The Activity Recognition Repository: Towards Competitive Benchmarking in Ambient Intelligence

Bostjan Kaluza, Simon Kozina, Mitja Lustrek
Department of Intelligent Systems, Jozef Stefan Institute
Jamova cesta 39, 1000 Ljubljana, Slovenia
{bostjan.kaluza, simon.kozina, mitja.lustrek}@ijs.si

Abstract

Rapid development in the area of ambient intelligence introduced numerous applications. One of the fundamental underpinnings in such applications is an effective and reliable context-aware system able to recognize and understand activities performed by a human, and context in which it happened. However, there are two pending issues: (i) transferability, i.e., a specific implementation is tightly interrelated with a selected algorithm, available sensors, and a scenario/environment where they are employed; and (ii) comparability, i.e., there is no established benchmark problem that would enable a direct comparison of the developed context-aware systems. This paper first reviews some recent initiatives that address the abovementioned problems and then proposes a centralized collection of resources related to design and evaluation of context-aware systems. The main idea is to establish an online repository of datasets accompanied with the task, result and applied approach. Ideally, the contributors will provide the dataset with short description of the data, task and results, relevant paper, and link to resources such as implementation of the approach, preprocessing tools, and filtering. This would allow the community to quickly start building upon the latest state-of-the-art approaches, to benchmark newly developed techniques, and ultimately, to advance the frontiers in ambient intelligence.

Introduction

Recent advances in the area of ambient intelligence (AmI) have shown a diverse range of applications that are sensitive and responsive to the presence of people. One of the fundamental underpinnings of such applications is an effective and reliable context-aware system able to recognize and understand activities performed by a human, and the context in which they occur. In order to achieve a reliable performance in real-life applications, development of such a system must be supported by rigorous tests and quality datasets, which are typically prepared by the authors of each system.

There are several AmI systems demonstrating good performance in various real-life scenarios (Park and Kautz, 2008; Dovgan et al., 2011), but comparing them is difficult due to differences in scenario, environment, sensors, and approach. Consider the problem of fall detection, where there are several studies implementing a wide variety of acceleration-based fall detection systems, each of which is tested each on its own dataset. There is no established benchmark problem that would enable a direct comparison of the developed methods.

Recently, several research groups have initiated the first steps to address the abovementioned problems. International Workshop on Frontiers of Activity Recognition (Dai et al., 2010), which was held in 2010, organized a competition oriented towards identifying challenges, gaps and opportunities in activity recognition, mostly from video data. The competition challenged participants with VIRAT video dataset (Oh et al., 2011), which was recorded with multiple cameras in natural, real-world scenarios with varying frame rates and different resolution, viewing angles and degrees of background clutter. The task was to design an activity recognition algorithm, which is capable of dealing with these issues. The following year, the Activity Recognition Competition Workshop (Davis and Hoogs, 2012) was held, where annotated data was added to the dataset.

Next, the Opportunity Activity Recognition Challenge (Opportunity, 2011) was organized in 2011, which is a part of the EU research project *Opportunity* that aims "to develop generic principles, algorithms and system architectures to reliably recognize complex activities and contexts despite the absence of static assumptions about sensor availability and characteristics in opportunistic systems" (Kurz et al., 2011). The dataset was recorded in sensor-rich environment comprising 72 sensors integrated in the environment and objects, and worn on the body. The dataset is annotated with both complex and atomic activities. The competition was divided into four tasks: (1) multimodal activity recognition from body-worn sensors; (2) automatic detection of time periods when no relevant action is performed; (3) hand gesture detection and recognition; and (4) hand gestures detection when additive and rotational noise was added to the testing dataset.

Finally, Evaluating AAL systems through competitive benchmarking (Chessa et al., 2011) was held in 2011. In contrast to the other competitions, the participants were required to provide both the hardware and the software components, which were evaluated on a testing scenario. The competition was divided into two tracks: indoor localization and tracking, where the goal was to find the best indoor localization system; and activity recognition that aimed to find the best activity recognition system. The competition will be also organized in 2012.

The aim of this paper is to propose a centralized collection of resources related to design and evaluation of context-aware systems. Our goal is to establish an online repository of datasets, accompanied with the tasks the dataset was used for, the results achieved so far, and the approach used. Ideally, the contributors will provide the dataset with short description of the data, task and results, relevant papers, and links to resources such as implementation of the approach, and preprocessing and filtering tools. This would allow the community to quickly start building upon the latest state-of-the-art approaches, to benchmark newly developed techniques, and ultimately, to advance the frontiers in ambient intelligence.

