Pattern Recognition 48 (2015) 1289–1301

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

Similar gait action recognition using an inertial sensor

Trung Thanh Ngo [a,*], Yasushi Makihara [b], Hajime Nagahara [a,1], Yasuhiro Mukaigawa [c], Yasushi Yagi [b]

[a] Faculty of Information Science and Electrical Engineering, Kyushu University, 744 Motooka, Nishiku, Fukuoka, 819-0395, Japan
[b] The Institute of Scientific and Industrial Research, Osaka University, 8-1Mihogaoka, Ibaraki, Osaka, 567-0047, Japan
[c] Graduate School of Information Science, Nara Institute of Science of Technology, 8916-5 Takayama-cho, Ikoma, Nara, 630-0192, Japan

ARTICLE INFO

Article history:
Received 14 January 2014
Received in revised form
12 August 2014
Accepted 9 October 2014
Available online 22 October 2014

Keywords:
Gait action recognition
Inertial sensor
Heel strike detection
Sensor orientation robustness
Interclass relationship

ABSTRACT

This paper tackles a challenging problem of inertial sensor-based recognition for similar gait action classes (such as walking on flat ground, up/down stairs, and up/down a slope). We solve three drawbacks of existing methods in the case of gait actions: the action signal segmentation, the sensor orientation inconsistency, and the recognition of similar action classes. First, to robustly segment the walking action under drastic changes in various factors such as speed, intensity, style, and sensor orientation of different participants, we rely on the likelihood of heel strike computed employing a scale-space technique. Second, to solve the problem of 3D sensor orientation inconsistency when matching the signals captured at different sensor orientations, we correct the sensor's tilt before applying an orientation-compensative matching algorithm to solve the remaining angle. Third, to accurately classify similar actions, we incorporate the interclass relationship in the feature vector for recognition. In experiments, the proposed algorithms were positively validated with 460 participants (the largest number in the research field), and five similar gait action classes (namely walking on flat ground, up/down stairs, and up/down a slope) captured by three inertial sensors at different positions (center, left, and right) and orientations on the participant's waist.

1. Introduction

With advances in micro-sensor and wireless communication technology, inertial sensors (accelerometer and/or gyroscope) are now low-power, small, accurate, and fast. They are increasingly being embedded in wearable and portable electronic devices such as smartphones, tablets, and smartwatches. As a result, many researchers have been employing a wearable inertial sensor in a variety of research topics such as human-machine interaction [1], user authentication [2], driving analysis [3], fall detection for medical alerts in the elderly [4], rehabilitation and therapy for patients [5], sport training support [6], and a user's daily life surveillance and monitoring [7]. Currently, recognizing a wearer's actions through an inertial sensor is one of the most attractive research topics.

Various actions with different levels of complexity have been investigated in this research field. They are mostly gestures, movements, behaviors, postures, transitions of postures, and sequence of movements of a participant such as sitting, standing, lying, walking, running, walking up/down a slope, falling, driving, cycling, dressing, working in an office, and cooking. Depending on the characteristics of the actions, such as their complexity, periodicity, and dynamicity, the optimal number of sensors and their placement, and recognition method has been decided. We refer readers to a number of recent reports and evaluations [8–14] for details.

There are two essential difficulties for inertial sensor-based action recognition methods: the segmentation of action signals and the relaxation of sensor attachment inconsistency between training and test stages. Particularly, in the case of recognizing similar action classes, an additional difficulty is low recognition accuracy.

Action signal segmentation is the first and most important step toward extracting a signal sequence from an action so that it can be classified. However, existing methods are sensitive to temporal and intensity variation of action signals such as when the participant changes their action speed or style.

The sensor attachment inconsistency problem occurs if locations and/or orientations of the sensor are different between training and testing stages. The existing methods can solve the orientation inconsistency between training and test stages; however, they have to pay a significant loss of signal information. For the details, they have to sacrifice some signal dimension to deal with this problem.

Existing methods have usually been evaluated for relatively different action classes, and hence there is no guarantee that they work well for very similar action classes. Although some authors

* Corresponding author. Tel.: +81 6 6879 8422; fax: +81 6 6877 4375.
E-mail addresses: trungbeo@gmail.com (T.T. Ngo), makihara@am.sanken.osaka-u.ac.jp (Y. Makihara), nagahara@limu.ait.kyushu-u.ac.jp (H. Nagahara), mukaigawa@is.naist.jp (Y. Mukaigawa), yagi@am.sanken.osaka-u.ac.jp (Y. Yagi).
[1] Tel.: +81 92 802 3595; fax: +81 92 802 3579.

http://dx.doi.org/10.1016/j.patcog.2014.10.012
0031-3203/

evaluated their methods with similar action classes such as gait action [15–17], there is no existing method that tentatively solves the problem of similarity of action classes.

