An Analysis of Eating Activities for Automatic Food Type Recognition

Hyun-Jun Kim and Mira Kim and Sun-Jae Lee and Young Sang Choi
Intelligent Computing Lab, Future IT Research Center
Samsung Advanced Institute of Technology(SAIT) Samsung Electronics Co., Ltd.
Nongseo-dong, Giheung-gu, Yongin-si, Gyeonggi-do, Korea 446-712
E-mail: {dannis.kim, mira1012.kim, sunjae79.lee, macho}@samsung.com

***Abstract*—Nowadays, chronic diseases such as type 2 diabetes or cardiovascular diseases are considered to be one of the most serious threats to healthy life. These kinds of diseases are primarily caused by an unhealthy lifestyle including lack of exercise, irregular meal patterns and abuse of addictive substances such as alcohol, caffeine and nicotine. Therefore, observing our daily lives is crucial in developing interventions to reduce the risk of lifestyle diseases. In order to manage and predict progression of diseases of a patient, objective measurement of lifestyle is essential. However, self-reporting questionnaires and interviews have limitation due to human errors and difficulty of conducting. In this paper, we analysed users' eating activities and comprising sub-actions for developing eating activity recognition system based on a tri-axial accelerometer embedded wrist band. By analysing actions in eating activities, we can improve the accuracy of the recognition of eating activities and also provide clues that identifying the type of foods.**

I. Introduction

As the demands for improving personal health without incurring high medical cost increases, mobile healthcare is growing as an emerging technology in the medical and the IT industries. Especially, chronic diseases that are typically caused by an unhealthy lifestyle such as unbalanced nutrition, irregular meal patterns, abuse of addictive substances and lack of exercise account for a large percentage of the cause of death in our society. Accordingly, technologies which were previously applied to remote patients with elderly or chronic diseases are now being employed to help ordinary people to manage good health and prolonged life. For example, Holter monitors were worn by patients to record the electrical activities of the cardiovascular system for the past decades. However, nowadays, smaller and smarter devices have been developed for continuous monitoring of users' physiological information, providing information of not only ECG but also blood pressure, glucose, SpO2 and pulse. Recently, many IT companies have released wellness devices such as fitness tracker, smart pedometer and sleep monitor which contain an accelerometer and several additional sensors to recognize users' activities. As shown in figure 1, most of the featured products focus on the user's weight management by monitoring physical activities such as walking, running and standing. Furthermore, some of the cutting-edge devices such as SensWear[1] and Zeo[2] provide in-depth monitoring by utilizing multiple sensors. By analysing the user's physical

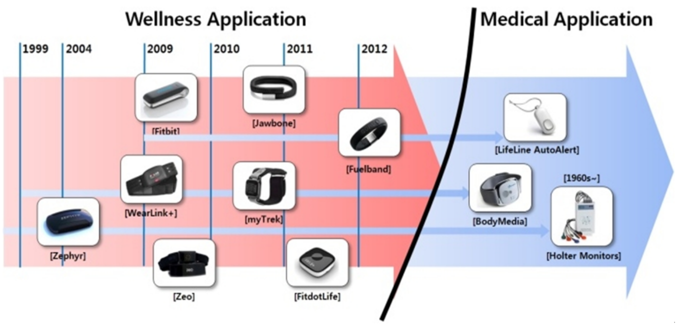

Fig. 1. Physical Activity Monitoring and Physiological Signal Sensing Devices

activity or physiological information, these devices facilitate users with a healthier lifestyle by providing suggestions for appropriate diet, offering warnings for the obesity risks and indicating the quality of sleep. On the other hand, in spite of the importance of food intake in measuring health status, researches in the area of eating activity recognition have not been sufficiently conducted. In many cases, researchers tried to recognize users' eating activities by using sensor network in a smart home environment or capturing eating activities with a video camera. These approaches show a relatively reasonable performance. However, the downside of these methods is that the recognition is possible only in the areas where the sensor network or video recording equipments are already installed such as smart home or laboratories.

