Evaluation of Tactile Feature Extraction for Interactive Object Recognition

Janine Hoelscher[1] Jan Peters[1,2] Tucker Hermans[3]

***Abstract*— Tactile sensing stands to improve the manipulation and perception skills of autonomous robots. Object and material recognition stand as two important tasks, where tactile sensing can aid robotics. While much work has been done on showing the applicability of specific sensors to recognition tasks, a comprehensive examination of the features used has not been performed. In this paper we thoroughly examine the different components of performing interactive object recognition with tactile sensing. We use a state-of-the-art multimodal tactile sensor, allowing us to compare features previously presented for a number of different platforms. We examine the statistical features, robot motions, and classification approaches used for performing object and material recognition. We show that by combining simple statistical features captured from five robot motions our robot can reliably differentiate between a diverse set of 49 objects with an average classification accuracy of $97.6 \pm 2.12\%$.**

I. Introduction and Motivation

Tactile sensing provides an exciting means for robots to identify objects and materials through touch. Object identification in robotics traditionally relies on visual and geometric information captured through cameras or lasers, however visual recognition is not always feasible. Tactile recognition offers a complementary method when vision is not available, along with benefits compared to using visual identification alone. For example during manipulation occlusion is inevitable. A robot can use tactile sensing in its end effector to recognize objects that its own arm occludes from a camera's view. Additionally, measurements received by the touch sensor can quickly be localized using the robot's forward kinematics ignoring scale and pose ambiguity issues that arise in visual recognition. Finally, visually similar materials may have very different haptic properties, which a tactile sensor can easily identify and use for discrimination.

A number of approaches for performing tactile recognition have been proposed in the literature [1–9], but no systematic identification of what features, interaction movements, and learning methods are best suited to the task have been examined. In this work we set out a straightforward examination of different feature extraction methods and interaction motions for tactile object recognition using a set of popular statistical classifiers. We have chosen to use the BioTac tactile sensor as it provides a number of different sensing modalities allowing us to compare features used across various tactile sensors in the literature. We set out to compare not only recognition of specific object instances, but also recognition of material classes, such as wood and plastic.

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007–2013) under grant agreement #610967 (TACMAN). [1]TU Darmstadt, Computer Science Department. janine.hoelscher@stud.tu-darmstadt.de, peters@ias.tu-darmstadt.de [2]MPI for Intelligent Systems. [3]School of Computing, University of Utah, thermans@cs.utah.edu

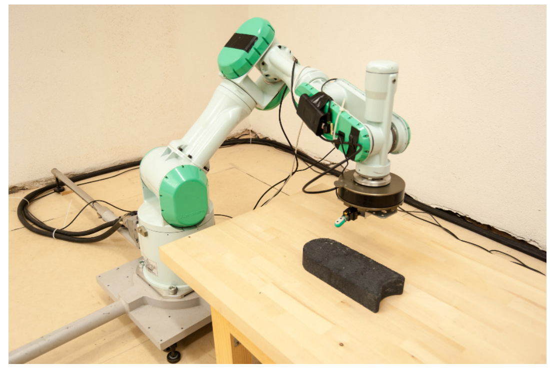


Fig. 1: The PA10 robot equipped with a Schunk force torque sensor and a BioTac finger.

Some approaches to tactile object recognition have used object shape and structure to identify objects [8, 9]. While shape can be helpful in recognizing a specific object, objects of the same (local) shape can be built from various materials. As such the discriminative power is low. We instead choose to focus on recognizing objects through their material properties.

