Applied Mathematics and Nonlinear Sciences

<https://www.sciendo.com>

Mathematical function data model analysis and synthesis system based on short-term human movement

Xuemei An^{1†}, Rui Yang¹, Daniyal M. Alghazzawi², Nympha Rita Joseph³¹ Hebei Institute of Mechanical and Electronic Technology, Xintai 054000, China² Department of Information Systems, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah Saudi Arabia³ Applied Science University, Al Eker, Kingdom of Bahrain

Submission Info

Communicated by Juan Luis García Guirao

Received June 17th 2021

Accepted September 24th 2021

Available online November 22th 2021

Abstract

The paper proposes a data model analysis algorithm for human motion function based on short-term behaviour. The algorithm uses a functional data analysis (FDA) method to perform Fourier fitting on the motion data and extract the fitted approximate single period data. Finally, the algorithm depicts the internal change in the motion in the low-dimensional space. The study found that the characteristic motion data obtained by the algorithm has smooth characteristics, and the relevant case analysis also verifies the algorithm's effectiveness.

Keywords: short-term behaviour, human motion recognition, mathematical function model, wearable motion capture system**AMS 2010 codes:** 93A30

1 Introduction

Human behaviour recognition is a method for computers to understand and describe human behaviour. It is a critical technology in related fields such as video surveillance, innovative nursing and human-computer interaction. Compared with traditional optical motion capture systems, wearable devices can collect data that are not affected by light, venue and occlusion [1]. Moreover, the device has the advantages of good portability and protection of the personal privacy of the collector. Therefore, the research on human behaviour recognition based on wearable motion capture systems is gradually becoming a hot topic nowadays.

Human behaviour actions are divided into periodic behaviours and non-periodic behaviours approximately. Approximately periodic behaviour refers to a movement behaviour that is executed multiple times continu-

[†]Corresponding author.Email address: xuemeian771@163.com

ously. The daily behaviour of the human body often has periodic characteristics, and the data sequence of daily behaviour collected by sensors also presents approximately periodic characteristics. This article considers converting this discrete motion data sequence into a periodic continuous function. This algorithm provides a theoretical basis for the study of the continuity characteristics of sports data and can accurately express the periodic characteristics of sports data [2]. Therefore, this article introduces the functional data analysis (FDA) algorithm to introduce the discrete data functional processing method. The basic idea of the algorithm is to treat the observed sample as a discrete data sequence but to treat the observed data as a whole and express it with a function. Then we use continuous data processing methods to analyse the samples.

2 Related theories and methods

2.1 FDA

The FDA method is a statistical method that has emerged in recent years. It is a development and extension of traditional statistical analysis methods. FDA is often used for economic data, meteorological data etc., but the application to sports data series is rare [3]. The changes of speed, acceleration and other variables in the process of human movement are continuous, while the corresponding data collected by sensors are all discrete. Because of this, the FDA's research on exercise data is reasonable. The general process of the FDA is as follows. For an observation data series, $y = (y_1, y_2, \dots, y_n)^T$, we build a model

$$y_i = x(t_i) + \epsilon_i, i = 1, 2, \dots, n \quad (1)$$

Among them, $x(t_i)$ represents the value of the function $x(t)$ of the observation data sequence at t_i , and ϵ_i represents the noise. To estimate the value of $x(t_i)$ in formula (1), we use the method of basis function expansion to select a set of basis functions $\Phi(t) = (\phi_1(t), \phi_2(t), \dots, \phi_K(t))$. At the same time, the discrete data is converted into a linear combination of basis functions,

$$x(t_i) = \sum_{k=1}^K c_k \phi_k(t_i), i = 1, 2, \dots, n \quad (2)$$

After the basis function is determined, a set of values of the coefficient vector $C = (c_1, c_2, \dots, c_K)^T$ uniquely determines a function [4]. The most direct method to solve the coefficient vector is the least-squares method, and the solution is as follows to minimise the residual sum of squares

$$SMSSE(y|c) = \sum_{i=1}^n \left[y_i - \sum_{k=1}^K c_k \phi_k(t_i) \right]^2 \quad (3)$$