To overcome this limitation, we propose an unobtrusive way of daily life monitoring by utilizing a tri-axial accelerometer embedded in the wrist band. Our approach recognizes each action of the eating activity such as 'picking up food with chopsticks' or 'eating rice with a spoon' and classifies along with the different types of food. With the use of an accelerometer and a video camera, we defined specific eating actions and analysed them for providing a platform for recognizing not only the eating activity but also the types and quality of food. Our approach will provide users with convenient and accurate daily life monitoring such as eating activities and has a potential to revolutionize health care in the near future. This paper is organized as follows: Section 2 summarizes the main contribution of the research and section 3 presents the related works to eating activities recognition. Section 4 describes the environment of the data collection and section 5 gives an analysis of actions in eating activities. And finally, we conclude with the summary and plans of future works in sections 6.

II. Contribution

The main contributions of this research are as follows: 1) We conducted a fundamental research of data analysis for recognizing eating activity that can be used for mobile healthcare devices.

2) We proposed a recognition approach based on the predefined action units of eating activity.

3) To the best of our knowledge, this is the first work that focused on the Asian eating style which is characterized by using spoon and chopsticks only with one dominant hand.

4) We also showed the potential for the proposed approach by recognizing the types of food.

III. Related Works

In general, researches in eating activity recognition have followed methods in physical activity recognition studies. Recent studies conducted in the area of physical activity recognition mostly proposed methods such as employing video analysis and sensors including RFID and accelerometer. A variety of areas in physical activity recognition have been researched including the calculation of the amount of exercise such as walking and running, fall detection, and daily life activities such as cooking, resting and sleeping for the purpose of health risk detection for elderly people. Nowadays, smartphones with an assortment of sensors embedded can enable various innovative researches including an accurate measurement of physical activities. Tapia et al. proposed a research of recognizing physical activities by detecting users' moving route and the duration of activity in home setting using 77 simple state-change sensors[1]. They executed a probabilistic reasoning of several sensors around the user's surroundings and successfully detected the daily round of activities such as preparing meal, toileting, watching TV and washing dishes. Additionally, some other approaches for improving the accuracy of the collected sensor data in specific places included methods for observing sequential states of individual sensor data such as semi-Markov model[2], finite state machine[6] or HMM[7] as well as utilizing RFID sensors[3], accelerometer sensors[4] or video clips all together[5]. Methods employing a sensor network produced a high accuracy rate but showed limitations; high cost due to the sensing environment and restrictions on areas for which sensors can detect. Another method for recognizing eating activities employed video footage detection. Kaiser et el. proposed a futuristic approach called 'Restaurant of the Future' for monitoring customers' behaviours that depended on the food or restaurant in a similar way as video surveillance system[8]. First of all, they detected specific objects such as a face, a table or a glass and probabilistically predicted the actions occurred by using state diagram. However, physiological activities recorded by video footage had a limitation that people have to exhibit unique characteristics. Alexandros et al. proposed a more individualized video footage detection method taking personal distinctiveness into account. By detecting user behaviours with a total of 8 cameras from various angles, they successfully improved the accuracy of eating and drinking activity detection[9]. But similar to the methods based on sensor networks, recognition based on video has restricted applicability and it is requiring fundamental techniques for object recognition and object tracking. Lastly, body worn sensors employing an accelerometer showed the highest potential for eating activity recognition. Bao et al. conducted an investigation in biaxial accelerometers worn simultaneously on different parts of the body for common everyday activities[12]. The accuracy varied by the different positions of the sensors worn, but showed a high overall accuracy. Nowadays, most smartphones are equipped with an accelerometer sensor resulting in an increasing number of research works in physical activity recognition[7][10][11]. Many existing research, utilized a triaxial accelerometer sensor embedded in smartphones to detect the transportation modes of users such as walking, running and stationary correlated to the sensor positions on the body. These studies achieved a considerably high accuracy and by using these techniques, numerous commercial applications like *Runkeeper*[1] have launched. Also, as shown in Fig.1, various types of body worn sensors are being marketed and are usually worn inside the pocket, on the wrist or on the chest to monitor users physical activities as well as sleeping habits. Tolstikov et al. used wrist worn accelerometer sensor under sensor network environment to detect eating activities[15]. They showed a potential for eating activities recognition through the use of accelerometer sensor. They used their work as an assisting tool for improving the accuracy for their existing work. Also, Zhang et al. proposed a method of adopting the Hierarchical Temporal Memory model to classify the extracted features of the eating and drinking activities using only the tri-axial accelerometer sensor worn on the wrist[16]. They achieved an accuracy over 85% for continuous eating activities. Similarly, there are studies that perform segmentation of arm gestures and body posture to recognize eating activities[6][17]. Several other approaches included using an implantable recording device to monitor the movement and the temperature of the gastrointestines[13], recording the temperature of neck and ears of users to monitor the eating behaviours[14]. However, these methods showed practical limitations for daily life monitoring due to the inconvenience of having to wear the devices at all times.

