Exercise classification based on motion sensor

Chu-Hui Lee and Jyun-Han Liou
Dept. of Information Management,
Chaoyang University of Technology,
Wufeng District, Taichung, Taiwan, R.O.C.
chlee@cyut.edu.tw and s10114620@gm.cyut.edu.tw

Abstract— **Walking and running are two common types of exercises of humans. These types of exercises are related to the subjective consciousness caused by an individual's cognition of speed. However, everyone has his own cognition of speed. It is difficult to define these types based on a fixed standard. For example, walking includes fast walking and slow walking, while running includes sprinting and jogging. Different cognitions lead to different classifications. Therefore, this study aimed to find the key factors from various constructs and detected the motion data of the subjects through a sensor. The accelerations along the x-axis were considered as the motion profiles, including walking-pattern ones and running-patter ones. Then the differences between walking and running were found. According to the results of the experiments conducted by this study, using the peak values obtained from the motion profile analysis as the thresholds, the two types of exercises could be distinguished properly. By this way, it would be possible to analyze walking or running behaviors objectively.**

Keywords- walking-pattern; running-pattern; motion profile; motion sensor.

I. INTRODUCTION

Forward moving exercises are classified into walking and running. Yet, the line between these two types is very subjective. The difference between these two types is speed of moving. While running, one's strides may get into the airborne phase [1]. The frequency gets higher as his speed gets faster. In the meantime, the swinging pattern of his arms gradually changes from pendulum swinging to swinging with bent elbow, and finally to horizontal piston motion. However, there have been very few studies specifically describing the breaking point between these two types of exercises. The reason is that an exercise is an activity involving all muscles of the body. The same muscles may work differently with different types of moving [1, 2]. Therefore, the relative location of the sensor for detection and the corresponding method for classification were directly related to this issue. Under these premises, this study also had to consider a suitable place on the subjects to equip the sensors so that exercises of the subjects wouldn't be influenced. This way, the kinetic data collected would be closer to those of true daily exercises.

As described above, this study collected the data of swinging accelerations of upper limbs while walking and running through detectors. Based on the exercise data of the subjects, the changes of the peaks were observed. Then, a proper threshold of the peak number was found (the data collected from each of the subjects include the data of the two types of exercises, walking and running). The threshold was founded based on the data using the sensor helped to determine the type of an exercise. Future studies regarding classifying motion data into walking data and running data can reference the method to determine threshold proposed by this study in order to make efficient judgments for classification.

The remainder of this paper is structured as follows. The section 2 reviews the related literature. The section 3 presents the structure of the data collected from the sensors and the rules of recording the packets. The section 4 covers the ideas and methods used in the experiment. Section 5 describes the process of the experiment. And the final section is the conclusion of the paper.

II. PREVIOUS WORK

A lot of issues in these two types of exercises had been further explored. For example, Jeen-Shing Wang et al. [3] placed a sensor above the right ankle of each of the subjects to collect motion data while the subjects walked. These data was called the gait cycle data. A gait cycle was composed of four phases. Then the differences in walking among the four gait phases were observed. This study used the acceleration data to determine the types of walking and estimate the distances walked. Another scholar, Laura Guidetti, used EMG to collect the motion data of running and observed the changes in the relations of the seven focused muscles while running [2]. Another study of observing the changes in how muscles were used while exercising was the study by Marnix G.J. Gazendam and At L. Hof [1]. They observed the performances of the muscles based on the profiles of running exercises of various speeds collected through EMG. They used a specially made running machine. When the subjects ran on that machine, the data of running in various speeds and the information of strides were analyzed in real time. In addition, Thyagaraju Damarla [4] used four types of sensors inside a fixed range to determine the types of movements. The feature of this study was that the exercise patterns were determined based on the signal detections.

According to the previous studies, locations of equipments may directly influence the conditions of the data collected through them [1,2,3]. Thus, in this study, the sensors were placed on the dominant hands of the subjects near the wrists to collect data. It was hoped that the data collected from various subjects of different exercising habits through the fixed sensors could be close to the data of motions in real life. Then, based on these data, this study tried to identify the types of exercises.

