This is a repository copy of *Multiple density maps information fusion for effectively assessing intensity pattern of lifelogging physical activity*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/150625/

Version: Accepted Version

Multiple Density Maps Information Fusion for Effectively Assessing Intensity Pattern of Lifelogging Physical Activity

Jun Qi*, Po Yang*, Martin Hanneghan, Stephen Tang

Department of Computer Science, Liverpool John Moores University, Liverpool, UK, L3 3AF

ARTICLE INFO

ABSTRACT

Physical activity (PA) measurement is a crucial task in healthcare technology aimed at monitoring the progression and treatment of many chronic diseases. Traditional lifelogging PA measures require relatively high cost and can only be conducted in controlled or semi-controlled environments, though they exhibit remarkable precision of PA monitoring outcomes. Recent advancement of commercial wearable devices and smartphones for recording one's lifelogging PA has popularized data capture in uncontrolled environments. However, due to diverse life patterns and heterogeneity of connected devices as well as the PA recognition accuracy, lifelogging PA data measured by wearable devices and mobile phones contains much uncertainty thereby limiting their adoption for healthcare studies. To improve the feasibility of PA tracking datasets from commercial wearable/mobile devices, this paper proposes a lifelogging PA intensity pattern decision making approach for lifelong PA measures. The method is to firstly remove some irregular uncertainties (IU) via an Ellipse fitting model, and then construct a series of monthly based hour-day density map images for representing PA intensity patterns with regular uncertainties (RU) on each month. Finally it explores Dempster-Shafer theory of evidence fusing information from these density map images for generating a decision making model of a final personal lifelogging PA intensity pattern. The approach has significantly reduced the uncertainties and incompleteness of datasets from third party devices. Two case studies on a mobile personalized healthcare platform MHA [1] connecting the mobile app *Moves* are carried out. The results indicate that the proposed approach can improve effectiveness of PA tracking devices or apps for various types of people who frequently use them as a healthcare indicator.

1. Introduction

Physical inactivity is a well-known severe health risk leading to a variety of chronic and obesity related diseases [2]–[6] in modern society. As an intuitive reflection of their underlying pathophysiology, continuous physical activity (PA) measurement in daily life is increasingly crucial to these patients for designing specific rehabilitation programs to promote an active lifestyle. Thus, the accuracy and stability of access to PA related information is of significant interest to the research community.

Traditionally, PA measurement recognizes the type, duration, and intensity of a broad range of activities and quantifies energy expenditure. For the purpose of assuring accuracy in accessing PA associated energy expenditure, typical PA measurement solutions require subjects to wear special devices in lab or clinical environments to acquire sensory signals, and then analyze them with advanced machine learning algorithms for recognizing different types of PA [7]–[9]. While these solutions deliver relatively high accuracy of PA measurement they are less feasible for long-term measurement in free-living conditions, often termed as 'life-logging' PA measures. The reason is because analysis of raw sensor data consumes too much energy on the portable/wearable device. Battery and storage capacity are key limiting factors when assessing one's PA pattern in a free-living environment [10], [11].

As PA measurement devices are becoming more affordable, lightweight and portable, the prevalence of commercial wearable devices and mobile apps with processed outcomes in smart healthcare fields [12]–

[16] make monitoring and tracking lifelogging PA associated information possible (such as walking, running, intensity, duration, etc.) to objectively ensure consecutive care for users. Particularly, the intensity of lifelogging PA data observed by such inertial sensors is often categorized into five levels: sedentary, light, moderate, vigorous and high intensity based on the metabolic equivalents (METs) cut-offs [17], which has been broadly adopted as a standard of PA levels for achieving healthcare life styles. However, the classification simply offers a generally instantaneous measure that exhibits deficiency of accumulated evaluation and assessment for lifelogging personal PA intensity patterns thereby restricting its usefulness.

Research as to how we can better take advantage of these scattered and heterogeneous data has become a critical issue when used in long-term observation for healthcare prevention and research purposes. The difficulty is that a certain amount of indicative PA data collected from existing inertial sensor-based wearable devices and mobile phones contains a variety of uncertainties. For instance, PA intensity observed by the app Moves [13] indicates that it often turns itself off to conserve energy for mobile devices. The wrist band Withings [14] produces erroneous PA recognition results occasionally since PA related human life patterns in free-living conditions are dynamic, diverse and noise-sensitive. Such results have a negative impact on the energy expenditure and PA intensity evaluations. The uncertainties are quite common but can prove challenging to eliminate. Nonetheless, based on existing works, to our best knowledge, almost none contributes such wrapped and scattered datasets from commercial devices to lifelogging PA intensity analysis for healthcare support. This work is the first attempt to address these uncertainties and, one step further, to improve the efficiency of low-cost wearable and mobile devices for one's PA intensity pattern in a long range effort.

The remainder of this paper is structured as follows. Section II presents the literature review of related work. Section 3 describes the proposed PA intensity pattern decision making approach. Section 4 reports two case studies for evaluating the proposed method in the MHA platform [1], [1]. Finally, the conclusions and future work are presented in Section 5.

2. Related work

Lifelogging, refers to the process of capturing one's entire life using digital devices for health and wellness, e.g. medical intervention or physical activity recommendation. In early attempts, lifelogging PA monitoring was preliminarily surveilled by image capturing via an external camera [19]–[21]. However, this approach could be deemed an invasion of privacy for the general public other than the subject and this has made it a less popular mechanism. Modern technology extends the definition of lifelogging into broader ranges. Wearable devices nowadays have been widely utilized to continuously track one's PA such as wearable camera, wristwatch and mobile phone [22]. The SenseCam wearable camera, a form of visual lifelogger, worn over one's neck, has been explored as an everyday activity data recorder in [23]–[25] by the means of analysis of a series of captured photos. Compared with traditional indoor/outdoor cameras, personal privacy of this wearable camera has a higher level of protection. Although there is general consensus that the device is appropriate for healthcare purpose, in most of cases, its cost is somewhat prohibitive for patients or researchers in a controlled lab environments. Recording and storage of a high volume of lifelogging pictures is also a big challenge for SenseCam.

In recent years, low-cost costumer wearable PA trackers with embedded inertial sensors are generating increasing public attention. Popular products, such as *Fitbit Flex* [26], *Nike+ Fuelband* [27], *Endomondo* [28] etc., are wristband devices that record PA information (e.g. steps, distance, and calories burnt) and other physiological information (e.g. heartbeat rate). Some third party Application Programming Interfaces (or APIs) of wearable devices have provided the functions to assess the intensity of PA walking speed. For instance, *Fitbit* [26] classifies the intensity of daily activities into very active, moderately active, lightly active and sedentary. Mobile apps, such as *Moves* [29] is based on smartphone 3D accelerometer data and GPS information which allows tracking the user's movements including location, distance and speed. *Moves* records a series of walking segments containing duration, distance and speed.

Evidently, customer PA monitors have addressed some practical issues such as storage, battery life and cost, especially mobile apps which are often free. Nevertheless, PA recognition results offered by mobile devices are widely divergent as a result of different places being carried by different users such as pocket or handbags [30], [31]. Furthermore, the diverse life pattern of an individual person may cause huge indeterminateness, as they perform PA in varying ways owning to age, gender, weight, etc. Hence, a specific PA tracking model that fits one group of user may not fit another one [32]. In addition to that, some applications often automatically switch off themselves for energy efficiency which has contributed to missing data. In general, the uncertainties of lifelogging PA from customer devices here is divided into two types as our previous work investigated [33]:

Irregular Uncertainty (IU): randomly and accidently occurs in lifelogging PA data. The causes of these uncertainties include device malfunctions or faults, breakdown of a third party server, misuse of devices or sudden change of personal circumstance. The occurrence of IU will appreciably impact the efficiency and accuracy of assessing personal health.