As mentioned above, there has been an increase in the amount of work for identifying eating activities. However, most of the studies proposed focus primarily on the methods of sensing, detection algorithm or improvement of accuracy rather than the in-depth analysis of the activities. Hence, in this paper we propose to establish the fundamentals for an efficient implementation of various methods by thoroughly analysing the eating activities.

[1] http://www.runkeeper.com

Fig. 2. Wrist Worn Type Accelerometer Sensor and Micro Camcorder

IV. Description of the Data

To analyse the eating activity, we collected accelerometer data from 13 subjects over 2 weeks classifying 2 different types of meal; rice and noodle.

A. Sensing Environment and Data Collection

The ez430-Chronos[2] accelerometer sensor was worn on the user's dominant hand and at the same time, the user's eating activity was recorded with a micro camcorder. The recorded video clips were analysed by 2 people in frames in milliseconds to label the specific actions during the eating activity. Accelerometer data was collected at a 30 Hz sampling frequency, with 128 windows overlapping by 50%.

B. Segmentation and Labeling

With the analysed data, we defined 29 types of eating actions using 3 types of utensils; spoon, chopsticks and bare hand. Taking efficiency of data training and classification into account, we isolated continuous actions and separated actions even for the same action. We maintained the 6 actions that did not appear from our collected data for the possibility of them appearing in the future data. In detail, we were able to show that out of the average eating time of 12:14 minutes[3] , 72.1% (8:45 minutes) showed eating actions, 27.9% involved non-eating related actions such as waiting while chewing, talking and using gestures while talking. Furthermore, from the 29 actions we defined, 58.6% (17 actions) were shown 107.3 times in each meal in average. Especially, the top 5 actions; S12, C1, C2, C7 and C10 occupied over 50%.

C. Limitations

Unlike western eating styles, Asian (especially Korean) eating styles differ by using a spoon and chopsticks alternately with the dominant hand. Eating activity varies depending on the cultural background and personal habits. In this paper, we did not try to overcome these limitations. Additionally, the decision criteria for labeling of the collected data can differ from individual to individual. For a more precise analysis and training of the actions, handling of these limitations is important but there may be a room for error since human intervention would be involved. For example, an undefined action may appear or an action may get interpreted differently by different individuals. Therefore, we will need further study for more accurate classification of actions and cross-validation for a better analysis of user activities. An Analysis of Eating Activities

TABLE I
Definition of Actions and Analysing Results

Type	Act.	Description	Cnt	Sec.	%
Spoon	S1	Taking a spoon	4	32.1	4.9
	S2	Taking & scoping up soup	3.5	29.1	4.4
	S3	Taking & scoping up rice	1.5	14.4	2.2
	S4	Scoping up side dish	0	0.0	0.0
	S5	Scoping soup & waiting	3	14.9	2.3
	S6	Scoping soup & putting in mouth	7.5	28.8	4.4
	S7	Scoping rice & waiting	5	11.9	1.8
	S8	Scoping rice & putting in eating	5.5	28.3	4.3
	S9	Scoping side dish & waiting	0	0.0	0.0
	S10	Scoping side dish & putting in mouth	0	0.0	0.0
	S11	Putting in mouth	8	14.2	2.2
	S12	Stirring	7	45.8	6.9
	S13	Whisking	1	1.1	0.2
Chopsticks	C1	Taking chopsticks	10	68.9	10.5
	C2	Taking & picking up toppings	3	41.7	6.4
	C3	Taking & picking up rice	2	10.0	1.5
	C4	Picking up side dish	4	33.4	5.1
	C5	Picking toppings & waiting	1	2.2	0.3
	C6	Picking toppings & putting in mouth	0	0.0	0.0
	C7	Picking rice & waiting	6	34.3	5.2
	C8	Picking rice & putting in eating	6.5	30.3	4.6
	C9	Picking side dish & waiting	5	25.0	3.8
	C10	Picking up fillet & stirring	15	95.1	22.3
	C11	Picking side dish & putting in mouth	8.3	37.3	5.7
	C12	Putting in mouth	19.6	31.6	4.8
	C13	Stirring	0	0.0	0.0
	C14	Whisking	0	0.0	0.0
Hand	H1	Eating with a hand	1.3	8.9	1.4
	H2	Cleaning with tissue	2.5	15.7	2.4