For periodic observation data, the Fourier series is often used as the basis function.

$$\phi_1(t) = 1, \phi_2(t) = \sin(\omega t), \phi_3(t) = \cos(\omega t), \dots, \phi_{2k}(t) = \sin(k\omega t), \phi_{2k+1}(t) = \cos(k\omega t) \quad (4)$$

Therefore, the Fourier expansion of the function $x(t)$ of the observation data sequence is

$$x(t) = c_1 + c_2 \sin(\omega t) + c_3 \cos(\omega t) + \dots + c_{2k} \sin(k\omega t) + c_{2k+1} \cos(k\omega t) \quad (5)$$

2.2 Fourier series fitting

2.2.1 The fitting model and the determination of the frequency of the fitting function

The algorithm in this paper is mainly aimed at the periodic behaviour data of the human body collected by the wearable motion capture system [5]. We need to functionalise the discrete motion data to obtain the

continuous characteristics of human motion. For this periodic discrete observation data, we use k-order Fourier series fitting and the value of k determines the complexity of the model. The model is more straightforward when the k value is small, but it is easy to cause the data to underfit. When the value of k is significant, the model may be more complicated, and over-fitting may occur. K can be estimated based on empirical values in this article, and $k = 4$. In addition, the frequency of the fitting function is denoted as ω_{func} in the formula. The frequency ω_{func} and period T_{func} of the fitting function should be as close as possible to the frequency ω_{data} and period T_{data} of the original data. The individual difference of each fitting function can be maintained.

Figure 2 shows two sets of original data sequences with approximate periodicity. We use Algorithm 1 to estimate the fundamental frequency ω to obtain the image of the 4th-order Fourier series fitting function. The formula is as follows:

$$x(t) = c_1 + c_2 \sin(\omega t) + c_3 \cos(\omega t) + \cdots + c_8 \sin(4\omega t) + c_9 \cos(4\omega t) \quad (6)$$

Experiments show that our use of Fourier series fitting can reflect the periodicity and local characteristics of the original data and remove outliers and abnormal points [6]. At the same time, compared with discrete motion data, the functionalised result reflects the continuity of human body motion and motion data. Moreover, it accurately selects a cycle starting point of the motion cycle.

Fig. 1 4th order Fourier series fitting.

2.2.2 Fourier series fitting applied to motion data

A movement behaviour often requires the use of multiple wireless sensors to collect data simultaneously, and each sensor is composed of multiple measurement units [7]. Therefore, each sensor generates a vector at each time and can be represented as follows:

$$a(t_i) = (a_1(t_i), a_2(t_i), \cdots, a_M(t_i)) \quad (7)$$

Among them, M represents the total number of sensor measurement units. A single sensor performs continuous sampling to generate an observation matrix

$$s = [a(t_1), a(t_2), \cdots, a(t_n)]^T \quad (8)$$

$a(t_i)$, $i = 1, 2, \cdots, n$ is the t_i sampling data of a single sensor. Therefore, Eq. (8) can be expressed as

$$s = \begin{bmatrix} a_1(t_1) & a_2(t_1) & \cdots & a_M(t_1) \\ a_1(t_2) & a_2(t_2) & \cdots & a_M(t_2) \\ \vdots & \vdots & \ddots & \vdots \\ a_1(t_n) & a_2(t_n) & \cdots & a_M(t_n) \end{bmatrix}_{n \times M} \quad (9)$$

Fourier series fitting is used for each column of the observation matrix. Then, we convert the data sequence of each measurement unit of a single sensor into a function about the sampling time t . Therefore, Eq. (9) is transformed into the following function vector:

$$s^* = (f_1(t), f_2(t), \dots, f_M(t)) \quad (10)$$

Multiple sensors collect the sample data at the same time, so there is the following observation function matrix:

$$s = \begin{bmatrix} f_1^1(t) & f_2^1(t) & \dots & f_M^1(t) \\ f_1^2(t) & f_2^2(t) & \dots & f_M^2(t) \\ \vdots & \vdots & \ddots & \vdots \\ f_1^N(t) & f_2^N(t) & \dots & f_M^N(t) \end{bmatrix}_{n \times M} \quad (11)$$

Note that N represents the total number of sensors and M represents the number of measurement units for each sensor.

