Research Article

Intelligent Analysis and Evaluation Method of Athletics Running Data Based on Big Data Statistical Model

Yushan Ge

Chongqing Industry Polytechnic College, Chongqing 401120, China

Correspondence should be addressed to Yushan Ge; geys@cqipc.edu.cn

Received 18 February 2022; Accepted 17 March 2022; Published 29 March 2022

1. Introduction

Athletics is a regular sporting event for many schools and local sporting bodies. Long distance running lap counting in athletics is a tedious task for referees and requires a large number of experienced and professional referees. Even so, there are still some problems in the process of officiating, and in serious cases, there are also misjudgements and omissions, leading to errors in the results of the competition and requiring a lot of time to correct them, which greatly affects the smoothness of the competition. In the information collection layer, the information of athletes is collected by means of voice recognition, which greatly reduces the errors that may be generated by manual lap counting; in the data transmission, it is directly uploaded to the cloud platform by means of GPRS, which facilitates the work of referees and improves the work efficiency. The system is low cost, easy to use and has a certain value and application prospect [1].

With the development of society and technology, a sedentary lifestyle with inadequate physical activity as the main characteristic due to the motorisation of transport, real-time communication and social networking is endangering the health of the general public [2]. Walking is one of the most fundamental forms of physical activity for humans, and its health benefits are becoming increasingly important. The easy and accurate measurement of walking is a key area of research in this field, as well as the most convenient and intuitive way for people to understand their physical activity levels on a daily basis [3]. From pedometers to triaxial accelerometers, there has been a constant quest for easier and more effective pedometry tools [4]. The advent of the smartphone has not only changed the single communication function of mobile phones, but has also changed the way people live and act, and it has become a new trend to use smartphone exercise APPs for fitness. At the same time, smart bracelets have also developed rapidly to guide healthy living by recording daily exercise data, which can be synchronised and shared with smart phones. To date, there have been reports on the development of smart bracelets and smartphone exercise applications (APPs) and gait recognition in China [5], but there is little literature on the accuracy of the pedometry function, which is assessed in two ways, namely at different walking speeds and on different walking surfaces [6].

The effectiveness of different pace counting functions based on the iPhone4siOS7 fitness app has been reported in the literature [7], but little has been reported on the effectiveness of different pace counting functions of smart bracelets and smartphone sports apps on different walking
surfaces (e.g. running on plastic athletic fields, hiking in the wild, climbing mountains and walking on city streets, etc.). This study is based on existing research in China and investigates the accuracy of the smartphone and smartphone sports app for different walking speeds and running surfaces.

2. Related Work

With the popularity of the pedometer function, the research results of step detection algorithms have increased. At present, the step detection algorithms based on acceleration sensors at home and abroad can be divided into the following eight categories, including: normalised autocorrelation, peak detection, finite state machine, etc.

In [8], a step detection algorithm is implemented based on autocorrelation coefficient analysis. The method detects the step count results of y-axis and z-axis acceleration data separately, and then selects the largest result as the final detection result. The final example of the method in the step detection algorithm research work of [9] was to exclude the influence of non-real walking data on the step counting results, and the authors matched the extracted nonrunning data template with the actual acceleration data for calculation based on autocorrelation coefficient analysis, thus ensuring the accuracy of the step counting. [10] used autocorrelation coefficient analysis to design a smartphone pedometer. Due to the uncertainty of the human stride period, the authors used an add-frame calculation to try to find the optimal stride period and validate the acceleration signal corresponding to the largest correlation coefficient result as a valid step. In [11], the same correlation coefficient analysis was used for step detection in the design of a Zee-based positioning navigation system, and invalid acceleration data was filtered by a threshold of signal variance. [12] set thresholds for acceleration maximum, minimum, and variance to trigger step counting in work investigating hybrid navigation systems [13]. [14] implemented a peak detection pedometry algorithm using a plus-minus score mechanism.