V. An Analysis of Eating Activities

In this section, we present a detailed analysis of eating activities from the labeled video clips. Fig. 2 shows the action types and the frequency of each action recorded for 4 subjects. With the exception of person 4, the rest shows the variance of actions even for the same food type 'Steamed Rice with Dishes'. In case of the food type 'Steamed Rice with Dishes', the average action types observed was 17 and the average number of actions was 110 for a meal. We did not find much variation in action types (σ=1.74) but the number of actions occurred showed more than double amount of variation (σ=3.72) during the meal time. It means that users used only a limited number of actions they are accustomed to and the actions during meal such as talking and watching TV can affect the range of fluctuations of the meal time and calorie intake.

Fig. 2 shows that even with the same food type, action frequency and patterns may vary by different individuals.

[2]http://processors.wiki.ti.com/index.php/EZ430-Chronos

[3]According to the study conducted by the Seoul National University Hospital in Bundang, Korea, 72% out of the 1,289 office worker reported that their average eating time for a meal was less than 15 minutes, 2008.

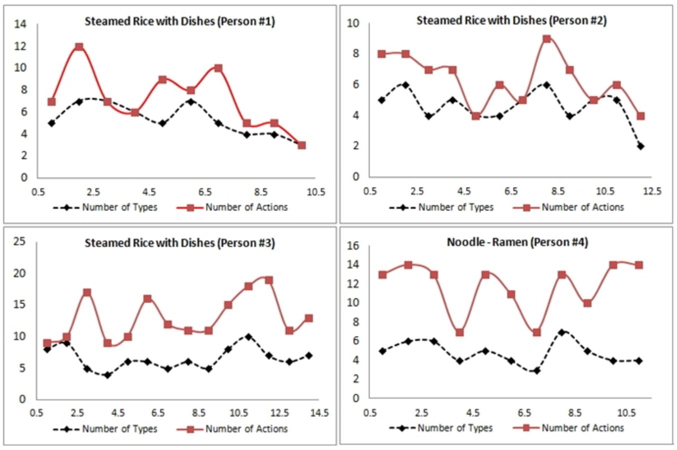

Fig. 3. Variation on Occurrences of the Action Types and Number of Actions

TABLE II
CLASSIFICATION RESULT ON EATING ACTIVITY

Classes	NaiveBayes	BayesNet	Boosting	Bagging	C4.5
3	0.62	0.65	0.68	0.75	0.72
29	0.19	0.18	0.19	0.28	0.20

Two graphs (in Fig. 2 and Fig. 3) of person 3 show that person 3 demonstrates a relatively regular eating pattern. Although person 4 ate a different food type (noodle) and the actions detected may vary from individual to individual, person 4's data shows different frequent actions compared with the actions of other subjects eating steamed rice with dishes. For example, since chopsticks are used for noodles, actions related with chopsticks are found more frequent. Fig. 3 shows the frequency and the occurrence time of actions. Similar to Fig.2, the 3 data tables for 'Steamed rice with dishes' show varied action patterns despite the identical meal type. Furthermore, the actions of the subject eating 'Noodle-Ramen' comprise of actions involving mostly chopsticks in the beginning. Towards the end, the actions consist of those involving a spoon since the subject put rice into soup as it is one of the common Asian (mostly Korean) eating habits. To conclude, eating pattern varies from individual to individual and also by the food types. Ultimately, with more data collection and analysis, we will be able to present a framework for a more precise eating pattern and food type detection.