In this study, we focus on similar gait action classes, which are the most frequent actions of humans in daily life. We provide solutions to the three above-mentioned problems in the case of classifying similar gait actions:

1. Step signal is detected and segmented employing a scale-space technique. The proposed step detection method can adaptively work with a large amount of variation even if the participant changes their walking speed or style.
2. To solve the practical sensor orientation inconsistency problem. First, we employ a gyroscope for the sensor tilt correction. Then, we apply an orientation-compensative matching algorithm [18] to solve the remaining relative sensor orientation angle between training and test signal sequences. As a result, the proposed method does not experience the information loss problem of existing recognition methods.
3. We propose an algorithm to deal with similar action classes. When action classes are similar, the relationship between one class and all others is more likely to have consistent and distinguished patterns as in the case of gait action. We utilize these relationship patterns to recognize gait action.

This paper is an extended version of our previous work [19]. First, while the previous work did not solve the sensor orientation inconsistency problem, the proposed method does. We employs both an accelerometer and a gyroscope sensors. The advantage of using a gyroscope is that we can fix the sensor tilt (represented by pitch and roll angles) in order to reduce the complexity before applying the orientation-compensative matching algorithm [18] to estimate the remaining orientation angle (yaw). Second, the robustness of step detection against sensor orientation inconsistency is realized and evaluated in this paper. Finally, the previous work evaluated performance using only an attachment location of a single accelerometer of 96 participants. Meanwhile, the proposed method is evaluated rigorously with three variations of sensor orientations and locations and a fourfold increase in the number of participants (460).

2. Related work

2.1. Action signal segmentation

A fixed-size sliding window has frequently been used [20–27]. However, a fixed-size window sometimes introduces errors since it may wrongly segment an action and cannot deal with temporal variation of an action due to speed or user difference. A dynamic window [28,29] has been proposed to solve the problem of the fixed-size window. These methods rely on signal events detected according to a fixed threshold of the signal intensity [28] or noise/signal separation theory to control the size and location of the window. The dynamic windows may, however, still fail when the signal intensity of an action also varies [28]. In the case of gait action recognition, there exist methods [30,31] that detect a gait period (or gait cycle of two consecutive steps) to construct a gait pattern; this is also considered to be using dynamic windows. However, these methods rely on local peak and valley detection, which is sensitive to variations in walking speed and/or style.

2.2. Gait period detection

In the field of inertial gait-based recognition, most existing methods try to detect gait period as a gait primitive, since they work better for the dynamicity of gait signals than those that use a fixed-size sliding window. In such cases, walking is a homogeneous and periodic action, it is hence possible to detect the period of the gait signal by dynamic programming [2], or matching with a sample primitive [32]. However, there is no such method in the field to cope with the situation where gait signal is drastically varied by a number of factors such as intensity, speed, and sensor orientation. The problem is more serious if these factors occur simultaneously.

2.3. Sensor attachment inconsistency

Most existing action recognition methods assume that the sensor is fixed at specific orientation and location on the participant's body. However, it is impractical and unnatural to fix the sensor at the same orientation and location all the time, particularly in daily life (e.g., the sensor orientation of a smartphone in a trouser pocket is subject to change). There are various methods that can be used to solve the sensor location inconsistency (or sensor displacement) problem such as unsupervised adaption [33,34], extracting invariant features from data of different sensor-locations [35], and employing heuristic knowledge [36]. The most popular approach to the sensor orientation inconsistency is to employ a 1D orientation-invariant signal [37,38], which is the magnitude of a 3D signal from an accelerometer or a gyroscope. Other researchers [39–41] use a 2D orientation-invariant signal, which relies on Mizell's research [42], to correct the sensor tilt using a 3D gait acceleration signal. However, these methods produce low performance because of the significant information loss by the dimension reduction of the signal. For the tilt correction, Mizell assumes that the average of 3D acceleration signal samples is the gravity vector in order to correct the sensor tilt. In fact, this assumption does not base on any theory. The averaging of the acceleration samples is performed ignoring the fact that the sensor is rotated when the human body moves. It is particularly incorrect for a short signal sequence that does not contain a natural number of gait periods or when the participant does not walk symmetrically.

A method that corrects the sensor orientation so that all the three dimensions of the signal can be used also exists. However, this method [43] relies on an assumption that the first principal component of the horizontal acceleration data corresponds to the forward (or backward) motion vector. This assumption is not always correct (e.g., when the participant turns), and hence the robustness of the method is reduced. There also exists a method that can estimate 3D relative orientation between a pair of acceleration signal sequences [18]. However this must be carried out for any pair of gallery and probe signal sequences, which is very time-consuming and only suitable for small database problem. In our research, taking the advantage of the gyroscope, we solve the 3D orientation by first estimating the absolute gravity vector to correct the sensor tilt at the pre-processing step and then employing [18] for only solving the remaining relative yaw angle. Consequently, the solution to the sensor orientation inconsistency problem in the proposed method is more advantageous in computational cost, robustness, and accuracy.

Ustev et al. [44] rely on a fusion of sensors to cope with the sensor orientation inconsistency. They use an accelerometer, a compass, and a gyroscope simultaneously to estimate the sensor orientation, hence the captured acceleration signal sequence can be corrected. The limitations of this approach are that the magnetic field is influenced by nearby electronic devices and the signal from the gyroscope is subject to the sensor drift. Moreover, their method needs to know the initial sensor orientation at the beginning of a capturing session of all the participants that limits the application of the method.