III. The Structure of the Data Collected through the Sensors

The data of every exercise were stored in a .bin file in the SD card. In order to obtain the original data, this study exploited the x-IMU GUI software package developed by x-io Technologies Limited to unpack the .bin file and retrieved 8 .csv files, which were, ordered by their file names, CalBattAndTherm (C.B.A.T.), CalInertialAndMag (C.A.M.), Commands, DateTime, EulerAngles, Quaternion, Registers, and RotationMatrix. These files were stored in different .csv files according to the specific recording order of the sensor through packets of the same series. The recording order is shown in Fig. 1, and purposes of file are shown in Table I.

First, when the sensor started recording, the first packet, packet No.0, was stored in the file DateTime.csv. Then, packets No.1~No.112 were stored in Register.csv. And then packet No.113 once again recorded the DateTime information. After that, each packet followed the order in Fig. 1 to record the kinetic information in different .csv files, until the stop command was given. The sensor then recorded that commend in the final packet in the Commands.csv file.

As shown in Fig. 1, the starting packet was packet No.0. It was noted by the upper-right corner of the top block of the flow chart. First, 112 packets were stored in the Registers file (with the number by the upper-right corner being thenumber of the repeated loops). Then, a packet was stored in CalBattAndTherm, CalInertialAndMag, and S.R. Then, with every two packets stored in C.A.M., one packet was stored in

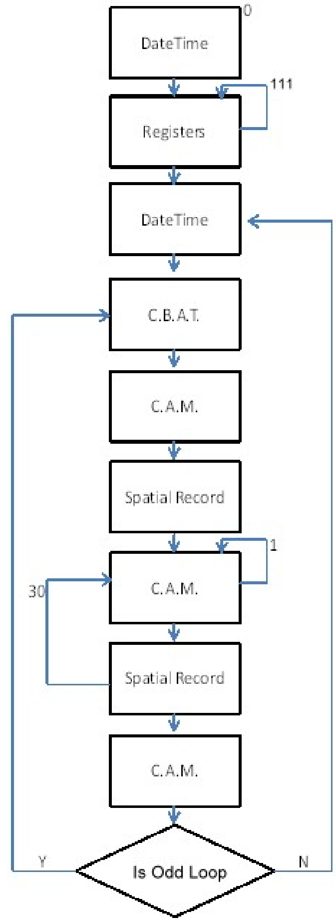

Fig. 1. The order of the packages being recorded, with the initial index value being 0 and the numbers by the corners being the number of repeats

TABLE I. THE NAMES AND CONTENT OF THE 8 FILES UNPACKED FROM THE BIN FILE

File name	Content recorded
CalBattAndTherm	Recording Battery voltage and Thermometer.
CalInertialAndMag	Recording gyroscope, accelerometer, and magnetometer data along the 3 axes.
Commands	Recording how a sensor record ends (e.g. sleep).
DateTime	Recording time duration of the corresponding kinetic data, including information of month, date, hour, minute, and second.
EulerAngles	Recording the Euler angles of the current packet.
Quaternion	Recording the quaternion of the current packet.
Registers	Recording the default settings of the kinetic data (i.e. DeviceID, ButtonMode, FirmwareVersionMajorNum, and GyroscopeSensitivity).
RotationMatrix	Recording the rotation matrix of the current packet.
The EulerAngles, Quaternion, and RotationMatrix files all recorded information regarding spatial changes. Therefore, in this study, they were considered of the same itemset. The three files together were called the Spatial Record (S.R.).	

S.R. After 30 repeats, one packet was stored in C.A.M. The next step depended on whether the number of the loop was odd. If yes, the same procedure was repeated; if not, a packet was stored in DatTime.csv before repeating the same procedure.

This study analyzed the kinetic data using CalInertial-AndMag along with the packets in DateTime. Based on the rules of recording, we found that, according to the information in DateTime, a total of 390 packets were stored within 1 second by the sensor. Among the 390 packets, 256 were stored in CalInertialAndMag.csv. We found that the intervals were all the same, a fixed amount of 256 packets, in the 7 sets of kinetic data collected.