VI. FEASIBILITY TEST

For the 29 actions defined, we conducted an experiment for the purpose of identifying valid actions for eating activities. First of all, we evaluated 2 types of classification performance on 3 actions based on the overall classification performance and action types. As shown in Table. 2, we experimented with 5 classification algorithms and obtained an average classification performance of 0.68 for 3 types of actions; spoon, chopsticks and hand. We achieved a classification performance of 0.59 and 0.79 on F-measure for spoon and chopsticks, respectively. However, we were not able to classify the hand actions.

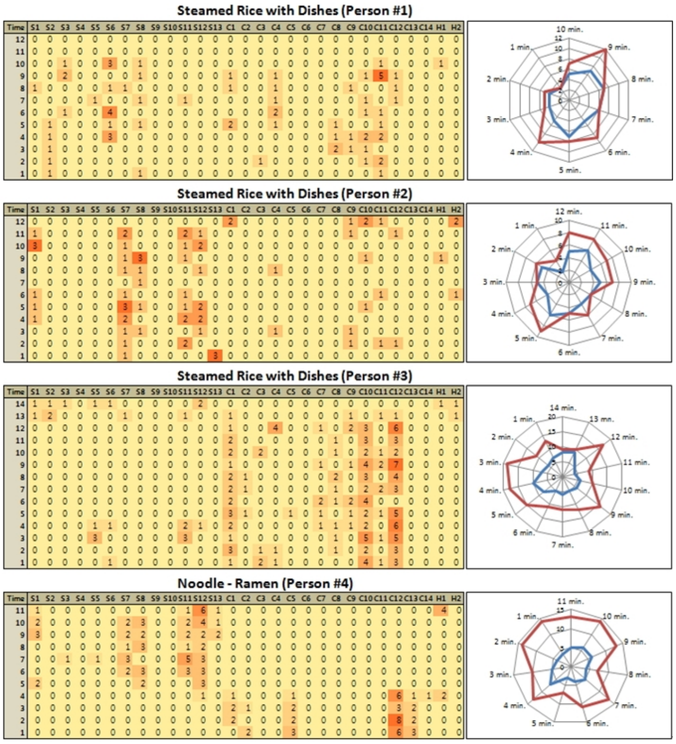

Fig. 4. Frequency of action types by the elapsed time

On the other hand, the classification performance of all 29 actions resulted in the average of 0.21 illustrating the difficulty of detecting all actions. However, the classification performance and clustering of 29 classes of actions may be utilized as a seed for identifying or redefining significant actions. In particular, the 5 most frequently occurring actions in the top 50 percentile in Table 1 did not produce high detection rate as shown in Table 3. That is, for an accurate detection of eating activities, we need detailed observation and analysis of frequently occurring actions and that the set of 29 predefined actions itself may not be suitable for the actual recognition of meals.

VII. CONCLUSION AND FUTURE WORKS

In this paper, we proposed a foundational approach of eating activity detection by using the accelerometer sensor data and video observation. By analysing the collected data, we defined 29 actions on Asian eating style of one dominant hand as the foundation. We believe our study will help recognizing eating activities more accurately for investigating various health risks of users by detecting types of food and other eating habits such as eating speed. The presented study was conducted on exploratory manner and has limitation on food type, place and eating time. In the future, we plan to analyse not only the meal types but also the eating patterns of users with sufficient amount of data collected in various sensing environments.

TABLE III
FEATURE EXTRACTED EATING ACTIONS AND CLASSIFICATION RESULT

Actions	Instances	Precision	Recall	F-Measure
S12	44	0.433	0.295	0.351
S7	72	0.338	0.361	0.349
S1	42	0.3	0.286	0.293
S2	28	0.265	0.31	0.286
C10	134	0.245	0.291	0.266
C1	95	0.234	0.274	0.252
C12	44	0.25	0.182	0.211
S8	20	0.13	0.15	0.14
C3	6	0.111	0.167	0.133
S6	23	0.136	0.13	0.133
C11	50	0.136	0.12	0.128
C8	28	0.108	0.143	0.123
C9	39	0.089	0.103	0.095
C4	52	0.074	0.077	0.075
S3	14	0	0	0
S5	13	0	0	0
H1	12	0	0	0
S11	12	0	0	0
H2	10	0	0	0
S13	0	0	0	0
C7	15	0	0	0
C2	19	0	0	0
C5	0	0	0	0

REFERENCES

[1] E. M. Tapia, S. S. Intille, and K. Larson, "Activity Recognition in the Home Using Simple and Ubiquitous Sensors," In Pervasive, pp. 158175, 2004.