The rest of the paper proposes an initial idea of the online repository and three sample datasets. The main goal is not to force the community to accept this idea, but to encourage discussion as how to design and what to include in the repository in order for it to serve its purpose. First, we discuss various types of datasets, which are categorized by the domain, sensor type, and task. Next, we propose the repository structure and describe the currently implemented version. Finally, we briefly describe the three sample datasets that were added to the repository.

Dataset Types

The area of ambient intelligence is wide, which results in a wide variety of tackled problems and the amount of created datasets. The datasets can be categorized by various criteria: by sensor type, by task tackled, by problem domain. Note that a dataset can be tagged with multiple labels from each category.

The first category specifies the hardware used for dataset recording. This category may include the following subcategories: (1) embedded sensors, which can be placed across the environment (e.g., pressure sensors on the ground) or in/on the objects (e.g., RFID tags in kitchen appliances); (2) visual hardware, such as cameras or infrared motion capture for localization; (3) body-worn sensors, such as location sensors, and a group of inertial sensors consisting of accelerometers, gyroscopes and magnetometers.

The second category, the task, may include the following: atomic activity recognition, complex activity/plan recognition, behavior recognition, activity context representation, activity analysis, energy expenditure estimation, etc. The third category describes the application domain, for example, health monitoring, disease recognition/detection, anomalous/suspicious behavior detection, (robotic) assistance for ADL, sport analysis, etc.

We should also consider the possibility to separate synthetically generate datasets from those recorded in physical environments.

Repository

We have implemented an initial version of the online repository (Activity Recognition Repository, 2012). The repository consists of the front page and the pages with descriptions of each dataset. The front page provides an overview of contributed datasets, which are organized by categories as introduced in the previous section. Note that labels are not final – additional labels will be included to describe datasets to be added to the repository. The first category includes the following labels: RFID tags, localization sensors, accelerometers, gyroscopes, magnetometers, infrared motion capture sensors and location sensors. The labels in the second category, i.e., the task, specify what kinds of problems are present in the dataset. Labels in this category are activity recognition, gesture recognition, posture recognition, gait recognition, fall detection and health analysis. Additionally, we included a category, which specifies the learning approach. The labels in this category divide the datasets into those which are annotated, so one can use supervised learning algorithms, and into those without annotations, which must be used only with unsupervised learning algorithms.

On the front page the user can obtain useful information such as a list of recently added datasets. There is also a part of the page which is dedicated to interesting links, for example, to web sites for activity recognition competitions, to conferences and workshops dedicated to ambient intelligence, and to other online repositories and resources relevant to ambient intelligence (e.g., Geib, 2012; Hu 2008).

The second part of the repository consists of the descriptions of the individual datasets. Each dataset is described with several sections. The first section contains downloadable links, where the users can download the dataset's description containing detailed information about the dataset, and a dataset file, which should be preferably in comma separated (.csv) format.

The second section contains a dataset abstract briefly describing the purpose and the structure of a dataset. The third section contains the labels of categories, i.e., sensor types, tasks, domain, and learning approach. Next, there are general, mostly numerical information about the dataset such as: (1) the number of instances, (2) the frequency of data acquisition, (3) the number of labels used for annotating the dataset, (4) information whether missing values are present in the dataset, (5) the total length of recordings in minutes, (6) the number of attributes, (7) the number of sensors, (8) the number of people recorded, (9) the date when dataset was donated, and (10) the number of visitors and downloads. In future, the dataset can be described by a complexity measure, for example, as proposed by Sahaf et al. (2011).

The fifth section is the experimental setup. In this section the dataset's contributor explains in detail the sensor placement, the scenario by which recordings were performed, and any other features, which can help a user to understand the dataset.

In the sixth section the dataset structure and tools are described. In dataset structure field the attributes are explained. The tools for manipulating the data should be provided by contributors in case the dataset is not in the standard format or if the data needs any preprocessing before it can be used in other applications.

In the seventh section the relevant papers are listed. This can be added by anyone who used the dataset in their experiments. For each added paper, the task tackled, the experimental settings (e.g., cross validation parameters, train and test data size, etc.), and the results achieved are briefly presented.

The last section is reserved for citation request. The contributor or the author may request citation of their selected paper when the dataset is used to produce the results for a new paper.

The current implementation of the repository does not allow automatic submission of new datasets. The contributors are asked to send an email to the authors of this paper. In future, however, we plan to implement this option, as well as interface for editing the information about the contributed dataset.

