What is That in Your Hand? Recognizing Grasped Objects via Forearm Electromyography Sensing

JUNJUN FAN, Institute of Software, Chinese Academy of Sciences, China and University of Chinese Academy of Sciences, China
XIANGMIN FAN, Institute of Software, Chinese Academy of Sciences, China
FENG TIAN*, Institute of Software, Chinese Academy of Sciences, China and University of Chinese Academy of Sciences, China
YANG LI, Google Research & Machine Intelligence, USA
ZITAO LIU, TAL AI Lab, China
WEI SUN and HONGAN WANG, Institute of Software, Chinese Academy of Sciences, China

Knowing the object in hand can offer essential contextual information revealing a user's fine-grained activities. In this paper, we investigate the feasibility, accuracy, and robustness of recognizing the uninstrumented object in a user's hand by sensing and decoding her forearm muscular activities via off-the-shelf electromyography (EMG) sensors. We present results from three studies to advance our fundamental understanding of the opportunities that EMG brings in object interaction recognition. In the first study, we investigated the influence of physical properties of objects such as shape, size, and weight on EMG signals. We also conducted a thorough exploration of the feature spaces and sensor positions which can provide a solid base to rely on for future designers and practitioners for such interactive technique. In the second study, we assessed the feasibility and accuracy of inferring the types of grasped objects via using forearm muscular activity as a cue. Our results indicate that the types of objects can be recognized with up to 94.2% accuracy by employing user-dependent training. In the third study, we investigated the robustness of this approach in a realistic office setting where users were allowed to interact with objects as they would naturally. Our approach achieved up to 82.5% accuracy in discriminating 15 types of objects, even when training and testing phrases were purposefully performed on different days to incorporate changes in EMG patterns over time. Overall, this work contributes a set of fundamental findings and guidelines on using EMG technologies for object-based activity tracking.

CCS Concepts: • **Human-centered computing → Ubiquitous and mobile computing**; • **Empirical studies in ubiquitous and mobile computing**;

*Corresponding author

Authors' addresses: Junjun Fan, State Key Laboratory of Computer Science and Beijing Key Lab of Human-Computer Interaction, Institute of Software, Chinese Academy of Sciences, Beijing, China, School of Computer and Control Engineering, University of Chinese Academy of Sciences, Beijing, China, junjun2015@iscas.ac.cn; Xiangmin Fan, State Key Laboratory of Computer Science and Beijing Key Lab of Human-Computer Interaction, Institute of Software, Chinese Academy of Sciences, Beijing, China, xiangmin@iscas.ac.cn; Feng Tian, State Key Laboratory of Computer Science and Beijing Key Lab of Human-Computer Interaction, Institute of Software, Chinese Academy of Sciences, Beijing, China, School of Computer and Control Engineering, University of Chinese Academy of Sciences, Beijing, China, tianfeng@iscas.ac.cn; Yang Li, Google Research & Machine Intelligence, Mountain View, California, USA, yangli@acm.org; Zitao Liu, TAL AI Lab, Beijing, China, zitao.liu@acm.org; Wei Sun, sunwei2017@iscas.ac.cn; Hongan Wang, State Key Laboratory of Computer Science and Beijing Key Lab of Human-Computer Interaction, Institute of Software, Chinese Academy of Sciences, China, wha@iel.iscas.ac.cn.

2474-9567/2018/12-ART161 $15.00
https://doi.org/10.1145/3287039

Additional Key Words and Phrases: Grasped Object Detection, Activity Tracking, Electromyography (EMG), Context-Aware Applications

ACM Reference Format:
Junjun Fan, Xiangmin Fan, Feng Tian, Yang Li, Zitao Liu, Wei Sun, and Hongan Wang. 2018. What is That in Your Hand? Recognizing Grasped Objects via Forearm Electromyography Sensing. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 4, Article 161 (December 2018), 24 pages. https://doi.org/10.1145/3287039

1 INTRODUCTION

Knowing the object in a user's hand can offer rich contextual information revealing both her current activity and future intention, which is crucial in designing intelligent interfaces for ubiquitous computing or accessible environments [1, 7, 23, 24, 33]. However, it is challenging to make such an approach robust or scalable due to challenges in variations of surrounding environments, the ambiguity of human activities, and the high cost of physical instrumentation on every object in a uniform style (e.g., with fiducial mark [32], RFID [4, 7, 25], acoustic barcode [18]). Furthermore, it may be hard to differentiate whether users are intentionally interacting with the objects or are just nearby.

Researchers have explored detecting uninstrumented objects by recognizing their characteristic signatures, such as electromagnetic noise [23], sound [41], magnetic field [27], and micro-vibration [23]. However, it is hard to identify a distinctive signature that can apply in each category of everyday objects. For example, EM-Sense [23] detects electromagnetic noise as the object signature, yet only the objects that produce such signals can be detected. Moreover, most of these techniques are sensitive to ambient noise due to the low signal-noise-ratio (SNR) of such signatures.

Compared with aforementioned intrinsic-oriented recognition where objects share identification information via either dedicated instrumentations or unique signatures, we propose a novel extrinsic-oriented approach which leverages forearm muscular activities to infer the type of grasped objects indirectly. Variations in object properties such as shape, size, weight, functionality, as well as intended use combined lead to different grip poses, muscle activation intensities and arm/hand movement patterns, which in turn result in characteristic muscle activation patterns (e.g., when and which muscle is activated and how active it is). Since electromyography (EMG) measures electrical potentials generated by muscle cells and reflects how active a muscle is at a given time point, we hypothesize that sensing and classifying EMG signals can enable a reliable detection of when an object is used and what type that object is. Compared with prior work, our method has three strengths. First, it relieves the costly need for target object instrumentation with identification technologies such as RFID tags [7, 25]. Second, it detects objects that people are actually interacting with, rather than simply nearby. Third, sensing human muscular activities allows the system to recognize a wider range of objects that may or may not produce a distinctive sound [41] or other specific signatures.

While EMG-based hand gesture/posture recognition has been well explored and validated in previous studies [2, 17, 19, 28, 34–36], we make the distinction that our goal is to recognize objects rather than grip postures. Relying on grip pose solely can be ambiguous for object detection, especially when different objects are interacted with similar grasp types. In comparison, our proposed method relies on not only grip poses, but also other factors such as muscle activation intensities or hand movement patterns during object manipulation. In other words, our approach detects objects based on a combination of characteristic factors rather than a single property such as grip.

The primary goal of this work is to determine whether forearm EMG can be used as a cue to recognize the object a user is interacting with, thus providing activity-related context. To achieve this goal, we conducted three studies to gain a fundamental understanding of both the potentials and limits of such approach. More specifically, in the first study, we investigated and quantified influences of multiple physical properties of objects including shape, size, weight, and all these factors combined on EMG signals. We demonstrate through empirical evidence that the uniqueness of EMG signal is caused by variations in all these factors together. We also conducted a thorough exploration of the feature space and sensor positions which can provide a solid base to rely on for future designers and practitioners for such interactive technique. In the second study, we assessed the feasibility and accuracy of inferring the types of grasped objects via using off-the-shelf EMG sensors. The results indicate that by employing user-dependent training, the proposed method can recognize 12 everyday objects with an average accuracy of 92.6% when holding the objects statically, and 95.8% when manipulating the objects dynamically. To give some real-world practicality to our problem, we performed the third study in a realistic office setting where users were allowed to interact with objects as they would naturally. Our approach achieved up to 82.5% accuracy in discriminating 15 types of objects, even when training and testing phrases were purposefully performed on different days to incorporate changes in EMG patterns over time. Overall, this work contributes a set of fundamental findings and guidelines on using EMG technologies for activity tracking in a fine-grained manner.

2 RELATED WORK

2.1 Object-based Activity Recognition

Traditional approaches rely on motion sensors (e.g., wearable accelerometers [3, 9, 21, 38]) to infer user activities. They have achieved high accuracies (e.g., 91.7 % in [21], 96.7% in [9], and 99.0% in [38]) in differentiating ambulatory motions and basic postures such as sitting, walking, or standing. However, such gross movements cannot fully reflect people's fine-grained everyday activities.

According to Activity Theory [20], activities have objectives and are accomplished by using tools and objects. Therefore, one can assume that we may be able to infer something regarding the activity that a user is currently engaged in if we know the object that the user is interacting with. Some approaches have been created to detect users' immediate activities in this manner. To achieve robust recognition of large quantity and variety of objects in dynamic and complex environments, most of them rely on intrinsic-oriented exploration where objects share identification information explicitly through sensor-based technologies such as fiducial marker [32], RFID tag [4, 7, 25], acoustic barcode [18], and NFC [16]. We must acknowledge that RFID-based approaches [4, 7, 25] can achieve finer-grained detection-they can discriminate every single item even though they are identical ones. In comparison, our approach is model-level detection-we can discriminate objects of different models (e.g., having different shapes, sizes, weights, etc.), rather than fully identical ones. However, such instrumentations on target objects are impractical for many scenarios in terms of cost, performance, reliability, and social acceptance of tag installations hindered by ethical concerns [40]. Therefore, it is difficult for such approaches to achieve widespread adoption. We believe that there is a tradeoff between detection granularity, cost, and efforts required to register objects in the system. Integrating different approaches together to achieve a better balance can be interesting future work but beyond the scope of this paper.

Researchers also explored detecting uninstrumented everyday objects via sensing unique object signatures and building machine learning algorithms for recognition [22, 23, 27, 33, 41, 42]. Camera-based detection methods differentiate objects based on visual features (e.g. color and texture in [33]). However, they may easily break when the lighting condition is bad or when occlusion exists. Laput et al. [23] utilized a low-cost, software-defined radio to sense the electromagnetic (EM) noise generated by an object's operation as the unique signature. They reported accurate results (i.e. 97.9%) in differentiating 23 everyday objects that spanned a wide range of contexts including home, office, workshop, large structural features and transportation. Similarly, Ward et al. [41] used accelerometers and microphone to recognize workshop tools via vitro-acoustic sound generated during operation. While choosing the right signature could help discriminate objects effectively, it limits the sensing scope. As in the aforementioned examples, neither [23] nor [41] can detect a book since it does not generate EM noise or vitro-acoustic sound.