Regular Uncertainty (RU): frequently and persistently occurs in lifelogging PA data. The causes resulting in these uncertainties are mainly from some regular influencing issues, like intrinsic sensors' errors, differentiation of personal physical fitness and changes of environment. The occurrence of regular uncertainty in physical activity data is inevitable so that it is impossible to completely eliminate these uncertainties.

Accordingly, these uncontrolled conditions are the key issues that cause relatively low accuracy of wearable device or mobile phone data logger compared with traditional non-naturalistic experiments. The encapsulated datasets, consequently, tend to be scattered, erroneous and unserviceable for long range healthcare studies.

To address the aforementioned challenge, our work attempts to take on these uncertainties for lifelogging PA intensity observation. To enhance device performance, we first use an ellipse fitting model for reducing IU of life-logging PA measures in an Internet of Things (IoT) environment. Secondly, hour-day density map images are constructed to represent the RU pattern on each month. We then propose Dempster-Shafer theory of an evidence based lattice model applying to these density map images fusion for determining a robust lifelogging PA intensity pattern. Unlike the five categories based on MET, in this work, only two basic standards of intensity (active and sedentary) are considered for long term observation. We believe that our work will help bring attention to the opportunities available for using datasets from commercial wearable

devices and mobile phones for the purpose of healthcare studies and will stimulate additional work in this area.

The ellipse fitting algorithm is represented as a circular form via projection to an image plane, which is often used to remove scattered or noisy data samples through setting points to the best fit or geometric fit [34], [35]. In comparison with a curve fitting function such as Gaussian fitting [36] or smoothing fitting [37], the ellipse fitting method is more suitable for the aggregative samples that belong to elliptical conic and excluding non-elliptical data [38]. Furthermore, the method has low computation cost and is easy to implement.

The density map is a visualization technique that uses different colors for presenting different activity levels in the image. In the work [39], an activity density map based visualization method is proposed for analyzing passive infrared motion sensor data for monitoring the elderly. Due to the similarity of motion sensor data and lifelogging inertial sensor data, density maps are used in our work to generate uncertainty patterns of lifelogging PA intensity measures.

The Dempster-Shafer Theory (DST) of evidence was proposed for the first time by Dempster in 1968 and improved by his student Shafer in 1971. The most important facet of this theory is the capability to deal with uncertainties from incomplete pieces of evidence in a way that multiple criteria facilitates an information fusion procedure, and then make a better decision by reducing this uncertainty and imprecision. The application of evidence theory has been widely regarded in various areas to effectively improve the overall performance by fused sources. For example, multi-sensor fusion based on DST has been applied to engine fault diagnosis [40] and activity recognition [41]. Similarly, multiple image fusion has been used for image restoration in [42], [43]. The applicability of this theory is relatively useful to our work and is used to create the dataset with RU from low-cost wearable and mobile devices.

3. Proposed approach

In this section, we propose a lifelogging PA intensity pattern decision making approach to assess the feasibility of specific mobile devices. The procedure is presented in Fig.1. Since lifelogging PA measures by mobile devices contain an amount of uncertainties including IU and RU as mentioned earlier, an Ellipse fitting model will firstly be used to preprocess the data and remove these IU through the projection of distribution of IU by defining a walking speed related score named as Daily Activity in Physical Space (DAPS). Secondly, a sequence of hour-day PA grey-levelled density maps will be manually constructed on a monthly basis. The features then can be extracted from the map for measures of the RU. If the result shows one's unstable states monthly, the histogram distribution of the map will be acquired, with capability of the pixel classification for grey levels corresponding to the images. In the end, the DST-based lattice model will be utilized to reduce the RU and thus produce a lifelogging PA intensity pattern.

3.1. Ellipse fitting model handle with irregular uncertainties (IU)

While IU occurs accidently and is hardly quantified by impacting factors, their occurrence frequency is relatively low over time. A statistical analysis in historical data can detect threshold parameters to filter them. Daily physical activity is mainly measured as *daily steps* (S_d), *daily walking distance* D_{dw} and *daily average walking speed* (V_{daw}). It is believed that the majority of daily steps and daily average walking speed have to be in a specific range. two threshold parameters which are T_S and
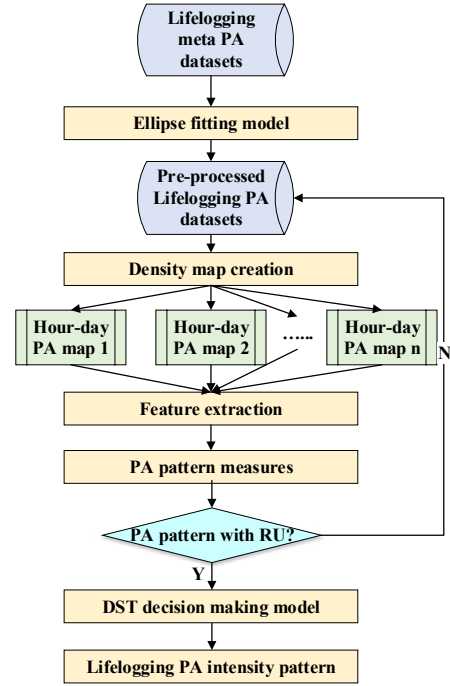

Fig.1. Flowchart of proposed approach

T_v are defined to filter the IU regarding a probabilistic distribution.

A benchmark is needed to represent a person's physical fitness from completed data sources. Here a walking speed related score is defined to represent a person's physical fitness, named as Daily Activity in Physical Space (DAPS). This score takes inspiration from the work [44] in which Herrmann et. al. proposes a *Movement* and *Activity* in Physical Space score as a functional outcome measurement for encompassing both PA and environmental interaction.
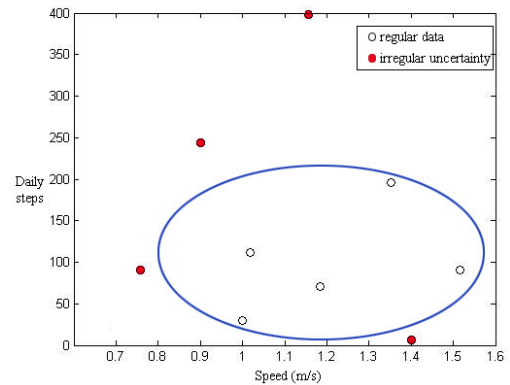

Fig.2. Distribution of Irregular Uncertainty (IU)

Here, we classify the intensity of daily PA into *N* levels in terms of the ranges of walking speeds ($V_1, V_2 \cdots V_n$). The DAPS formula is created by summing these different level walking speeds:

$$DAPS = \sum_1^N V_t \qquad (1)$$

Using the data of *DAPS* and *Daily Steps*, we can calculate V_{daw}, and plot S_d and V_{daw} in with Ellipse fitting model presented as Fig.2.

