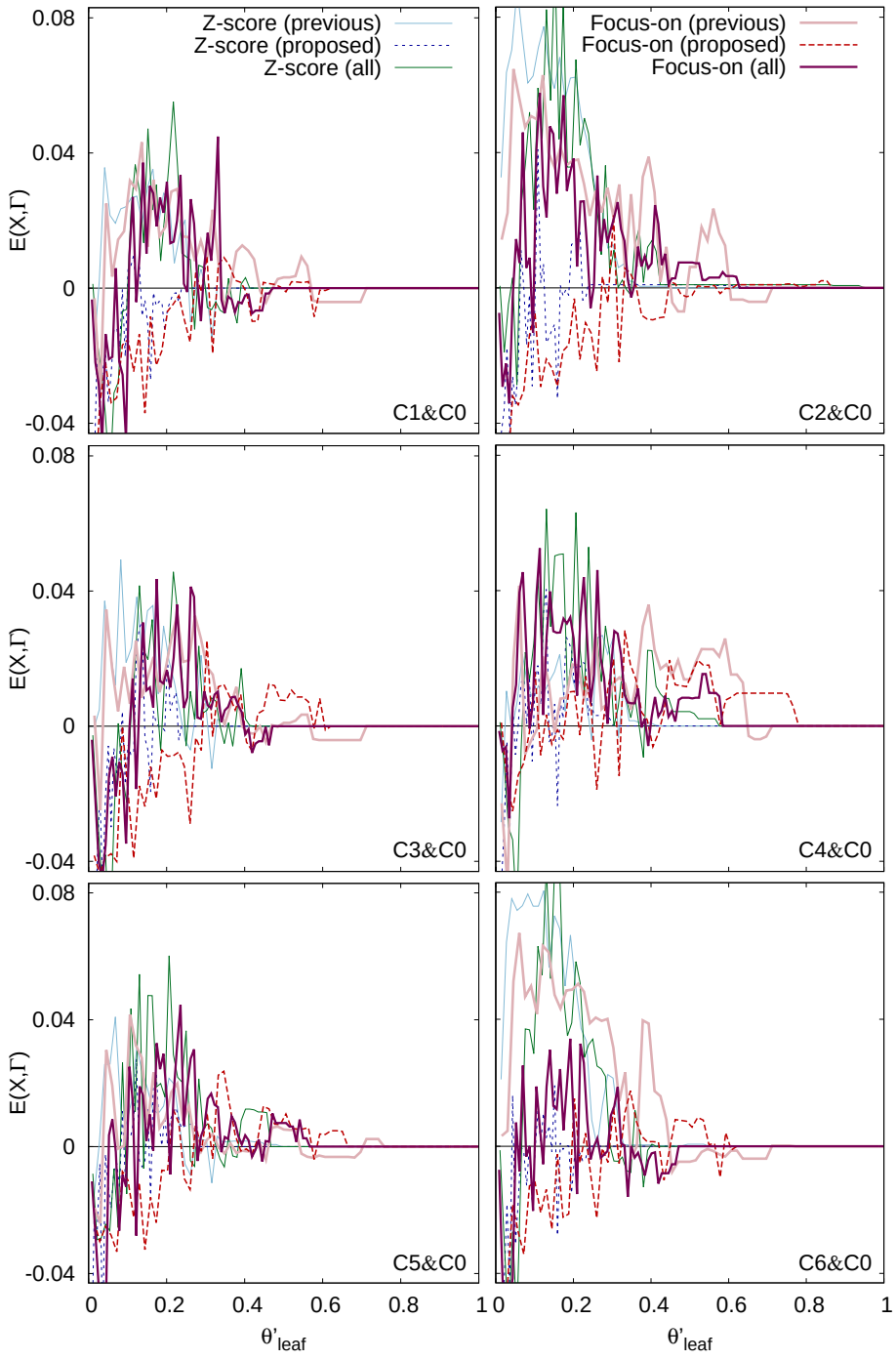


Fig. 20 The evaluation measure $E(\mathcal{X}, \Gamma)$ in terms of the threshold θ'_{leaf}: the results of the experiments (1) (on the dataset of the person **C**). There is an irregular case in which fall risks appear as lower walking speed and smaller strides, thus the previous features are accidentally as effective as the proposed features.