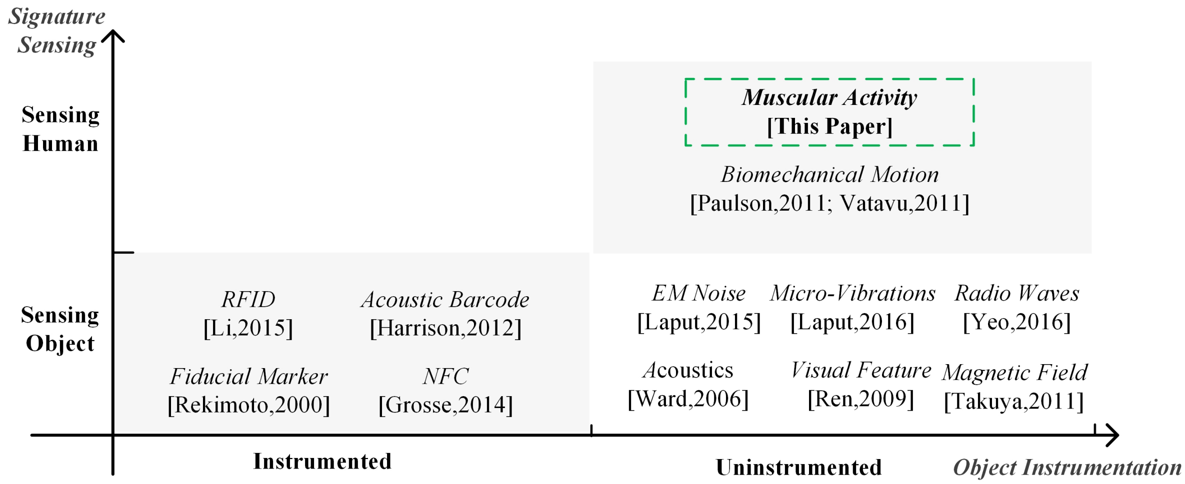

Fig. 1. The design space of in-use object detection.

Besides sensing object signatures directly, another line of research detects in-use objects indirectly via sensing and classifying physiological signatures of human body. Such approaches are extrinsic-oriented where human body serves as a specialized sensing device that extracts information from the object that a user is interacting with. The underlying rationale is that grasped objects varying in physical properties or functionalities could result in different physiological responses when being manipulated. For example, Vatavu et. al [40] recognized the physical properties (e.g., 6 basic shapes, each had 3 levels in size) of the grasped objects with over 90.0% accuracy via measuring postures of the grasping hand through glove sensors. Similarly, Paulson et al. [30] achieved 97.0% accuracy in differentiating 12 objects in an office environment. However, wearing a glove directly on the dominant hand is intrusive and cumbersome, which interferes with normal finger activities. Besides, these methods rely on grip posture information solely via measuring the grasping hand. When different objects are interacted with similar grasp types, detection ambiguities might occur. We argue that other characteristic factors such as the weight of the object or the hand/arm movement patterns during manipulation can contribute to the discerning power, while such information cannot be fully reflected in hand posture measurements (e.g., positioning and orientation of the fingers and palm).

In this work, we explored recognizing grasped objects via sensing and decoding forearm muscular activities. The variations in physical design, functionality, and hand/arm movement pattern when interacting with different types of object could lead to differences in 1) the subset of muscles involved; 2) how active they are; 3) the temporal regularity of combining 1 and 2. Since EMG signals reflect muscular activities in high-fidelity, we hypothesize that sensing and classifying EMG in real time could enable a quick and robust detection of the object in use. Compared with prior work relying on hand posture solely [30, 40], our proposed method does rely on, but is not limited to hand postures; it also relies on other factors (e.g., muscle activation intensities, hand movement patterns, etc.) which might be determined by the weight or functionality of the grasped object. These factors can be reflected in forearm EMG rather than grip measurements. Moreover, our method is relatively unobtrusive compared with wearing data gloves [30]. To summarize, we use a design space (Figure 1) covering two dimensions (i.e. object instrumentation and signature sensing) to organize the state-of-the-art techniques of fine-grained object recognition, which demonstrates the relationship between our approach with the existing ones as well. We hope the design space could inspire new ideas for future research.

2.2 EMG Sensing and Its Applications

Humans make skeletal movements through muscle contractions. The brain first initiates a contraction process by sending an electrical signal through the nervous system. When the signal reaches the target muscle, a subset of its motor units will be activated, and then the corresponding muscle fibers contract to make body movements. EMG senses muscular activities via measuring the electrical potentials between sensor and ground electrodes [8].

Traditionally, EMG was frequently used to assess muscle functions [14] and to control prosthetics [13] in clinical settings. Recent work demonstrated the feasibility of using EMG as an input modality in muscle-computer interfaces [2, 12, 15, 17, 19, 26, 28, 34–36, 39]. For example, [34] achieved 78.0% accuracy in classifying wrist, finger, and combined wrist and finger flexion via forearm EMG sensing. [35] moved beyond gross movement classification to detect finger gestures in real-time with high accuracies (79.0% for pinching, 85.0% while holding a travel mug). Following this line, one most recent advance [19] achieved 82.9% accuracy in classifying fine-grained thumb gestures including left swipe, right swipe, tap, and long press.

Although highly related, the essential distinction is that our goal is to recognize objects rather than hand gestures. Most prior work focused on recognizing dynamic gestures which typically include a continuous sequence of hand postures and refer to the changes in both finger/palm positioning and hand orientation. Actually, our method is more relevant to hand posture recognition, which refers to the static finger/palm positioning and hand orientation. Researchers have also explored using EMG sensing and classification to detect hand postures (a.k.a. static gestures) [12]; however, as we discussed earlier, relying on posture solely can be ambiguous for object detection. Our approach detects objects based on a combination of characteristic factors including shape, size, weight, and functionality of an object. Variations in these factors combined can result in unique EMG signatures. Our purpose is to investigate whether forearm EMG can yield a fine enough resolution to reliably determine the object that a user is currently interacting with.

3 STUDY 1: UNDERSTANDING THE INFLUENCES OF OBJECT PROPERTIES ON FOREARM EMG

3.1 Experiment Design

The goal of this study is to gain a thorough understanding of the influences of various object properties (e.g., shape, size, weight) on forearm EMG. We believe that it is rigorous to start with such fundamental research questions which once answered construct the base for more future work on the use and applications of such technique.

We start our investigation from surveying the existing taxonomies of the grasping hand based on object properties. It is difficult to create a complete taxonomy due to the huge variations in both physical geometry and intended use of everyday objects. A very early and simple yet effective one summarized by Schlesinger [37] characterizes natural human grasping grips consisting of the following six categories:

- Cylindrical: holding cylindrical objects (e.g., water bottle) with open fist grip
- Spherical: holding spherical objects (e.g., ball) with spread fingers and arched palm
- Tip: gripping small and sharp objects (e.g., pen)
- Hook: supporting heavy objects (e.g., toolbox)
- Palmar: holding flat and thick objects (e.g., tablet) between thumb and other fingers
- ateral: grasping flat and thin objects (e.g., card) primarily between thumb and index finger

Informed by the classification of prehensile postures of Schlesinger [37], we started by selecting six everyday objects (e.g., bottle, ball, tablet, pen, card and toolbox) in accordance with the above taxonomy. Note that discriminating these six objects can be trivial since they are different in multiple aspects, such as shape, size and weight, but nevertheless we can get some preliminary insights and references to easy cases. Then we conducted controlled studies to investigate the influences of object shape, size, and weight on EMG signals while controlling other factors. Specifically, we selected three basic shapes and included four different sizes (i.e. control weight)

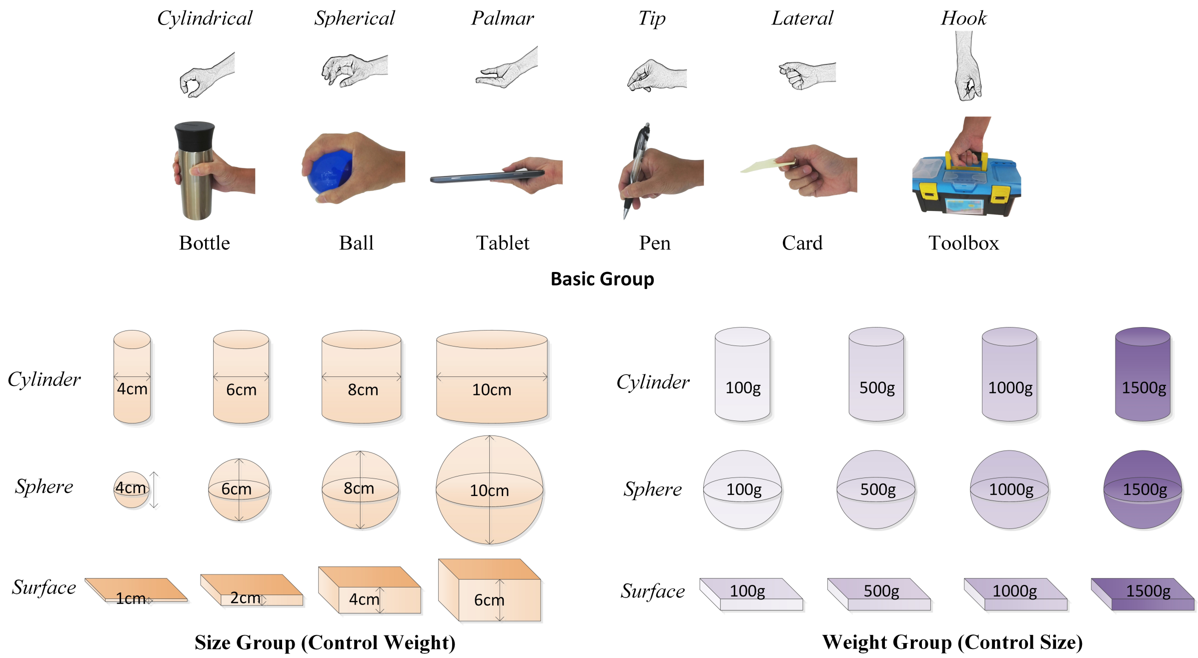

Fig. 2. Objects used in Study 1. In the basic group, we selected six everyday objects in accordance with the Schlesinger taxonomy [37]. In the size group (control weight), there are four levels in size for each shape and all objects have the same weight of 100g. In the weight group (control size), objects sharing the same shape have the same size while there are four levels in weight. The objects in the latter two groups were 3D printed. Weights were controlled by adding iron sand inside.

and four different weights (i.e. control size) for each shape. The properties of the objects used for this study are illustrated in Figure 2. In the following, four different sizes were labeled for our reference as small, medium, large and largest while four different weights were labeled as light, medium, heavy and heaviest. We specifically introduced objects with basic geometries and similar physical properties to investigate whether such minimal variations can be captured by EMG sensing. We believe such design can help us better understand the limits of this approach.

3.2 Participants and Apparatus

We recruited 10 participants (4 female) from a local university for this study. They ranged from 16 to 31 years of age with an average of 25. There was no muscular condition or skin allergy reported. All of them were right-handed. The study took around 1 hour and each participant got a $15 gift card after completion. We used the BIOPAC MP150 [5] system as the EMG sensing device. The system supports up to 16 sensing channels and we used eight channels in this study, following the configurations in prior literature [34]. Each channel had two electrodes in the ends and one ground electrode in the middle, and was sampled at 1000 Hz. The eight channels are placed in an approximately uniform ring around each participant's upper forearm (Figure 3). The average time for setup was less than 10 minutes.

