An 'Object-Use Fingerprint': The Use of Electronic Sensors for Human Identification

Mark R. Hodges and Martha E. Pollack

Computer Science and Engineering
University of Michigan, Ann Arbor, MI, USA
{hodgesm,pollackm}@umich.edu

Abstract. We describe an experiment in using sensor-based data to identify individuals as they perform a simple activity of daily living (making coffee). The goal is to determine whether people have regular and recognizable patterns of interaction with objects as they perform such activities. We describe the use of a machine-learning algorithm to induce decision-trees that classify interaction patterns according to the subject who exhibited them; we consider which features of the sensor data have the most effect on classification accuracy; and we consider ways of reducing the computational complexity introduced by the most important feature type. Although our experiment is preliminary, the results are encouraging: we are able to do identification with an overall accuracy rate of 97%, including correctly recognizing each individual in at least 9 of 10 trials.

1 Introduction

A body of recent work has focused on the use of sensors to recognize the performance of particular human activities [1,2,3,4,5,6]. This paper describes work that also uses sensors to monitor human activity, but towards a different end: our goal is to identify individuals from their behavior. More specifically, we seek to determine whether individual people have regular and recognizable patterns of interaction with the objects they use in performing daily activities and whether these patterns create an "object-use fingerprint" that can be used for human identification.

The primary motivation for this work is a scientific one: it is interesting in its own right to know whether people interact in regular ways with objects and, if so, what these regularities are. However, there are also potential pratical implications of the work. As one example, imagine wanting to gather information about the use of various devices (a new model refridgerator or a newly designed copy machine, for instance). One way to do this would be to place the device in an open area (an office kitchen or mailroom) and gather data as multiple people interact with the device. If people indeed have "object-use fingerprints," then it should be possible to distinguish amongst different users in the collected data without having to actually identify who those users are.

J. Krumm et al. (Eds.): UbiComp 2007, LNCS 4717, pp. 289–303, 2007.

A second motivation for this work comes from our long-term goal of employing sensor-based activity monitoring to track the performance of individuals at risk for cognitive decline, so as to detect changes that may indicate a change in cognitive status. As a preliminary step, we sought to determine the degree to which people behave in regular and predictable ways while performing common activities. Of course, the existence of identifiable patterns of object use by individuals is by itself neither necessary nor sufficient for our larger goal. However, we feel that establishing that individuals have some degree of regularity in their object-use patterns makes it more likely that one can observe trends of deviation from those regularities, and in turn, learn deviations that may indicate cognitive decline.

To test our hypothesis–that individuals have "object-use fingerprints," making it possible to identify them on the basis of the way in which they interact with objects during the performance of routine daily activities–we conducted a series of experiments in a controlled laboratory setting. Specifically, we had individuals wear an RFID reader affixed to a glove while they made coffee in a kitchen in which objects were instrumented with RFID tags. We then applied machine-learning techniques to classify interaction patterns. Key research questions included determining the features of the activity that were most predictive of an individual's identity, analyzing the computational time required for the learning algorithm to process each type of feature, and developing strategies for reducing reliance on the most computationally expensive features. This was an initial investigation with a number of simplifying assumptions that we describe in section 3.4. Our results are therefore preliminary but encouraging nonetheless.

The next section briefly reviews prior work on sensor-based activity recognition. Section 3 discusses the methodology of our experiment, including the selection of the task, selection of the sensing technology, the experimental setup, and the limitations of the experiment. Section 4 describes the machine-learning techniques used to analyze the collected data. Section 5 presents the experimental results, which are then discussed in Section 6, which also presents avenues for future research.

2 Background

Automated activity recognition encompasses techniques using different types of sensors that detect activities at a wide range of granularity. One approach is to use a few extremely data-rich sensors such as video cameras or microphones. For example, Ben-Arie, et al. employ video cameras to distinguish among eight basic actions including jumping, walking, and picking an object up [7]. For many potential applications, however, data-rich sensors, and especially video cameras, may be problematic, in part because they have been shown to provoke strong privacy concerns [8].