In order to enclose points $P:\{P_1, P_2, \cdots, P_n\}$ in the 2D plane, we use an Ellipse ε to cover all the regular points P_n. The Ellipse with centre point (m, n) and semiaxes a and b are defined in equation (2):

$$\frac{(x-m)^2}{a^2} + \frac{(x-n)^2}{b^2} = 1 \tag{2}$$

Where:

h : Average daily walking speed

k : Average daily walking steps

m : Error range of average daily walking speed

n : Error range of average daily walking steps

In Fig.2, the red dots fall out of the Ellipse circle represents the IU, the hollow dots are the regular physical activity data covered by the Ellipse algorithm. A noticeable issue here is that we only consider the lower limits of walking steps and the upper limits of walking speeds as threshold parameters. On some days users might walk distinctly more steps than usually, while the other days might be more sedentary. The threshold parameters are represented in equation (3)

$$T_y = m + a;$$
$$T_s = n - b \tag{3}$$

Thus, the strategy for removing IU follows following steps:

- To calculate the parameters S_d, D_{dw}, V_{daw} with raw data.
- To plot the data of S_d, D_{dw}, V_{daw} and calculate the value of T_s and T_y with an Ellipse filtering equation to cover data with a confidence interval of 95%-98%.
- To use T_s and T_y for removal of IU.
- To iterate the above process in another time period with updated raw data.

3.2. Determine Regular uncertainties (RU)

3.2.1 Density map visualization

We use a density map proposed in [39] to construct longitudinal PA monitoring data distribution. An example of a density map derived from the mobile device is presented as Fig. 3. Here we only select steps (e.g. walking, jogging and running) as the validation standard for the individual's PA density and represent this on the density map, since they are highest frequency PA recorded by the mobile trackers. Such PA density is shown in Eq.(4),

$$Intensity = \frac{Steps}{maxSteps} \times pixel \tag{4}$$

Where steps refer to the total steps value of each hour of a day; maxSteps refer to the maximum steps value of the month; image pixel ranges from 0 to 255, which means the image is in the grey level. In the density map, the vertical axis indicates 24 hours a day from 0:00 to 23:00, while horizontal axis indicates days of each month from 1st to 30th or 31st. PA intensity ranges from light colour to dark colour, where dark colour denotes very high intensity and vice versa. The white part in the map represents sedentary and uncertain patterns.

3.2.2 Feature extraction

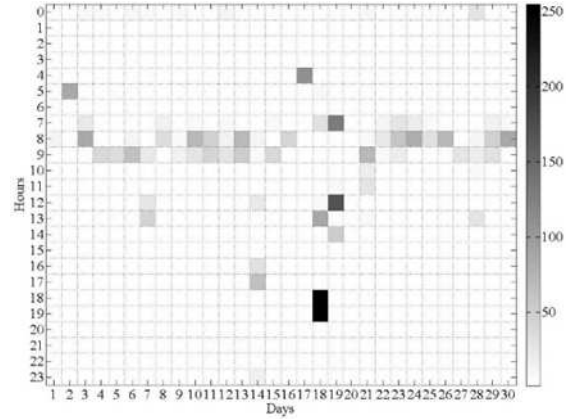A grey-levelled histogram is able to explicitly reveal accociated grey

Fig.3. An example of PA density map from the mobile device's dataset
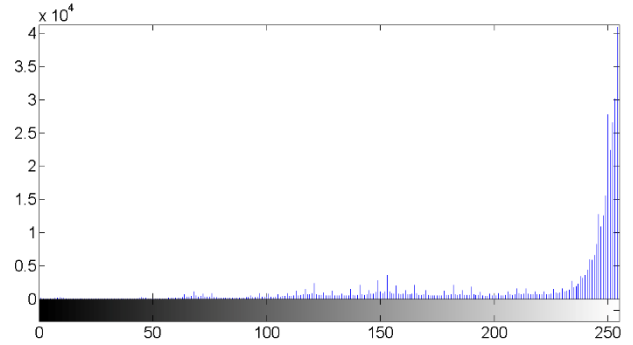

Fig. 4. An example of histogram distribution for the density map

levels of the image, and to cover abundant statistical distribution information that can well reflect the percentage and frequency of appearance for pixels in a image. Fig. 4 is a typical example of the histogram distribution of PA density map recorded by mobile device. We selected six features extracted from it: mean, variance, skewness, kurtosis, energy and entropy. The fomula are shown as the following:

$$Mean(x_1 \cdots x_N) = \frac{1}{N} \sum_{j=1}^{N} x_j \tag{5}$$

$$Var(x_1 \cdots x_N) = \frac{1}{N-1} \sum_{j=1}^{N} (x_j - \bar{x})^2 \tag{6}$$

$$Skew(x_1 \cdots x_N) = \frac{1}{N} \sum_{j=1}^{N} \left[\frac{x_j - \bar{x}}{\sigma}\right]^3 \tag{7}$$

$$Kurt(x_1 \cdots x_N) = \left\{\frac{1}{N} \sum_{j=1}^{N} \left[\frac{x_j - \bar{x}}{\sigma}\right]^4\right\} - 3 \tag{8}$$

$$Eng = \sum_{j=1}^{N} \left(\frac{x_j}{N}\right)^2 \tag{9}$$

$$Ent = \sum_{j=1}^{N} \frac{x_j}{N} \log_2 \frac{x_i}{N} \tag{10}$$

The mean, variance and skewness present the degree of average, discreted and asymmetrical distribution in a grayscale histogram, respectively. Kurtosis measures the realtive peakness or flatness of the distirbution to a normal distribution. Energy and entropy represent the average degree of gray distribution.

3.2.3 Distance measure

In order to assess the subject's PA intensity state, Euclidean distance is adopted to measure the dissimilarity among density maps, as Eq. (11).

$$distance_{i,j} = \sqrt{\sum_{i=2, j=1}^{k} (x_{i,j} - x_{i-1,j})^2} \tag{11}$$

Where x represnts the vectors of extracted features. $x = [x_1 \dots x_k]$. The smaller the distance, the similar the two map images, and vice versa.
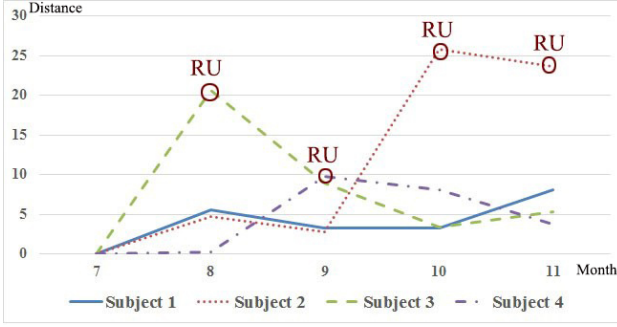

Fig.5. Euclidean distance measures dissimiliary of density map with RU for four subjects of five months dataset from the mobile device

The validation dataset features four randomly selected healthy individuals using the mobile devices for 5 months. The subjects investigated are staff and reseach students at a university. They are working 6 to 8 hours in front of a computer every work day, whilst intensity and time for workout are relatively stable each month, thus distance of density maps should be highly similar with small fluctuations. However, as we can see in Fig. 5, only the first subject present a normal PA pattern shown as the solid blue line, while others suffer from large distance changes, which demonstrates that there are major RU of PA types and durations recorded by the mobile devices.

3.3. DST decision making model

Due to the uncertainties and incompleteness of the existing dataset, the intensity and quantities emergent on the density map are considerably unstable. Using DST therefore, different prior knowledge from different density maps will arrive at a degree of belief that takes into consideration all the available evidence as well as to reduce the RU. Moreover, based on the concept of DST, we construct a three layer hierarchical lattice model to conduct a decision making process for the subject's long-term PA intensity pattern.

In Fig. 6, the three layers within the model are a visualization layer, data fusion layer and decision making layer. In the visualization layer, PA density, duration and category of each subject's wearable/mobile device's datasets are produced as density maps. DPA and SPA sets are also extracted as individual evidence presented as dash line round node. The data fusion layer composes of combinations of the information presented as a dash line oval node in Fig.6. The decision making layer provides the final outcomes of the subject's lifelogging PA intensity decision from prior knowledge revealed on density map and mass values from combination rules.