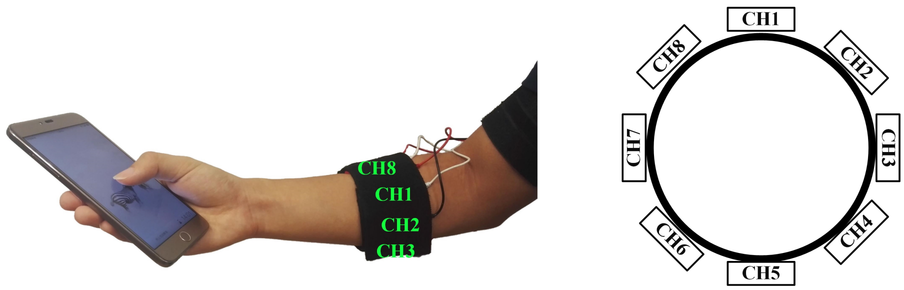

Fig. 3. (Left) EMG sensors worn on the upper forearm of a participant. (Right) Positions of the eight sensors/channels.

3.3 Procedure

We divided the data collection into four rounds in each of the three object groups. Each round consisted of a single repetition of each object in a randomized order. In each trial, a participant held an object firmly in a stable grip for 3 seconds. The participants could pause and had a break any time during the study. We collected 1,200 3-second data samples in total, including 240 (6 objects x 4 rounds x 10 participants) in the basic group, 480 (12 objects x 4 rounds x 10 participants) in the size group (control weight), and 480 (12 objects x 4 rounds x 10 participants) in the weight group (control size).

3.4 Results

We first explored and quantified the influence of object properties on the output values of EMG sensors. We further investigated such influence on features generated based on the sensor output values. Then we assessed the discriminatory power of various features in the feature space by comparing their classification performances. Lastly, we analyzed the relative independence of each sensing position and suggested the ones that might be more informative for classification.

3.4.1 Influence of Object Properties on EMG Sensor Outputs. We first report whether and how variations in physical properties of the grasped objects influence the forearm EMG regarding the data from each sensing channel. For each 3-second data sample, we calculated the root mean square of the raw values generated by each sensor. We conducted one-way ANOVA (i.e. F test) in each object group to analyze if the null hypothesis, that samples in all conditions are drawn from populations with the same mean values, can be rejected. Figure 4 shows the results obtained where statistical significance noted as: $p< 0.001$(***), $p< 0.01$(**) and $p < 0.05$(*). In the basic group, variations in shape, size and weight together exhibited significant effects on all sensing channels (e.g., all with significance $p<0.001$). In the size group (control weight), variations in object size exhibited significant effects on 5, 5, 6 out of 8 channels when grasping cylinders, spheres, surfaces, respectively. In the weight group (control size), variations in object weight exhibited significant effects on all sensing channels when grasping all the objects in this group (e.g., all with significance $p<0.001$). Therefore, it is safe to draw the conclusion that variations in either a single property or a combination of multiple properties can both significantly influence the output values of EMG sensors, and thus contribute to the uniqueness of EMG signatures.

3.4.2 Discriminatory Power of EMG Features. In the past decade's literature on EMG sensing and classification, the success of achieving a high classification performance depends almost entirely on the selection of EMG features [31, 43]. A comprehensive comparison among various features can give us insights about the usefulness and relevance of certain features in this task. EMG features can be separated into three groups including time domain (TD) features, frequency domain (FD) features, and time-frequency domain (TFD) features based on the

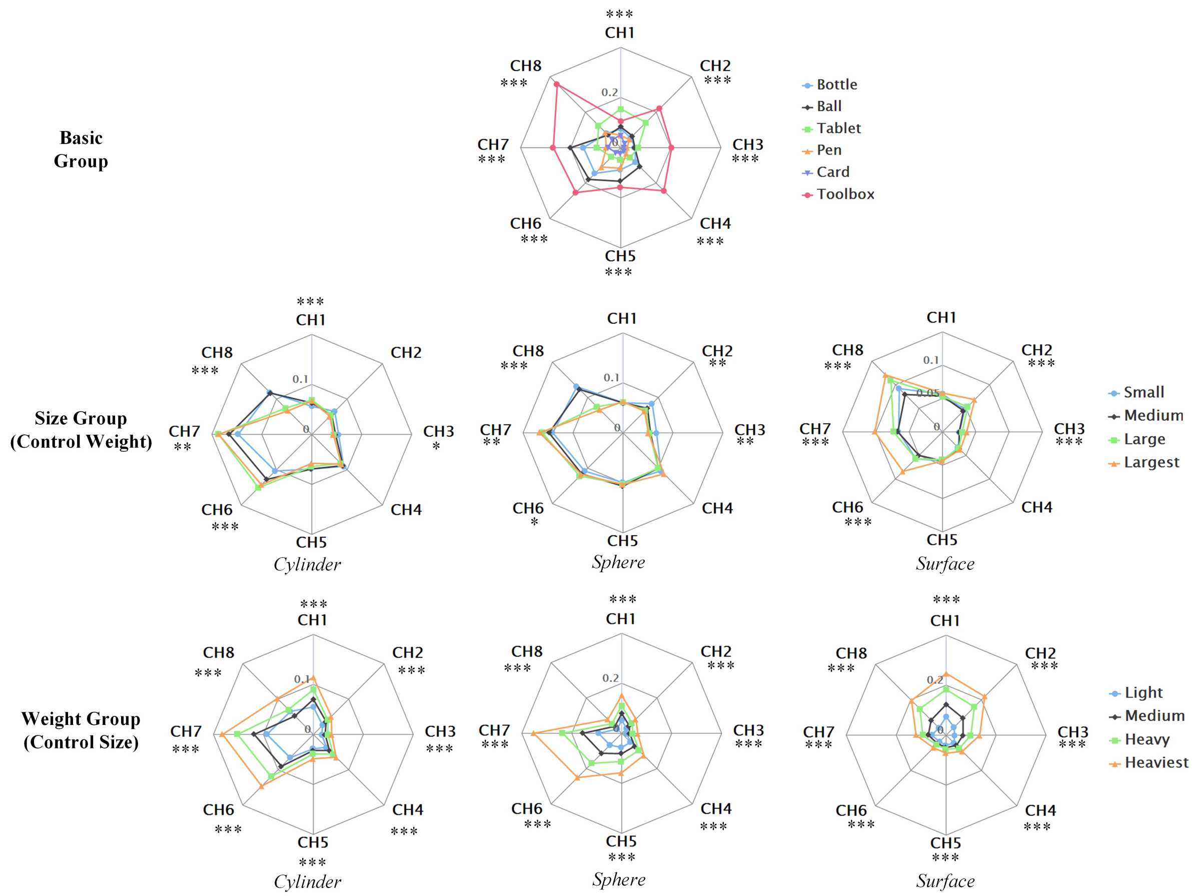

Fig. 4. Averaged output values of eight sensors when grasping different objects. Statistical significances are noted as: p< 0.001(***), p< 0.01(**) and p < 0.05(*).

literature [29]. We explored 12 types of features (e.g., 7 TD, 4 FD, 1 TFD) and all of them have been previously used in the analysis of surface EMG signals. Note that different features can have different numbers of dimension in the feature vector (e.g., Mean Absolute Value has 8 dimensions corresponding to the 8 sensors while Energy Total has only 1 dimension). In total there are 169 dimensions in the feature vector. The optimal parameters for calculating these features are based on the suggestion of related works and our preliminary experiments. Table 1 shows the features explored in this study.

Following the analysis in the previous section, we first investigate whether and how physical properties of objects influence the features in the future space. Within each object group, we conducted pairwise comparisons between objects in different levels regarding the value of each feature in the feature space. This included a total of 8,619 comparisons (2,535 in the basic group, 3,042 in each of the other two groups). Figure 5 shows the results of the pairwise comparisons regarding the statistical differences achieved. We found that within each of the three object groups, 84.8%, 35.3%, and 72.0% among all comparisons exhibited significant differences, respectively. This suggests that variations in shape, size, and weight combined (e.g., in the basic group) can make EMG more

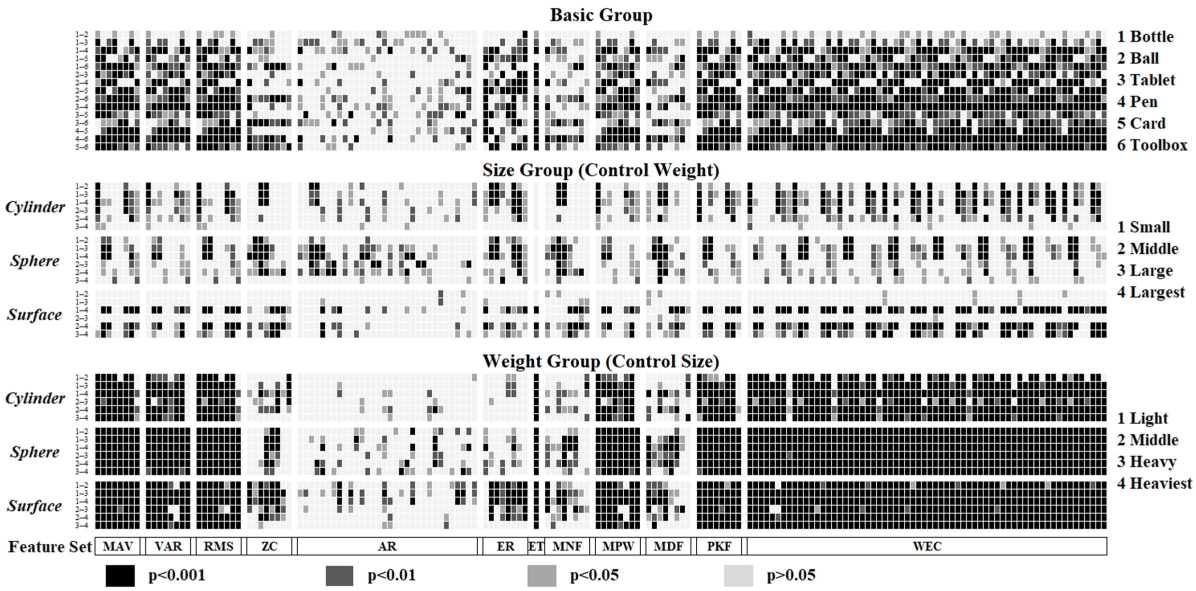

Fig. 5. The color-coded results of pairwise comparisons made in this study.

unique compared with variations in size (e.g., in the size group) or weight (e.g., in the weight group) solely. Moreover, variations in weight have a more significant effect on EMG than variations in size-this finding is consistent regardless of object shape (e.g., 63.9% vs. 39.2% for cylinders, 75.4% vs. 30.4% for spheres, 76.7% vs. 35.3% for surfaces). We also compared the statistical significances achieved in different features in the feature space. Overall, higher percentages of significant difference were achieved on MAV (77.3%), WEC (76.7%), and RMS (76.1%), which suggests that these features can be more sensitive to the variations in object properties.