2.3 Cycle start point selection

Different sample function matrices are obtained according to different types of behaviour data. However, the morphological curve characteristics of the corresponding functions in different sample function matrices are different [8]. This difference in morphology is the basis for the classification of sample sets, and this difference can be reflected in one movement cycle as shown in Figure 2. Therefore, we can distinguish between different behaviours by extracting only one period of data.

Fig. 2 Differences in data curves of different sports behaviours.

We select one element $f(t)$ from the sample function matrix of Eq. (11) and find its characteristic points. This paper selects the extreme point set $f(t)$ as the feature point set of the function and then selects the point in the feature point set that makes $f(t)$ to obtain the minimum value the starting point of the motion cycle [9]. Finally, each element function in the sample function matrix uses this starting point as the starting point of the motion cycle.

We choose the minimum extreme point as the starting point of the cycle and achieve data alignment. The minimum extreme point of the acceleration function in the vertical direction of the foot is the point where the acceleration value is the smallest during the movement. This time point often corresponds to when the person's foot is raised to the highest point. All the periodic behaviours based on footsteps include this movement process [10]. Therefore, we use this point as the starting point of the movement, which can align all sample functions of the same action here and also align periodic functions of different behaviours at this starting point.

3 Experiment and analysis

3.1 Experimental environment and data sources

This article uses Windows7 64-bit operating system and Matlab software for experiments. The data set we used comes from the wearable sensor behaviour recognition database (WARD) provided by the University of Berkeley in the United States. This data set binds sensors to 5 parts of the body namely left wrist, right wrist, waist, left ankle and right ankle. Each sensor consists of a three-axis accelerometer and a two-axis gyroscope. The sampling frequency is 30 Hz. The data set includes 13 types of daily behaviours. When collecting data, 20 collectors are required to do each action 5 times. Therefore, a total of 1,300 samples are included in the data set. The 1000 samples of 10 types of dynamic behaviours selected are shown in Table 1.

Table 1 Behaviour description in WARD

Numbering	Behaviour category	Behaviour description
1	Walk normally	Walk forward for >10 s
2	Walk counterclockwise	Go counterclockwise for >10 s
3	Walk clockwise	Go clockwise for >10 s
4	To the left	Turn left on the spot for >10 s
5	Turn right	Turn right on the spot for >10 s
6	Go up the stairs	Go up >10 stairs
7	Down the stairs	Go down >10 stairs
8	Jogging	Jogging lasts >10 s
9	Jump	Jump in place >5 times
10	Push wheelchair	Push the wheelchair for >10 s

WARD, wearable sensor behaviour recognition database.

3.2 Data preprocessing and cycle extraction

WARD uses 5 sensors to collect data and each sensor has 5 measurement units. Therefore, the values of N and M are 5. So the sample function matrix is stated as follows:

$$S = \begin{bmatrix} f_1^1(t) & f_2^1(t) & f_3^1(t) & f_4^1(t) & f_5^1(t) \\ f_1^2(t) & f_2^2(t) & f_3^2(t) & f_4^2(t) & f_5^2(t) \\ f_1^3(t) & f_2^3(t) & f_3^3(t) & f_4^3(t) & f_5^3(t) \\ f_1^4(t) & f_2^4(t) & f_3^4(t) & f_4^4(t) & f_5^4(t) \\ f_1^5(t) & f_2^5(t) & f_3^5(t) & f_4^5(t) & f_5^5(t) \end{bmatrix} \quad (12)$$

When sampling frequency $F = 30$ Hz, 45 points in the sensor sequence have been obtained with good experimental results. Note that 1.5 s is close to the time to complete an action. In our experiment, the time for an ordinary person to complete an exercise cycle is $T_{data} = 1.5$ s. In this paper, 45 points are used as the data length of a period. From Eq. (9), the length of the feature vector is $N \times M \times [l/d] = 1125$.