IV. Proposed Method

In the experiment, the axes of the sensor were defined as: when the subject's both hands were dropped vertically with the palms facing inside, the horizontal motion was along the x-axis, the vertical motion was along the y-axis, and the in-out motion was along the z-axis. However, when a man walks, his arms usually swing naturally. When this study collected the motion data, the subjects were told to swing their upper limbs naturally, so that the data collected could be closer to the data of daily life motions. While walking, the subjects' arms swing forward and backward slowly. While running, the subjects usually clenched their fists or swung their forearms faster. Thus, our focus was on the changes along the forward-moving axis, the x-axis. The observation results based on the x-axis data were indeed consistent with the swinging speed changes of the subjects while doing the exercises.

Then, this study considered that the subjects' motions in the first and final 5 seconds might be related to pressing the record button before starting to run or pressing the stop button after running. The data from these periods of time were deleted, so that the data left were more likely related to the motion patterns we wanted to study. For the convenience of observation, we randomly selected motion data of continuous 40 seconds of each exercise as our observation samples. With each sample, we randomly selected 30% of the total packages as the initial package indexes. The sections we observed were the sampled initial package indexes plus 256. According to the acceleration-related data, the difference in the numbers of peaks of accelerations along the x-axis between the walking pattern and the running patter was significant (the definition of a peak in this study: a local maximum of a function).

When the subjects moved, they had the habit of swinging their arms. While swinging arms, their bodies swayed. Therefore, the range of every swing was different. However, the speed of swinging arms while walking was slower than that while running. This fact largely increased the situations false stillness while recording the packages. Stillness might, due to noise interferences or base line displacements of a certain degree, slightly influence the acceleration data, leading to an increasing number of peaks. In other words, in the process of upper limb swinging in a large range, false stillness might be caused by the sensor recording packages with small time intervals, leading to more chances of creating peaks along the x-axis. In cases of running, the speeds were higher, and the ranges of upper limb swinging were smaller, high frequencies of arm swinging were less likely to cause speed changes. Therefore, the acceleration changes were small, and relatively the numbers of peaks along the x-axis would be less. We will use this phenomenon to distinguish two types of exercises.

Fig. 2 The directions of the three axes and position of the sensor

Sensors with acceleration meters may easily be influenced by noises of high frequencies and base line drifting. Thus, acceleration in a condition of stillness may still have ups and downs, increasing the chances of peaks. Although people are used to different speeds and ranges of upper limb swinging, these two interferences can influence the data of both walking and running. In this study, the number of peaks in the pattern was used to determine whether that exercise was walking or

TABLE II. THE VOTES AND PRECISIONS BASED ON DIFFERENT THRESHOLDS. (WITH THE SAMPLE SIZE BEING 30% OF THE TOTAL SAMPLE SIZE)