To further verify these findings, we assessed the classification performances of using different features. Compared with other classifiers such as random forest, K-nearest-neighbor (KNN), and decision tree, we found support vector machine (SVM) classifiers with linear kernel achieved higher accuracies in our preliminary studies. Please note that SVMs are only directly applicable for two-class classification tasks. We applied the LIBSVM Toolkit [11] in our experiments. In their implementation, the problem of multiclass classification is reduced to multiple binary classification problems. A total of N(N-1)/2 binary classifiers are built to distinguish between every pair of the N classes (i.e. one-versus-one). Classification is done by a max-wins voting strategy [6], in which every classifier assigns the unseen instance to one of the two classes, the class with the most votes gets predicted by the combined classifier. We applied half-half cross-validation for each participant (user-dependent training) and Figure 6 shows the corresponding averaged accuracies. We found that using all features together can result in significant higher accuracies than using each single feature solely in each group. Besides using all features together, the highest three overall classification accuracies were achieved when using WEC (64.4%), RMS (56.2%), and MAV (55.6%) only, which confirmed our previous finding that these features can be more discriminative than others when the physical properties of the grasped objects vary.

3.4.3 Degree of Independence of Sensor Positions. In this section we report results on correlation analysis for muscle activations in order to detect shared variance between sensor outputs. Forearm muscles co-activate to perform hand and finger movements [2]-knowing the independence of each sensor position can potentially inform

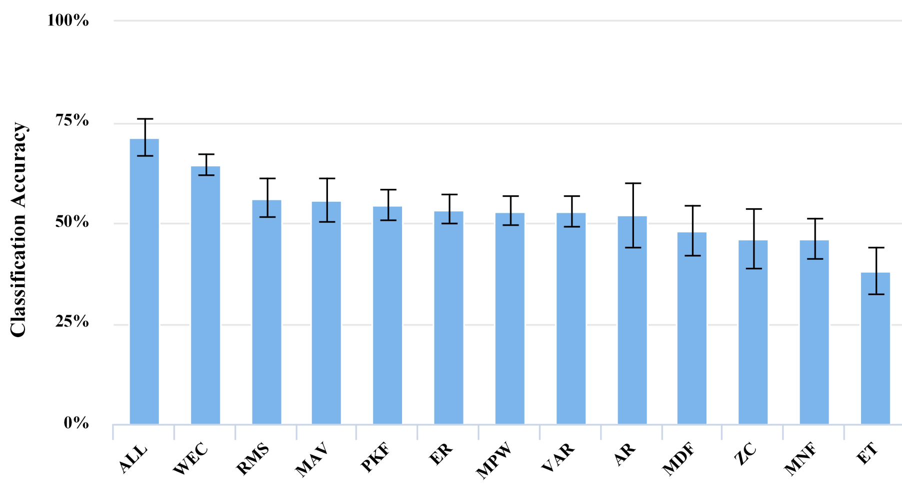

Fig. 6. Classification accuracies of using all features together (ALL) and using each single feature solely. Sorting is done from high to low regarding the overall averaged accuracies across three object groups.

better designs of both the sensing device (e.g., using less sensors to reduce information overlap or redundancy) and the classification techniques (e.g., adopting a weighting scheme based on information gain by each sensor). Figure 7 (Left) illustrated the color-coded Pearson correlation coefficients computed between all eight sensors on the entire dataset in this study (N = 9,600). All coefficients were significant at $p<0.05$ (2-tailed). Sensor locations are shown earlier in Figure 3 (Right).

For each sensor we calculated the average correlation with all the other seven sensors and define the degree of independence for each sensor as the complement of the averaged correlation value with respect to one. The sensors with highest four degrees of independence are CH5 (0.49), CH1 (0.47), CH2 (0.42), and CH7 (0.42). Sensors with lower degrees of independence include CH6 (0.41), CH8 (0.40), CH4 (0.40), and CH3 (0.37). We further assess the classification performances of using the output values from sensors varying in degree of independence (Figure 7 Right). We found that the classification accuracy of using all eight sensors is significantly higher than using four high-independence sensors ($p<0.01$) and using four low-independence sensors ($p<0.01$). We also found that using high-independence sensors can achieve higher accuracy than using low-independence sensors while the difference is marginal ($p=0.07$).

3.5 Brief Summary

In this study, we investigated and quantified the influences of multiple physical properties of objects including shape, size, weight, and all these factors together on EMG signals. We also conducted a thorough exploration of the feature spaces and sensor positions regarding their discriminatory power. Our empirical findings include: 1) variations in either an individual property or a combination of multiple properties can both significantly influence the EMG sensor outputs as well as the values of the features, but the degrees of impact vary; 2) variations in all factors combined can make EMG more unique compared with variations in size or weight solely; 3) variations in weight have a more significant effect on EMG than variations in size and such finding is consistent regardless of object shape; 4) the MAV, WEC, and RMS features are more sensitive to the variations in object properties and are relatively more informative for the classification; 5) high-independence sensor positions (CH1, 2, 5, 7) can result in higher classification accuracies than low-independence sensor positions (CH3, 4, 6, 8) while the

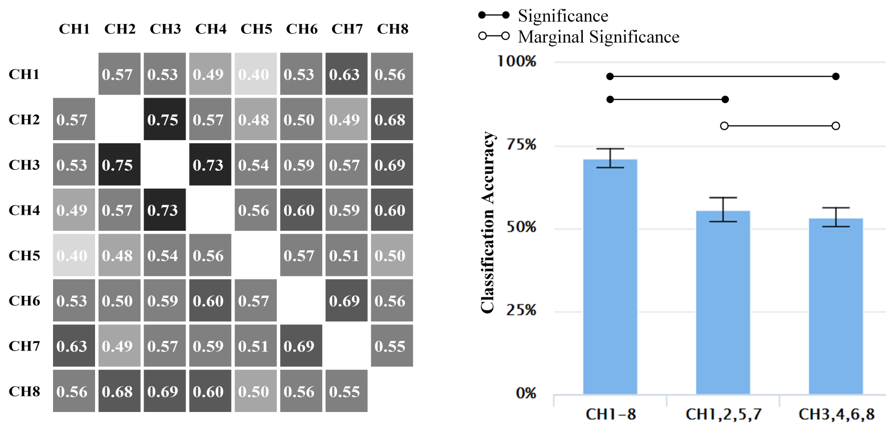

Fig. 7. Correlation analysis between the output values generated by sensors at various positions: (Left) Pearson correlation coefficients computed on the whole dataset (N=9,600, all coefficients were significant at $p<0.05$); (Right) classification accuracies of using all 8 sensors vs. using 4 more independent sensors vs. using 4 less independent sensors.

difference is marginal. We believe that such foundational findings can help the community gain a comprehensive understanding about the relationship between physical properties of grasped objects and EMG signals.

4 STUDY 2: EXPLORING THE FEASIBILITY OF RECOGNIZING EVERYDAY OBJECTS VIA FOREARM EMG SENSING

In Study 1, we show that variations in physical properties, either individually or together, can significantly influence the forearm EMG signals. In this study, we explore the feasibility and accuracy of detecting everyday objects that users are interacting with based on EMG signals. We tried to make our dataset more representative for grasping objects in daily life by 1) extending the detection targets from objects with basic shapes (in Study 1) to everyday objects that are frequently and actually used; 2) including object manipulation condition where a large variety of postures or finger configurations would presumably be employed. This new scenario was specifically introduced to provide a better coverage of daily scenarios and therefore to better understand the feasibility.

4.1 Object Selection

We have three criteria in choosing the objects: 1) they should be frequently used in our daily life; 2) they should cover different grip poses in the Schlesinger taxonomy [37]; 3) we should also include objects sharing similar grip poses to see if this causes classification confusion, which may help us better understand the limits of EMG sensing in this task. According to our criteria, we first conducted a poll with 33 participants (ages ranged from 21 to 45, came from a variety of occupations) on their frequently used everyday objects. Each participant could name up to ten objects. Then we ranked the objects and chose two most frequently mentioned items for each grip type. The following is the final list (Figure 8) along with the numbers of vote: smartphone (31), water bottle (24), pen (23), mouse (18), electric toothbrush (15), key (13), ID card (10), handbag (9), tablet (9), portable charger (7), spoon (6), and toolbox (2).

In this object set, we purposefully included highly confusable objects such as those sharing similar grip poses (e.g., objects in the same column in Figure 8), weights (e.g., key vs. ID card, etc.), and operation patterns (e.g., toolbox vs. handbag, etc.), which can provide answers to questions such as: whether and to what extent this

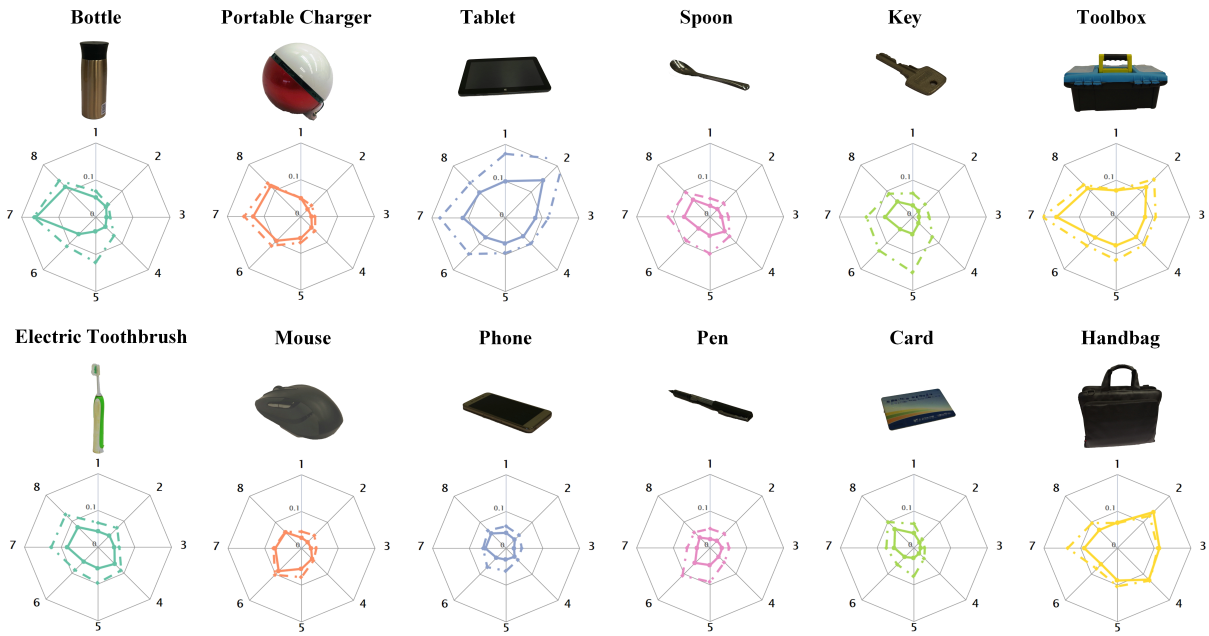

Fig. 8. Objects used in Study 2. The averaged sensor outputs when holding the objects statically and manipulating the objects dynamically are visualized in the radar charts, using solid lines and dash lines, respectively. Objects in the same column share similar grip poses.

approach can discriminate objects with/without similar grip poses; whether it can differentiate objects sharing similar grips and weights while being manipulated differently (e.g., key vs. ID card), etc. Although investigating techniques for classifying all possible objects is clearly important, as the first empirical study on using EMG for object detection, our goal at current stage is to acquire a fundamental understanding of the discerning power of EMG signals which can lay a foundation for detecting a broader range of objects in the future.