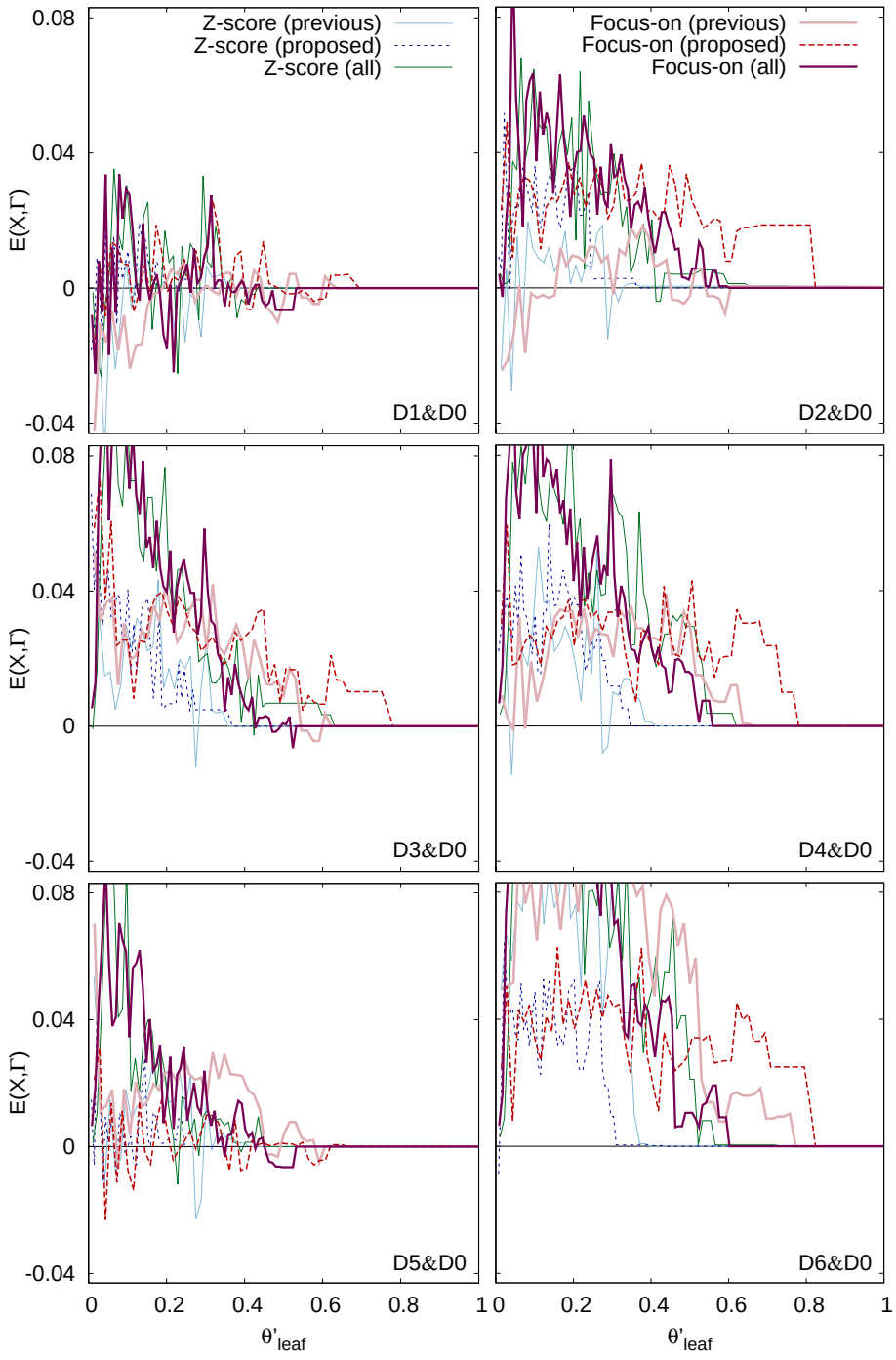


Fig. 21 The evaluation measure $E(\mathcal{X}, \Gamma)$ in terms of the threshold θ'_{leaf}: the results of the experiments (1) (on the dataset of the person **D**). In four out of the six plots, the combination of the focus-on transformation and the proposed features shows high performance while θ'_{leaf} exhibits high values. Thus our method is effective to discriminate the normal and the abnormal gaits for the person such as **D**.

forward relatively, thus δ_{HC} exhibits a lower value. When the target human walks with shorter strides, his/her ankles are close to the center of his/her shoulders on the horizontal plane, thus α_{W} exhibits a lower value. Therefore the previous features δ_{HC} and α_{W} accidentally discriminate the normal and the abnormal gaits. However, δ_{HC} and α_{W} reflecting slower walking speed and shorter strides exhibit similar values extracted from standing postures. Thus the focus-on transformation does not emphasize the differences of δ_{HC} and α_{W} in **C2**&**C0**, **C4**&**C0** and **C6**&**C0** and induces negative evaluation values.

In the datasets taken from the person **B** except **B6**&**B0**, the method using the focus-on transformation with the proposed features exhibits negative values for evaluation measure $E(\mathcal{X}, \Gamma)$. Skeletons in the datasets are too noisy for the proposed features, thus the methods using the proposed features do not show high performance. Unfortunately, the focus-on transformation emphasizes the effect of noise. Since skeletons are noisy, all methods show low performance in several datasets such as **B1**&**B0** and **B5**&**B0**.

6.3 Using Normal Gait Dataset of Another Person for the Subject

It is difficult to prepare the normal gait dataset of the target elderly as $\mathcal{D}_{\mathrm{stable}}$ except several cases in which the gait of the target is monitored periodically for a long term, e.g., every month. Therefore, it is useful to employ the normal gait dataset of another healthy person as $\mathcal{D}_{\mathrm{stable}}$. The experiments (2) employed a normal gait dataset of another person as $\mathcal{D}_{\mathrm{stable}}$. For example, we calculated $\Delta_{\mathcal{X},\Gamma}(\mathbf{A0}, \mathbf{B0})$, $\Delta_{\mathcal{X},\Gamma}(\mathbf{A0}, \mathbf{C0})$ and $\Delta_{\mathcal{X},\Gamma}(\mathbf{A0}, \mathbf{D0})$ and used the average value of these similarities as $\Delta_{\mathcal{X},\Gamma}(\mathbf{A0}, \mathcal{D}_{\mathrm{stable}})$.

The results of the experiments are shown in Figs. 22-25. As the same to the experiments (1), the right three plots in each figure tend to show higher performance than the left plots. In the datasets taken from **A** and **D**, in which fall risks appear in postures as instabilities, our method using the focus-on transformation with the proposed features outperforms the other methods. For example in **A2**&**A0**, our method achieves $E(\mathcal{X}, \Gamma) = 0.030$ in $\theta'_{\mathrm{leaf}} = 0.72$, but the other methods exhibit no more than 0.003 for $E(\mathcal{X}, \Gamma)$ while θ'_{leaf} exhibits more than 0.59.

In Fig. 23 except **B2**&**B0**, all methods show low performance. This is because the noisy skeletons in **B0** are dissimilar to skeletons in each normal gait of the other three persons, thus the positive term $\Delta_{\mathcal{X},\Gamma}(\mathcal{D}_{\mathrm{normal}}, \mathcal{D}_{\mathrm{stable}})$ of $E(\mathcal{X}, \Gamma)$ (refer to Equation (5.3.2)) becomes low. In the datasets taken from **C**, the method using the focus-on transformation with the previous features exhibits negative values for $E(\mathcal{X}, \Gamma)$ around $\theta'_{\mathrm{leaf}} = 0.40$. In contrast to the experiments (1), the previous features are considered to be ineffective to detect fall risks in **C**. As we explained in Section 6.2, the previous features δ_{HC} and α_{W} exhibit lower values reflecting lower walking speed and smaller strides. However, the values are similar to the values which are extracted from skeletons in another normal gait, thus the methods using the previous features do not show high performance as same as in the experiments (1). On the other hand, though our method using the focus-on transformation with the proposed features does not exhibit high evaluation values in the datasets taken from **C**, our method tends to outperform the other methods. For example in **C1**&**C0**, the proposed features inhibit negative evaluation values. The same effect also appears in **C3**&**C0**. For example in **C4**&**C0**, the combination