Pattern Type	**Thresholds**											
	26		***27***		***28***		***29***		***30***		***31***	
1st group walking	0.0023	7	0.0065	20	0.0182	56	0.0277	85	0.0378	116	0.0439	135
	0.9977	3065	0.9935	3052	0.9818	3016	0.9723	2987	0.9622	2956	0.9561	2937
1st group running	0.9264	2846	0.9486	2914	0.9629	2958	0.9684	2975	0.9648	2964	0.9645	2963
	0.0736	226	0.0514	158	0.0371	114	0.0316	97	0.0352	108	0.0355	109
2nd group walking	0.0000	0	0.0000	0	0.0000	0	0.0000	0	0.0000	0	0.0000	0
	1.0000	3072	1.0000	3072	1.0000	3072	1.0000	3072	1.0000	3072	1.0000	3072
2nd group running	1.0000	3072	1.0000	3072	1.0000	3072	1.0000	3072	1.0000	3072	1.0000	3072
	0.0000	0	0.0000	0	0.0000	0	0.0000	0	0.0000	0	0.0000	0
3rd group walking	0.0000	0	0.0000	0	0.0000	0	0.0000	0	0.0000	0	0.0000	0
	1.0000	3072	1.0000	3072	1.0000	3072	1.0000	3072	1.0000	3072	1.0000	3072
3rd group running	0.9821	3017	0.9873	3033	0.9980	3066	0.9993	3070	1.0000	3072	1.0000	3072
	0.0179	55	0.0127	39	0.0020	6	0.0007	2	0.0000	0	0.0000	0
4th group walking	0.0000	0	0.0000	0	0.0000	0	0.0000	0	0.0000	0	0.0000	0
	1.0000	3072	1.0000	3072	1.0000	3072	1.0000	3072	1.0000	3072	1.0000	3072
4th group running	1.0000	3072	1.0000	3072	1.0000	3072	1.0000	3072	1.0000	3072	1.0000	3072
	0.0000	0	0.0000	0	0.0000	0	0.0000	0	0.0000	0	0.0000	0
5th group walking	0.0000	0	0.0000	0	0.0000	0	0.0000	0	0.0000	0	0.0000	0
	1.0000	3072	1.0000	3072	1.0000	3072	1.0000	3072	1.0000	3072	1.0000	3072
5th group running	1.0000	3072	1.0000	3072	1.0000	3072	1.0000	3072	1.0000	3072	1.0000	3072
	0.0000	0	0.0000	0	0.0000	0	0.0000	0	0.0000	0	0.0000	0
6th group walking	0.0039	12	0.0088	27	0.0195	60	0.0293	90	0.0397	122	0.0654	201
	0.9961	3060	0.9912	3045	0.9805	3012	0.9707	2982	0.9603	2950	0.9346	2871
6th group running	1.0000	3072	1.0000	3072	1.0000	3072	1.0000	3072	1.0000	3072	1.0000	3072
	0.0000	0	0.0000	0	0.0000	0	0.0000	0	0.0000	0	0.0000	0
7th group walking	0.0000	0	0.0000	0	0.0000	0	0.0000	0	0.0000	0	0.0000	0
	1.0000	3072	1.0000	3072	1.0000	3072	1.0000	3072	1.0000	3072	1.0000	3072
7th group running	0.9941	3054	0.9990	3069	1.0000	3072	1.0000	3072	1.0000	3072	1.0000	3072
	0.0059	18	0.0010	3	0.0000	0	0.0000	0	0.0000	0	0.0000	0
Average Accuracy	0.9926		0.9943		**0.9945**		0.9936		0.9920		0.9897	

running. Thus, the stable influences of the noises caused larger and more significant differences in the numbers of peaks between these two exercises.

V. EXPERIMENTS

In this study, a piece of Velcro was pasted on the back of the sensor to make sure the sensor wouldn't come off during movements and to avoid the strong shaking of the sensor caused by upper limb swinging which might influence the precisions of the data collected. In the process of collecting data, the data were stored in the SD card. This way, the subjects could swing their arms more freely, the spaces for movements of their limbs were larger, and the data obtained would be closer to those from our real daily exercises as we expected.

This study invited 7 people to participate in this experiment and help to provide the required kinetic data. 4 of the 7 participles were male, while 3 were female. Their limbs could all perform functions properly and they were all able to continuously walk and run on the 400-meter track. They only performed one type of exercise at a time so that the kinetic data collected were for one type of exercise, making observations easier. If there were obstacles or other people in front of the participants and their speeds might be influenced, what they did was to slowly move right or left and move back to the pre-defined track after passing the obstacles/people. The track used in this experiment for the kinetic data collection was the 6th track of the sports field in the Chaoyang University of Technology in Taichung, Taiwan.