A filtering method is also proposed in the pedometry algorithm study of [15], which normalises the peak features so that an acceleration point with a value of 1 after filtering represents the generation of a step. In the work on the pedometry algorithm of [16], the authors extracted features such as the number of relationships and the degree of deviation from positive and negative values of the acceleration data, and then set thresholds for each feature; only acceleration peaks that reached the threshold setting were selected as candidate peaks. In the 3D positioning system research work of [17], the authors implemented an adaptive peak counting algorithm that improves the adaptability of the peak detection pedometry algorithm under multiple movements by setting different peak thresholds for different movements. [18] designed an adaptive peak detection pedometry algorithm using different threshold ranges based on the characteristic that running and walking produce different ranges of acceleration peaks, improving the adaptiveness of the peak detection method for running and walking movements. [19] implemented a step detection algorithm based on finite state machines, in which the authors decomposed the one-step acceleration signal into six states. [20] used a Kalman filter to preprocess the acceleration signal when designing the finite state machine based step counting algorithm, making the acceleration signal smoother and therefore allowing the process of noise processing to be removed in the finite state machine, effectively improving the step counting efficiency of the finite state machine based step counting algorithm by reducing the number of state transitions.

3. Human Running and Acceleration Signals

3.1. Analysis of the Human Running Process. In this paper, the following phases are distinguished from the other phases: supported phase, single foot lift forward phase, upright phase and heel landing phase. By analysing the changes in the body as the pedestrian completes a two-step manoeuvre.

As can be seen from Figure 1, during the actual continuous walking process, one step is actually completed during the double-legged support phase. After the support phase, the body gradually leans forward while lifting one foot forward, and the entire foot of the unlifted foot gradually comes into full contact with the ground until the lower limb of the body is approximately perpendicular to the ground, at which point the zero speed phenomenon, as indicated in the zero speed correction method occurs.

3.2. Starting Point Marking Method. The start point is the moment in the acceleration data that may indicate the start of a step, and each start point will correspond to the sampling time of a specific acceleration sampling point. When the acceleration sensor is located in the hand, trouser pocket and chest area of the body in Figure 1, the collected acceleration signal will exhibit the acceleration signal characteristics respectively. Although noise is present in the acceleration data, the acceleration signal for normal walking maintains a certain periodicity when viewed as a whole. It is assumed that the linear synthetic acceleration collected and calculated over a given time period is $A(n,t)$, where $n$ denotes the number of sampling points, $t$ denotes the sampling time and $a_k(t_k)$ denotes the sample point of linear synthetic acceleration with the value $k$ acquired and calculated at the $t_k$ time, where $k = 1, 2, \ldots, n$. In this paper, the following method is used to mark the starting points and record them as a set $TP_s$:

\[
TP_s = \{ t_k | a_k(t_k) \geq Th_s \cap a_{k-1}(t_{k-1}) < Th_s, k = 1, 2, \ldots, n \},
\]

where $Th_s$ is the starting point detection threshold. It should be noted that the linear synthetic acceleration calculated in this paper has lost the vector characteristics of the three axes and the gravity component, directly reflecting the acceleration changes generated by the sensor at a fixed part of the
body during walking. The theoretical output value of the acceleration sensor at rest is \( g = (0, 0, 9.8) \text{m/s}^2 \), so the linear synthetic acceleration magnitude should be zero under the ideal resting action. Due to the process material and measurement accuracy of the sensor itself, this can lead to errors in the sampling value at rest. The sensor used in this paper reads around \( g = (0.2, 0.2, 10.1) \text{m/s}^2 \) at rest, and is therefore set here \( Th_r = 0.5 \text{m/s}^2 \). Figure 2 shows the starting point in the acceleration signal detected by the accelerometer as it swings with the arm.

The detection of the start point allows a better representation of the periodicity of the data, and the recorded \( TP_s \) and linear acceleration series \( A(n,t) \) will be used in the next stage of the pacing algorithm. The effectiveness of this method is demonstrated in the literature [10], where it is used as a threshold for starting point detection, i.e. to filter errors.