In the past, different methods have been proposed to recognize objects through measuring surface features. Dallaire et al. [1] used a turntable with a three-axis accelerometer to distinguish between 28 disks made of different materials. A similar sensor built into an artificial fingernail was used by Sinapov et al. [2] to discriminate objects via five different exploratory scratching movements. Fishel and Loeb [3] present a method that distinguishes between 117 textures with a 95.4% success rate using the BioTac. They use a Bayesian active learning approach that selects from 36 different sliding movements. This was done in a well calibrated environment with sample texture swatches. In contrast we are interested in examining recognition in a more natural setting using real-world objects. Lepora et al. [7] also propose an active perception approach that performs tapping motions in order to improve knowledge of the examined object. They are able to distinguish from a small set of different diameter bars, while also localizing the finger with respect to the environment. Tanaka et al. [6] also use an active learning approach where a latent variable estimation is performed to learn individual object models presented to the robot. The robot chooses exploratory motions by modifying the parameters to dynamic movement primitive. Apart from sliding there are other stereotypical exploratory procedures human beings perform, such as static contact or enclosing on an object [10]. Chu et al. [4] focus on testing such finger movements in probing an object's surface. They not only recognize objects, but also use the collected data to learn and predict haptic adjective, given by humans that also touched the object. Xu et al. [5] provide a set of three physically motivated features possible to distinguish objects. They are able to reliably classify 10 objects using the three features relating to temperature, compliance, and texture of the object's surface. Beyond object recognition, tactile sensing has been used to perform other qualitative assessments of objects. Chitta et al. determine if an object contains liquid [11], while Bhattacharjee et al. detect if objects are deformable or rigid as well as fixed or movable [12].

We organize the rest of the paper as follows. Section II explains our experimental setup, followed by a description of the recognition methods used in Section III. We give a detailed presentation of results in Section IV and summary of conclusions and future work in Section V.

II. Experimental Setup

In this section, we explain our experiments in detail. First, we describe the hardware used and environment setup. Then, we detail the movements the robot performs in capturing object data. Finally, we present the objects used and the criteria for choosing them.

A. Hardware

For all experiments we use the BioTac sensor, a fingertip-sized, multimodal, tactile sensor. It consists of a rigid core enveloped by an elastic skin held fixed by a plastic fingernail. An incompressible conductive fluid fills the skin. The skin's surface structure has ridges similar to a fingerprint. The BioTac provides three sensing modalities. Nineteen electrodes attached to the core measure the impedance through the conductive fluid which changes as the skin deforms during contact events. A thermistor embedded in the finger measures the fluid temperature. As the BioTac's core temperature rests approximately 10° C above room temperature, heat flow into a touched object can also be measured by calculating the signal derivative. A hydrophone measures the fluid pressure in order to detect vibrations, such as those caused by texture when the BioTac slides over a surface [3] [13]. This produces five types of tactile signals: 19 electrical impedance electrodes (Fig. 2), a low-frequency fluid pressure (PDC) signal, a high-frequency fluid vibration (PAC) signal, temperature (TDC), and temperature flow (TAC). This results in a total of 23 distinct data channels. The PAC channel is sampled at 2200 Hz, the others at 100 Hz. Producing 4400 total measurements per second.

We mount the BioTac finger on a Mitsubishi PA10 robot arm with seven degrees of freedom. An inverse manipulator Jacobian controller drives the desired end effector velocities of the robot. We allow the BioTac to warm up for approximately 20 minutes prior to data collection to reach its core temperature [5]. The object of interest rests on a table in front of the robot. The robot chooses a random starting position for each movement execution to avoid overfitting and create a more robust dataset for learning. The complete setup can be seen in Figure 1.

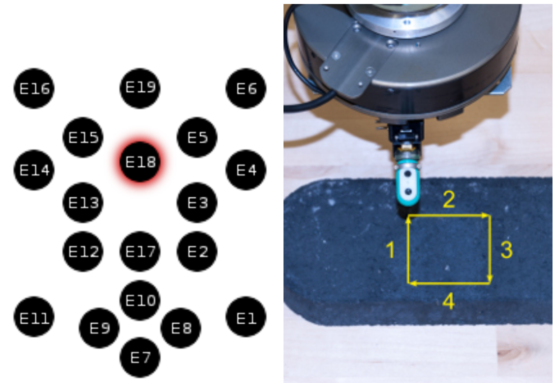


Fig. 2: (Left) Layout of BioTac sensing electrodes. Electrode 18 is highlighted in red. (Right) Arrows indicate the sliding movements the finger executes on the object surface.

