Challenges for Activity Recognition with Real Data

Sozo Inoue
Faculty of Engineering, Kyushu Institute of Technology
1-1 Sensui-cho, Tobata-ku, Kitakyushu, 804-8550, Japan
sozo@mns.kyutech.ac.jp

ABSTRACT
In this paper, we introduce a large-scale activity data collection, and discuss the challenges for activity recognition with real data. We address the requirements of annotation for sequential activities and of utilizing device status, and also address research directions for unsupervised methods and complementary information.

Author Keywords
Human activity recognition, three-axis accelerometer, smart phone

INTRODUCTION
Recent deployment of smart phones equipped with accelerometers will make it possible to recognize activities of the users. In this paper, we introduce a large-scale activity data collection system with smart phones. From the result of training activity recognition upon sampled data from gathered over 35,000 activity data, we discuss the challenges for more realistic activity recognition. We address the requirements of annotation for sequential activities and of utilizing device status, and also address research directions for unsupervised methods and complementary information.

ALKAN SYSTEM
To collect activity data efficiently, we used a large-scale activity gathering system named ALKAN[2, 3]. In ALKAN, to achieve accuracy of annotations, we introduce the idea of "mission". A *mission* is a sequence of choosing an activity so-called *activity class*, choosing the position of the device on the body, and performing the activity. Using this method, users can record activities anytime they want, and the annotation is accurately stamped within deviations of few seconds. And, for usability, we adopted smart phones as mobile sensor devices. Most smart phones are equipped with 3-axis accelerometers, storage, and wireless communication, which enable recording activity data anytime. The data can be uploaded to the server when it is connected to the network. Smart phone client software is easy to scale up by installing client software through application deploying services. On the other hand, the server can be scaled up by existing distributed web technology.

The ALKAN system consists of mobile device clients and a server. A user records missions using the mobile device client. The information is uploaded to the server when it is online and accumulated in the server database. The user can view statistical information of the uploaded data, such as a calendar of activity history and rankings, by connecting to the web server through the smart phone or another web browser on a PC.

Lessons Learned
We have delivered 216 iPodTouchs as smart phones to university students and staff. We asked users a favor to collect activity data once a day on average. As a result, we gathered 35,310 missions during about 14 months.

Upon the sampled data, we applied the well-known activity recognition method by Bao and Intile[1]. Surprisingly, the result was worse than shown in the single sensor case in [1]. The following are considered as the reasons:

- A mobile sensor is not firmly fixed to the body, but shaken in the pocket.
- Activity classes are similar to each other. As we can imagine, similar activity pairs such as "eat.sit"–"sit" and "sit"–"train.sit" are often mis-recognized.
- Actual activities may have varieties. Since users have performed activities in their own situations, environments could differ greatly on each trial.
- Labels are ambiguously understood by users. Since we do not have a method to verify activities, users are even possible to lie performing activities.

Although these factors will decrease the recognition accuracy, they can produce a more challenging data set for activity recognition since these situations are more realistic than traditional laboratory settings.

CHALLENGES FOR REAL DATA ACQUISITION
In this section, we discuss the challenges for obtaining real data for training activity recognition methods.

Annotation for sequential activities

UbiComp '11, Sep 17-Sep 21, 2011, Beijing, China.

In ALKAN, we adopted to obtain activities and their annotation one by one, by the concept of 'mission' to keep annotation correct on time axis. However, if we challenge real data acquisition, we cannot avoid obtaining sequential activities, in which the occurrence probabilities of each activity can be provided, and also new activities we do not know beforehand are included.

When we obtain sequential activity data, annotation becomes more difficult. In the following, we discuss the difficulty from several aspects: segmentation and multi-level activity classes.

Segmentation

Segmentation, which is to divide the sequential data into several activity parts, can be considered to be as a first part of annotation for sequential data, and to be the major reason why manual annotations are inaccurate. If we do offline segmentation afterwards, many activity information will be lost, and the timing will be largely mis-focused. If we do online, a user has to do additional action, or often forgets to do segmentation. If we engage an observer to annotate the user's activity, it often costs too much to scale up the data.

Multi-level activity classes

A single activity can often be captured by observers as different activities. For example, 'cooking' consists of 'cutting', 'boiling', and so on, and each of which consists of more low-level activities such as 'moving hands', and 'standing'.

One of the countermeasure for it is to structure annotations for activities as multi levels. An activity can consist of several activities, and it can also be a part of others. By this, we can expect to reduce the space for annotation words by knowing the relationship among them, and the recognition of a high-level activity which consist of a sequence of lower-level activities becomes easier.

Knowing device status

From the result of ALKAN, the variation of device status might affect the recognition accuracy. Device status consists of such as the direction, position on the body, and clothes types. Since these could provide different sensor data, or provide noises for activity recognition, knowing them is important for refining activity recognition.

Research directions

Since the challenges of annotations for sequential activities and of the device status critically affect the recognition accuracy, solutions for them are highly important. Here, we discuss two direction for solutions: unsupervised method and complementary information.

Unsupervised method

When annotation cannot be trusted, we can consider using unsupervised method. Unsupervised methods do not require annotations in training data, and only analyze the relationship among sensor data items. Then, we can extract information about activities, which is independent from annotations.

Recently, several methods utilizing unsupervised methods are proposed such as [4, 5]. These approaches should be more discovered.

Of course, unsupervised method must be combined with some supervised methods with annotations. However, performing basic parts such as segmentations and knowing device status with unsupervised methods as much as possible, and assigning the rest to supervised methods are a hopeful direction, since it can minimize the inaccuracy of manual annotations.

Complementary information

The other direction is to use complementary information along with the target sensor data. Sound data is nowadays easy to be recorded with the same device with accelerometers. Videos can be also useful when there are such environments, or observers. Kinetic information is recently become easy to be captured using widely-spread consumer devices. Since these data are objective, they can help obtaining annotations for activities if the timing is synchronized correctly.

Moreover, several information can be obtained from application systems. For example, location information can be also obtained in location-based services. Moreover, in healthcare systems, body features and lifestyle information of each user will be stored in the system database.

The challenges for these complementary information reside not only in data analysis, but also in system design, where these complementary information should be obtained with less stresses of users, no privacy risk, and lower costs.

CONCLUSION

In this paper, we summarized large-scale activity data collection and discussed the challenges for training activity recognition with real data. ALKAN data are open and free to use. Large-scale data is still important for accomplishing the challenges discussed in this paper.

ACKNOWLEDGEMENTS

This work is supported by Grant-in-Aid for Young Scientists (A) (21680009) of JSPS.

REFERENCES

1. L. Bao and S. Intille, "Activity Recognition from User-Annotated Acceleration Data", Proc. Pervasive 2004, pp. 1–17.
2. ALKAN web site, (Anonymized)
3. (Anonymized)
4. H. Bayati, J. Millan, R. Chavarriaga, "Unsupervised adaptation to on-body sensor displacement in acceleration-based activity recognition", Proc. ISWC2011, pp.71-78.
5. T. Maekawa, S. Watanabe, "Unsupervised Activity Recognition with User's Physical Characteristics Data", Proc. ISWC2011, pp. 89-96.