To calibrate the starting point of the cycle, it is necessary to select one of the element functions in the sample function matrix of Eq. (10) as the cycle's starting point. We use this period starting point as the starting point for the resampling of each element function in the sample function matrix. Since the acceleration function of the human body in the vertical direction can better reflect the difference of different behaviours and steps, the starting point of the calibration of this function is when a person's foot is raised to the highest point during exercise [11]. Therefore, we use the vertical acceleration function of the left ankle sensor in the WARD data set as the function of the starting point of the calibration cycle to be more accurate.

Figure 3 shows a sample of the collector walking normally. We extract the acceleration data of the sensor at the left ankle in the vertical direction and periodically extract the results. From Figure 3, it can be seen that the algorithm used by us achieves a more accurate function of the motion data – ground cycle extraction.

Fig. 3 Period extraction results of a single measurement unit.

3.3 Experimental results

In the simulation experiment, the K-CV method is adopted to select test samples and training samples. In the experiment, the data set is divided into $K = 20$ groups. We conducted a total of 20 experiments and they show that the algorithm used by us has a recognition rate of 97.5% for periodic human behaviour based on the WARD dataset. The specific results are shown in Table 2.

Table 2 Confusion matrix of 10 dynamic behaviour categories.

	1	2	3	4	5	6	7	8	9	10	Recognition rate (%)
1	97	0	0	0	0	1	0	0	0	2	97
2	0	97	0	1	0	0	0	0	0	2	97
3	0	0	98	0	1	0	1	0	0	0	98
4	0	0	0	100	0	0	0	0	0	0	100
5	0	0	0	0	100	0	0	0	0	0	100
6	0	0	0	0	0	97	1	1	1	0	97
7	0	0	0	0	0	2	94	1	0	3	94
8	0	0	0	0	0	0	0	100	0	0	100
9	0	0	0	0	0	1	5	0	94	0	94
10	1	0	0	0	0	0	0	0	1	98	98
Total recognition rate											97.5

3.4 Algorithm comparison

Table 3 and Figure 4 show the simulation results of different algorithms in the same environment using the WARD data set in Table 1. Table 3 and Figure 4 show that the recognition rate of this algorithm is higher than other algorithms [12]. The choice of the classifier is significant for behaviour classification. Table 4 results from a classification using different classifiers after extracting feature vectors according to algorithm used by us. It can be seen from Table 4 that the feature vector extracted by the algorithm can achieve a higher recognition rate.

Table 3 Comparison of algorithm results based on WARD

Window size	Feature extraction method	Classifier	Classifier time (s)	Recognition rate (%)
Literature [6]	200 time-frequency domain features	DT	<1	93.7
Literature [11]	45 LPP	DSC	N/A	87.05
Literature [12]	40 RP	SRC	N/A	82.1
Algorithm	45 Periodic extraction of calibration starting point	SVM	<1	97.5

LPP, partially preserved projection; RP, random projection; WARD, wearable sensor behaviour recognition database.

Table 4 Comparison of classification results of different classifiers

Classifier	Recognition rate (%)
Support vector machines	97.5
Decision tree	89.5
Naive Bayes	86.9
K-nearest neighbour	93.9

Fig. 4 Summary of classification results based on WARD-based methods and traditional methods. WARD, wearable sensor behaviour recognition database.

3.5 Experiments based on self-collecting database

3.5.1 Introduction to self-collecting database

To further verify the rationality and scalability of the algorithm, a motion capture system is used to record data under actual conditions. First, it applies the algorithm to the data set [13]. The database in this paper collects eight daily behaviours of 8 collectors of motion capture data. It extracts from the data of the left wrist, right wrist, waist, left ankle and right ankle. The data samples from the self-collected database are shown in the Figure 5. It is worth noting that the data sampling frequency of the WARD data set is low (30 Hz), while the motion capture system used by us has a high sampling frequency (90 Hz). It can be seen from Eq. (12) that if the resampling interval d of the fitting function is kept at one at this time, the period data length will be longer, and the feature vector dimension will be too large. To solve this problem, the sampling interval d in formula (12) is set to 3. Then the fitting function is resampled, and the feature vector is extracted. This adjustable resampling method reflects the applicability of this algorithm to the existing high-frequency motion capture system and reflects the advantages of operational data.