4. Design of a Pedometry Algorithm Based on Autocorrelation Coefficient Analysis

Combined with the results of the start point detection in Figure 2, it can be seen that the data period in this case is different from the signal period when the accelerometer is in the trouser pocket and chest area. In addition to this, in terms of peak variation, it can be seen from Figure 2 that there is a peak in some cycles, some of which have a small amplitude, which can cause large pacing errors if the peak detection method is used for detection processing. Therefore, the design of the step counting algorithm can be considered from the perspective of calculating the similarity of the acceleration signal. In view of these factors, a step counting algorithm based on autocorrelation analysis was implemented to process the step detection of this action.

4.1. Design of a Pedometry Algorithm Based on Peak Detection.

When the action detection result is a relatively stable action, this paper uses a pacing algorithm based on peak detection, which requires a threshold variable \( Th_{peak} \), and in order to improve the adaptive nature of the peak, the threshold is set at the beginning of the algorithm execution \( Th_{peak} = 0.2g \) \( (g = 9.8 \text{m/s}^2) \), after which the \( Th_{peak} \) threshold will change in the following way:

\[
Th_{peak} = \frac{1}{M} \sum_{j=1}^{M} peak_j,
\]

where \( peak_j \) is the peak acceleration in the current window that has been validated and \( M \) is the number of validated peaks in the analysis window, the meaning of the above equation is that after a period of time, \( Th_{peak} \) is set to the mean value of the validated peaks. Suppose that the algorithm is processing a starting point in \( TP_s \), denoted \( TP_{s,\text{ur}} \) [21].

4.2. Step Detection Compensation Strategy.

When a pedestrian makes a transition during walking, this can cause significant noise interference in the acceleration signal. It is important to emphasise that the analysis window for action recognition in this paper is also the analysis window for the pedometry algorithm, so when just one transition action occurs within an analysis window, the system should recognise the acceleration signal within this analysis window as the acceleration signal under the transition action. For the case that the action recognition result belongs to the transition action, the acceleration data in one running cycle may produce several pseudo-peaks, and the frequency characteristics and variance characteristics of the whole acceleration signal are also less regular, for this situation, this paper tries to use a compensation method to detect the number of steps in the transition state as far as possible, the compensation method is as follows:

\[
\text{CSC} = \frac{TP_{s,\text{new}} - TP_{s,\text{old}}}{LS_{last}},
\]

where \( TP_{s,\text{new}} \) and \( TP_{s,\text{old}} \) denote the most recent and latest moments in the set of starting points \( TP_s \) of the transition.
movement acceleration sequence and $LS_{last}$ denotes the period of the last valid step. The strategy adopted in the above equation is to default the period of the run under the transition action to be consistent with the period of the previous step or steps, and to update the CSC value to the TSC in time after each compensated pacing is completed. The final TSC result will consist of the autocorrelation coefficient analysis pacing algorithm, the pacing calculation for peak detection and the compensated pacing result.

5. Case Studies

5.1. Research and Test Subjects. The subjects of this study are smart bracelets and smartphone sports APPs with pedometer function, three smart bracelets from Xiaomi, Lexin and SmartHealth, and four sports APPs from Goudong, Yueyun Circle, Yidong GPS and Dynamic as test subjects.

5.2. Pedometer Validity Tests on Different Surfaces. The test was conducted outdoors on four different running surfaces: plastic, dirt, concrete and hills. During the test, the participant wore an arm bag with an iPhone 5s phone (on the outside of the upper arm) and began to adapt to low speed, normal speed and fast speed for 3 minutes each (low speed, normal speed and fast speed were determined by the participant’s own subjective perception of speed). Afterwards, the Pro-Active App (which was determined to be more effective in the first part of the study) was activated, the phone was placed in the arm bag and the Heart of Joy bracelet was worn on the left wrist at the same time. The participant then walked 300 steps at a low, normal and fast pace (the number of steps was calculated by the participant), stopping at the end and recording the number of steps recorded by the APP at this time, 5 times for each venue. To avoid fatigue affecting the test results, there was a 5 min break between each exercise. During the test, a staff member held a video camera to follow the test and the participant counted 300 steps after the start of the walk and then finished the test.