B. Movements

According to Xu et. al [5] humans perform three typical exploratory movements that require only cutaneous information: applying pressure, static contact and lateral sliding. In contrast to these movements, others such as enclosure or hefting also need proprioceptive information [10]. As such we decided to execute static contact and lateral sliding in multiple directions yielding a total of five movements.

Static contact: The robot moves its end effector (the mounted BioTac) down until it detects contact and stops. We detect contact using a simple threshold (10 bits) on the change in absolute pressure (PDC). The robot maintains contact for 15 seconds. Experiments show (see Figure 5a) that this timespan is necessary to reach a thermal equilibrium with the object [5]. Afterwards, the robot retracts the finger to its initial position.

Lateral sliding: For the sliding movements the robot establishes contact using the same method as for static contact. Subsequently, the robot executes sliding movements in 4 different directions. Each movement translates a distance of 10 cm at a constant velocity of 2.5 cm/s. As can be seen in Figure 2, the finger moves in the positive y-axis (robot base frame) direction, the positive x-axis direction, the negative y-axis direction, and then the negative x-axis direction one after another, before retracting the finger upwards. In the remainder of this paper these movements are referred to as *backward*, *left*, *forward* and *right*. The control used by the robot attempts to keep the height (z-position) of the finger constant, ignoring any change in pressure or object height.

C. Objects

In choosing the objects to analyze we tried to fulfill several requirements to help ensure similar conditions for data collection and evaluation. First, each object needs to be comprised of a single primary material and stand on a table without any further support. All objects need to have a relatively flat surface to simplify the interaction with the sensor and the surface needs to be greater than 15×15cm to ensure permanent contact with the BioTac finger during data collection and to allow some variation in the position of the measurements. As the BioTac's silicone skin can be damaged easily, objects with sharp edges had to be avoided. Furthermore, we desired a set of objects belonging to a few material classes to examine the difficulty in identifying objects of the same class. At the same time we desired objects varying in terms of thermal conductivity, texture, roughness, and compliance. We found five objects for each of the eight material classes: wood, ceramic, stone, plastic, sponge, paper, metal and fabric. We chose nine additional objects that do not belong to any of the classes or have very different surface properties. This results in a set of forty-nine objects; these are shown in Figure 3. The objects are all common home, garden, or office items.

III. Object Recognition Methods

In this section we describe the technical approaches we evaluated in performing object recognition. We first discuss processing of the raw sensory data followed by the different types of features extracted. We then explain the classification methods used for recognition.

A. Data Processing

For a given interaction with a specific object we split the sensor data into five segments corresponding to each executed movements. We trim the beginning of each motion in order to avoid noisy sensor signals due to contact establishment. This is especially important for sliding movements (see Figure 4), because the elastic skin is pushed to the opposite direction of the current movement. This occurs both in the case of first establishing contact, as well as after a direction change.

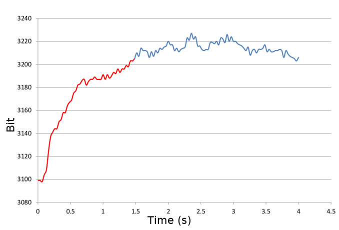


Fig. 4: Sample electrode data.

We collected data over a time period of a few days. We noticed minor environmental (temperature) changes resulted in small shifts in the sensor signals. We subtract the mean of the first 50 samples from each channel to calibrate the incoming data to the current setting.

B. Feature Extraction

We compare seven different methods of extracting features from the processed data. The first two methods were specifically designed for the BioTac sensors [4, 5]. The third and fourth method are simple features computed from the sensory signals. Two more methods are motivated by other examples from the literature reported on different tactile sensors [1, 11]. The final feature extraction method was developed based on preliminary results we found during data processing.