A contrasting approach is to use a large number of very basic sensors that each capture only limited information. Accelerometers and other related sensors that are worn by individuals may be used to differentiate actions. For instance, Bao and Intille used five biaxial accelerometers, positioned at key places on subjects’ bodies, to identify with high accuracy an activity from among twenty types of activities including vacuuming, reading, and watching TV. Their description of this work also provide a good overview of past work in activity recognition using accelerometers [9]. Wren and Tapia use motion detectors and hierarchical learning to recognize actions in a work environment such as walking in straight line, turning, and “joining” (coming into the same geographic region as another person) [6]. Liao et al. take an different approach and use GPS and Relational Markov Networks to recognize the locations of activities such as “home,” “work,” and “shop” [2,10].

The use of RFID readers as a device for recgonizing activities has also become popular. What is intersting about this approach is that it focuses on recognizing the objects with which a person interacts, rather than on monitoring the person’s movements directly. Thus, Philipose and his colleagues have subjects wear a glove or bracelet with an attached RFID reader, which can sense objects in the environment that have RFID tags affixed to them [3,4,8]. Their sensing technology–an RFID reader worn on the hand–served as the inspiration for the RFID glove used in this paper. Tapia et al. also use RFID technology and naive Bayesian classifiers to perform in-home activity recognition [5].

The work just described has all addressed the question of identifying an activity, not of identifying a person. To the best of our knowledge, there has not been work done on identifying individuals based on their object-interaction patterns. However, prior work has been done on identification of individuals using other biometrics. Keystroke dynamics have been studied as a way to provide additional computer security, with a particular focus on using keystroke dynamics to harden password security. Peacock provides a good overview of the work in this area [11]. Vision-based identification techniques have also garnered significant interest to supplement the physical security of environments such as airports and banks. These efforts have focused mainly on gait recognition, which uses patterns of movement for identification [12], as well as on automatic face recognition using a photograph or a series of photographs to identify individuals [13].

3 Methodology

3.1 Selection of Task

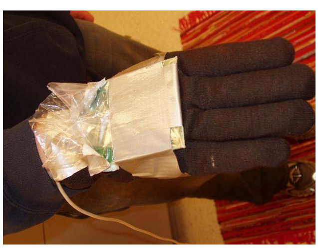
Several criteria were used in the selection of a task for subjects to perform. Most obviously, we sought a task that was an activity of daily living, and was one that that is performed by many people on a regular basis. In addition, an ideal task would be relatively constrained in terms of the ways in which it might be performed, but would also have some natural variance in performance (not as broad as “prepare dinner,” but not so narrow as “pour a glass of milk”). Finally, it should be possible to perform the task in an instrumented laboratory. The task of making a cup of coffee was chosen for this experiment since it is an excellent fit for all of these criteria.

3.2 Selection of Technology

As noted earlier, Radio Frequency Identifier (RFID) technology has been used successfully in several activity recognition projects and we thus chose to use it here as well. RFID equipment consists of tags, which can be placed throughout an environment, and readers which detect nearby tags. A key advantage of this type of sensing is that RFID has a 0% false positive rate. In addition, tags are inexpensive (less than US$0.20) and small in size (approximately the size of a postage stamp). There are two types of RFID tags, active and passive. Active tags are extremely accurate but require a power source. Passive tags, on the other hand, are not detected as reliably, but do not require a power source, instead harvesting energy from the reader to sense and communicate [14]. For this reason, they can be placed throughout an environment without a need for cords or batteries that will need to be replaced.

In earlier work done at Intel laboratories, an RFID reader placed on a glove or bracelet was used to detect detect tags that are in close proximity to the hand (within 10cm) [15,3,4,8,16]; detected objects are assumed to be ones with which the user is interacting. This form factor also has added value with regards to privacy. If a user wishes to prevent the system from observing her, she may simply remove the glove or bracelet containing the RFID reader. In addition, the short range of the reader makes it possible to observe fine-grained patterns in the way the object is held, i.e. whether an object is held from the side or from the bottom. This information is potentially valuable in identifying people from the object interactions. Because Intel's wireless iBracelet was not available in time for use in this study, a wired system was used, consisting of an off-the-shelf RFID reader and tags created by Phidgets, Inc. ®.