5.3. Study Results. Table 1 shows the average of the recorded and actual step counts and the dispersion (standard deviation) for each of the five speed levels of 3.2 km/h, 4.8 km/h, 6.4 km/h, 8.0 km/h and 9.6 km/h for the three smart bracelets. As can be seen from the table, the most significant increase in the cumulative number of steps in 5 min occurred when the step rate was increased from 3.2 km/h to 4.8 km/h. Thereafter, the increase in the cumulative number of steps in 5 min slowed down with each 1.6 min increase. As the speed increases, the cumulative number of steps in 5 min decreases between the different rings and the difference between the actual number of steps decreases.

As seen in Table 2, similar to the three smart bands, the most significant increase in the cumulative number of steps in 5 min was observed when the four sports APPs were increased from 3.2 km/h to 4.8 km/h. Thereafter, the increase in the cumulative number of steps in 5 min slowed down with each 1.6 min increase. Again, as the speed increases, the cumulative number of steps in 5 min decreases between the different rings and the difference between the actual number of steps decreases. At the lowest speed level of 3.2 km/h, there was a highly significant difference between the four sports APPs in terms of actual steps ($p < 0.01$).

From the scatter diagram of linear regression analysis (Figures 3, 4 and 5, where x-axis is the number of samples), it can be seen that with the increase of running speed, the actual steps of the three bracelets show a good linear correlation with the steps recorded by the three bracelets, $R^2$ being 0.982, 0.998 and 0.998 respectively. With the improvement of running speed, the consistency between Xiaomi bracelet and the actual steps is slightly weaker than the other two bracelets. As can be seen from Figures 6 and 7 (where x-axis is the number of step), with the increase of walking speed, the actual steps of the four sports apps also show a good linear correlation with the steps recorded by the three bracelets, $R^2$ being 0.997, 0.997, 0.994 and 0.997 respectively.

As the speed increases, the three smart bracelets with the best stability of the step-keeping function are the Le Xin bracelet, as analysed in Figures 2 to 4 and Table 3. From Figures 5 to 7, as the speed increases, the 4 sports APPs with the best stability of the step recording function are Yidong and Goudong. Therefore, the effectiveness of the pedometer on different running surfaces was tested with the Le Xin Bracelet and Yidong APPs respectively. As shown in Figure 5, with the increase of running speed, the Le Xin smart bracelet was significantly different from the actual step count on concrete and mountainous terrain at low walking speed, with p-values <0.01 and <0.05 respectively, while the Yidong APP was also significantly different from the actual step count on concrete and mountainous terrain, with p-values <0.05.

6. Discussion

Based on three smartbands and four smartphone movement apps, this study investigated the effectiveness of the smartbands/phone movement apps in counting steps at different walking speeds and on different walking surfaces. Firstly, the single factor ANOVA showed that the measured step counts of the three smartbands were only significantly different from the actual step counts at the speed class of 3.2 km/h ($p < 0.01$) for all five speed states, while the other four speed classes did not show significant differences. This result is consistent with previous research literature on pedometers [11]. The same results were also found for the four smartphone sports apps, where the measured number of steps in the five walking speed states was only very significantly different from the actual number of steps in the speed class of 3.2 km/h ($p < 0.01$). The reason for this may be related to the pacing principle of the smart bracelet and
Table 1: Analysis of the results of the different speed tests of the 3 smart bracelets.