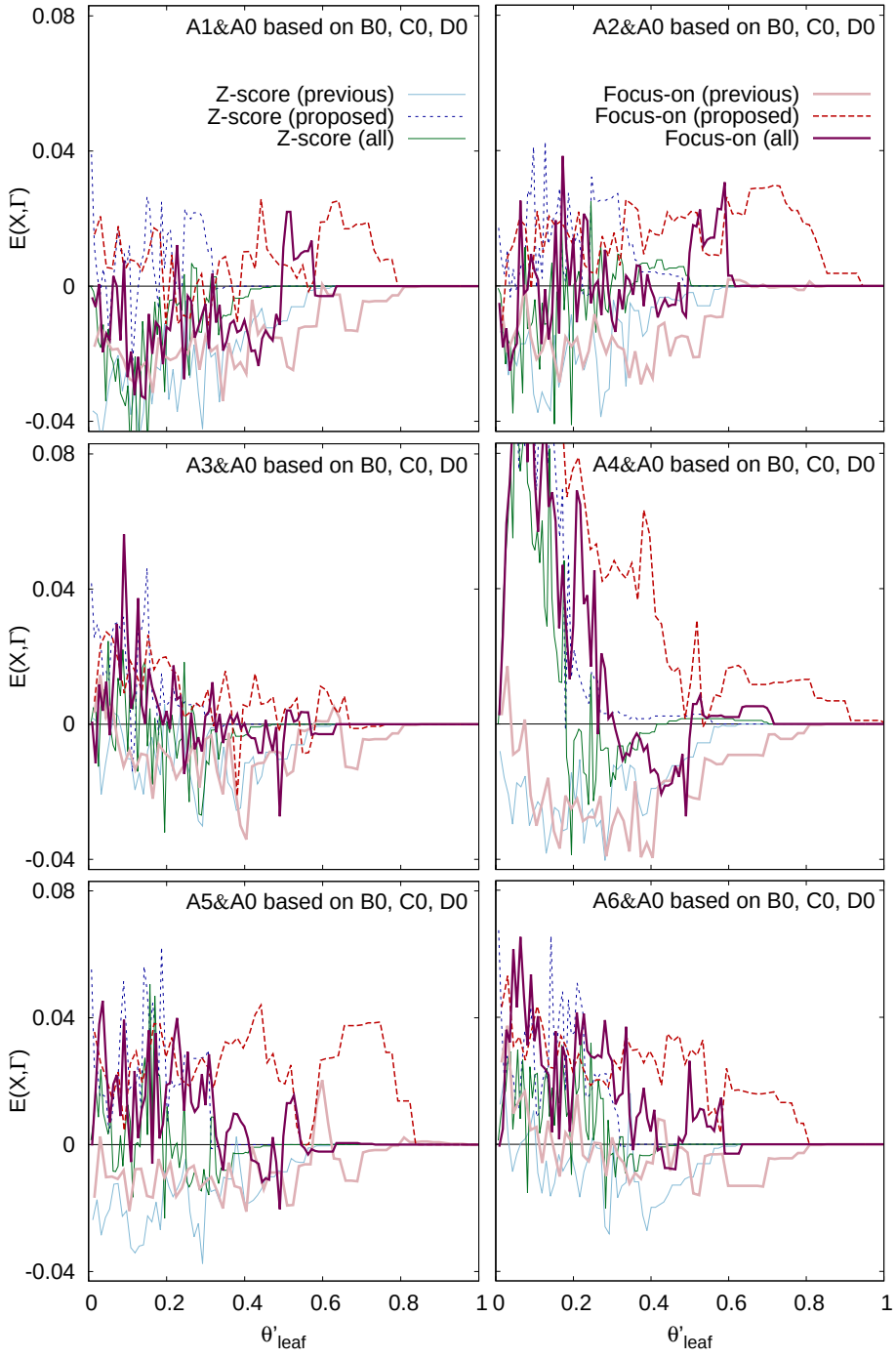


Fig. 22 The evaluation measure $E(\mathcal{X}, \Gamma)$ in terms of the threshold θ'_{leaf}: the results of the experiments (2) (using the abnormal gait datasets of the person **A**). In five out of the six plots, the method using the focus-on transformation with the proposed features exhibits high values for $E(\mathcal{X}, \Gamma)$ while θ'_{leaf} exhibits high values.

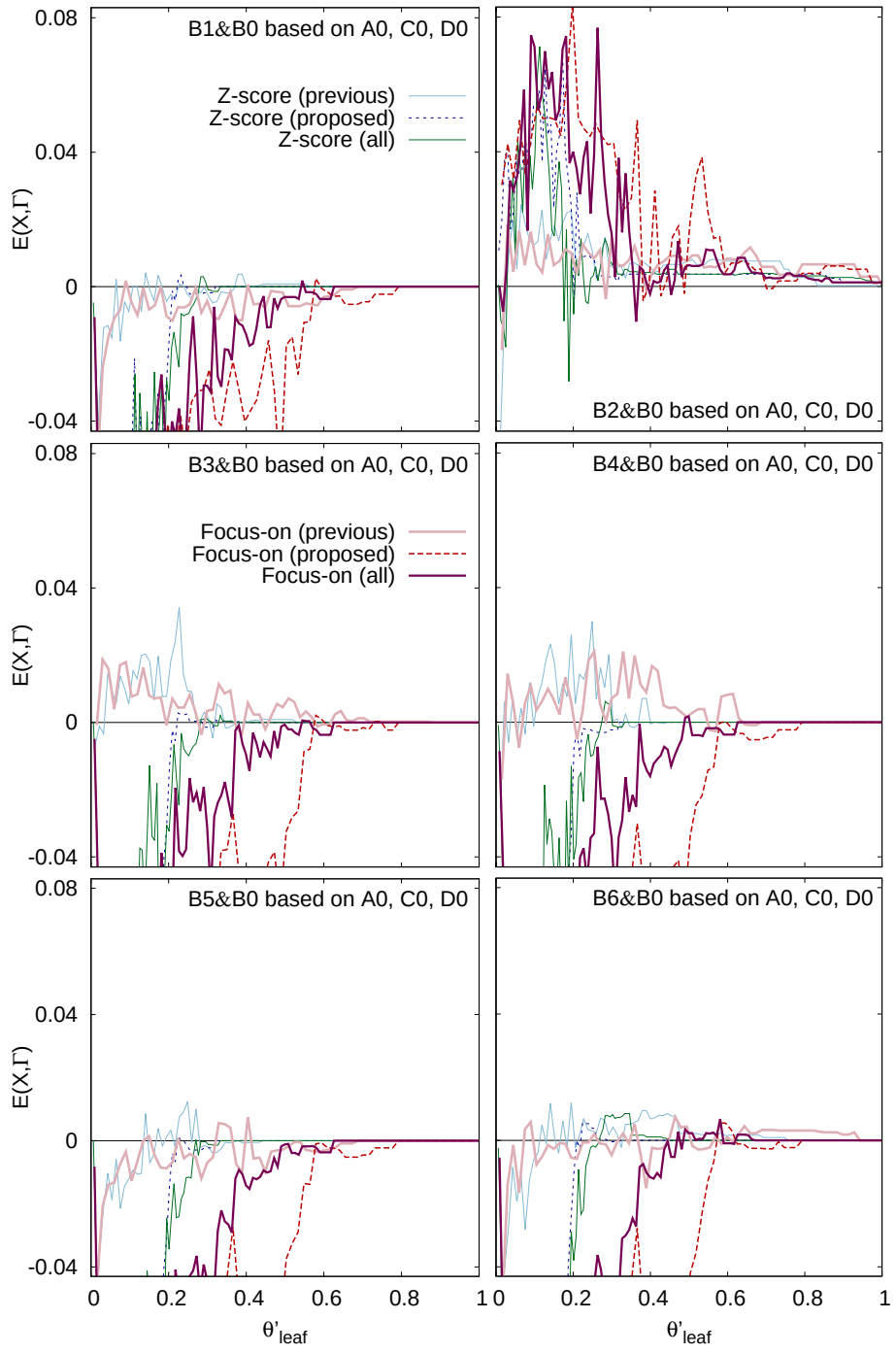


Fig. 23 The evaluation measure $E(\mathcal{X}, \Gamma)$ in terms of the threshold θ'_{leaf}: the results of the experiments (2) (using the abnormal gait datasets of the person **B**). All methods show low performance because **B0** is dissimilar to the normal gait datasets of the other three persons due to noise.