Current Datasets

Currently, we contributed three sample datasets to the repository (Dovgan et al., 2011). They are focused on activity recognition from body-worn sensors. Figure 1 shows statistical data of the available datasets in the repository. The first two datasets use the same type of sensors (3D coordinates extracted from visual markers), but differ in the recorded scenarios. The last dataset uses wireless location tags and scenarios similar to those in the second dataset.

Activities and Falls, SMART, Phase I

The dataset contains recordings of short scenarios suitable for activity recognition, fall detection and the detection of limping. They consist of the coordinates of 12 markers attached to the bodies of three volunteers. The recordings were made with the Smart infrared motion capture system (Smart, 2012) consisting of six infrared cameras and infrared light sources. Three volunteers were equipped with markers reflecting infrared light. The markers were

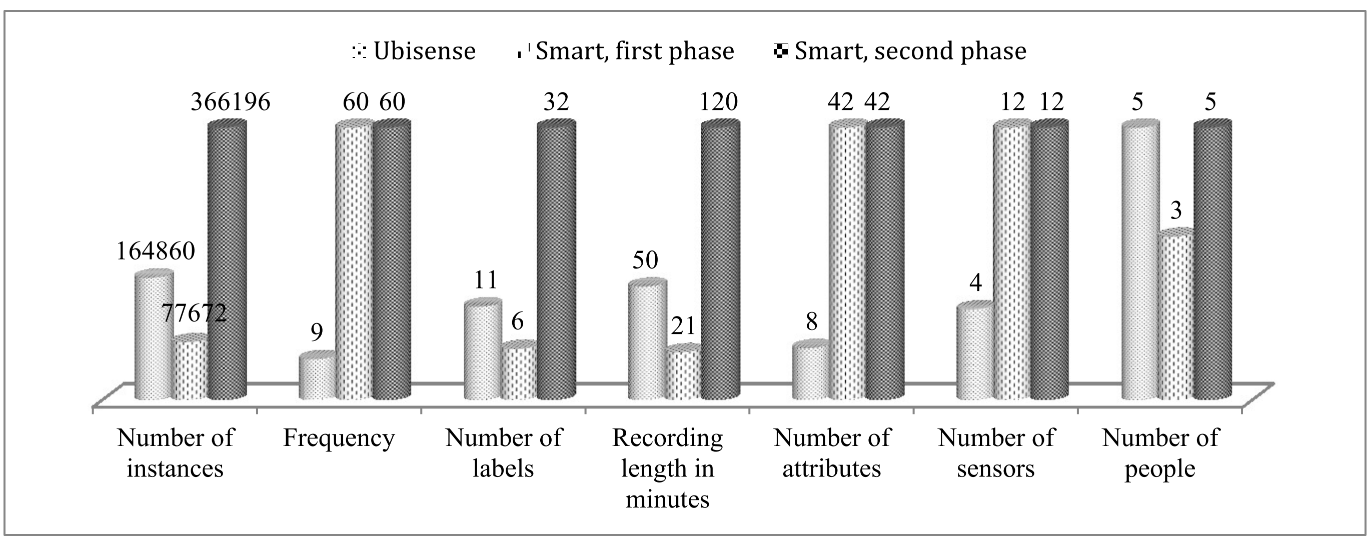

Figure 1: Statistical data presentation of datasets currently available in the repository.

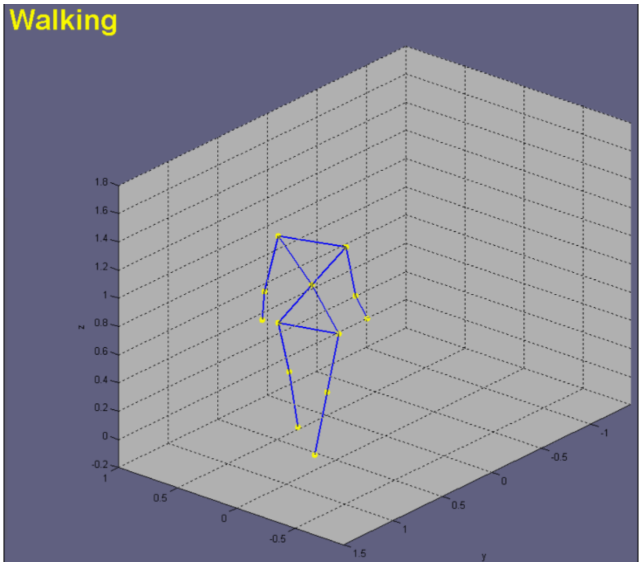

Figure 2: Reconstruction of body markers when person is walking.

attached to both ankles, knees, hips, shoulders, elbows and wrists. They were tracked with the cameras, and their 3D coordinates were estimated with roughly 1 mm accuracy. Reconstruction of body markers while person is walking is shown in Figure 2.