3.3.1 Dempster-Shafer theory of evidence (DST)

DST allows for direct representation and reasoning of uncertainties in a way that fuses accumulative evidence and changing prior knowledge in the presence of new evidence, where the input can be an imprecise or incomplete set or an interval, while the output is also a set or an interval. DST assumes that there are all possible elements denoted as Θ. A basic probability assignment (BPA) is represented by mass function $2^\Theta \subseteq [0,1]$, where $m(\emptyset) = 0, \sum_{A \subset U} m(A) = 1$. Set mass functions m_1, m_2 are the BPA in Θ.
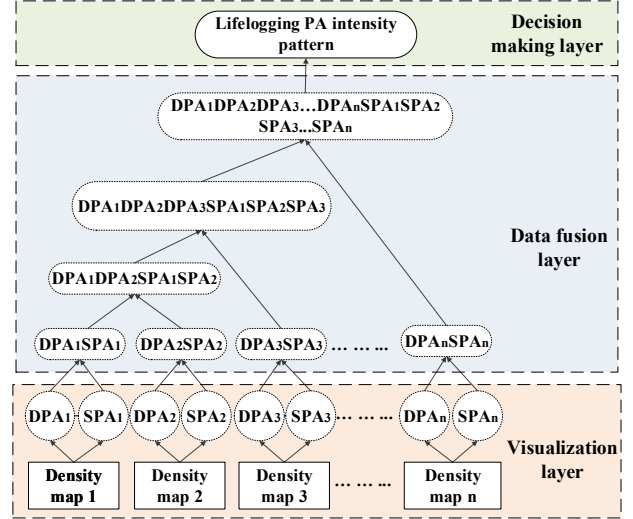

Fig.6. Lattice model of DST for personal lifelogging PA pattern decision (DPA-Dynamic Physical Activity, SPA- Sedentary Physical Acitivity)
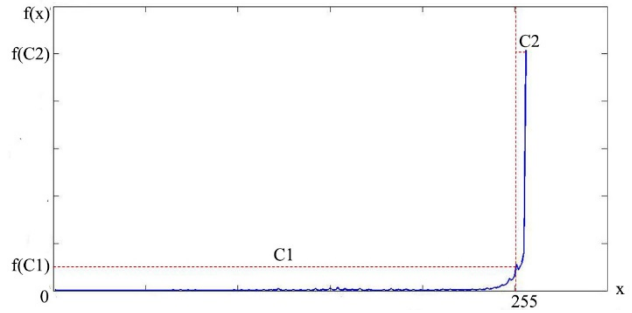

Fig.7 Histogram distribution with the threshold setting

Upper and lower bounds of an interval are defined as belief function and plausibility function, expressed as *Bel()* and *Pls()*, denotes as:

$$Bel(A) = \sum_{B \subset A} m(B) \tag{12}$$
$$Pls(A) = \sum_{B \cap A \neq 0} m(B) \tag{13}$$

Where $A \in 2^\Theta$ and $B \in 2^\Theta$, $0 \leq Bel(A) \leq Pls(A) \leq 1$.

The plausibility is defined as the degree of objection or no objection towards A, which denotes as:

$$Pls(A) = 1 - Bel(A) \tag{14}$$

Therefore, the degree of belief is in the interval $[Bel(A), Pls(A)]$, in terms of the total interval $[0,1]$.

Dempster's combination rule [45] is a way to aggregate information from uncertain context whether data is a single source or multiple sources. The formula of fusion independent sources is defined as:

$$m(C) = \begin{cases} \frac{\sum_{A \cap B = C} m_1(A) m_2(B)}{1-K}, & \forall C \subseteq \Theta, C \neq \emptyset \\ 0, & C = \emptyset \end{cases} \tag{15}$$

Where $m(C)$ determines the final mass of $m_1(A)$ and $m_2(B)$, and $K = \sum_{A \cap B = \emptyset} m_1(A) m_2(B)$. $1 - K$ is a normalization factor which is constant for all subsets whilst has no impact on the behaviour of the operator. $K = 1$ represents m_1 and m_2 is completely contradictory, so combination cannot be executed. When $K = 0$, it refers to that two

evidence are completely compatible. When $0 < K < 1$, it represents that two evidence are partly compatible.

Furthermore, multiple mass functions also can be combined to obtain the final result in terms of the orthogonal sum defined in Dempster's combination rule, shown as Eq.(16).

$$m_1 \oplus m_2 \oplus \dots \oplus m_n = (((m_1 \oplus m_2) \oplus \dots) \oplus m_n) \qquad (16)$$

3.3.2 Mass function definition in density map

A crucial issue in DST inference is to define a mass function for each evidence set based on the sourced information, which are the density maps from the visualization layer in our work. From the grey levelled histogram distribution of the image, as shown in Fig. 7, two classifications can be clearly determined based on the threshold $f(C_1)$ presented as the red vertical dash lines in terms of the grey levels: in the interval [0,254] are grey pixel numbers represented the PA data acquired from wearable/mobile device (e.g., steps, distance and duration), whilst the value of sedentary and unknown PA are presented as the white pixel equal to 255. As there are a majority of blank uncertain information on the map, in order to better express its uncertainties, we assume that two stationary PA (working and sleeping) duration of an individual is prior knowable. Thus the categories can be finally defined as three types:

- *DPA*: dynamic physical activity, e.g. walking, jogging or running, shown as grey and black part on the density map and denoted as a set $\{DPA\}$.
- *SPA*: sedentary physical activity, e.g. sitting, standing or lying, shown as white part on the density map and denoted as a set $\{SPA\}$.
- *Ambiguity*: not sure DPA or SPA, shown as white part on the density map and denoted as a set $\{DPA, SPA\}$.

According to Fig. 7, therefore, the initial mass function based on the pixel numbers is defined as:

$$m_{DPA}^i = \frac{f(C_1^i)}{f(x^i)} \qquad (17)$$

$$m_{SPA}^i = \frac{f(C_2^i) - f(C_1^i)}{f(x^i)} \times \frac{Duration(sleep+work)}{TotalDuration} \qquad (18)$$

$$m_{DPA,SPA}^i = 1 - m_{DPA}^i - m_{SPA}^i \qquad (19)$$

Where i is the density map numbers. $f(C_1)$ and $f(C_2)$ are the grey levelled pixel numbers of $\{DPA\}$ and $\{SPA\}$ on the density map, respectively.

3.3.3 Evidence combination of density maps

In the evidence fusion layer, the combination m of two maps of information is acquired from Dempster's rule of combination presented in table 1 according to the Eq. (15). Some information may support the same conclusion, while others may conflict with each other. As such, different evidence will be computed a way to obtain the final mass function m. The intervals of belief and plausibility of sets $\{DPA\}$ and $\{SPA\}$ will also be ultimately achieved in the decision making layer based on the Eq. (16). Meanwhile, fusion of multiple density maps is possible to reduce the RU of an individual's PA records.