Physically motivated features (Xu): Xu et. al [5] provide one simple, physically-motivated feature per sensor modality. The first feature attempts to measure compliance: $\log(\Delta\text{joint angle}/\Delta\text{E18})$. (See Figure 2 for position of E18.) In their setup the angle between the BioTac and surface could easily be measured, as we do not have access to such a measurement, we instead examine only the electrode data giving $\log(\Delta E18)$. The second feature stands as a proxy for the surface texture: $\log\left(\text{var}\left(\text{PAC}_{\text{filtered}}\right)\right)$, where the PAC signal has been band-pass filtered between 20 and 500Hz. And the final feature examines the thermal conductivity of the object: $\Delta\text{TAC} = \max(\text{TAC}) - \min(\text{TAC})$.

Temporal BioTac features (Chu): Chu et al. [4] suggest a more complex featureset, which also takes the signal changes over time into account. They first compute statistics on the pressure channel: maximum PDC, mean PDC, and the greatest change in PDC, computed from a smoothed (Hanning window) version of the data signal. An energy spectral density (ESD) is created from the high frequency pressure data (PAC) and the following values were computed as features: area under the ESD curve, weighted average over ESD, spectral variance, spectral skewness, and spectral kurtosis. Thermal features consisting of area under the TAC curve and the time constant of an exponential function fit of TDC time series data. Finally Principal Component Analysis (PCA) was computed on the electrode data and fifth order polynomials were fit to the first two principal components over the time series. The polynomial coefficients of the fit functions serve as additional features. The authors also proposed a set of proprioceptive features which are not available in our hardware setup.

PCA Raw Data: We concatenate the values for all 23 data channels over time to produce one single vector, where every sensory signal at each timestep is a feature. We reduce the dimension of this feature vector using PCA. We found using eight dimensions worked best, although it only maintained $\approx 54\%$ of the variance from the original data.

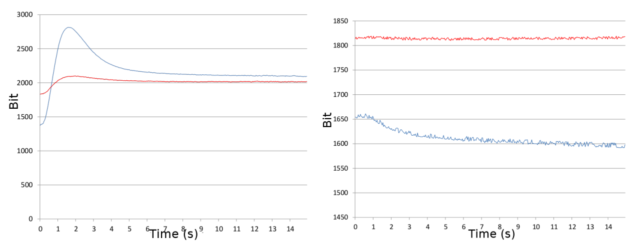
Mean features: We calculate the mean of each signal channel over time, which results in a total of 23 features. This differs from the features of Tanaka et al. [6], which were computed on a subset of the BioTac signals. Reducing these dimensions further with PCA helped improve learning. We could capture $\approx 98\%$ of the variance by keeping only the first eight dimensions.

Pressure features: Dallaire et al. [1] defined a set of features to pick up vibratory information from a MEMS accelerometer. The features are: variance, skew, kurtosis, fifth central moment, sum of the variation over time, number of times 20 uniformly separated thresholds are crossed, and the sum of upper half of the amplitude spectrum. We compute the features for the PAC time series signal of the BioTac, as it most closely represents the behavior of an accelerometer.

Electrode features: We extract electrode features following Chitta et al. [11], who designed features for use on an array of tactile pressure cells in the gripper of the PR2 robot. We first filter the electrode data using a first-order Butterworth filter and then calculate the Euclidean Norm to combine all tactile signals into a single signal: $f(t) = \left(\sum_i^N \text{signal}_i^2\right)^{1/2}$ We take the mean and variance of this signal as two features.

Temperature features: Based on our experience with the above mentioned features we developed a set of features considering only temperature data mainly for use with the static contact movement. Distinguishing values for the analog derivative of temperature (TAC) can be seen after about 6 seconds of contact (Figure 5a). After an initial spike the temperature signal (TDC) remains nearly constant. We take

Fig. 3: Forty-nine objects grouped in columns by their material class. From left to right: plastic, metal, paper, fabric, ceramic, stone, wood, and sponge. The last two columns of objects do not belong to any single material class.