Activities, Falls and Other Health Problems, SMART, Phase II

This dataset was captured with the same hardware as the first one. The differences between the datasets are in recording scenarios and in the number of recordings. This dataset includes additional recordings of specific diseases such as Parkinson, hemiplegia, pain in the back, epilepsy, etc.

Activities and Falls, Ubisense

This dataset contains recordings of short scenarios suitable for activity recognition and fall detection. The recordings were made with an Ultra-Wideband technology Ubisense (Ubisense, 2012). They consisted of the coordinates of 4 tags attached to the bodies of five volunteers. The tags were attached to the chest, belt and both ankles. Each tag returns the coordinates of their position with approximately 6-9 Hz.

Conclusion

This paper proposed an online repository of datasets that would allow the community to compare newly developed approaches and to quickly start building upon the latest advances in ambient intelligence. Prospective dataset contributors are encouraged to participate and donate their datasets as well as additional resources such as implemented methods and preprocessing tools. We also appeal to the community to provide feedback and ideas as how to broaden the repository in order for it to serve its purpose.

References

Activity recognition repository 2012. http://dis.ijs.si/ami-repository/.

Chessa, S., Gulliem, S.. Evaluating AAL Systems through Competitive Benchmarking 2011. http://evaal.aaloa.org/.

Davis, L., Hoogs, L. Activity Recognition Competition Workshop 2011. http://www.umiacs.umd.edu/conferences/cvpr2011/ARC/.

Dai, L., Desai, M., Soatto, S., Hoogs, A. International Workshop on Frontiers of Activity Recognition , http://vision.ucla.edu/ActionWorkshop.html, 2010.

Dovgan, E., Luštrek, M., Pogorelc, B., Gradišek, A., Burger, H., and Gams, M. 2011. Intelligent elderly-care prototype for fall and disease detection. *Slovenian Medical Journal* 80(11): 824-831.

Geib, C.W. Resources for Plan, Activity, and Behavior Recognition Research 2012. http://homepages.inf.ed.ac.uk/cgeib /PlanRec/Resources.html.

Hu, D.H. Activity Recognition: Datasets, Bibliography and others, http://www.cse.ust.hk/~derekhh/ActivityRecognition/ index.html, 2008.

International Workshop on Frontiers of Activity Recognition 2012. http://vision.ucla.edu/ActionWorkshop.html.

Kurz M., Hölzl G., Ferscha A., Calatroni A., Roggen D., Tröster G., Sagha H., Chavarriaga R., Millán J. R., Bannach D., Kunze K., and Lukowicz P. 2011. The OPPORTUNITY Framework and Data Processing Ecosystem for Opportunistic Activity and Context Recognition. *International Journal of Sensors, Wireless Communications and Control* 1: 102-125.

Oh, S., Hoogs, A., Perera, A.G.A., Cuntoor, N.P., Chen, C., Lee, J.T., Mukherjee, S., Aggarwal, J.K., Lee, H., Davis, L., Swears, E., Wang, X., Ji, Q., Reddy, K.K., Shah, M., Vondrick, C., Pirsiavash, H., Ramanan, D., Yuen, J., Torralba, A., Song, B., Fong, A., Roy-Chowdhury, A., Desai, M.: A large-scale benchmark dataset for event recognition in surveillance video (VIRAT video dataset). In CVPR 2011 , pp. 3153-3160, 2011.

Opportunity. Opportunity Activity Recognition Challenge, 2011. http://opportunity-project.eu/challenge.

Park, S., Kautz, H. Privacy-preserving recognition of activities in daily living from multi-view silhouettes and rfid-based training. In AAAI Symposium on AI in Eldercare: New Solutions to Old Problems, 2008.

Sahaf Y., Krishnan N. C., Cook D. J. 2011. Defining the Complexity of an Activity. AAAI Workshops. Available at: http://www.aaai.org/ocs/index.php/WS/AAAIW11/paper/view/39 46.

Smart infrared motion capture system 2012. http://www.btsbioengineering.com/BTSBioengineering/Kinemati cs /BTSSMARTDX/BTS_SMARTDX.html.

Ubisense real-time location system 2012. http://www.ubisense.net/en/rtls-solutions/.