3.3.4 Decision making rules

Once all the available density maps have been fused, the decision

Table 1

Dempster's rule of combination

m_2	m_1		
	$\{DPA\}$	$\{SPA\}$	$\{DPA, SPA\}$
$\{DPA\}$	$\{DPA\}$	$\emptyset$	$\{DPA\}$
$\{SPA\}$	$\emptyset$	$\{SPA\}$	$\{SPA\}$
$\{DPA, SPA\}$	$\{DPA\}$	$\{SPA\}$	$\{DPA, SPA\}$

belief intervals can be consequently achieved by provided evidence, which are denoted as:

$$EI(DPA) = [Bel(DPA), 1 - Bel(SPA)];$$
$$EI(SPA) = [Bel(SPA), 1 - Bel(DPA)] \qquad (20)$$

Generally speaking, the more evidence fusion, the smaller value of RU. As shown in Fig. 8, the lifelogging PA intensity pattern decision making strategy is based on whether the value of RU is eliminated. With consideration of the two conditions, the lifelogging PA style takes advantage of following rules:

- When $RU = 0$

$$\begin{cases} If\ Bel(DPA) > Bel(SPA), \\ then\ one's\ lifelogging\ PA\ Intensity\ pattern\ is\ active; \\ If\ Bel(DPA) < Bel(SPA), \\ then\ one's\ lifelogging\ PA\ Intensity\ pattern\ is\ sedentary. \end{cases} \qquad (21)$$

- When $RU \neq 0$

$$\begin{cases} If\ EI(DPA) > EI(SPA), \\ then\ one's\ lifelogging\ PA\ Intensity\ pattern\ is\ active; \\ If\ EI(DPA) < EI(SPA), \\ then\ one's\ lifelogging\ PA\ Intensity\ pattern\ is\ sedentary. \end{cases} \qquad (22)$$
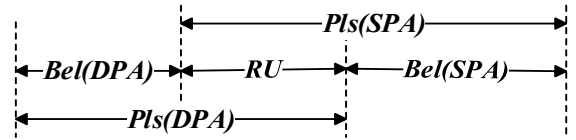

Fig.8. Intervals of Belief (Bel) and Plausibility (Pls) of *DPA* and *SPA*

When RU has been removed, the precise belief value would be derived, so the decision is taken through the comparison between $Bel(DPA)$ and $Bel(SPA)$. When RU has been reduced but still exists, the result is determined by the comparison of values of interval. The higher confidence is assigned to the hypothesis with maximum belief value in the interval. For instance, if the belief interval of the set $\{DPA\}$ is [0.3, 0.6], and the belief interval of the set $\{SPA\}$ is [0.5, 0.9], the decision will be assigned to the set $\{SPA\}$ with the maximum value for both lower limit and higher limit.

4. Case study

In order to better illustrate how the proposed approach using wearable or mobile devices is applied in the assessment of lifelogging PA intensity patterns, two case studies based on the two hypothesises defined earlier are introduced next. In the first case study, we use multiple density maps fusion to illustrate the long term PA intensity decision making procedure under the condition of removal of RU, represented with a certain value

Bel. In the second example, two months density maps with confliting envidence are fused to state the other hypothesis that the RU still exsits, represented as intervals *EI*. Both subjects used the mobile app *Moves* [5] and its datasets are collected from the mobile personalized healthcare platform MHA [6].

4.1. Case study 1

4.1.1 Handling IU

The subject is a research student at the university, female, aged 29. The proposed Ellipse fitting model operates on the individual's PA distribution for pre-processing. Fig. 9 shows the subject's daily steps and speed. The confidence value of Ellipse fitting is 0.98 for the individual, which means that 98% of samples fall inside the defined region based on a Gaussian distribution.
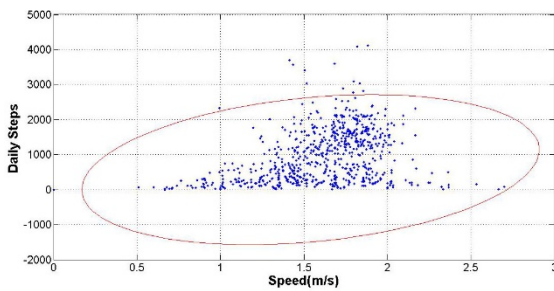

Fig.9. Case study 1: Ellipse fitting model for IU removal of the subject's *Moves* dataset (c= 0.98)

In Fig. 9, 98% of data falling within the oval are below 3000 steps per hour. The speed or walking intensity is between 0.5 to 2.7m/s for this individual.

4.1.2 RU confirmation

Fig. 10 (a) shows the individual's density map of PA distribution of *Moves* data record from July to Oct., 2015. Although the features of mean, variance, energy and entropy from images are quite close, the Euclidean distance measures that significant dissimilarity among Sept. and Oct, as shown in Fig. 10 (b). This is due to the features of large kurtosis difference. Density maps of Sept. and Oct. display more intensive distribution of the grey part, which means they have higher kurtosis, tends to have distinct peaks near the mean. This also implies that the mobile phone recorded more DPA data in most hour cells during this period than Jul. and Aug. Nevertheless, the subject's healthcare condition and activity frequency were stable within this period. Furthermore, DPA datasets are relatively complete each month, since they are emergent almost every day on the images. Hence we can infer that the subject constantly kept her phone and the app on when performing DPA, the RU may come from the app's internal error such as inaccurate activity recognition, the app's mechanism (e.g., automatically off for energy saving) or data transmission failure, etc.
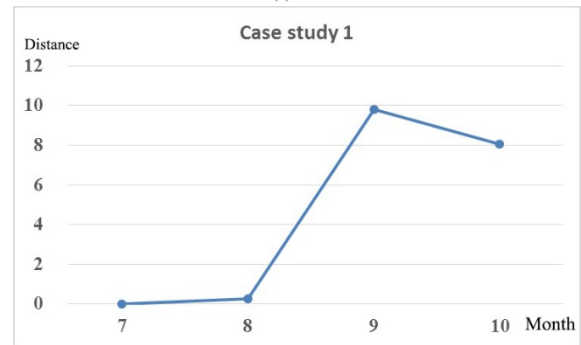
4.1.3 Lifelogging PA intensity pattern inference

To analyse the individual's life-long PA intensity pattern, the DST is utilized in this situation. The steps are shown as below.
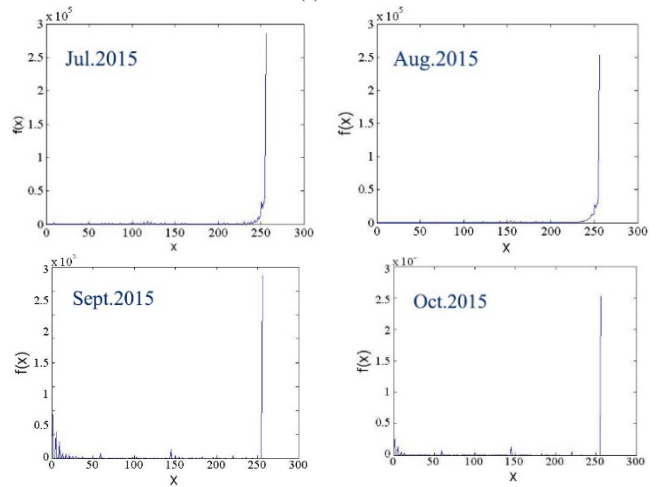*(1) Define the initial mass functions*
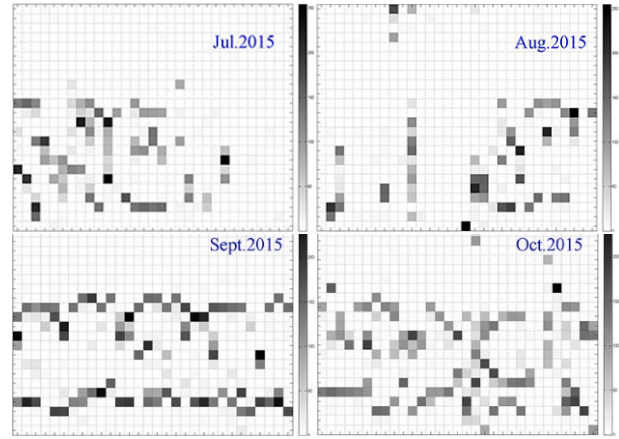
(a)

(b)

(c)

Fig.10. Case study 1. (a) Density maps of four consecutive months' datasets from *Moves*; (b) Euclidean distance measures dissimiliary of the density maps in (a) displaying RU; (c) histogram distribution of four months

A histogram regarding grey and white parts are distinctly determined into two classes for the consecutive four months, as presented in Fig. 10 (c). During this period, the subject performed routine activities. As determined by our face-to-face survey, she usually sat in front of the computer for 7 hours working every workday and sleeping 8 hours per day, and thus the information is regarded as the known SPA duration.