4.2 Participants and Apparatus

We recruited 12 participants (4 female) from a local university for this study. They ranged from 18 to 35 years of age with an average of 26. There was no muscular condition or skin allergy reported. All of them were right-handed. The study took around 2 hours and each participant got a $20 gift card after completion. The apparatus was the same with Study 1.

4.3 Procedure

We collected data in two conditions, i.e. C1) holding the objects statically and C2) manipulating the objects dynamically. Putting them together provides a better coverage of our everyday usage of the objects. To enhance the classifier generality and avoid over-fitting, we encouraged participants to apply various hand poses in C1 and to mimic their daily usage patterns as much as they could in C2. The order of these two conditions was counter-balanced across participants.

We divided the data collection into 20 rounds in both C1 and C2. Each round consisted of a single repetition of each object in a randomized order. Each trial (i.e. one-time hold/manipulation of an object) lasted 3 seconds. This design is important as it avoided the inherent similarities if the same object was manipulated back-to-back, thus

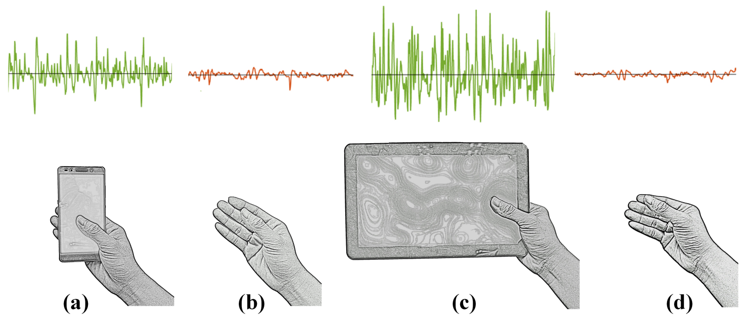

Fig. 9. Examples of actual grips (a, c) and the corresponding simulated grips (b, d).

could help prevent overfitting of the model. The participants could pause and have a break any time during the study. In total, we collected 5,760 data samples (12 objects x 20 rounds x 2 conditions x 12 participants) in C1 and C2.

In addition, we added a "hand-free" class and collected data when participants were 1) performing simulated grips of each object without truly having the object in hand (Figure 9); 2) performing free-style gestures as background data. This allows us 1) to validate the resistance of our approach to false positives (i.e. some object is detected as "being handled" while the user is not interacting with anything); 2) to analyze the difference in EMG signals between actual grips and simulated grips with the same hand posture, which may help us answer questions such as: is the uniqueness of EMG signal caused by the grip posture or by the object-in-hand?

4.4 Results

While we observed huge variability in how different users interact with the same objects (e.g., difference in grip postures, finger configurations, hand movement patterns, etc.), we applied half-half cross-validation for each participant (user-dependent training) and reported the corresponding classification results. We also report the results of user-independent training at the end of this section.

4.4.1 Actual Grip vs. Simulated Grip. We first investigated whether and to what extent the EMG signals are different between actual grips (i.e. having the object in hand) and simulated grips (i.e. not having the object in hand) with the same grip posture for each object (Figure 9, a vs. b, c vs. d). With two-class SVMs, most of the objects reached an accuracy of 100% in discriminating these two types except the followings: mouse (90.0%), pen (75.0%), key (90.0%) and ID card (85.0%). We further compared the object classification accuracy by using EMG signals of actual grips vs. simulated grips. T-test showed that the actual grips were more informative and can lead to significantly higher object classification accuracy (94.2% vs. 75.0%, $p<0.01$).

These results suggested that: 1) even with visually similar grips, having the object truly in hand can make the EMG signal highly different with that of a simulated grip; 2) the posture of grip plays an important role in object classification (i.e. achieved 75.0% accuracy when only relying on posture itself); 3) other factors that are not reflected directly in hand posture (e.g., object weight) also contribute to the uniqueness of EMG signals, which supported a conclusion that our method detected objects relying on, but not limited to grip pose; it also relied on other factors as well.

4.4.2 Classification Confusion. The classifier achieved an overall accuracy of 94.2% (SD=4.5%) across the 13 classes (12 objects + "hand-free" class) on the whole dataset of this study. Figure 10 shows the confusion matrixes in both static holding and dynamic manipulation conditions. Smaller bounding boxes show the confusions

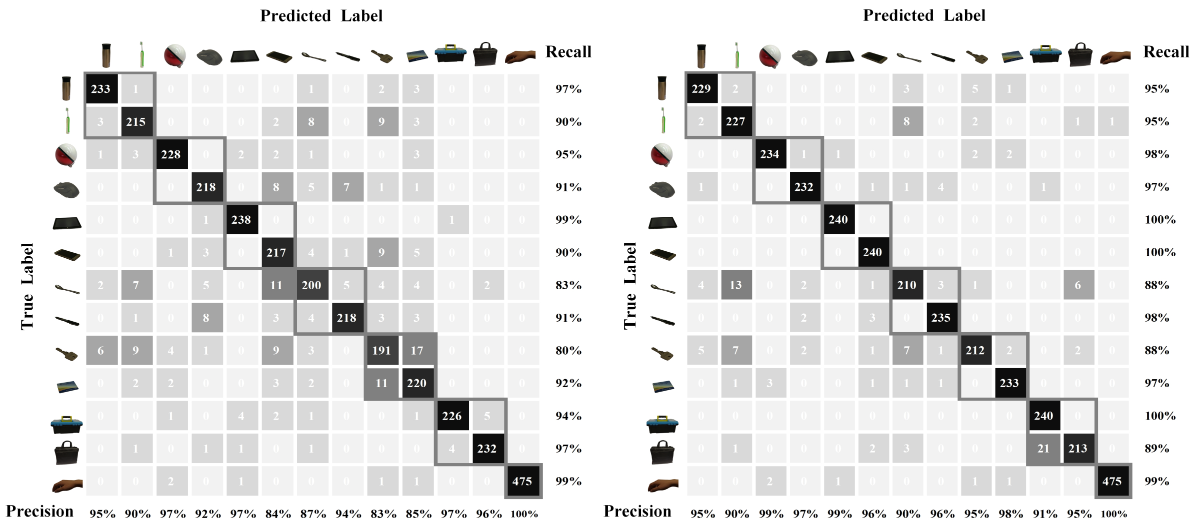

Fig. 10. Object confusion matrixes of holding the objects statically (Left) and manipulating the objects dynamically (Right). Smaller bounding boxes show the confusions of objects sharing similar grip poses.

of objects sharing similar grip poses. Darker areas that occur off the diagonal indicate more confusions. The detection accuracies reached the level of 90% on 11 out of the 13 classes. We also found that the false positive rate is low (3.8%), despite the inclusion of simulated grips in the "hand-free" class.

Among all the confusions, 21.4% involved objects sharing similar grips while 78.6% involved objects with dissimilar ones. However, please note the unbalance in object quantity of these two classes-in other words, 21.4% of the confusions occurred on 1 object with a similar grip while 78.6% occurred on a total of 11 objects with dissimilar grips. We further calculated the probabilities of confusing two objects with similar vs. dissimilar grips, and t-test showed that the former probability was significantly higher than the latter (1.8% vs. 0.6%, $p<0.05$), which indicated that differentiating objects with dissimilar grip postures can be easier than those with similar grip postures.

We found that the classification accuracy was higher when objects were manipulated dynamically (95.8%) than they were held statically (92.6%). T-test showed that the difference was significant ($p=0.005$), which suggested that dynamic manipulation could provide richer information (e.g., hand/arm movement patterns, dynamics of muscle activation, etc.) and help to make the EMG signatures more unique. For example, the classifier made 28 confusions when differentiating key and card while users grasped them statically; in comparison, the classifier only made 2 confusions when they are manipulated dynamically. We attribute such improvement to the variations in hand movement patterns when interacting with these two objects.

By ranking and dividing the 12 objects into two groups based on their weights (6 heavy items vs. 6 light items, weight averaged 1611.67g and 46.33g, respectively), t-test showed the average recognition accuracy of heavy objects was higher while the difference was marginal (95.2% vs. 91.5%, $p=0.07$). One potential reason might be that it was easier for EMG devices to capture the details of more intense muscular activities when handling heavier items.

4.4.3 User-dependent System vs. User-independent System. In this study, we also explored using user-independent classifiers via leave-one-subject-out training, and the overall detection accuracy was 47.9% (chance was 7.7%). We attributed this to the huge variety of the ways between different users who interact with the same objects, which can introduce additional interference. Considering that our classifier achieved an overall accuracy of 94.2% with a total of 2 minutes recording when handling each object, it is still promising in that the requirement on training examples can be low.

4.5 Brief Summary

In this study, we investigated the feasibility and accuracy of inferring the types of grasped objects via using off-the-shelf EMG sensors. Our empirical findings include: 1) handled object recognition via sensing and classifying forearm EMG signals is accurate (94.2%), despite purposeful inclusion of highly confusable objects with visually similar postures and simulated grips; 2) we confirmed the finding presented in Study 1 that the uniqueness of EMG signals is caused not only by the grasping grips, but also by other characteristic factors of the objects that users are interacting with; 3) compared with static holding, manipulating objects dynamically could provide richer information which is useful for classification (95.8% vs. 92.6%); 4) it is more accurate to detect heavier objects than lighter objects (95.2% vs. 91.5%). Collectively, these findings shed light on the feasibility of leveraging muscle activity as a cue for detecting object-based interactions.

5 STUDY 3: GRASPED OBJECT RECOGNITION IN A REALISTIC OFFICE SETTING

To investigate the practicality of the proposed approach, we conduct another study in a realistic office setting where users were allowed to interact with objects as they would naturally. Moreover, prior literature indicated that a practical issue of EMG-based approach is the changes in EMG patterns over time, even on two consecutive days [31]. Therefore, we aimed to explore these practical issues and assess the robustness of this method in a naturalistic environment.