We placed the sensor on the exerciser's right forearm, a place near the wrist, and fixed it using the Velcro on the back of it (the battery side), as shown in Fig. 2. When the arm was raised with the right palm facing the front, the 3 axes were shown on the other side of the palm. Moving left meant moving along the positive direction of the x-axis of the sensor. This study collected the data of the exercises of the 7 participants. Also, 3 people who did not participate in the experiment (2 males and 1 female) were invited to provide the data of walking and running exercises. On the next day, the same group of people was asked to provide the kinetic data of fast walking and fast running (200 meters) for more observations. After integrating the fast walking data and the fast running data, we found that: a higher speed led to a smaller number of peaks in the pattern (with the number still within the range). This fact proved that in frequent and fast swinging, the changes of acceleration were too large to create peaks due to noises. Thus, although there were noises in the patterns we analyzed, those noises did not influence the identification of the two types of exercises. On the contrary, they made the differences between the two types of exercises even more significant.

There are seven people in the experiment. Therefore, 7 sets of data of walking and running patterns were obtained. For each pattern, 30% of the total packets were sampled and used as the initial packets. All the initial packets had to be different. Then, we added 256 (this interval equaled to 1 second of the recorder) to each initial packet. This was defined as a sample set for the each sample. Based on these sample sets, we found that the numbers of peaks of the 7 different walking patterns were between 28 and 79, and those of the 7 different running patterns were between 6 and 29. Although the ranges were overlapped (28~29), the probability of this situation was very low. In order to extract the threshold between the two types exercises more precisely, we performed the threshold precision comparisons with the overlapped numbers (28 and 29). The experiment found that the precision was the highest

TABLE III. THE VOTES AND PRECISIONSS BASED ON DIFFERENT THRESHOLDS (WITH THE SAMPLE SIZE BEING 50% OF THE TOTAL SAMPLE SIZE).

pattern type	Thresholds											
	26		*27*		*28*		*29*		*30*		*31*	
1st group walking	0.0023	12	0.0053	27	0.0162	83	0.0279	143	0.0430	220	0.0471	241
	0.9977	5108	0.9947	5093	0.9838	5037	0.9721	4977	0.9570	4900	0.9529	4879
1st group running	0.9211	4716	0.9447	4837	0.9586	4908	0.9609	4920	0.9635	4933	0.9691	4962
	0.0789	404	0.0553	283	0.0414	212	0.0391	200	0.0365	187	0.0309	158
2nd group walking	0.0000	0	0.0000	0	0.0000	0	0.0000	0	0.0000	0	0.0000	0
	1.0000	5120	1.0000	5120	1.0000	5120	1.0000	5120	1.0000	5120	1.0000	5120
2nd group running	1.0000	5120	1.0000	5120	1.0000	5120	1.0000	5120	1.0000	5120	1.0000	5120
	0.0000	0	0.0000	0	0.0000	0	0.0000	0	0.0000	0	0.0000	0
3rd group walking	0.0000	0	0.0000	0	0.0000	0	0.0000	0	0.0000	0	0.0000	0
	1.0000	5120	1.0000	5120	1.0000	5120	1.0000	5120	1.0000	5120	1.0000	5120
3rd group running	0.9830	5033	0.9898	5068	0.9961	5100	1.0000	5120	1.0000	5120	1.0000	5120
	0.0170	87	0.0102	52	0.0039	20	0.0000	0	0.0000	0	0.0000	0
4th group walking	0.0000	0	0.0000	0	0.0000	0	0.0000	0	0.0000	0	0.0000	0
	1.0000	5120	1.0000	5120	1.0000	5120	1.0000	5120	1.0000	5120	1.0000	5120
4th group running	1.0000	5120	1.0000	5120	1.0000	5120	1.0000	5120	1.0000	5120	1.0000	5120
	0.0000	0	0.0000	0	0.0000	0	0.0000	0	0.0000	0	0.0000	0
5th group walking	0.0000	0	0.0000	0	0.0000	0	0.0000	0	0.0000	0	0.0000	0
	1.0000	5120	1.0000	5120	1.0000	5120	1.0000	5120	1.0000	5120	1.0000	5120
5th group running	1.0000	5120	1.0000	5120	1.0000	5120	1.0000	5120	1.0000	5120	1.0000	5120
	0.0000	0	0.0000	0	0.0000	0	0.0000	0	0.0000	0	0.0000	0
6th group walking	0.0037	19	0.0111	57	0.0172	88	0.0311	159	0.0545	279	0.0639	327
	0.9963	5101	0.9889	5063	0.9828	5032	0.9689	4961	0.9455	4841	0.9361	4793
6th group running	1.0000	5120	1.0000	5120	1.0000	5120	1.0000	5120	1.0000	5120	1.0000	5120
	0.0000	0	0.0000	0	0.0000	0	0.0000	0	0.0000	0	0.0000	0
7th group walking	0.0000	0	0.0000	0	0.0000	0	0.0000	0	0.0000	0	0.0000	0
	1.0000	5120	1.0000	5120	1.0000	5120	1.0000	5120	1.0000	5120	1.0000	5120
7th group running	0.9920	5079	0.9982	5111	1.0000	5120	1.0000	5120	1.0000	5120	1.0000	5120
	0.0080	41	0.0018	9	0.0000	0	0.0000	0	0.0000	0	0.0000	0
Average Accuracy	0.9921		0.9940		**0.9944**		0.9930		0.9904		0.9899	