The ratio presented in table 3 denotes the percentage of duration of SPA against to the whole duration of the month (720 or 744 hours). Thus, the initial mass functions assigned for each density map based on knowable sedentary duration (table 2) and the pixel number proportion (table 3) are as below

Table 2

The subject's known SPA durations and their ratio per month

SPA types	Duration (hours)			
	Jul.	Aug.	Sept.	Oct.
Work	161	147	140	154
Sleep	248	248	240	248
Ratio	0.55	0.53	0.53	0.54

Table 3

Pixel numbers for density maps of four months

Classes	Pixel numbers			
	Jul.	Aug.	Sept.	Oct.
C_{DPA}	401955	678916	1127961	1077651
C_{SPA}	879293	734699	403211	440297
$C_{DPA} \cup C_{SPA}$	2060670	2065140	1888737	1893015

Table 4

Evidence fusion for Jul. and Aug.

	$m_{Jul}(DPA)$	$m_{Jul}(SPA)$	$m_{Jul}(DPA,SPA)$
$m_{Aug}(DPA)$	0.07	0.14	0.12
$m_{Aug}(SPA)$	0.08	0.15	0.13
$m_{Aug}(DPA,SPA)$	0.07	0.13	0.12

$m_{Jul}(DPA) = 0.22$ $\qquad m_{Jul}(SPA) = 0.43$
$m_{Jul}(DPA,SPA) = 0.35$;
$m_{Aug}(DPA) = 0.33$ $\qquad m_{Aug}(SPA) = 0.36$
$m_{Aug}(DPA,SPA) = 0.31$;
$m_{Sept}(DPA) = 0.6$ $\qquad m_{Sept}(SPA) = 0.21$
$m_{Sept}(DPA,SPA) = 0.19$;
$m_{Oct}(DPA) = 0.57$ $\qquad m_{Oct}(SPA) = 0.23$
$m_{Oct}(DPA,SPA) = 0.2$;

(2) Using Dempster's rule to combine them

According to the DST rule defined as the table 1, we can obtain the following new mass functions one by one based on the Eq.(15) and (16). Therefore, we can obtain a new mass assignment for Jul. and Aug. shown in table 4:

$m_{Jul,Aug}(DPA) = 0.07 + 0.12 + 0.07 = 0.28$;
$m_{Jul,Aug}(SPA) = 0.15 + 0.13 + 0.13 = 0.41$;
$m_{Jul,Aug}(DPA,SPA) = 0.12$;

As such, the new mass will be then combined with initial mass function from Sept. as:

$m_{Jul-Sept}(DPA) = m_{Jul,Aug}(DPA) \oplus m_{Sept}(DPA) = 0.29$;
$m_{Jul-Sept}(SPA) = m_{Jul,Aug}(SPA) \oplus m_{Sept}(SPA) = 0.18$;
$m_{Jul-Sept}(DPA,SPA) = m_{Jul,Aug}(DPA,SPA) \oplus m_{Sept}(DPA,SPA) = 0.02$;

Finally, we can obtain another mass functions combining with the evidence of Oct. as:

$m_{Jul-Oct}(DPA) = m_{Jul-Sept}(DPA) \oplus m_{Oct}(DPA) = 0.24$;
$m_{Jul-Oct}(SPA) = m_{Jul-Sept}(SPA) \oplus m_{Oct}(SPA) = 0.08$;
$m_{Jul-Oct}(DPA,SPA) = m_{Jul-Sept}(DPA,SPA) \oplus m_{Oct}(DPA,SPA) = 0$.

(3) Determine lifelogging PA intensity pattern

With the fusion of four density maps as a consequence of the above,
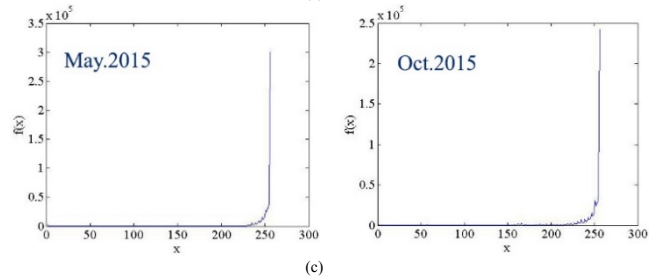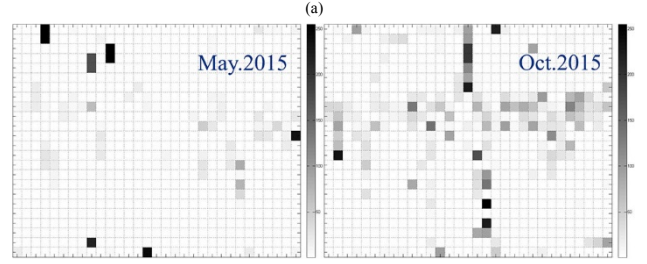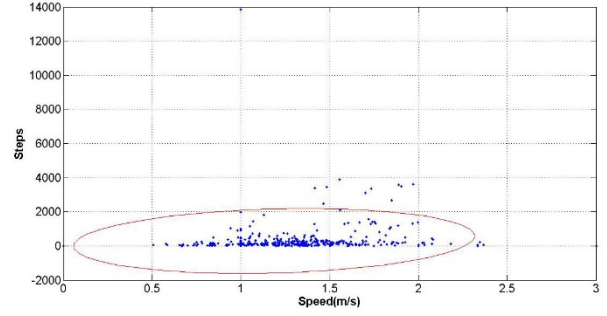
Fig.11. Case study 2. (a) Ellipse fitting model for IU removal of the subject's *Moves* dataset (c= 0.95); (b) density maps of two inconsecutive months' datasets from *Moves*; (c) histogram distribution of the density maps in (b).

the RU is almost removed. The final mass functions are next normalized based on the Eq. (15) and thus achieve the belief and plausibility value for *DPA* and *SPA* as:

$Bel_{Jul-Oct}(DPA) = Pls_{Jul-Oct}(DPA) = 0.29$;
$Bel_{Jul-Oct}(SPA) = Pls_{Jul-Oct}(SPA) = 0.1$;

Therefore, it can be inferred that the individual's PA intensity pattern for the consecutive four months is *active* with the higher degree of confidence according to the rule defined in Eq.(21).

4.2. Case study 2

The second subject is a healthy person, male, aged 35. Similarly, he worked in a sedentary fashion in front of computer 7 hours per workday, and sleeping duration was 7 hours per day on average. The *Moves* datasets are collected for two inconsecutive months of Jun. and Oct., 2015. The Ellipse fitting model is firstly conducted to remove IU. The confidence degree in this case is set to 0.95, somewhat smaller than in case 1, as most of the samples are scattered intensively between 0 to 1000 steps. A few samples are still in the normal range such as 2000 to 4000 steps per hour (Fig. 11 (a)), but here it is only taken into account estimation of the best fit of samples for this individual.

As presented in the density maps in Fig.11 (b), a smaller amount of data has been collected on May. The subject either tend to be more

sedentary or rarely record DPA data with mobile phone that the conclusion of *active* is impaired by the evidence on May. On the other hand, more information can be acquired from Oct., which may support this decision result.