| Speed (km/h) | Millet         | Happy heart    | Smarthealth   |
|--------------|----------------|----------------|---------------|
| 3.2          | 453.6 ± 49.17  | 482.8 ± 8.59   | 479.8 ± 12.21 |
| 4.8          | 552.6 ± 8.27   | 549.6 ± 9.2    | 544.2 ± 9.68  |
| 6.4          | 585.4 ± 8.49   | 548.8 ± 9.75   | 577 ± 10.29   |
| 8.0          | 724 ± 9019     | 720.2 ± 9.65   | 729 ± 6.20    |
| 9.6          | 738.8 ± 6.47   | 738.8 ± 8.94   | 738.5 ± 6.54  |

Table 2: Analysis of the results of the 4 sports APPs in the lab running table.

| Speed (km/h) | Move         | Yidong        | Yue Pao       | Gudong       |
|--------------|--------------|---------------|---------------|
| 3.2          | 444 ± 3.123  | 444 ± 3.125   | 445.8 ± 2.247 | 444 ± 3.162  |
| 4.8          | 540.2 ± 8.075| 540.2 ± 8.337 | 540.8 ± 12.09 | 540.2 ± 8.075|
| 6.4          | 625 ± 21.79  | 625.6 ± 27.415| 628.8 ± 27.635| 625.6 ± 21.789|
| 8.0          | 732.41 ± 7.085| 732.6 ± 7.257 | 732.6 ± 15.587| 732.4 ± 7.045|
| 9.6          | 750 ± 5.241  | 750 ± 5.241   | 753.2 ± 7.147 | 750 ± 5.241  |

Figure 3: Linear regression scatter plot of the number of steps recorded and the actual number of steps taken by Xiaomi’s bracelet at different speed levels.

Figure 4: Scatter plot of the number of steps recorded by the Smarthealth bracelet versus the actual number of steps taken at different speed levels.
smart phone movement APP. The acceleration generated during walking causes the electrons in the sensor to move, resulting in a change in electrode position, and eventually the change in capacitance difference is integrated by the chip and the voltage value is output, resulting in a count. The “threshold value” of the pedometer is not sufficient to affect

Figure 5: Linear regression scatter plot of recorded steps and actual steps for different speed levels.

Figure 6: Linear regression scatter plot of the number of steps recorded by Kinetic APP and the actual number of steps at different speed levels.

Figure 7: Linear regression scatter plot of the number of steps recorded at different run speed.
the recording [12, 13]. The disadvantages of the method are that the period of the acceleration signal is not taken into account, and that the thresholds controlling the state transitions are fixed, making the method less adaptive overall.

7. Conclusions

This paper first analyses the track and field running data and further analyses the start point marking method adopted in this paper. The start point marker is critical to the size of the analysis window in this paper. In the actual system implementation process, when the moment of the currently detected start point meets a time length, the system will input the sampled data within this length into the action detection module for action recognition, and then apply the corresponding step counting algorithm according to the recognition result. The specific pacing algorithms are then described in detail, including the autocorrelation coefficient based pacing algorithm, the peak detection based pacing algorithm and the compensation algorithm. The pacing algorithm based on autocorrelation coefficient analysis gives better pacing results when the action recognition result is a swinging motion M2.

Data Availability

The experimental data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declared that they have no conflicts of interest regarding this work.