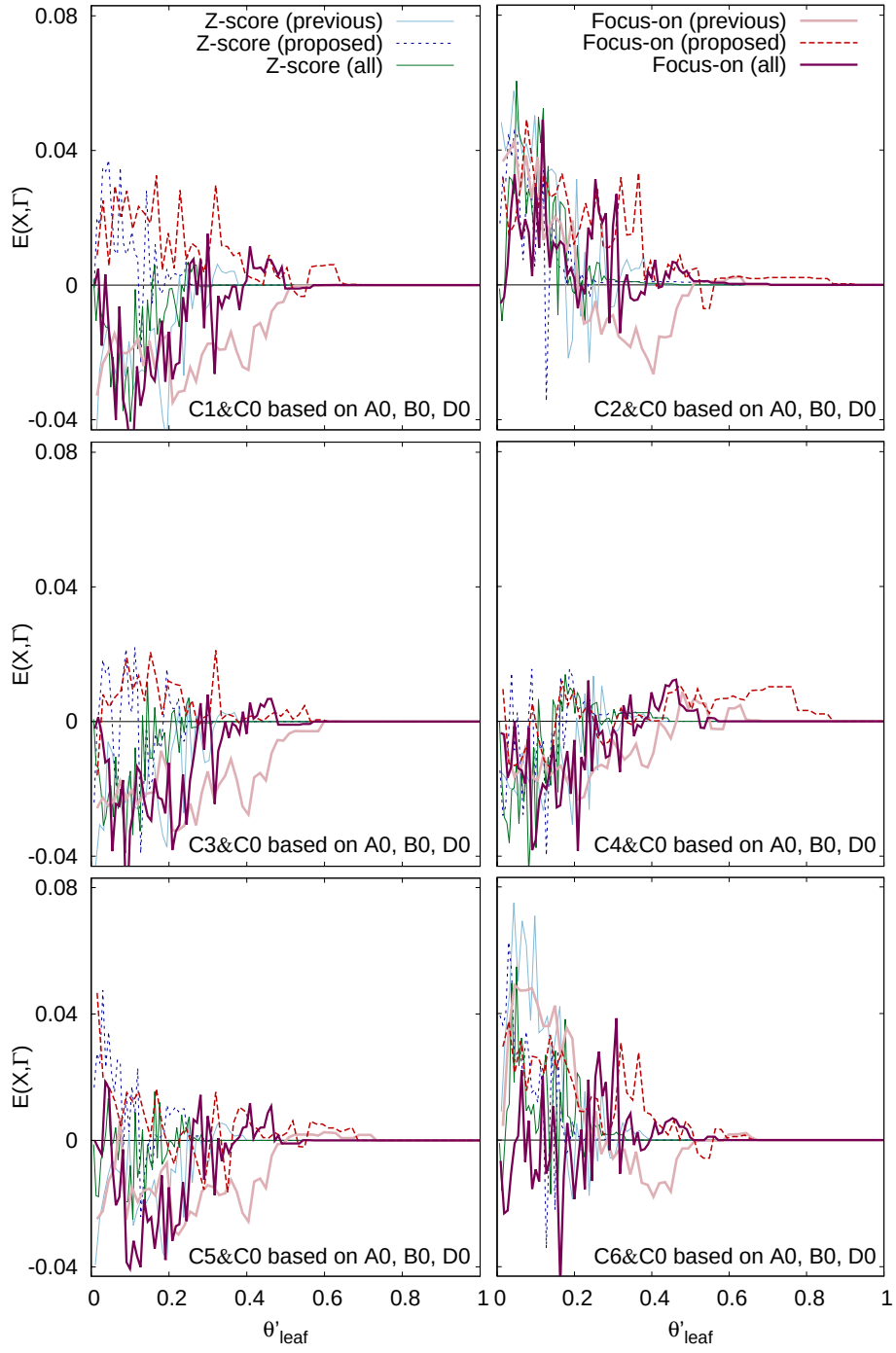


Fig. 24 The evaluation measure $E(\mathcal{X}, \Gamma)$ in terms of the threshold θ'_{leaf}: the results of the experiments (2) (using the abnormal gait datasets of the person **C**). In contrast to the first experiments, the previous features are not effective because fall risks appear in skeletons as slower walking speed and shorter strides, but not instabilities.

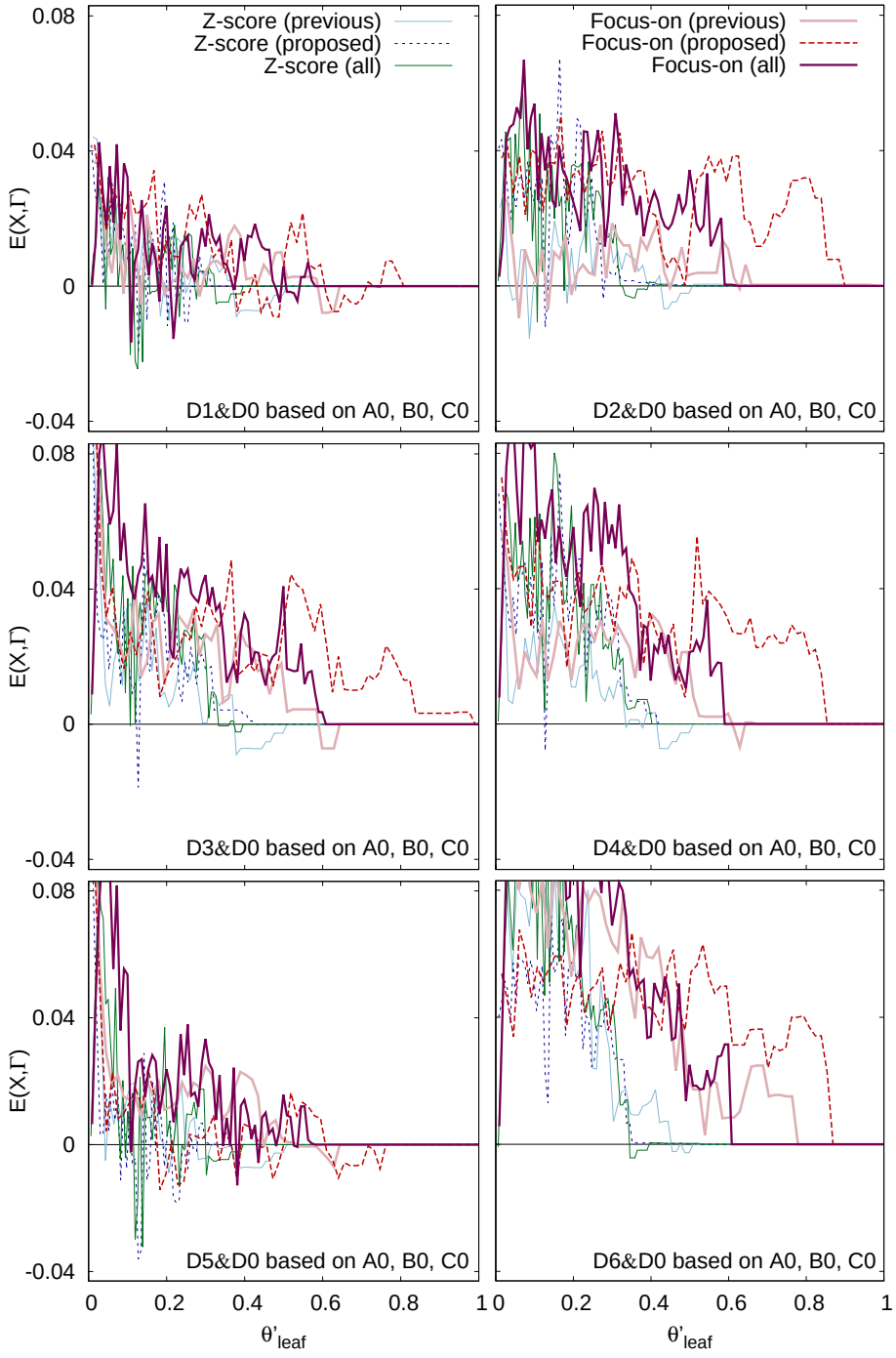


Fig. 25 The evaluation measure $E(\mathcal{X}, \Gamma)$ in terms of the threshold θ'_{leaf}: the results of the experiments (2) (using the abnormal gait datasets of the person **D**). In four out of the six plots, the method using the focus-on transformation with the proposed features shows high performance while θ'_{leaf} exhibits high values.

of the focus-on transformation and the proposed features exhibits $E(\mathcal{X}, \Gamma) = 0.01$ in $\theta'_{\text{leaf}} = 0.77$. Thus the method keep a high ability to discriminate safe and unsafe skeletons when θ'_{leaf} becomes high.