The sensor glove is depicted in Fig. 1 (l), while one of the tagged objects-a coffee grinder-is shown in Fig. 1 (r). Obviously, the glove as used in the experiment would not be appropriate for actual use in a home setting–and not only because of the attached wire! Nonetheless, it was satisfactory for collecting the data we needed for these experiments.



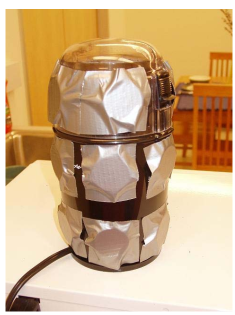


Fig. 1. (l) The glove with an RFID reader attached. (r) The coffee grinder with several RFID tags attached.

3.3 Experimental Setup

Ten subjects were recruited to participate in the experiments. For each trial, the subject was instructed to make a cup of coffee as if about to drink it, including adding sugar and creamer as preferred. Subjects wore a glove outfitted with an RFID reader on their right hand, but were told to ignore it as best they could and use both hands as they typically would. Each subject participated in ten trials, spaced out with generally at most one per day, so that the trials would reflect normal patterns of use, rather than artificial patterns created by performing trials repeatedly one after another[1].

Subjects were given a brief tour of the instrumented lab before their first trial, and those who did not know how to make coffee were given basic instructions. These instructions were as general as possible. For example, subjects were told to "put water in the reservoir at the back of the coffee-maker," rather than indicating exactly how the water should be put there, so they would choose for themselves whether to use the coffee cup or coffee carafe to transport the water from the sink. No physical demonstration of the coffee-making process was given.

The experimental set-up consisted of a coffee maker, one cup, one spoon, a coffee bean grinder, and a cabinet containing a bag of filters, a bag of coffee grounds, a bag of coffee beans, and a canister each of cream and sugar. Each item was tagged with multiple RFID tags and before each trial was put in the same place, facing the same direction. (The bag of filters did not have an obvious front and thus may have been reversed between trials).

3.4 Experimental Limitations

This study is an initial investigation and, as such, had a number of design simplifications. Possibly most significantly, the task was performed in a controlled setting–our laboratory–rather than in a naturally occurring environment. The environment was very regular, so that each time a subject began a task, all the objects he or she might use were in the same location and aligned in the same direction, without the natural variation that would occur in a real-world setting, especially in an environment shared by several users. The users performed only a single task–making coffee–and this task was not interleaved with other tasks nor was it interrupted by outside influences or distractions such as a ringing telephone. Finally, the problem of segmenting tasks was avoided by starting the trial immediately before the task began and stopping it immediately after the task was finished.

There is no question that these limitations may affect the generality of the results. Nonetheless, we believe that the results of our preliminary study, as presented below, are encouraging enough to support follow-on work that would determine their generality.

[1] In some cases, the availability of the subject required more than one trial per day; five subjects performed two trials on the same day at least once and one of those performed six trials on the last day of the subject's availability.

4 Machine Learning Approach

The data collected by the sensors during each subject's performance of the coffee-making task was input to a machine-learning algorithm for classification. We used a decision-tree induction algorithm for this purpose, primarily because it is the simplest form of classification algorithm and thus provided a reasonable starting point for our investigation. A decision tree is a classifier, taking a set of properties as input and outputting a "decision" by following paths on a tree, starting at the root and working to a leaf node. Internal nodes in this tree are a test of the value of a property, and the branches from that node represent the possible values of the test. The leaf node is the decision reached by the tree. In our study, we used the C4.5 decision-tree induction system, which is based on the ID3 algorithm. C4.5 particularly attempts to avoid overfitting, the condition where an overly complex tree is created that is less accurate than a simpler one would be [17].

A key question then is what features of the sensor data should be used as input to the classifier. We investigated two types of features for sensor data: observation granularity and interaction measure.