Table 5

Lifelogging PA pattern comparisons of four subjects

	Subject1	Subject2	Subject3	Subject4
$Bel(DPA)$	0.32	0.2	0.25	0.22
$Bel(SPA)$	0.21	0.26	0.08	0.05
Bel comparison	$Bel(DPA)$ $>Bel(SPA)$	$Bel(DPA)$ $<Bel(SPA)$	$Bel(DPA)$ $>Bel(SPA)$	$Bel(DPA)$ $>Bel(SPA)$
Intensity pattern of lifelogging PA	Active	Sedentary	Active	Active

Likewise, two classes are also defined associated with histogram distribution of the density maps in Fig.11 (c). With the respect of pixel numbers in two classes as well as the duration of known sedentary PA (working and sleeping), the mass functions are assigned as:

$$m_{May}(DPA) = 0.15; \qquad m_{May}(SPA) = 0.42;$$
$$m_{May}(DPA, SPA) = 0.43;$$
$$m_{Oct}(DPA) = 0.4; \qquad m_{Oct}(SPA) = 0.3;$$
$$m_{Oct}(DPA, SPA) = 0.3.$$

Applying to the evidence fusion rule, we can obtain that,

$$Bel(DPA) = 0.28; \quad Bel(SPA) = 0.38.$$

Thus, the interval for $\{DPA\}$ and $\{SPA\}$ are [0.28, 0.62] and [0.38, 0.72], respectively. According to the decision making rules in Eq. (22), it can be concluded that the subject's PA intensity pattern of is *sedentary*.

4.3. Lifelogging PA intensity pattern comparisons

The belief values of four subjects' PA styles we investigated for five months are presented in the table 5. The subject 1, 3 and 4 are more active in DPA, as they were often walking around within most of each hour's cells during the day. Apart from walking, for example, subject 3 often spent two hours a day playing sports and running. Subject 1 also performed jogging at least one hour a day. On the contrary, the subject 2 either tend to be more sedentary or rarely use a mobile tracker to log his DPA that the conclusion is impaired by all the evidence.

5. Discussion and conclusion

In this paper, an approach of lifelogging PA intensity pattern measures using datasets of third party devices has been proposed. The approach firstly employs an ellipse-fitting model to remove obvious errors (IU) from processed PA data based on the best fit of the samples for the individual. It then builds day-hour grey-levelled density maps for each month to represent the intensity for each hour cell. Finally, in order to reduce regular uncertainties (RU), a DST evidence lattice model was applied to lifelogging PA decision determination in light of information fusion from multiple density images. The decision conclusion is categorized into two types which are active and sedentary. The results, based on the analysis of *Moves* [29] data sets from the mobile personalized healthcare platform MHA [1], have demonstrated that the proposed approach has a strong ability to achieve lifelogging life styles of

intensity for different individuals who successively take the mobile tracker as an evaluation criterion for healthcare. In this paper, we limit discussion of PA to only use 'steps'. The reason for this is firstly the step is the most significant lifelogging PA variable. Generally, from the perspective of longitudinal duration, individuals spend more time walking, jogging and running than other activities such as swimming or cycling. Secondly, some non-step activities are rarely automatically detected owing to the limitations of existing hardware devices (swimming and yoga are particularly problematic for example). We believe the proposed approach has extensibility with the technological progress of activity recognition techniques and device capability. Furthermore, this work typically evaluated healthy subjects. Nonetheless, this principle can be applied for detection of PA changes for elderly people and patients with chronic disease using such low-cost devices or free apps. Since the paper only assessed a few months' data sets from limited devices, in the next stage, we will be collecting larger amounts of data over a longer period of time from more subjects for evaluation. More wearable commercial devices (e.g., *Fitbit* [26], *Withings* [14]) will be also taken into consideration as a part of resource for the information fusion.

References

[1]   "MHA." [Online]. Available: http://www.myhealthavatar.eu/. [Accessed: 10-Oct-2014].

[2]   J. Yang, B. Chen, J. Zhou, and Z. Lv, "A Low-Power and Portable Biomedical Device for Respiratory Monitoring with a Stable Power Source," pp. 19618–19632, 2015.

[3]   Z. Lv, A. Tek, F. Da Silva, C. Empereur-mot, M. Chavent, and M. Baaden, "Game On , Science - How Video Game Technology May Help Biologists Tackle Visualization Challenges," vol. 8, no. 3, 2013.

[4]   Z. Lv, J. Chirivella, and P. Gagliardo, "Bigdata Oriented Multimedia Mobile Health Applications," *J. Med. Syst.*, 2016.

[5]   Z. Lv, V. Penades, S. Blasco, J. Chirivella, and P. Gagliardo, "Neurocomputing Evaluation of Kinect2 based balance measurement," *Neurocomputing*, pp. 1–9, 2016.

[6]   P. Fergus, I. Idowu, A. Hussain, and C. Dobbins, "Neurocomputing Advanced arti fi cial neural network classi fi cation for detecting preterm births using EHG records," vol. 188, pp. 42–49, 2016.

[7]   A. Vahdatpour, N. Amini, and M. Sarrafzadeh, "Toward unsupervised activity discovery using multi-dimensional motif detection in time series," *IJCAI Int. Jt. Conf. Artif. Intell.*, pp. 1261–1266, 2009.

[8]   J. Yang, "Toward Physical Activity Diary : Motion Recognition Using Simple Acceleration Features with Mobile Phones," *Proc. 1st Int. Work. Interact. Multimed. Consum. Electron.*, pp. 1–9, 2009.

[9]   T. van Kasteren, A. Noulas, G. Englebienne, and B. Kröse, "Accurate Activity Recognition in a Home Setting," pp. 1–9, 2008.

[10]   M. Shoaib, S. Bosch, O. Incel, H. Scholten, and P. Havinga, "A Survey of Online Activity Recognition Using Mobile Phones," *Sensors*, vol. 15, no. 1, pp. 2059–2085, 2015.

[11]   H. Martín, A. M. Bernardos, J. Iglesias, and J. R. Casar, "Activity logging using lightweight classification techniques in mobile devices," *Pers. Ubiquitous Comput.*, vol. 17, no. 4, pp. 675–695, 2013.

[12]   "Fitbit Flex." [Online]. Available: http://www.fitbit.com/uk. [Accessed: 10-Oct-2014].

[13]   "Moves." [Online]. Available: https://www.moves-app.com/. [Accessed: 14-Oct-2014].

[14]   "Withings." [Online]. Available: http://www.withings.com/uk/. [Accessed: 15-Oct-2014].

[15]   "Fitbit One." [Online]. Available: http://www.fitbit.com/uk. [Accessed: 10-Oct-2014].

[16]   "JawboneUp." [Online]. Available: https://jawbone.com/up. [Accessed: 10-Feb-2014].

[17]   K. Norton, L. Norton, and D. Sadgrove, "Position statement on physical activity and exercise intensity terminology ," vol. 13, pp. 496–502, 2010.

[18]   E. G. Spanakis, "MyHealthAvatar : personalized and empowerment health services through Internet of Things technologies MyHealthAvatar : personalized and empowerment health services through Internet of Things technologies," no. August, 2015.

[19]   D. A. Epstein, F. Cordeiro, E. Bales, J. Fogarty, and S. A. Munson, "Taming Data Complexity in Lifelogs : Exploring Visual Cuts of Personal Informatics Data," pp. 667–676, 2014.

[20]   A. Jalal, Z. Uddin, and T. Kim, "Depth Video-based Human Activity Recognition System Using Translation and Scaling Invariant Features for Life Logging at Smart Home," pp. 863–871, 2012.