6.4 Evaluation from Practical Viewpoints

In the experiments (3.1) and (3.2), we evaluated each method from more practical viewpoints: how many skeletons are necessary to identify an abnormal gait and how useful information are included in the obtained clusters to inspect fall risks. In the experiments (3.1), we used the datasets taken from the person **D** and normal gait datasets of the other three persons in the same way as the experiments (2). The normalized thresholds θ'_{leaf} which we used are shown in Table 4. The results are shown in Fig. 26.

There are six plots which are different from each other in terms of the combination of a preprocessing method Γ and an instability feature set $\mathcal{X}$. The vertical axis represents the similarity $\Delta_{\mathcal{X},\Gamma}(\mathcal{D}, \mathcal{D}_{\text{stable}})$ between two datasets $\mathcal{D}$ and $\mathcal{D}_{\text{stable}}$ in terms of the occurrence frequency of skeletons with respect to the obtained clusters. Here $\mathcal{D}$ was selected from the dataset taken from **D** and $\Delta_{\mathcal{X},\Gamma}(\mathcal{D}, \mathcal{D}_{\text{stable}})$ was an average value of the similarities each of which used one of **A0**, **B0** and **C0** as $\mathcal{D}_{\text{stable}}$. The horizontal axis represents the number of skeletons from $\mathcal{D}$ and $\Delta_{\mathcal{X},\Gamma}(\mathcal{D}, \mathcal{D}_{\text{stable}})$ was calculated when every ten skeletons of $\mathcal{D}$ were input as data points into the CF tree constructed from the skeletons in $\mathcal{D}_{\text{stable}}$.

After about 500 skeletons were input, the values for $\Delta_{\mathcal{X},\Gamma}(\mathcal{D}, \mathcal{D}_{\text{stable}})$ in Fig. 26 became stable and the difference between **D0** and another dataset became large. Therefore, our approach needs more than 500 skeletons to discover subtle fall risks effectively in these settings. To calculate the skeletons per second in Table 5, we regarded a computation time per skeleton longer than two seconds as the time that the both robots lost sight of the target and excluded them. In the worst case, 4.43 skeletons were observed per second, so it takes about two minutes without missing

Table 4 Normalized thresholds θ'_{leaf} which is used to estimate that how many skeletons are necessary for each method to discriminate the normal and the abnormal gaits effectively (the experiments(3.1)).

	Z-score normalization			Focus-on transformation		
	previous	proposed	all	previous	proposed	all
θ'_{leaf}	0.094	0.086	0.091	0.419	0.549	0.236

Table 5 Skeletons per second in each dataset

$\mathcal{D}$	skeletons/s	$\mathcal{D}$	skeletons/s	$\mathcal{D}$	skeletons/s	$\mathcal{D}$	skeletons/s
A0	5.52	**B0**	5.59	**C0**	6.87	**D0**	7.77
A1	4.74	**B1**	4.71	**C1**	6.94	**D1**	7.88
A2	4.80	**B2**	5.09	**C2**	7.26	**D2**	6.21
A3	4.43	**B3**	6.91	**C3**	7.50	**D3**	6.08
A4	4.58	**B4**	5.65	**C4**	7.20	**D4**	7.55
A5	4.44	**B5**	8.03	**C5**	6.94	**D5**	5.44
A6	4.71	**B6**	7.29	**C6**	6.32	**D6**	8.54

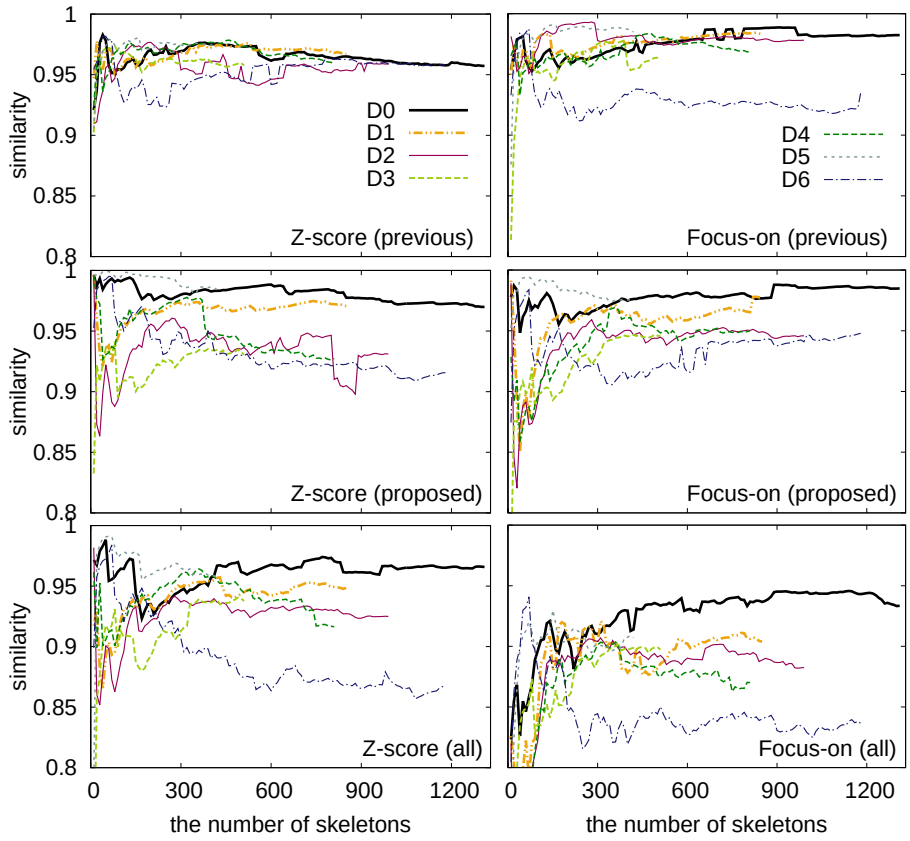


Fig. 26 Similarity in Section 5.3 in terms of the number of skeletons taken from the person **D**. Before about 500 skeletons are input, the similarity is highly unstable.

time to observe 500 skeletons. We believe that two minutes are short enough not to impose heavy burden on the target elderly.

In the results using the proposed features, the larger the number of skeletons becomes, the lower the similarity $\Delta_{\mathcal{X},\Gamma}(\mathbf{D5}, \mathcal{D}_{\text{stable}})$ becomes. If our monitoring system observed more skeletons of **D5**, it may be possible to discover fall risks more effectively. In the results using the focus-on transformation with the proposed features, the similarity $\Delta_{\mathcal{X},\Gamma}(\mathbf{D1}, \mathcal{D}_{\text{stable}})$ becomes high when the number of skeletons is about 800. Such a case was considered as irregular, in which the person **D** became familiar with the abnormality on his eyes and the method succeeded in reflecting the change.