Observation granularity has to do with how abstract our observations are: do we need to provide information to the machine learning algorithm about the fact that a subject touched the coffeepot or the fact that she touched the left side of the lid (or both)? Many of the objects used in the study had multiple tags affixed to them, and we thus considered observations of interactions at three layers of abstraction:

1. Tag: Detected interaction with an individual tag affixed to an object;
2. Group: Detected interaction with any of a group of tags that are equivalent except for the orientation of the object (e.g., the tag on the left side of the coffee grounds and the tag on the right); and
3. Object: Detected interaction with an object (e.g., with any of the tags that are affixed to the coffeepot).

Table 1 shows the objects we used, and the tag groups associated with each object. Note that for some objects, like the mug, there is only one tag, so the tag, group, and object are all the same.

Independent of the level of abstraction, there are also different ways in which we can measure interactions; we explored five types of features that measure interactions, applying them to each of the levels of granularity:

1. Detected: A binary feature that is positive iff there was any interaction with a tag, group, or object.
2. Count: A scalar feature that records the number of interactions with a tag, group, or object.
3. Total Duration: A scalar feature that records the total amount of time of interaction with a tag, group, or object.

Table 1. List of Tag Groups and Objects

Objects	Groups
Coffee Maker	Lid, power switch
Carafe	Carafe
Mug	Mug
Spoon	Spoon
Coffee Grinder	Top row of tags, middle row, bottom row
Left Cabinet Door	Left cabinet door
Right Cabinet Door	Right cabinet door
Coffee Grounds	Top tags on front and back, bottom tags on front and back, tag on bottom, tags on sides
Coffee Beans	Top tags on front and back, bottom tags on front and back, tag on bottom, tags on sides
Filters	Tags on sides, tag on bottom
Creamer	Tags on top row, tags on bottom row
Sugar	Tags on top row, tags on bottom row
Faucet	Faucet

4. Average Duration: A scalar feature representing the average time of interaction with a tag, group, or object: this is a computed feature, equal to Total Duration divided by Count.
5. Order: A binary feature that is positive iff an arbitrary two- or three-tag, group, or object ordering is observed.

The order feature deserves a little more explanation. It determines whether a specific ordering of interactions is observed in a trial. This ordering may consist of two or three tags, groups, or objects, but within a specific ordering only one level of granularity is used. Because there are 70 tags, 25 groups, and 13 objects in the experiment, over 300,000 possible orderings exist. As a result, considering all possible orderings of tags comes at a significant cost in performance. This performance cost will be discussed further in Sections 5 and 6.

Although the RFID reader and tag system provides accurate and generally reliable results, an individual tag is sometimes found and lost in quick succession, either when it is near the maximum distance from the reader at which it can be sensed, or if the reader moves rapidly. In order to smooth the data, a pre-processing step can be performed on each trial prior to analysis. The pre-processing step looks for consecutive accesses of the same tag, group, or object within 0.5 seconds. When this is found, the records of the two accesses are merged into one, hopefully providing a more accurate model of the subject's actual behavior. This means that when a subject quickly draws her hand away from a tag and then puts her hand back on the tag, the action will be interpreted as one continuous interaction, and that when a subject moves her hand over several tags on the same object, the action will be interpreted as one continuous interaction at the object level (analysis on the tag feature level will not be impacted in this case). In Section 5.5 we describe the effect of this pre-processing.

We performed one additional type of pre-processing on the collected data. The task of making coffee was selected in part due to the natural variation in how people perform the task. Although more obvious indicators, such as whether people put cream or sugar in their coffee, might be considered valid differences in behavior, it is a more interesting question to ask if an individual can be determined without that information. For that reason, in most of our analyses, we removed from the data all information about cream and sugar tags. (We indicate places where this is not true.) Similarly, a subject's choice of grinding whole beans verses using grounds can be used to distinguish amongst subjects. Because removing information from the tags for the grounds, beans, and grinder from the trial would remove one-fifth of the data collected, we simply note that only one user used whole beans, and that user used whole beans in every trial. For that reason, one of the ten users can be distinguished very easily from the others.