when the threshold was set to 28, as shown in Table II. Then we could use this threshold to identify if an exercise was walking or running based on the corresponding data collected.

Table II shows the overlapped numbers of peaks of the two types of exercises, and the numbers next to the two ends were used as candidate values to perform classification of exercises. The values on the top of the table were the thresholds, from 26 to 31. The left side lists the types of exercises and the 7 groups. Each group had data of walking and running. Take the first group's walking data with the threshold of 26 for example, the number in the top-right grid of the small 2x2 cross table is 7, meaning among the 3072 different samples (randomly selected, 30% of the packets during the 40 seconds) 7 were identified as running. The number in the bottom-right grid, 3065, means that among 3065 samples were identified as walking. The number in the top-left grid is the estimated probability for running, and that in the bottom-left grid is the estimated probability for walking. The numbers of peaks in a sample were used to estimate whether the exercise performed during the 40 second was walking. And the result was compared with the real exercise performed. As the table shows, the precisions of classifying the 7 sets of exercise patterns (14 patterns) were very high, especially when the threshold used was 28. Yet, we also found that the precisions corresponding to the thresholds of 27 and 28 were very close with the 40-second samples. Thus, we increased the percentage of sampling for the 40-second samples from 30% to 50%. The results are summarized in Table III. The results showed that precision corresponding to the threshold of 28 was still the highest. Though the number of packages sampled was increased, the precision corresponding to the threshold of 28 remained stable compared with those corresponding to other thresholds. Therefore, by using 28 as the threshold of numbers of peaks, we were able to achieve a higher precision for classifying unknown patterns into the two types of exercises.

VI. CONCLUSION

This study tried to find the differences between walking and running using the data of sensor. Through the changes of accelerations, this study proposed a classification method different from the general one based on subjective perception of speed. Using the threshold obtained by this study, computers could precisely determine whether an unknown pattern was a walking pattern or a running one based on the corresponding acceleration data. Thus, for future studies regarding daily movements which require separating walking from running, the method proposed by this study would be a very effective choice for classifying movements with very high precision. In addition, with more training and testing samples, the obtained threshold for classification would be more and more precise.

REFERENCES

[1] Marnix G.J. Gazendam and At L. Hof, "Averaged EMG profiles in jogging and running at different speeds," Gait & Posture, Vol.5 Issue 4, pp. 604-614, 2007.

[2] Laura Guidetti, Gianfranco Rivellini, and Francesco Figura, "EMG Patterns During Running: Intra- and Inter-individual Variability," Journal of Electromyography and Kinesiology, Vol. 6, NO. 1, pp. 37-48, 1996.

[3] Jeen-Shing Wang, Che-Wei Lin, Ya-Ting C. Yang, and Yu-Jen Ho, "Walking Pattern Classification and Walking Distance Estimation Algorithms Using Gait Phase Information," IEEE Transactions on Biomedical Engineering, Vol. 59, No. 10, 2012.

[4] Thyagaraju Damarla and James Sabatier, "Classification of people walking and jogging/running using multi-modal sensor signatures," Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense X, Vol. 8019, 2011.