In the experiments (3.2), we show several postures in the obtained clusters from a pair of the datasets **A0** and **D1** (or **D6**) using the focus-on transformation with the proposed features. In this clustering, the numbers of input skeletons in **D1** and **D6** are limited to 500, respectively and we used the value for the normalized absorption thresholds θ'_{leaf} of BIRCH shown in Table 4. Several postures in the obtained clusters are shown in Figs. 27 and 28. All postures shown in Figs. 27 and

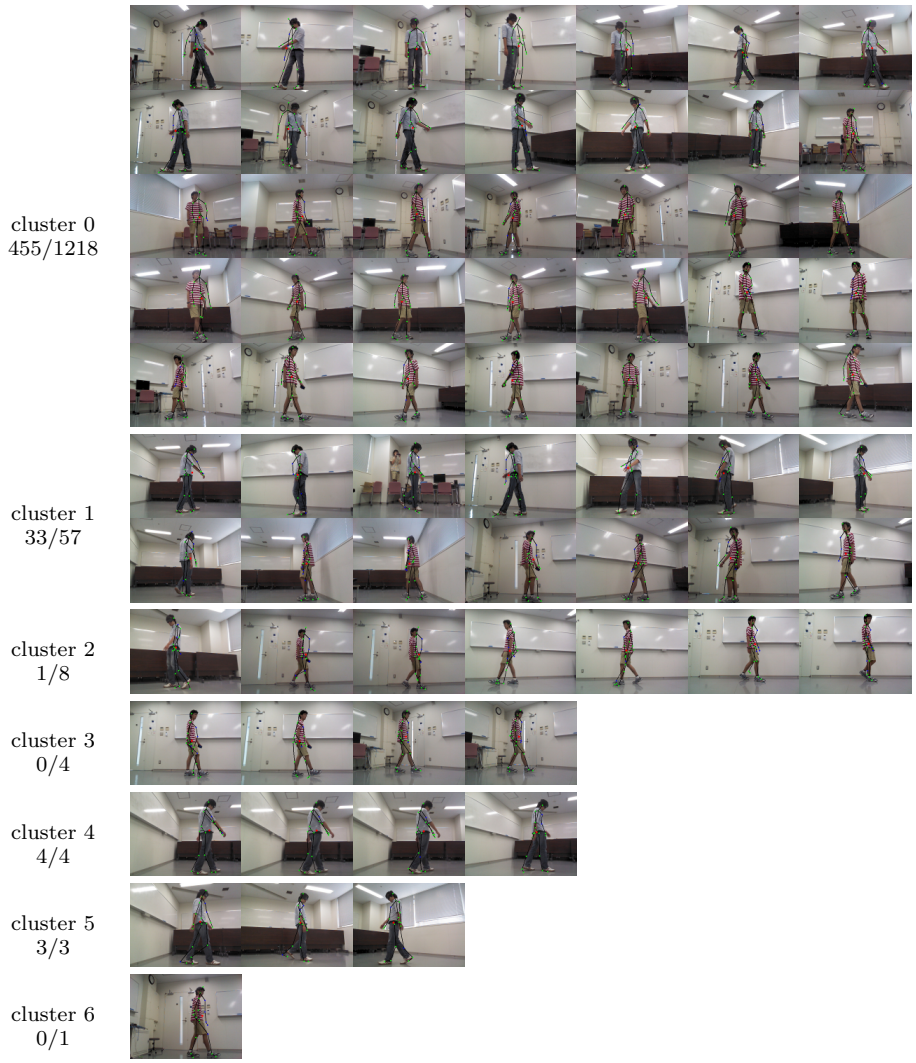


Fig. 27 Examples of the obtained clusters from **D1** with **A0**: $n1/n2$ represents that there are $n2$ skeletons including $n1$ skeletons from **D1** in the cluster. cluster 1, cluster 4 and cluster 5 mostly consist of skeletons from **D1** and these clusters are considered to be useful in inspecting the fall risks.

28 look like safe at a glance. However, several clusters mostly consist of postures from **D1** or **D6** which look more unsafe than other postures. In these postures, the subtle fall risks which are included in **D1** and **D6** are more apparent. Therefore, we believe that our method using the focus-on transformation with the proposed features can provide useful information in inspecting the subtle fall risks.