By using ten subjects who each perform ten trials, we obtained 100 cases for analysis. We used a ten-fold cross-validation process, repeatedly using 90 of the trials as training data for C4.5 and reserving the remaining 10 trails as test data. In each iteration, the training data contained 9 trials for each subject, with the tenth reserved for testing data; however, our learning system did not use the information that there is exactly one trial per subject during the classification process.

5 Results

5.1 Full Feature Set

We begin by describing the results obtained when C4.5 is run using the full set of 15 features (5 feature types, applied to each of the three layers of abstraction) and including the cream and sugar information. In this case, the system is extremely accurate, correctly recognizing the subject in 97 of the 100 trials (again, under a 10-fold cross validation experiment). Two of the incorrectly-recognized trials were performed by the same subject, meaning that the system correctly identified at least 8 of each user's 10 trials. 8 of the 10 users were identified accurately in all 10 of their trials, with the remaining participant correctly identified in 9 of the 10 trials. Table 2 shows the confusion matrix for this feature set.

The ten decision trees produced here have an average of 10.8 internal nodes and an average maximum depth of 7.3. Well over half the internal nodes consider order features and every observation granularity appears in at least one tree, while count is the only interaction measure that does not appear in any of the trees. Figure 2 shows one of the trees produced. In this case, all but one of the ten internal nodes use order.

Perhaps surprisingly, the system has the same accuracy when using the full set of 15 features, but now ignoring the use of cream and sugar, with the subjects again being correctly identified in 97 of the 100 trials. In this case, all three of the incorrectly identified trials were from different users, meaning three users were each recognized correctly in 9 of their 10 trials, and the other seven were

Table 2. Confusion matrix of full feature set using cream and sugar

	Inferred Identity									
Truth	A	B	C	D	E	F	G	H	I	J
A	10									
B		10								
C			10							
D				10						
E				1	9					
F	1			1		8				
G							10			
H								10		
I									10	
J										10

```
Detected Tag: Coffee Grinder 1-2
├──true: Subject I
└──false: Order Object: Coffee Maker - Grounds - Coffee Maker
   ├──true: Subject I
   └──false: Order Tag: Coffee Maker Lid 4 - Grounds
      ├──true: Subject B
      └──false: Order Object: Right Cabinet Door - Mug - Right Cabinet Door
         ├──true: Order Object: Carafe - Mug - Carafe
         │  ├──true: Order Object: Filters - Right Cabinet Door - Filters
         │  │  ├──true: Subject C
         │  │  └──false: Subject G
         │  └──false: Subject A
         └──false: Order Object: Sugar - Mug
            ├──true: Subject C
            └──false: Order Group: Carafe - Filters Bottom - Carafe
               ├──true: Subject E
               └──false: Order Object: Grounds - Carafe
                  ├──true: Order Tag: Carafe - Right Cabinet Door
                  │  ├──true: Subject J
                  │  └──false: Subject D
                  └──false: Subject F
```

Fig. 2. One of the decision trees produced with the full feature set using cream and sugar

recognized in all of their trials. The confusion matrix for this feature set is given in Table 3. We hypothesize that the sugar and cream did not improve performance because of a ceiling effect: classification was so high (97%) even without the specific information about sugar and cream that there was no room for improvement.

The ten trees produced here are slightly larger and slightly deeper than those produced when cream and sugar is considered. They have an average of 12.2 internal nodes and an average maximum depth of 7.8. Again, well over half the internal nodes consider order features, and in this case every type of interaction measure is used in at least one tree as is every observation granularity.

Table 3. Confusion matrix of full feature set without cream and sugar

	Inferred Identity									
Truth	A	B	C	D	E	F	G	H	I	J
A	10									
B		10								
C			10							
D				10						
E					9					1
F				1		9				
G							10			
H	1							9		
I									10	
J										10

Although the system is very accurate when using the full feature set, performance is not ideal. With the cream and sugar data deleted, the system takes an average of 8 minutes to compute the features and learn one tree, then perform subject identification for ten trials. Including the cream and sugar data degrades performance significantly, requiring an average of 40 minutes for the same task. Since performing identification is still very quick, taking less than one second, this performance may be considered acceptable. However, as we describe further in section 5.4, even learning time must be bounded; moreover, the performance problems are likely to be exacerbated by more complex environments.