First of all, we conducted a field study to identify the types of object interactions that can be performed in an office environment. Informed by the observations, we recruited 12 participants and asked them to perform these activities while their forearm EMG signals were collected. In the rest of this section, we report the detailed procedure as well as the experimental results.

5.1 Object Selection

We conducted a field study to identify the different types of object interactions that can be performed an office setting. The study was located at a local IT office where 20 employees (7 females) worked. Each employee had a single-person open cubical and basic office equipment including a desktop computer, a telephone, pens and notebooks, etc. We observed and recorded their object-based activities in one-hour time intervals over the course of two consecutive weekdays, with a total time period of approximately 16 business hours. We observed the following types of object interactions that were frequently performed by the employees:

- Typing: putting the hands on the keyboard while pressing the keys
- Using a Mouse: moving, clicking or scrolling a mouse
- Making a Phone Call: holding and dialing on a smartphone or holding the headset of a telephone
- Reading: holding a book to read
- Drinking: grasping a mug or a bottle to drink
- Eating: using a utensil to eat
- Writing: gripping a pen or pencil to write
- Organizing Paperwork: cutting papers with a scissor, assembling papers with a stapler or glue

Informed by the observations in this field study, we chose a total of 15 types of objects for this study, including mouse, keyboard, smartphone, telephone, book, paper, pen, mug, knife, water sprayer, apple, scissor, stapler, glue bottle and glass kettle (Figure 11).

Fig. 11. The setup of the cubicle in this study. Types of objects include: (1) mouse, (2) keyboard, (3) smartphone, (4) telephone, (5) water sprayer, (6) book, (7) pen, (8) paper, (9) stapler, (10) scissor, (11) glue bottle, (12) knife, (13) apple, (14) glass kettle, and (15) mug.

5.2 Participants

We recruited 12 participants (5 female) for this study. They ranged from 25 to 35 years of age with an average of 28. There was no muscular condition or skin allergy reported. All of them were right-handed. The EMG apparatus was the same with Study 1 & 2. Figure 11 shows the setup of the cubicle used in this study. Each participant got a $20 gift card after completion.

5.3 Procedure

Participants completed this study over the course of two adjacent days. We collected training data on the first day and collected testing data on the second day. This design was informed by a recent finding that EMG data measured in one day can be relatively different from that in another day even on the same subject [31]. Therefore, collecting training data and testing data on different days can provide a better illustration of robustness while considering the changes in EMG patterns over time.

On the first day, participants were asked to interact with the above 15 objects that may be typically found around the desk in an office. Each participant manipulated each object for 4 times and each trial lasted around 15 seconds. In total we collected 180-minute (15 seconds x 4 rounds x 15 objects x 12 participants) EMG data for the training dataset.

On the second day, participants were asked to freely manipulate and explore the objects on the table for 10 minutes. Note that we only provided a recommended activity list (e.g., typing, using a mouse, making a phone call,

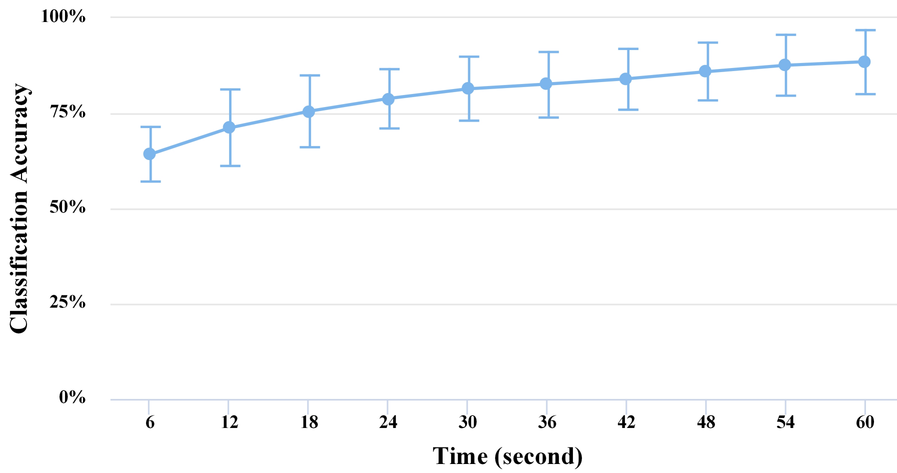

Fig. 12. Classification accuracies with various amounts of training data per object class.

reading, drinking, writing, eating, and organizing paperwork) for their reference without any explicit instruction on how or when they should perform these activities. We believe such design can simulate the real situations quite closely and enable enough variety in the testing dataset as well. The study was videotaped from which two experimenters manually generated the ground truth labels.

5.4 Results

We segmented the training and testing data into 3-second samples and excluded hand-free intervals and the ones involving object interaction transitions. In total we got 3,600 training samples and 1,621 testing samples. We first explored user-dependent approaches where the classifier was trained using only data from a given test user. We assessed the classification accuracies with an increasing amount of training data. For each level, certain amount of random samples in each object class were used for training and all the samples in the testing dataset were used for testing. The results of each level were averaged across 10 folds. Figure 12 shows the averaged classification accuracies with various amounts of training data per object class. Not surprisingly, the average accuracy improved as more examples were used for training. As the amount of training data increases to 60 seconds per object class, the classifier reached a maximum average accuracy of 82.5% (chance was 6.7%). The minimum accuracy for one user was 75.0%, while the maximum accuracy for another user was 89.3%.

Figure 13 shows the color-coded confusion matrix of the user-dependent approach using all the training data of the given test user. Dark areas that occur off the diagonal indicate high confusion while lighter areas indicating little confusion. Figure 14 plots both the actual (above) and predicted (below) grasped objects at each point in time for each participant, which provides more detailed information about where and when the confusions occurred. We found that the most notable confusions occurred when discriminating knife vs. glue bottle (21), apple vs. book (14), and apple vs. paper (10). By localizing these confusions in Fig. 14, we found that the vast majority of each type of confusion were associated with one particular user. For example, 15 out of 21 knife-glue bottle confusions were associated with P9; 10 out of 10 apple-paper confusions were associated with P8; 10 out of 14 apple-book confusions were associated with P7. We further inspected the recorded videos of these three users and found that they behaved quite differently when interacting with the aforementioned objects on the two days. For instance, P9 gripped the knife statically when collecting training examples on the first day but she used the knife to cut

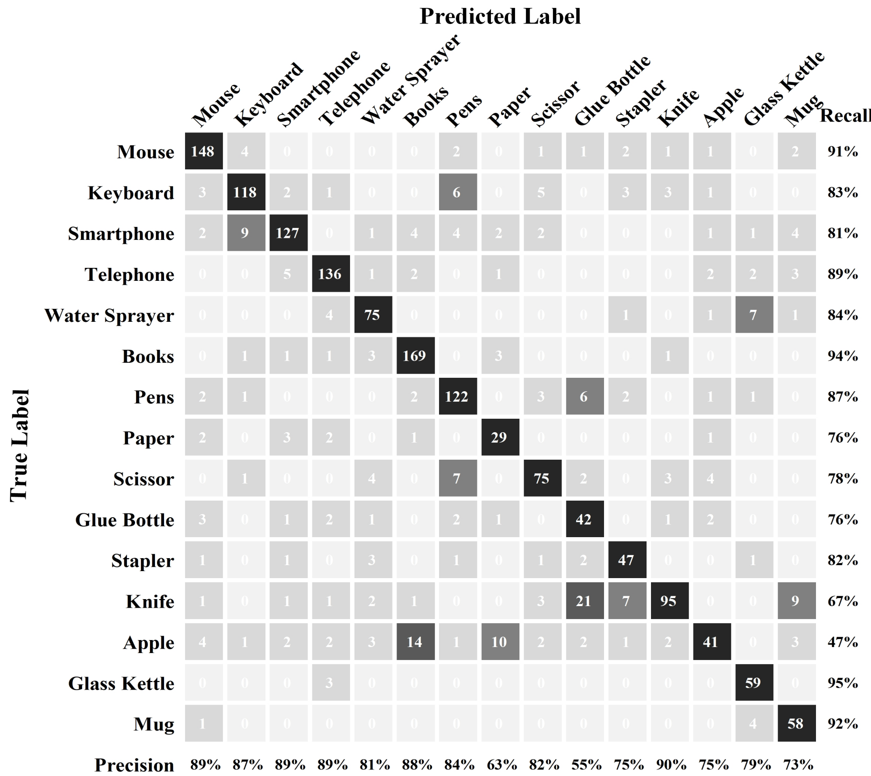

Fig. 13. Confusion matrix of objects by employing user-dependent training. Dark areas that occur off the diagonal indicate high confusion while lighter areas indicating little confusion.

the apple when collecting testing examples on the second day. P7 and P8 grasped the apple firmly on the first day while they actually ate the apple on the second day. We believe that the classification performance can be further enhanced by incorporating more in-situ training examples.

We also applied leave-one-subject-out cross validation to test the user-independent performance and the average accuracy was 45.1% (chance was 6.7%). Note that given a 6-second piece of training data of a test user, the accuracy of user-dependent method (61.1%) is already higher than the accuracy of the best user-independent method. As discussed earlier in Study 2, we attributed this to the high degree of variation in the way that users interacted with the same objects, especially in a natural and unconstrained environment.

Detecting Non-registered Objects. In practice, it is not possible to obtain supervision for all the relevant everyday objects. Therefore, it is important to validate the detection performance with the interference of non-registered objects. In addition to the standard multi-class classification settings that we reported above, we also explored a two-stage approach where the first stage is detecting non-registered (and registered) objects and the second stage is discriminating the objects that are predicted as registered objects in the prior stage. Essentially, detecting non-registered objects in this task is an anomaly detection problem which has been researched extensively within diverse research areas and application domains. We applied a multi-class classification based anomaly detection technique [10] to recognize non-registered objects. Specifically, for each registered object class, a discriminative boundary was learned using a one-class classification algorithm (e.g., one-class SVMs [11] were used in our