References

[1] R. Rein and D. Memmert, “Big data and tactical analysis in elite soccer: Future challenges and opportunities for sports science,” SpringerPlus, vol. 5, no. 1, pp. 1410–1413, 2016.
[2] R. Agarwal and V. Dhar, "Editorial-Big data, data science, and analytics: The opportunity and challenge for IS research," Information Systems Research, vol. 25, no. 3, pp. 443–448, 2014.
[3] S. Shafqat, S. Kishwer, R. U. Rasool, J. Qadir, T. Amjad, and H. F. Ahmad, "Big data analytics enhanced healthcare systems: A review," The Journal of Supercomputing, vol. 76, no. 3, pp. 1754–1799, 2020.
[4] B. Gupta, M. Goul, and B. Dinter, “Business intelligence and big data in higher education: Status of a multi-year model curriculum development effort for business school undergraduates, MS graduates, and MBAs,” Communications of the Association for Information Systems, vol. 36, no. 1, p. 23, 2015.
[5] G. Vossen, "Big data as the new enabler in business and other intelligence,” Vietnam Journal of Computer Science, vol. 1, no.1, pp. 3–14, 2014.
[6] J. Li, L. Xu, L. Tang, S. Wang, and L. Li, “Big data in tourism research: A literature review,” Tourism Management, vol. 68, pp. 301–323, 2018.
[7] G. Liu, J. Yang, Y. Hao, and Y. Zhang, “Big data-informed energy efficiency assessment of China industry sectors based on K-means clustering,” Journal of Cleaner Production, vol. 183, pp. 304–314, 2018.
[8] C. Cao, Y. Tang, D. Huang, W. Gan, and C. Zhang, “IIIE: an improved identity-based encryption algorithm for WSN security,” Security and Communication Networks, vol. 2021, p. 8, Article ID 8527068, 2021.
[9] M. Cantarella, R. Martinez-Espana, B. Ayuso, J. A. Yáñez, and A. Muñoz, “Analysis of student behavior in learning management systems through a Big Data framework,” Future Generation Computer Systems, vol. 90, pp. 262–272, 2019.
[10] P. P. Maglio and C.-H. Lim, “Innovation and big data in smart service systems,” Journal of Innovation Management, vol. 4, no. 1, pp. 11–21, 2016.
[11] S. Batistić and P. van der Laken, “History, evolution and future of big data and analytics: A bibliometric analysis of its relationship to performance in organizations,” British Journal of Management, vol. 30, no. 2, pp. 229–251, 2019.
[12] V. C. Storey and I.-Y. Song, “Big data technologies and management: What conceptual modeling can do,” Data & Knowledge Engineering, vol. 108, pp. 50–67, 2017.
[13] S. Sharma, U. S. Tim, J. Wong, S. Gadia, and S. Sharma, “A brief review on leading big data models,” Data Science Journal, vol. 13, pp. 138–157, 2014.
[14] M. Haghighat, H. Rastegari, N. Nourafza, N. Branch, and I. Esfahan, “A review of data mining techniques for result prediction in sports,” Advances in Computer Science: An International Journal, vol. 2, no. 5, pp. 7–12, 2013.
[15] J. Wang, Y. Yang, T. Wang, R. S. Sherratt, and J. Zhang, “Big data service architecture: A survey,” Journal of Internet Technology, vol. 21, no. 2, pp. 393–405, 2020.
[16] E. E. Cust, A. J. Sweeting, K. Ball, and S. Robertson, “Machine and deep learning for sport-specific movement recognition: A systematic review of model development and performance,” Journal of Sports Sciences, vol. 37, no. 5, pp. 568–600, 2019.
[17] Y. Chen, H. Chen, A. Gorkhali, Y. Lu, Y. Ma, and L. Li, “Big data analytics and big data science: A survey,” Journal of Management Analytics, vol. 3, no. 1, pp. 1–42, 2016.
[18] C. Gurrin, A. F. Smeaton, and A. R. Doherty, “Lifelogging: personal big data,” Foundations and trends in information retrieval, vol. 8, no. 1, pp. 1–125, 2014.
[19] A. Abbasi, S. Sarker, and R. H. Chiang, “Big data research in information systems: Toward an inclusive research agenda,” Journal of the Association for Information Systems, vol. 17, no. 2, p. 3, 2016.
[20] F. Liang, W. Yu, D. An, Q. Yang, X. Fu, and W. Zhao, “A survey on big data market: Pricing, trading and protection,” IEEE Access, vol. 6, pp. 15132–15154, 2018.
[21] H. Hu, Y. Wen, T.-S. Chua, and X. Li, “Toward scalable systems for big data analytics: A technology tutorial,” IEEE Access, vol. 2, pp. 652–687, 2014.