5.2 Influence of Observation Granularity

Next, we consider the system's success when only certain subsets of the features are used in the identification process. There are three reasons for doing this. First, as noted above, the performance of the system is relatively slow, and so it is important to determine whether it can be improved by using fewer features without sacrificing accuracy. In particular, there are a huge number of order features, so it is valuable to understand their impact on performance. Second, it is useful to know if the number of tags in the environment can be reduced, both for aesthetic reasons and so that we can anticipate the effects of tags falling off during use. Third, the time spent identifying groups and objects may not be necessary if performance does not degrade when only tag-level features are used.

To begin, we consider the influence of features at different layers of abstraction; the results are shown in Table 4, which provides the accuracy of the system in performing subject identification when using only tag-level observations, group-level observations, and object-level observations. In this analysis we make use of all interaction measures, and we delete the cream and sugar data.

As can be seen, when we restrict the algorithm to object-level data, accuracy is essentially unchanged from that using the full feature set, but it degrades noticeably when using only tag or group level observations. Interestingly, this is due primarily to the inability of the system to classify a particular subject

Table 4. Comparison of feature levels

Observation Granularity	Accuracy
Tag	87%
Group	88%
Object	97%

(subject F). Using only interactions at the tag level, the system correctly identifies that subject only once, while using interaction at the group level it correctly identifies the same subject just twice. That subject’s other trials at the tag and group level are all misidentified as subject D, indicating that those two may be very difficult to distinguish (although subject D is never misidentified). The accuracy of both feature sets on the other subjects remains very high at 95.6%.

5.3 Influence of Interaction Measure

We next consider the importance of each of the types of interaction measures; again, in all our analyses here, we omit the cream and sugar data. Table 5 gives the results, showing both the overall accuracy when using only a single interaction measure, and a list of subjects who were correctly identified in 5 or fewer of their 10 trials.

Table 5. Comparison of individual features

Interaction Measure	Accuracy	Subjects Correctly Identified in $\leq$ 5 Trials
Accuracy Detected	73%	A, C, E, J
# of Times Detected	75%	A, C, E
Total Duration	85%	F
Average Duration	72%	E, F, H
Order	93%	

Most single interaction measures result in difficulty in identifying some subjects, but no problems in identifying others. Subject F, the one consistently misidentified in the tag and group feature sets, is never identified correctly in the Total Duration feature set, and only once in the Average Duration feature set. However, the same subject is correctly identified 9 of 10 times in the # of Times Detected feature set, 8 of 10 times in the Detected feature set, and every time in the Order feature set.

5.4 “All But Order”

As described in section 5.1, the time to learn one tree and perform subject identification for ten trials can take up to forty minutes. Because the actual identification is still performed in under one second, this performance may be considered acceptable since computing the features and learning the tree would not need to happen in real time. However, even learning time must be bounded; moreover, the performance problems are likely to be exacerbated by more complex environments. Observing subjects performing larger and more complex tasks may require several times as many sensors as were used in this experiment. Additionally, allowing a user to interleave actions from multiple tasks may prevent a system from simplifying the learning process by only considering the sensors relevant to a single task.

While the number of features computed for other interaction measures grow linearly, order undergoes cubic growth since using order involves generating and considering a large number of possible two- and three-step sequences. We thus repeated the analysis, using the full feature set except for the order features. This analysis is also important to answering the question of what features are important in performing identification.

As expected, processing time decreases significantly in this case: it takes, on average, less than five seconds to learn one tree and perform identification ten times, a speed up of two orders of magnitude from the eight minutes needed when all features are considered. Unfortunately accuracy also decreases, as shown in Table 6 (which again omits the cream and sugar data). Subjects correctly identified in 5 or fewer of their 10 trials are also listed, and again we see that specific subjects are consistently misidentified while the others are identified with high accuracy.