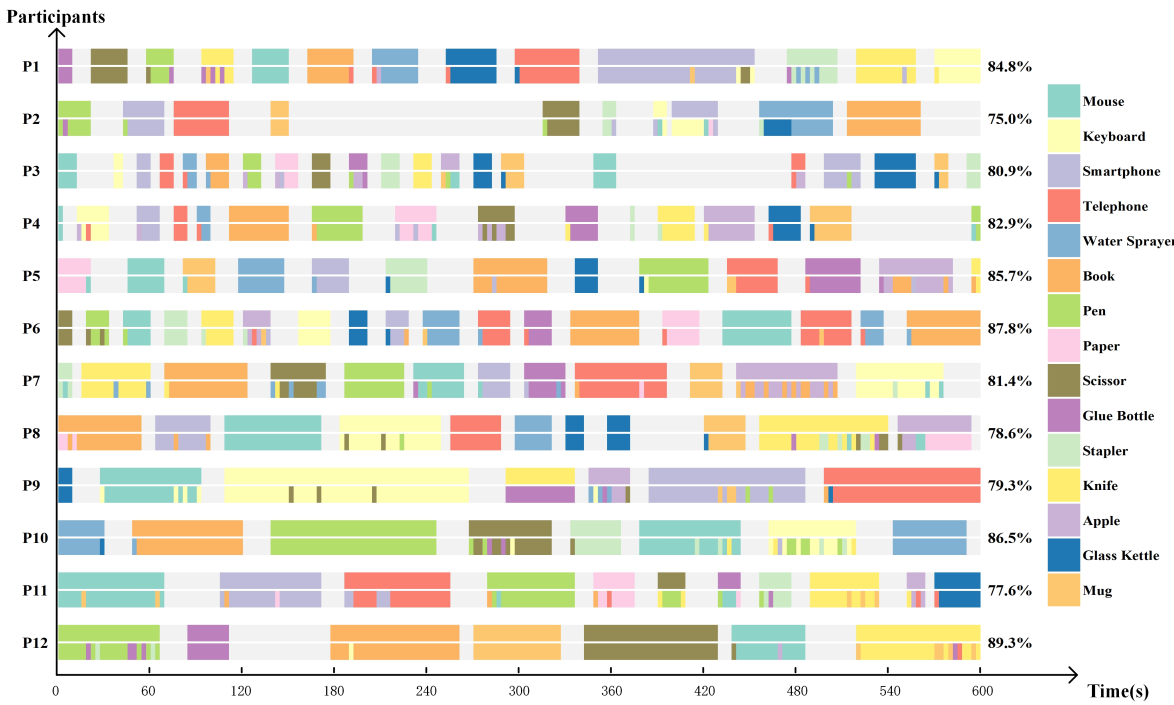

Fig. 14. Visualization of the actual (above) and predicted (below) grasped objects at each point in time for each participant. The light gray color indicates hand-free intervals. (We recommend readers to access the web version of this article for interpretation of the color coding.)

implementation). An object is considered as non-registered if it does not fall within the learned boundary of any registered object class. In our analysis, for each time five objects were randomly selected as non-registered objects and the other ten objects were then considered as registered objects. We validated the detection performance in this setting and repeated this process for 100 times. Please note that for each time, the data of the non-registered objects was unseen by the system-for each testing sample, the system first determined whether it was anomalous. If yes, the system predicted it as "non-registered"; otherwise the system assigned it to a registered object class. In this setting, our approach achieved an average overall accuracy of 81.7% (compared with 82.5% accuracy achieved in the standard multi-class classification settings), which suggest the robustness of our approach with the interference of non-registered objects.

Robustness over time. Based on the literature, EMG signals may fluctuate over time due to electrode location shift across sessions [2, 19] and variations in muscle contraction effort (e.g., muscle fatigue) over time [31]. Some of the prior studies collected training and testing data from different sessions on one or a few days [2, 19, 34]. Following the tradition, we also collected our training and testing data from two separate days to validate the cross-session robustness of this approach. However, we believe that it is still necessary to validate the robustness against the long-term effect of fluctuating EMG signals. Therefore, we conducted another experiment five months after we collected the training data-we recruited 11 out of the 12 participants in this study again and asked them to freely manipulate and explore the objects again for 10 minutes. The procedure and apparatus were the same with Day 2 of this study. The overall detection accuracy was 80.6% (chance was 6.7%). The maximum accuracy for one user was 87.1%, while the minimum accuracy for another user was 69.1%. There was no significant difference between the results of the two sessions conducted five months apart (80.6% vs. 82.5%, p=0.66). The results suggest that this approach can be robust over time.

6 DISCUSSIONS AND FUTURE WORK

This paper presents a thorough exploration of the feasibility, accuracy, and robustness of recognizing the uninstrumented object in a user's hand by sensing and decoding her forearm muscular activities via off-the-shelf electromyography (EMG) sensors. We conducted three studies and present a set of useful findings to advance our understanding about both potentials and limits of EMG-based grasped object recognition. With the advent of commodity EMG devices, such as Myo arm band device (http://www.thalmic.com) , we hope this work could inspire future researchers to move beyond gesture recognition and to push the limits of what can be done with EMG sensing.

We should emphasize that the proposed approach detects objects based on a combination of characteristic factors including shape, size, weight, and functionality of an object, rather than relying on grip pose solely. Variations in these factors together result in unique EMG signatures. For example, the classification accuracy in Study 2 was 75.0% based on grip alone, which suggested that grip (mostly affected by object size and shape) was important but other factors (e.g., weight and usage pattern) were also essential to achieve the final accuracy of 94.2%. Looking at the combination of these factors that come with daily objects enables us to detect a reasonable range of objects via EMG. However, we should also mention that this approach cannot differentiate items that are identical in all the dimensions at the same time: grip, weight and usage pattern (e.g., hanging a handbag or a toolbox with the same weight), although we feel this kind of multi-dimensional identical case can be rare. Compared with existing approaches which rely on either 1) physical instrumentation on every object (e.g., RFID [4, 7, 25]) or 2) sensing certain signatures generated directly by the objects (e.g., EM-noise [23]), our approach can achieve a better balance among cost, accuracy and sensing scope.

In addition to user-dependent models, we also validated user-independent models in our studies. Being able to do this has implications for the potential of creating systems that require little or no user-specific training. Our approach achieved an overall accuracy of 47.9% (chance was 7.7%) in Study 2 and 45.1% (chance was 6.7%) in Study 3 via leave-one-user-out cross validations. We attribute the degradation of the accuracy to inter-individual differences in how they interact with a same object, as well as differences in the positioning of electrodes across users. While most of the prior works on using EMG for Human Computer Interaction adopted user-dependent settings [2, 19, 34–36], such accuracy degradations have been reported as well by the literature [2, 34]. The generalization of the classifier across subjects is still an open research question in this research direction and remains our future work. On the other side, we also notice that the classifiers performed considerably better than chance, suggesting that there exists potential for building user-independent classifiers in this domain. Moreover, considering that our classifier achieved an overall accuracy of 94.2% with a total of 2-minute recording when handling each of the 12 objects in Study 2, and achieved 82.5% accuracy with a total of 1-minute recording when handling each of the 15 objects in Study 3, it is still promising in that the requirement on training examples can be low.

Based on the detection granularity, existing methods that detect in-use objects are either item-level detection or model-level detection. Item-level approaches such as IFID-based method [4, 7, 25] can discriminate every single item, even though they are identical objects (i.e. finest grained detection). Our approach is model-level, which means that we can discriminate objects of different models (e.g., having different shapes, sizes, weights, etc.), even though they may belong to the same category. Compared with item-level detection, the benefit is that we do not need to register every single target object in the system; instead, we only need to register one of the identical items and the system can detect the other ones afterwards. It is interesting to explore whether this method can further generalize to detect a range of different objects/models in a particular category while some objects/models are unseen by the system (i.e. category-level detection). We believe that it could be promising when the EMG database of known objects grows until having diverse examples in each category, as users add newly encountered objects over time. Exploring and quantifying such generalizability can be interesting future work.

One of the major goals of this research is to fundamentally explore the influence of physical properties of objects such as shape, size, and weight on EMG signals, which can gain our understanding of the opportunities that EMG brings in object interaction recognition. To achieve this goal, we need to collect high-fidelity EMG signals to avoid the potential interference introduced by the sensing device due to low sampling rates so that we can uncover the truth as much as possible. Thus, we chose the current research device which has 1k Hz sampling rate in our studies and we believe that such device might become ubiquitous in the near future. Moreover, the high-density recordings of EMG signals (dataset available at: url) can serve as a common ground for future research and performance comparison of different approaches on EMG based wearable interfaces.

One limitation regarding object selection in this work is that all chosen objects were rigid or semi-rigid. In the future, we plan to explore detecting soft and deformable objects as well as detecting simultaneous grasps of multiple objects. Moreover, the current approach detects the manipulation of objects as a list of discrete actions. It would be interesting to detect the stage (e.g., reaching, starting, holding, releasing, etc.) and the strength of the grip (e.g., weak, medium, strong) to further improve the interactions supported. A large amount of data were acquired for this work. In order for other researchers to replicate and advance our results, we decide to release the dataset as well as the source code used for the analysis under a BSD license. Please contact the first author to make a request.

7 CONCLUSION

The primary goal of this work is to determine whether forearm EMG can be used as a cue to recognize the object that a user is interacting with, thus providing some form of activity-related context. We present results from three studies to gain a fundamental understanding of both potentials and limits of such approach. In the first study, we investigated and quantified the influences of multiple sources such as object shape, size, weight, and all these factors combined on EMG signals. We demonstrate through empirical evidence that the uniqueness of EMG signal is caused by variations in all these factors together. We also conducted a thorough exploration of the feature space and sensor positions which can provide a solid base to rely on for future designers and practitioners for such interactive technique. In the second study, we assessed the feasibility and accuracy of inferring the types of grasped objects via using forearm muscular activity as a cue. Our results indicate that the types of objects can be recognized with up to 92.6% accuracy in firm grasp conditions and up to 95.8% in object manipulation conditions by employing user-dependent training. To investigate the robustness of this approach, we performed the third study in a realistic office setting where users were allowed to interact with objects as they would naturally. Our approach achieved up to 82.5% accuracy in discriminating 15 types of objects, even while the training and testing phrases were purposefully performed on different days to incorporate changes in EMG patterns over time. This work pushes the boundaries of EMG-based detection to a new level and we hope it can lead HCI developments in fine-grained activity tracking towards the future world of ubiquitous computing.

ACKNOWLEDGMENTS

The work is supported by the National Key Research and Development Plan under Grant No.: 2016YFB1001402, Key Research Program of Frontier Sciences, CAS under Grant No.: QYZDY-SSW-JSC041, and partially supported by the CAS 100-Talent Program.

REFERENCES

[1] Jake K Aggarwal and Michael S Ryoo. 2011. Human activity analysis: A review. *ACM Computing Surveys (CSUR)* 43, 3 (2011), 16. https://doi.org/10.1145/1922649.1922653

[2] Christoph Amma, Thomas Krings, Jonas Böer, and Tanja Schultz. 2015. Advancing muscle-computer interfaces with high-density electromyography. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 929–938. https://doi.org/10.1145/2702123.2702501

[3] Ling Bao and Stephen S Intille. 2004. Activity recognition from user-annotated acceleration data. In *International Conference on Pervasive Computing*. Springer, 1–17. https://doi.org/10.1007/978-3-540-24646-6_1

[4] Eugen Berlin, Jun Liu, Kristof Van Laerhoven, and Bernt Schiele. 2010. Coming to grips with the objects we grasp: detecting interactions with efficient wrist-worn sensors. In *Proceedings of the fourth international conference on Tangible, embedded, and embodied interaction*. ACM, 57–64. https://doi.org/10.1145/1709886.1709898

[5] Inc BIOPAC Systems. 2017. MP System Hardware Guide. Retrieved March 31, 2017 from https://www.biopac.com/wp-content/uploads/MP-Hardware-Guide.pdf

[6] Christopher M Bishop. 2006. Pattern Recognition and Machine Learning (Information Science and Statistics). 4 (2006), 049901.