[21]   Q. Liu, W. Zhang, H. Li, and K. N. Ngan, "Hybrid human detection and recognition

in surveillance," *Neurocomputing*, vol. 194, pp. 10–23, 2016.

[22] Z. Deng, P. Yang, Y. Zhao, X. Zhao, and F. Dong, "Life-Logging Data Aggregation Solution for Interdisciplinary Healthcare Research and Collaboration," *2015 IEEE Int. Conf. Comput. Inf. Technol. Ubiquitous Comput. Commun. Dependable, Auton. Secur. Comput. Pervasive Intell. Comput.*, pp. 2315–2320, 2015.

[23] C. Gurrin, Z. Qiu, M. Hughes, and N. Caprani, "The Smartphone As a Platform for Wearable Cameras in Health Research," *AMEPRE*, vol. 44, no. 3, pp. 308–313, 2013.

[24] A. R. Doherty, N. Caprani, C. Ó. Conaire, V. Kalnikaite, C. Gurrin, A. F. Smeaton, and N. E. O. Connor, "Computers in Human Behavior Passively recognising human activities through lifelogging," *Comput. Human Behav.*, vol. 27, no. 5, pp. 1948–1958, 2011.

[25] P. Wang and A. F. Smeaton, "Using visual lifelogs to automatically characterize everyday activities," *Inf. Sci. (Ny).*, vol. 230, pp. 147–161, 2013.

[26] "Fitbit Flex." [Online]. Available: http://www.fitbit.com/uk. [Accessed: 10-Apr-2015].

[27] "Nike+ Fuelband." [Online]. Available: http://www.nike.com/gb/en_gb/c/nikeplus-fuelband. [Accessed: 10-Apr-2015].

[28] "Endomondo." [Online]. Available: https://www.endomondo.com/. [Accessed: 10-Apr-2015].

[29] "Moves." [Online]. Available: https://www.moves-app.com/. [Accessed: 14-Apr-2015].

[30] J. Qi, P. Yang, D. Fan, and Z. Deng, "A Survey of Physical Activity Monitoring and Assessment using Internet of Things Technology," *2015 IEEE Int. Conf. Comput. Inf. Technol. Ubiquitous Comput. Commun. Dependable, Auton. Secur. Comput. Pervasive Intell. Comput.*, pp. 2353–2358, 2015.

[31] J.-L. Reyes-Ortiz, L. Oneto, A. Samà, X. Parra, and D. Anguita, "Transition-Aware Human Activity Recognition Using Smartphones," *Neurocomputing*, vol. 171, pp. 754–767, 2015.

[32] O. D. Lara and M. a. Labrador, "A Survey on Human Activity Recognition using Wearable Sensors," *IEEE Commun. Surv. Tutorials*, vol. 15, no. 3, pp. 1192–1209, 2013.

[33] P. Yang, M. Hanneghan, J. Qi, Z. Deng, F. Dong, and D. Fan, "Improving the Validity of Lifelogging Physical Activity Measures in an Internet of Things Environment," *2015 IEEE Int. Conf. Comput. Inf. Technol. Ubiquitous Comput. Commun. Dependable, Auton. Secur. Comput. Pervasive Intell. Comput.*, pp. 2309–2314, 2015.

[34] W. Gander, G. H. Golub, and R. Strebel, "Fitting of Circles and Ellipses Least Squares Solution," 1994.

[35] G. H. Golub and R. Strebel, "Least-Squares Fitting of Circles and Ellipses ∗," vol. 34, pp. 558–578, 1994.

[36] K. Okada, D. Comaniciu, S. Member, and A. Krishnan, "Robust Anisotropic Gaussian Fitting for Volumetric Characterization of Pulmonary Nodules in Multislice CT," vol. 24, no. 3, pp. 409–423, 2005.

[37] "A New Method of Interpolation and S m o o t h Curve Fitting Based on Local Procedures," vol. 17, no. 4, pp. 589–602, 1970.

[38] A. Fitzgibbon, M. Pilu, and R. B. Fisher, "Direct least square fitting of ellipses," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, no. 5, pp. 476–480, 1999.

[39] S. Wang, M. Skubic, and Y. Zhu, "Activity Density Map Visualization and Dissimilarity Comparison for Eldercare Monitoring," vol. 16, no. 4, pp. 607–614, 2012.

[40] O. Basir and X. Yuan, "Engine fault diagnosis based on multi-sensor information fusion using Dempster – Shafer evidence theory q," vol. 8, pp. 379–386, 2007.

[41] J. Liao, Y. Bi, and C. Nugent, "Using the Dempster-Shafer theory of evidence with a revised lattice structure for activity recognition.," *IEEE Trans. Inf. Technol. Biomed.*, vol. 15, no. 1, pp. 74–82, 2011.

[42] G. Simone, "Image fusion techniques for remote sensing applications," vol. 3, pp. 3–15, 2002.

[43] A. Bendjebbour, Y. Delignon, L. Fouque, V. Samson, and W. Pieczynski, "Multisensor Image Segmentation Using Dempster – Shafer Fusion in Markov Fields Context," vol. 39, no. 8, pp. 1789–1798, 2001.

[44] S. D. Herrmann, E. M. Snook, M. Kang, C. B. Scott, M. G. Mack, T. P. Dompier, and B. G. Ragan, "Development and validation of a movement and activity in physical space score as a functional outcome measure.," *Arch. Phys. Med. Rehabil.*, vol. 92, no. 10, pp. 1652–8, Oct. 2011.

[45] I. Bloch, "ation Combination Operators for Data Comparative Review with Classi," vol. 26, no. 1, pp. 52–67, 1996.

Jun Qi is currently pursuing her Ph.D. in the department of Computing Science at Liverpool John Moores University, UK. She received the M.Sc. degree and B.Sc. degree in Computer Science and technology at Changzhou University, China, in 2010 and 2013, respectively. She was research associate in the department of Computing and Mathematics at University of Ulster, UK, in 2014. Her research interests include machine learning, expert system, image processing, lifelogging and software development.

Dr Po Yang is a Senior Lecturer in the department of Computing Science at Liverpool John Moores University, UK. He received the PhD degree in Electronic Engineering from the University of Staffordshire, UK, in 2010, and the M.Sc. degree in Computer Science from the University of Bristol, UK, in 2006, and the B.Sc. degree in Computer Science from Wuhan University, China, in 2004. Dr Yang has a wide range of research interests, including Internet of Things, RFID and indoor localisation, Pervasive Health, Image Processing, GPU and Parallel Computing. He holds a strong tracking of high quality publications and research experiences. From 2006, he participated and took a technical leadership of 6 EU funded projects and 3 EPSRC/TSB funded projects. He has published over 40 papers. He has served as reviewer for over 10 international journals (including 6 IEEE/IET transactions) and as the committee member of 5 international conferences and workshops.

Dr Martin Hanneghan is Associate Director of the Department of Computer Science at Liverpool John Moores University. He holds a B.Sc. (Hons) and Ph.D. in Computing. His research interests are in software engineering, Internet of Things and Serious Games and has published widely in these areas.

Dr. Stephen Tang (B.Sc., M.Sc., PGCertLTHE, Ph.D., SFHEA) is a Principal Lecturer in Digital Creative Technologies and Programme Leader for B.Sc. (Hons)/ MComp (Hons) for Computer Games Development in the Department of Computer Science at Liverpool John Moores University. He received his Ph.D. in Model Driven Game Engineering applied to serious games and his research focus is on visualisation, user experience, serious games, gamification and game-based learning. His current research is on the smart serious games (serious games and IoT), application of games in pharmacy and computing education, interactive architectural visualisation and application of virtual reality in education and therapy.