Table 6. Comparison of 'all but order' feature sets

Observation Granularity	Accuracy	Subjects Correctly Identified in ≤ 5 Trials
All Levels	86%	D, E
Tag	86%	D, E
Group	80%	D, E
Object	80%	E, H

The removal of the order features results in a pattern of observation granularity that is the inverse of that seen when order features are included: now the use of only object-level observations produces the lowest, rather than the highest level of accuracy, suggesting that there is an important interaction between order and object-level observations.

5.5 The Influence of Pre-processing

A final analysis concerns the pre-processing step that we described earlier to smooth the data by combining consecutive interactions that take place with the same tag, group, or object in rapid succession. The results are shown in Table 7 (again, cream and sugar deleted).

In turns out that our pre-processing technique does not have the intended effect: in some cases it decreases accuracy slightly, and it only increases accuracy when we look just at average duration. Because consecutive usage of the

Table 7. Effects of pre-processing

Interaction Measure	Accuracy Before Pre-Processing	Accuracy After Pre-Processing
All Features	97%	95%
Detected	73%	73%
# of Times Detected	75%	76%
Total Duration	85%	81%
Average Duration	72%	81%
Order	93%	93%

same item should not affect the Detected or Order features, their accuracy, by definition, remains unchanged. The increased accuracy of average duration is expected because it is the feature most affected by the rapid finding and losing of a tag. The decreased accuracy of total duration is a surprise, however, since using pre-processing should increase how accurately duration is measured by filling gaps in detected usage that are probably not gaps in actual usage. We are uncertain at this time as to how to explain that result.

6 Discussion

The motivation for this work was to determine whether individuals have predictable object-use patterns as they carry out activities of daily living, and to determine whether these patterns could be used for identification. Although the experiment presented here is preliminary, in that it only involves 10 subjects and the performance of a single type of task in a controlled environment, it is nonetheless promising in suggesting the existence of such regularities. There are no subjects that the machine learning algorithm has trouble identifying with the best-performing feature sets. Additionally, there are no trials that are misidentified by all four of the best feature sets. However, when the computationally costly order features are omitted, the level of accuracy varies with subjects. Thus, a challenge for future research is to automatically learn which features sets are best at identifying the regularities of a given individual's behavior.

One way to address this challenge would be to reduce the number of order features required. In general, the number of possible orderings grows quickly with the number of items (tags, groups, or objects), ($O(n^3)$). However, only a small subset of these orderings are important to subject identification. Thus, rather than include all or none of the possible orderings as input to the machine-learning algorithm, the application of domain-specific knowledge to identify interesting orderings may be viable. In the trials run with every feature at every feature level and ignoring cream and sugar, order was used an average of 8.4 times per tree (84 times total). Several orderings were used multiple times, though, including one that was used 9 different times and there were only 35 unique orderings used. If one were able to identify most of these “relevant” orderings, then classification time could be significantly reduced without a penalty in accuracy.

Finally, we note that the use of object-level observations provided the highest accuracy of the three layers of abstractions, but every individual interaction measure except order performed better using tag-level observations. This suggests the added resolution of using a large number of short-range tags may be more beneficial than placing a smaller number of longer-range tags on each object. It also implies much of the benefit of using order can be gained by just considering order at the object-level, and this may also greatly reduce the computational costs of processing. In addition, the process of identifying relevant orderings may be simplified by focusing on those that occur at the object level.

A key area for future work involves replicating the experiment described here on more and different types of subjects, on a broader range of activities, and in naturalistic settings, so as to validate the generality of our preliminary results.