[7] Michael Buettner, Richa Prasad, Matthai Philipose, and David Wetherall. 2009. Recognizing daily activities with RFID-based sensors. In *Proceedings of the 11th international conference on Ubiquitous computing*. ACM, 51–60. https://doi.org/10.1145/1620545.1620553

[8] Baptiste Caramiaux, Marco Donnarumma, and Atau Tanaka. 2015. Understanding gesture expressivity through muscle sensing. *ACM Transactions on Computer-Human Interaction (TOCHI)* 21, 6 (2015), 31. https://doi.org/10.1145/2687922

[9] Graeme S Chambers, Svetha Venkatesh, Geoff AW West, and Hung Hai Bui. 2002. Hierarchical recognition of intentional human gestures for sports video annotation. In *Pattern Recognition, 2002. Proceedings. 16th International Conference on*, Vol. 2. IEEE, 1082–1085. https://doi.org/10.1109/ICPR.2002.1048493

[10] Varun Chandola, Arindam Banerjee, and Vipin Kumar. 2009. Anomaly detection: A survey. *ACM computing surveys (CSUR)* 41, 3 (2009), 15. https://doi.org/10.1145/1541880.1541882

[11] Chih-Chung Chang and Chih-Jen Lin. 2011. LIBSVM: a library for support vector machines. *ACM transactions on intelligent systems and technology (TIST)* 2, 3 (2011), 27. https://doi.org/10.1145/1961189.1961199

[12] Xun Chen and Z Jane Wang. 2013. Pattern recognition of number gestures based on a wireless surface EMG system. *Biomedical Signal Processing and Control* 8, 2 (2013), 184–192. https://doi.org/10.1016/j.bspc.2012.08.005

[13] Christian Cipriani, Christian Antfolk, Marco Controzzi, Göran Lundborg, Birgitta Rosén, Maria Chiara Carrozza, and Fredrik Sebelius. 2011. Online myoelectric control of a dexterous hand prosthesis by transradial amputees. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 19, 3 (2011), 260–270. https://doi.org/10.1109/TNSRE.2011.2108667

[14] Catherine Disselhorst-Klug, Thomas Schmitz-Rode, and Günter Rau. 2009. Surface electromyography and muscle force: Limits in sEMG–force relationship and new approaches for applications. *Clinical biomechanics* 24, 3 (2009), 225–235. https://doi.org/10.1016/j.clinbiomech.2008.08.003

[15] Nadine Fligge, Holger Urbanek, and Patrick van der Smagt. 2013. Relation between object properties and EMG during reaching to grasp. *Journal of Electromyography and Kinesiology* 23, 2 (2013), 402–410. https://doi.org/10.1016/j.jelekin.2012.10.010

[16] Tobias Grosse-Puppendahl, Sebastian Herber, Raphael Wimmer, Frank Englert, Sebastian Beck, Julian von Wilmsdorff, Reiner Wichert, and Arjan Kuijper. 2014. Capacitive near-field communication for ubiquitous interaction and perception. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 231–242. https://doi.org/10.1145/2632048.2632053

[17] Faizan Haque, Mathieu Nancel, and Daniel Vogel. 2015. Myopoint: Pointing and clicking using forearm mounted electromyography and inertial motion sensors. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 3653–3656. https://doi.org/10.1145/2702123.2702133

[18] Chris Harrison, Robert Xiao, and Scott Hudson. 2012. Acoustic barcodes: passive, durable and inexpensive notched identification tags. In *Proceedings of the 25th annual ACM symposium on User interface software and technology*. ACM, 563–568. https://doi.org/10.1145/2380116.2380187

[19] Donny Huang, Xiaoyi Zhang, T Scott Saponas, James Fogarty, and Shyamnath Gollakota. 2015. Leveraging dual-observable input for fine-grained thumb interaction using forearm EMG. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*. ACM, 523–528. https://doi.org/10.1145/2807442.2807506

[20] Kari Kuutti. 1996. Activity theory as a potential framework for human-computer interaction research. *Context and consciousness: Activity theory and human-computer interaction* 17 (1996).

[21] Jennifer R Kwapisz, Gary M Weiss, and Samuel A Moore. 2011. Activity recognition using cell phone accelerometers. *ACM SigKDD Explorations Newsletter* 12, 2 (2011), 74–82. https://doi.org/10.1145/1964897.1964918

[22] Gierad Laput, Robert Xiao, and Chris Harrison. 2016. Viband: High-fidelity bio-acoustic sensing using commodity smartwatch accelerometers. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*. ACM, 321–333. https://doi.org/10.1145/2984511.2984582

[23] Gierad Laput, Chouchang Yang, Robert Xiao, Alanson Sample, and Chris Harrison. 2015. Em-sense: Touch recognition of uninstrumented, electrical and electromechanical objects. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*. ACM, 157–166. https://doi.org/10.1145/2807442.2807481

[24] Matthew L Lee and Anind K Dey. 2015. Sensor-based observations of daily living for aging in place. *Personal and Ubiquitous Computing* 19, 1 (2015), 27–43. https://doi.org/10.1007/s00779-014-0810-3

[25] Hanchuan Li, Can Ye, and Alanson P Sample. 2015. IDSense: A human object interaction detection system based on passive UHF RFID. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 2555–2564. https://doi.org/10.1145/2702123.2702178

[26] Minas V Liarokapis, Panagiotis K Artemiadis, and Kostas J Kyriakopoulos. 2013. Task discrimination from myoelectric activity: a learning scheme for EMG-based interfaces. In *Rehabilitation Robotics (ICORR), 2013 IEEE International Conference on*. IEEE, 1–6. https://doi.org/10.1016/j.jelekin.2013.03.009

[27] Takuya Maekawa, Yasue Kishino, Yasushi Sakurai, and Takayuki Suyama. 2011. Recognizing the use of portable electrical devices with hand-worn magnetic sensors. In *International Conference on Pervasive Computing*. Springer, 276–293. https://doi.org/10.1007/978-3-642-21726-5_18

[28] Jess McIntosh, Charlie McNeill, Mike Fraser, Frederic Kerber, Markus Löchtefeld, and Antonio Krüger. 2016. EMPress: Practical hand gesture classification with wrist-mounted EMG and pressure sensing. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 2332–2342. https://doi.org/10.1145/2858036.2858093

[29] Mohammadreza Asghari Oskoei, Huosheng Hu, et al. 2008. Support vector machine-based classification scheme for myoelectric control applied to upper limb. *IEEE Trans. Biomed. Engineering* 55, 8 (2008), 1956–1965. https://doi.org/10.1109/TBME.2008.919734

[30] Brandon Paulson, Danielle Cummings, and Tracy Hammond. 2011. Object interaction detection using hand posture cues in an office setting. *International journal of human-computer studies* 69, 1-2 (2011), 19–29. https://doi.org/10.1016/j.ijhcs.2010.09.003

[31] Angkoon Phinyomark, Franck Quaine, Sylvie Charbonnier, Christine Serviere, Franck Tarpin-Bernard, and Yann Laurillau. 2013. EMG feature evaluation for improving myoelectric pattern recognition robustness. *Expert Systems with applications* 40, 12 (2013), 4832–4840. https://doi.org/10.1016/j.eswa.2013.02.023

[32] Jun Rekimoto and Yuji Ayatsuka. 2000. CyberCode: designing augmented reality environments with visual tags. In *Proceedings of DARE 2000 on Designing augmented reality environments*. ACM, 1–10. https://doi.org/10.1145/354666.354667

[33] Xiaofeng Ren and Matthai Philipose. 2009. Egocentric recognition of handled objects: Benchmark and analysis. In *Computer Vision and Pattern Recognition Workshops, 2009. CVPR Workshops 2009. IEEE Computer Society Conference on*. IEEE, 1–8. https://doi.org/10.1109/CVPRW.2009.5204360

[34] T Scott Saponas, Desney S Tan, Dan Morris, and Ravin Balakrishnan. 2008. Demonstrating the feasibility of using forearm electromyography for muscle-computer interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 515–524. https://doi.org/10.1145/1357054.1357138

[35] T Scott Saponas, Desney S Tan, Dan Morris, Ravin Balakrishnan, Jim Turner, and James A Landay. 2009. Enabling always-available input with muscle-computer interfaces. In *Proceedings of the 22nd annual ACM symposium on User interface software and technology*. ACM, 167–176. https://doi.org/10.1145/1622176.1622208

[36] T Scott Saponas, Desney S Tan, Dan Morris, Jim Turner, and James A Landay. 2010. Making muscle-computer interfaces more practical. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 851–854. https://doi.org/10.1145/1753326.1753451

[37] Georg Schlesinger. 1919. Der mechanische aufbau der künstlichen glieder. In *Ersatzglieder und Arbeitshilfen*. Springer, 321–661.

[38] Isabel Suarez, Andreas Jahn, Christoph Anderson, and Klaus David. 2015. Improved activity recognition by using enriched acceleration data. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 1011–1015. https://doi.org/10.1145/2750858.2805844

[39] Marian Theiss, Philipp M Scholl, and Kristof Van Laerhoven. 2016. Predicting grasps with a wearable inertial and EMG sensing unit for low-power detection of in-hand objects. In *Proceedings of the 7th Augmented Human International Conference 2016*. ACM, 4. https://doi.org/10.1145/2875194.2875207

[40] Radu-Daniel Vatavu and Ionuţ Alexandru Zaiţi. 2013. Automatic recognition of object size and shape via user-dependent measurements of the grasping hand. *International Journal of Human-Computer Studies* 71, 5 (2013), 590–607. https://doi.org/10.1016/j.ijhcs.2013.01.002

[41] Jamie A Ward, Paul Lukowicz, Gerhard Troster, and Thad E Starner. 2006. Activity recognition of assembly tasks using body-worn microphones and accelerometers. *IEEE transactions on pattern analysis and machine intelligence* 28, 10 (2006), 1553–1567. https://doi.org/10.1109/TPAMI.2006.197

[42] Hui-Shyong Yeo, Gergely Flamich, Patrick Schrempf, David Harris-Birtill, and Aaron Quigley. 2016. Radarcat: Radar categorization for input & interaction. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*. ACM, 833–841. https://doi.org/10.1145/2984511.2984515

[43] Mahyar Zardoshti-Kermani, Bruce C Wheeler, Kambiz Badie, and Reza M Hashemi. 1995. EMG feature evaluation for movement control of upper extremity prostheses. *IEEE Transactions on Rehabilitation Engineering* 3, 4 (1995), 324–333. https://doi.org/10.1109/86.481972

Received May 2018; revised August 2018; accepted October 2018