Fig. 5 Single sample data sequence.

3.5.2 Experimental results

Analogous to the experimental approach of the WARD data set, we use the K-CV method to establish a training sample and test sample library, and finally the support vector machine is used for classification and recognition. It is found that the recognition rate reaches 98.75%.

4 Conclusion

This paper uses the FDA method to functionalise the discrete data compared with the traditional discrete motion data processing method. This paper expresses the continuity of human motion data. First, we accurately calibrated the starting point of the motion cycle according to the function of the motion data. We then extracted the periodic sequence of the motion data approximately according to the starting point. The method in this paper has certain advantages over traditional window processing methods. Then, we use SVM and the periodic data sequence extracted by this algorithm to classify the human body's dynamic behaviour better to obtain a recognition rate comparable to the traditional method. The algorithm's effectiveness in this paper has been verified on the public data set WARD and the self-collecting database.

References

- [1] Booranawong, A., Jindapetch, N., & Saito, H. Adaptive filtering methods for rssi signals in a device-free human detection and tracking system. *IEEE Systems Journal.*, 2019. 13(3): 2998-3009.
- [2] Baker, L. B. Physiology of sweat gland function: The roles of sweating and sweat composition in human health. *Temperature.*, 2019. 6(3): 211-259.
- [3] Shrestha, D. C., & Acharya, S. Time Dependent Mathematical Model of Thermoregulation in Human Dermal Parts During Sarcopenia. *Journal of Nepal Mathematical Society.*, 2021. 4(1): 41-53.
- [4] Rahaman, H., Kamrul Hasan, M., Ali, A. & Shamsul Alam, M. Implicit Methods for Numerical Solution of Singular Initial Value Problems. *Applied Mathematics and Nonlinear Sciences.*, 2020. 6(1): 1-8.
- [5] Sulaiman, T., Bulut, H. & Baskonus, H. On the exact solutions to some system of complex nonlinear models. *Applied Mathematics and Nonlinear Sciences.*, 2020. 6(1): 29-42.
- [6] Chen, X., Zhang, A., Wang, H., Gallaher, A., & Zhu, X. Compliance and containment in social distancing: mathematical modeling of COVID-19 across townships. *International Journal of Geographical Information Science.*, 2021. 35(3): 446-465.
- [7] de Freitas, E., Ferrara, F., & Ferrari, G. The coordinated movements of collaborative mathematical tasks: The role of affect in transindividual sympathy. *ZDM.*, 2019. 51(2): 305-318.
- [8] Krishnan, R., Björzell, N., Gutierrez-Farewik, E. M., & Smith, C. A survey of human shoulder functional kinematic representations. *Medical & biological engineering & computing.*, 2019. 57(2): 339-367.
- [9] Trenchev, I., Mavrevski, R., Traykov, M., & Zajmi-Rugova, I. A mathematical model of movement in virtual reality through thoughts. *International Journal of Electrical and Computer Engineering (IJECE).*, 2020. 10(6): 6592-6597.
- [10] Reissner, L., Fischer, G., List, R., Giovanoli, P., & Calcagni, M. Assessment of hand function during activities of daily living using motion tracking cameras: A systematic review. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine.*, 2019. 233(8): 764-783.
- [11] Ali, A., Hussain, M., Ghaffar, A., Ali, Z., Nisar, K. S., Alharthi, M. R., & Jamshed, W. Numerical simulations and analysis for mathematical model of avascular tumor growth using Gompertz growth rate function. *Alexandria Engineering Journal.*, 2021. 60(4): 3731-3740.
- [12] Saxe, A. M., McClelland, J. L., & Ganguli, S. A mathematical theory of semantic development in deep neural networks. *Proceedings of the National Academy of Sciences.*, 2019. 116(23): 11537-11546.
- [13] Wada, S., Matsunaga, N., & Tamai, I. Mathematical modeling analysis of hepatic uric acid disposition using human sandwich-cultured hepatocytes. *Drug Metabolism and Pharmacokinetics.*, 2020. 35(5): 432-440.

This page is intentionally left blank