References

1. Barger, T., Alwan, M., Kell, S., Turner, B., Wood, S., Naidu, A.: Objective remote assessment of activities of daily living: Analysis of meal preparation patterns (poster presentation)
2. Liao, L., Fox, D., Kautz, H.: Location-based activity recognition. In: Weiss, Y., Schölkopf, B., Platt, J. (eds.) Advances in Neural Information Processing Systems 18, pp. 787–794. MIT Press, Cambridge, MA (2006)
3. Patterson, D.J., Fox, D., Kautz, H., Philipose, M.: Fine-grained activity recognition by aggregating abstract object usage. In: Gil, Y., Motta, E., Benjamins, V.R., Musen, M.A. (eds.) ISWC 2005. LNCS, vol. 3729, pp. 44–51. Springer, Heidelberg (2005)
4. Philipose, M., Fishkin, K.P., Perkowitz, M., Patterson, D.J., Fox, D., Kautz, H., Hahnel, D.: Inferring activities from interactions with objects. IEEE Pervasive Computing 3(4), 50–57 (2004)
5. Tapia, E.M., Intille, S.S., Larson, K.: Activity recognition in the home using simple and ubiquitous sensors. In: Ferscha, A., Mattern, F. (eds.) PERVASIVE 2004. LNCS, vol. 3001, pp. 158–175. Springer, Heidelberg (2004)
6. Wren, C.R., Tapia, E.M.: Toward scalable activity recognition for sensor networks. In: Hazas, M., Krumm, J., Strang, T. (eds.) LoCA 2006. LNCS, vol. 3987, pp. 168–185. Springer, Heidelberg (2006)
7. Ben-Arie, J., Wang, Z., Pandit, P., Rajaram, S.: Human activity recognition using multidimensional indexing. IEEE Trans. Pattern Anal. Mach. Intell. 24(8), 1091–1104 (2002)
8. Fishkin, K.P., Philipose, M., Rea, A.: Hands-on rfid: Wireless wearables for detecting use of objects. In: ISWC 2005, pp. 38–43. IEEE Computer Society, Los Alamitos (2005)
9. Bao, L., Intille, S.S.: Activity recognition from user-annotated acceleration data. In: Ferscha, A., Mattern, F. (eds.) PERVASIVE 2004. LNCS, vol. 3001, pp. 1–17. Springer, Heidelberg (2004)
10. Patterson, D.J., Liao, L., Gajos, K., Collier, M., Livic, N., Olson, K., Wang, S., Fox, D., Kautz, H.A.: Opportunity knocks: A system to provide cognitive assistance with transportation services. In: Davies, N., Mynatt, E.D., Siio, I. (eds.) UbiComp 2004. LNCS, vol. 3205, pp. 433–450. Springer, Heidelberg (2004)
11. Peacock, A., Ke, X., Wilkerson, M.: Typing patterns: A key to user identification. IEEE Security and Privacy 2(5), 40–47 (2004)

12. Wang, L., Tan, T., Ning, H., Hu, W.: Silhouette analysis-based gait recognition for human identification. IEEE Trans. Pattern Anal. Mach. Intell. 25(12), 1505–1518 (2003)
13. Arandjelovic, O., Shakhnarovich, G., Fisher, J., Cipolla, R., Darrell, T.: Face recognition with image sets using manifold density divergence. In: CVPR ’05: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05), vol. 1, pp. 581–588. IEEE Computer Society Press, Washington (2005)
14. Finkenzeller, K.: RFID Handbook: Fundamentals and Applications in Contactless Smart Cards and Identification. John Wiley & Sons, Inc., New York, NY, USA (2003)
15. Fishkin, K.P., Jiang, B., Philipose, M., Roy, S.: I sense a disturbance in the force: Unobtrusive detection of interactions with rfid-tagged objects. In: Davies, N., Mynatt, E.D., Siio, I. (eds.) UbiComp 2004. LNCS, vol. 3205, pp. 268–282. Springer, Heidelberg (2004)
16. Smith, J.R., Fishkin, K.P., Jiang, B., Mamishev, A., Philipose, M., Rea, A.D., Roy, S., Sundara-Rajan, K.: Rfid-based techniques for human-activity detection. Commun. ACM 48(9), 39–44 (2005)
17. Quinlan, J.R.: C4.5: Programs for Machine Learning. Morgan Kaufmann, San Francisco (1993)