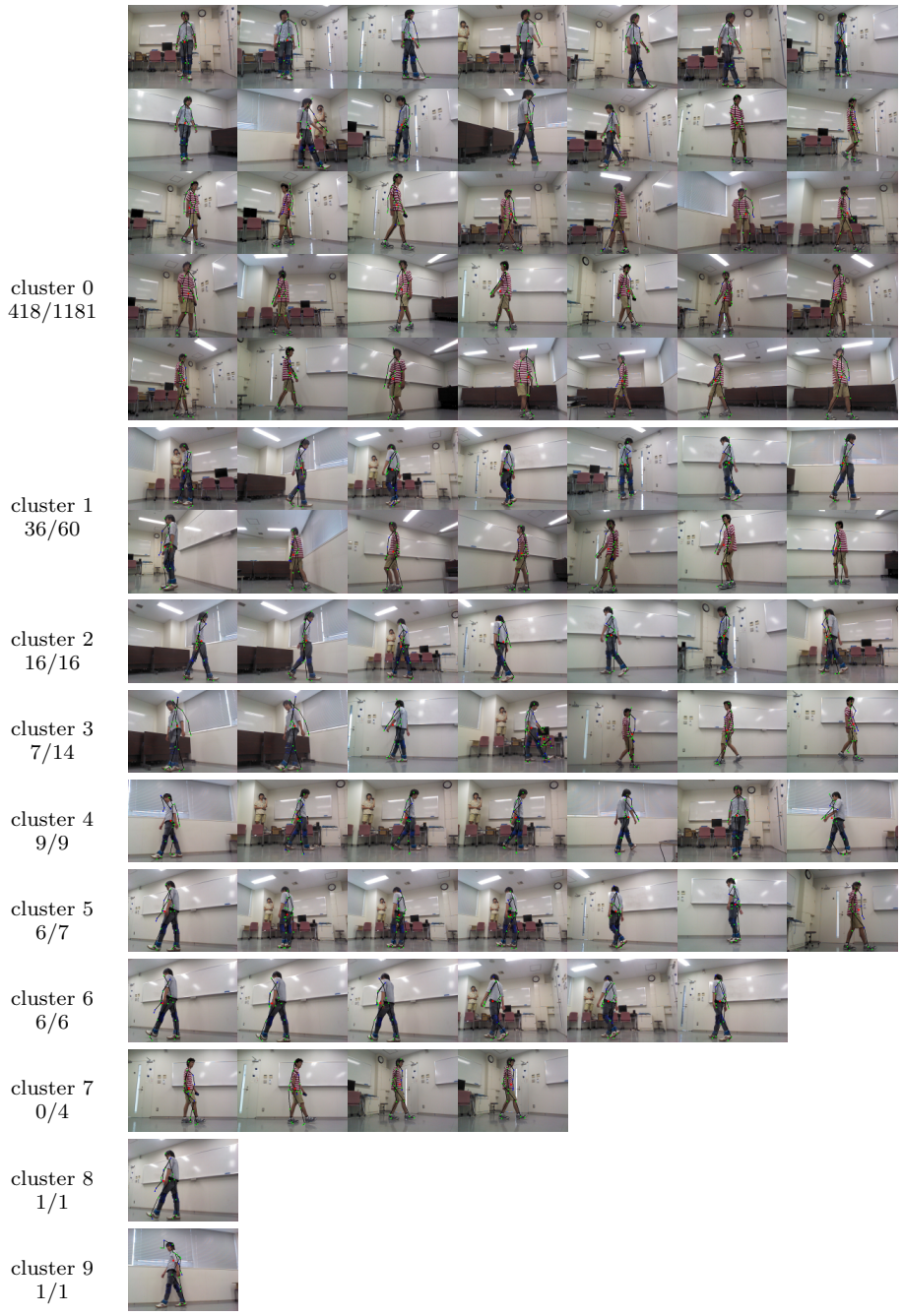


Fig. 28 Examples of the obtained clusters from **D6** with **A0**: $n1/n2$ represents that there are $n2$ skeletons including $n1$ skeletons from **D6** in the cluster. cluster 2, cluster 4, cluster 6, cluster 8 and cluster 9 mostly consist of skeletons from **D1** and these clusters are considered to be useful in inspecting fall risks.

7 Conclusions

We developed an autonomous multi-robot monitoring system and tackled the problem of discovering fall risks: to discriminate normal and abnormal gaits and to obtain useful information in inspecting fall risks. In this problem, we aim to prevent dangerous situations in which an elderly person is likely to fall. Since unsafe skeletons, in which fall risks appear as slight instabilities, are similar to safe skeletons, our problem is challenging. It is difficult to reflect small instabilities of unsafe skeletons, thus we need features which are sensitive to the instabilities. There are no obvious boundary between safe and unsafe skeletons, thus we need a method enabling clustering to focus on the difference between safe and unsafe skeletons. It is difficult to find unsafe skeletons at a glance, thus we need an evaluation method which does not need a class label for each skeleton.

To discriminate safe and unsafe skeletons, we proposed two instability features and a preprocessing method for skeleton clustering. The features, the horizontal deviation of the upper and lower bodies and the curvature of the back, are sensitive to instabilities which are caused by subtle fall risks. The preprocessing method, the focus-on transformation for an instability feature, emphasizes the difference between safe and unsafe skeletons. To identify an abnormal gait, we defined a similarity between two datasets in terms of the occurrence frequencies of skeletons with respect to the obtained clusters. By using the similarity, we proposed an evaluation method which uses two normal and an abnormal gait datasets.

To evaluate our method, we conducted three kinds of experiments by using seven kinds of gait datasets of four persons. The datasets of each person consist of a normal gait dataset and six kinds of abnormal gait datasets. In the first experiments, we assumed that a situation that an elderly person is monitored periodically, e.g., every month to confirm whether he/she walks with fall risks. Thus our evaluation method used the datasets of a singular person for calculating two kinds of similarities. In the second experiments, we assumed that a normal gait dataset of the target person were not available for identifying whether the gait which is observed as having fall risks. Thus we used a normal gait dataset of another person for calculating the similarity between two datasets. The third experiments were conducted from more practical viewpoints: how many skeletons are necessary to identify an abnormal gait and whether there are several clusters which mostly consist of unsafe skeletons.

The first and second experiments exhibited that our method using the focus-on transformation with the proposed features mostly outperformed the other methods when fall risks appeared in skeletons as instabilities. In an irregular case, fall risks appeared as slower walking speed and shorter strides, The previous features reflected the difference between a normal gait and an abnormal gait which had such fall risks. Thus the previous features were as effective as the proposed features in the first experiments. However, the previous features were ineffective in the second experiments because the features exhibited similar values extracted from standing postures.

In the third experiments, the similarity between two datasets became stable and the difference between a normal and an abnormal gait datasets became large after about 500 skeletons were input. By using our monitoring system, it took about two minutes without missing time to observe 500 skeletons. In the obtained clusters, there were several clusters which mostly consist of postures from the

abnormal gait dataset and these postures look more unsafe. Thus these clusters were considered to be useful in inspecting fall risks. Therefore, our method using the focus-on transformation with the proposed features is effective to discover fall risks when more than 500 skeletons are observed and fall risks appear in skeletons as instabilities.

Future work includes an investigation on the effective range of the evaluation method in terms of its robustness as well as alternative theoretical methods which are valid for a wider range of problems.

Acknowledgements A part of this research was supported by a Bilateral Joint Research Project between Japan and France funded from Japan Society for the Promotion of Science (JSPS) and Centre National de la Recherche Scientifique (CNRS/JSPS PRC 0672), and JSPS KAKENHI 24650070 and 25280085.

References

[Ardiyanto and Miura(2014)] Ardiyanto I, Miura J (2014) Cameraman robot: dynamic trajectory tracking with final time constraint using state-time space stochastic approach. In: Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference on, IEEE, pp 3108–3115

[Bayer and McCreight(1972)] Bayer R, McCreight EM (1972) Organization and maintenance of large ordered indices. Acta Informatica 1:173–189, DOI 10.1007/BF00288683

[Bhattacharyya(1946)] Bhattacharyya A (1946) On a measure of divergence between two multinomial populations. Sankhyā: The Indian Journal of Statistics pp 401–406

[Clauser et al(1969)Clauser, McConville, and Young] Clauser CE, McConville JT, Young JW (1969) Weight, volume, and center of mass of segments of the human body. Tech. rep., AMRL-TR-69-70, Wright Patterson Air Force Base, Ohio, USA

[Coradeschi et al(2013)Coradeschi, Cesta, Cortellessa, Coraci, Gonzalez, Karlsson, Furfari, Loutfi, Orlandini, Palumbo, Pecora, von Rump, Š
Coradeschi S, Cesta A, Cortellessa G, Coraci L, Gonzalez J, Karlsson L, Furfari F, Loutfi A, Orlandini A, Palumbo F, Pecora F, von Rump S, Štimec A, Ullberg J, Östlund B (2013) GiraffPlus: combining social interaction and long term monitoring for promoting independent living. In: Human System Interaction (HSI) 2013

[Deguchi(2014)] Deguchi Y (2014) Construction of a human monitoring system based on incremental discovery by multiple mobile robots. Master thesis, Graduate School of ISEE, Kyushu University, (in Japanese)

[Deguchi and Suzuki(2014)] Deguchi Y, Suzuki E (2014) Skeleton clustering by autonomous mobile robots for subtle fall risk discovery. In: ISMIS 2014. LNCS, vol. 8502, pp 500–505, DOI 10.1007/978-3-319-08326-1_51

[Deguchi et al(2014)Deguchi, Takayama, Takano, Scuturici, Petit, and Suzuki] Deguchi Y, Takayama D, Takano S, Scuturici VM, Petit JM, Suzuki E (2014) Multiple-robot monitoring system based on a service-oriented DBMS. In: Seventh ACM International Conference on Pervasive Technologies Related to Assistive Environments (PETRA 2014)