[2] T. van Kasteren, G. Englebienne and B. Krose, "Activity recognition using semi-markov models on real world smart home datasets," Journal of Ambient Intelligent. Smart Environment, vol. 2, no. 3, 2010.

[3] P. Rashidi, D. Cook. "Mining Sensor Streams for Discovering Human Activity Patterns over Time," ICDM Conference, 2010.

[4] N. Zouba, F. Bremond and M. Thonnat, "Multisensor Fusion for Monitoring Elderly Activities at Home," In Proc. of the 6th IEEE Int. Conf. on Advanced Video and Signal Based Surveillance, pp. 98-103, 2009.

[5] K. Sim, G-E. Yap, C. Phua, J. Biswas, A.P.W. Aung, A.Tolstikov, W. Huang, P. Yap, "Improving the Accuracy of Erroneous-Plan Recognition System for Activities of Daily Living," In Proc. of the 12th IEEE International Conference on e-Health Networking Applications and Services (Healthcom), pp. 28-35. 2010.

[6] T. Teixeira, J. Deokwoo, G. Dublon, and A. Savvides. "Recognizing activities from context and arm pose using finite state machines," In Proc. of the 3rd ACM/IEEE International Conference on Distributed Smart Cameras, pp. 1-8, 2009.

[7] Y. Lee, "Activity Recognition Using Hierarchical Hidden Markov Models on a Smartphone with 3D Accelerometer," Lecture Notes in Computer Science, Hybrid Artificial Intelligent Systems, Springer, Vol.6678, pp.460-467, 2011.

[8] M. Kaiser, D. Arsic, B. Hornler, M. Hofmann and G. Rigoll, "The Noldus database: Automated Recognition of Restaurant related Activities for the Restaurant of the Future," In Proc. of the WIAMIS, April 2011.

[9] I. Alexandros, T. Anastasios and P. Ioannis, "Person Specific Activity Recognition using Fuzzy Learning and Discriminant Analysis," In Proc. of the 19th European Signal Processing Conference, 2011.

[10] L. Hong, Y. Jun, L. Zhigang, D. L. Nicholas, C. Tanzeem and T. C. Andrew, "The jigsaw continuous sensing engine for mobile phone applications," In Proc. of the 8th ACM Conference on Embedded Networked Sensor Systems, 2010.

[11] J.R. Kwapisz, G. M. Weiss, and S. A. Moore. "Activity recognition using cell phone accelerometers," In Proc. of the 4th International Workshop on Knowledge Discovery from Sensor Data, pp. 10-18, 2010.

[12] L. Bao, and S. S. Intille, "Activity recognition from user-annotated acceleration data," In Proc. of the 2nd International Conference on Pervasive Computing, pp. 117, 2004.

[13] R. Aviv, C. P. Sanmiguel, A. Kliger, S. Policker, W. Haddad, M. Hagiike and E. E. Soffer, "The use of gastric electrical signals for algorithm for automatic eating detection in dogs," Neurogastroenterol Motil.Vol. 20(4), pp.369-76, 2008.

[14] O. Amft and G. Trster, "On-body sensing solutions for automatic dietary monitoring," IEEE Pervasive Computing., vol. 8, no. 2, pp. 6270, 2009.

[15] A. Tolstikov, J. Biswas, C.-K. Tham, and P. Yap, "Eating activity primitives detection - a step towards ADL recognition," In Proc. of HealthCom, pp. 3541, 2008.

[16] S. Zhang, J. Ang, H. Marcelo, and oth., "Detection of activities by wireless sensors for daily life surveillance: eating and drinking," Sensors, vol. 9, no. 3, pp. 14991517, 2009.

[17] O. Amft, H. Junker, G. Trster, "Detection of eating and drinking arm gestures using inertial body-worn sensors," In Proc. of the 9th IEEE International Symposium on Wearable Computers, pp. 160163, 2005.