[Diraco et al(2010)Diraco, Leone, and Siciliano] Diraco G, Leone A, Siciliano P (2010) An active vision system for fall detection and posture recognition in elderly healthcare. In: Design, Automation and Test in Europe, pp 1536–1541

[Dunbabin and Marques(2012)] Dunbabin M, Marques L (2012) Robots for environmental monitoring: significant advancements and applications. IEEE Robotics & Automation Magazine 19(1):24–39, DOI 10.1109/MRA.2011.2181683

[Fischinger et al(2013)Fischinger, Einramhof, Wohlkinger, Papoutsakis, Mayer, Panek, Koertner, Hofmann, Argyros, Vincze, Weiss, and Gisi
Fischinger D, Einramhof P, Wohlkinger W, Papoutsakis K, Mayer P, Panek P, Koertner T, Hofmann S, Argyros A, Vincze M, Weiss A, Gisinger C (2013) Hobbit - the mutual care robot. In: ASROB-2013 in conjunction with IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), (in conjunction with IROS 2013)

[Freund and Schapire(1997)] Freund Y, Schapire RE (1997) A decision-theoretic generalization of on-line learning and an application to boosting. Computer and System Sciences 55(1):119–139, DOI 10.1006/jcss.1997.1504

[Gjoreski et al(2012)Gjoreski, Lustrek, and Gams] Gjoreski H, Lustrek M, Gams M (2012) Context-based fall detection using inertial and location sensors. In: Ambient Intelligence - Third International Joint Conference, AmI 2012, pp 1–16, DOI 10.1007/978-3-642-34898-3_1

[Gripay et al(2010)Gripay, Laforest, and Petit] Gripay Y, Laforest F, Petit JM (2010) A simple (yet powerful) algebra for pervasive environments. In: EDBT 2010, 13th International Conference on Extending Database Technology, pp 359–370, DOI 10.1145/1739041.1739086

[Kondo et al(2014)Kondo, Deguchi, and Suzuki] Kondo R, Deguchi Y, Suzuki E (2014) Developing a face monitoring robot for a desk worker. In: AmI 2014. LNCS, vol. 8850, Springer-Verlag, pp 226–241

[Koshmak et al(2013)Koshmak, Linden, and Loutfi] Koshmak GA, Linden M, Loutfi A (2013) Evaluation of the android-based fall detection system with physiological data monitoring. In: the 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC'13), IEEE, pp 1164–1168

[Leibrandt et al(2014)Leibrandt, Marcus, Kwok, and Yang] Leibrandt K, Marcus HJ, Kwok KW, Yang GZ (2014) Implicit active constraints for a compliant surgical manipulator. In: 2014 IEEE International Conference on Robotics and Automation, ICRA 2014, Hong Kong, China, May 31 - June 7, 2014, pp 276–283, DOI 10.1109/ICRA.2014.6906622

[Martinson and Yalla(2014)] Martinson E, Yalla V (2014) Guiding computational perception through a shared auditory space. In: Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference on, IEEE, pp 3156–3161

[Neumann et al(2012)Neumann, Asadi, Lilienthal, Bartholmai, and Schiller] Neumann PP, Asadi S, Lilienthal AJ, Bartholmai M, Schiller JH (2012) Autonomous gas-sensitive microdrone: wind vector estimation and gas distribution mapping. IEEE Robotics & Automation Magazine 19(1):50–61, DOI 10.1109/MRA.2012.2184671

[Pal et al(2014)Pal, Saha, and Konar] Pal M, Saha S, Konar A (2014) A fuzzy c means clustering approach for gesture recognition in healthcare. Enhanced Research in Science Technology & Engineering 3:87–94

[Palunko et al(2012)Palunko, Cruz, and Fierro] Palunko I, Cruz P, Fierro R (2012) Agile load transportation : safe and efficient load manipulation with aerial robots. IEEE Robotics & Automation Magazine 19(3):69–79, DOI 10.1109/MRA.2012.2205617

[Pippin and Christensen(2014)] Pippin C, Christensen HI (2014) Trust modeling in multi-robot patrolling. In: 2014 IEEE International Conference on Robotics and Automation, ICRA, pp 59–66, DOI 10.1109/ICRA.2014.6906590

[Rougier et al(2011)Rougier, Auvinet, Rousseau, Mignotte, and Meunier] Rougier C, Auvinet E, Rousseau J, Mignotte M, Meunier J (2011) Fall detection from depth map video sequences. In: Toward Useful Services for Elderly and People with Disabilities - 9th International Conference on Smart Homes and Health Telematics, ICOST 2011, pp 121–128, DOI 10.1007/978-3-642-21535-3_16

[Rubenstein(2006)] Rubenstein LZ (2006) Falls in older people: epidemiology, risk factors and strategies for prevention. Age and Ageing 35(Suppl 2):ii37–ii41

[Scuturici et al(2012)Scuturici, Surdu, Gripay, and Petit] Scuturici VM, Surdu S, Gripay Y, Petit JM (2012) UbiWare: web-based dynamic data & service management platform for AmI. In: ACM (ed) 13th International Middleware Conference (Middleware'12), ACM, pp 11:1–11:2, DOI 10.1145/2405153.2405164

[Sixsmith and Johnson(2004)] Sixsmith A, Johnson N (2004) A smart sensor to detect the falls of the elderly. IEEE Pervasive Computing 3(2):42–47, DOI 10.1109/MPRV.2004.1316817

[Suzuki et al(2012)Suzuki, Matsumoto, and Kouno] Suzuki E, Matsumoto E, Kouno A (2012) Data squashing for HSV subimages by an autonomous mobile robot. In: Discovery Science (DS 2012), LNAI 7569, Springer-Verlag, pp 95–109

[Suzuki et al(2014)Suzuki, Deguchi, Takayama, Takano, Scuturici, and Petit] Suzuki E, Deguchi Y, Takayama D, Takano S, Scuturici VM, Petit JM (2014) Towards facilitating the development of a monitoring system with autonomous mobile robots. In: Information Search, Integration and Personalization, Springer, pp 57–70

[Takayama et al(2014)Takayama, Deguchi, Takano, Scuturici, Petit, and Suzuki] Takayama D, Deguchi Y, Takano S, Scuturici VM, Petit JM, Suzuki E (2014) Multi-view onboard clustering of skeleton data for fall risk discovery. In: AmI 2014. LNCS, vol. 8850, Springer-Verlag, pp 258–273

[Tinetti et al(1986)Tinetti, Williams, and Mayewski] Tinetti ME, Williams TF, Mayewski R (1986) Fall risk index for elderly patients based on number of chronic disabilities. The American Journal of Medicine 80(3):429–434

[United Nations, Department of Economic and Social Affairs, Population Devision(2013)] United Nations, Department of Economic and Social Affairs, Population Devision (2013) World population ageing 2013. Tech. rep., ST/ESA/SER.A/348

[Wang(2010)] Wang Z (2010) Comparison of four kinds of fuzzy c-means clustering methods. In: Information Processing (ISIP), 2010 Third International Symposium on, IEEE, pp 563–566

[Zhang et al(1997)Zhang, Ramakrishnan, and Livny] Zhang T, Ramakrishnan R, Livny M (1997) BIRCH: a new data clustering algorithm and its applications. Data Mining and Knowledge Discovery 1(2):141–182, DOI 10.1023/A:1009783824328