(a) TAC data (b) TDC data

Fig. 5: Temperature signals over time for static contact on the metal box (blue curves) and on the beige sponge (red).

the mean for both TAC and TDC value for the remaining 9 seconds. We augment this with change in TAC signal after the initial peak (a slight modification of ΔTAC from above), which we found to be more discriminative than computing ΔTAC on the entire signal. To avoid issues with numerical instability and balance the importance of different features for learning, we normalize each feature channel to have a mean of 0 and standard deviation of 1 on the training data.

C. Learning Methods

We used several common supervised learning methods to train object recognition classifiers. We did so to examine the robustness of the features across different learners to find the best overall performing features. We chose two generative and two discriminative classifiers to compare.

The first generative model we used is a Naive Bayes classifier. The naive Bayes classifier models each feature as being generated from an independent Gaussian distribution conditioned on the object class. This produces a classifiers that is very efficient to learn, but can miss important correlation in the data [14]. As such we chose a Gaussian classifier, which fully models the covariance of the features, conditioned on the object class, as our second generative classifier.

We chose two popular discriminative classifiers to compare, Support Vector Machines and Random Forests. A support vector machine is a distribution-free, discriminative model, which attempts to find the single decision boundary that maximizes the margin between two classes [14]. We evaluate both a standard linear SVM as well as a nonlinear SVM with an exponential kernel $K(x_i, x_j) = \exp(-\lambda \|x_i - x_j\|^2)$ where λ is the kernel bandwidth. In order to identify optimal parameters for the kernel function and the penalty for missclassified samples, we executed a grid search [15] on a validation set at training time.

A random forest classifier discriminates between object classes using a set of binary decision trees. These trees are independently trained classifiers using different random feature subsets for the training data. A single classification result is chosen by finding the average classification over all trees in the forest [16]. We trained 100 trees to forest with a maximal depth of 3 splitting nodes. We use the scikit-learn [17] implementation of random forests and the libSVM [18] SVM implementation.

IV. Results

We collected data on each of the 49 objects in our data set, executing the 5 exploratory actions on each object (4 sliding movements and 1 static contact). We performed each motion ten times on every object, resulting in a total of 2450 samples. We outline in detail the effects of the different features, classifiers, and motions used.

A. Object Classification

We evaluate the various components of the recognition methods using leave-one-out cross validation. We separate the data based on the different movement and randomly choose one test sample for each object. The remaining nine examples are used as training samples. This results in 49 test samples and 441 training samples per classifier. We repeat this split procedure 100 times. Figure 6 displays the classification accuracy based on feature and movement type, averaged over all classifiers. We note that the mean features obtain the highest accuracy for all movements separately.

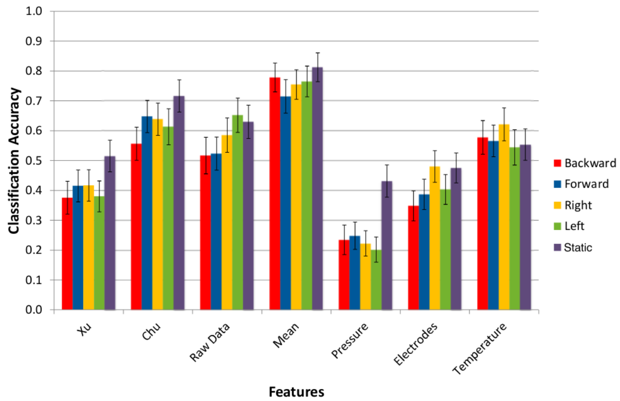


Fig. 6: Classification accuracy for each feature method and each movement, averaged over all classification methods, with 9 training samples per object. Error bars represent one standard deviation.

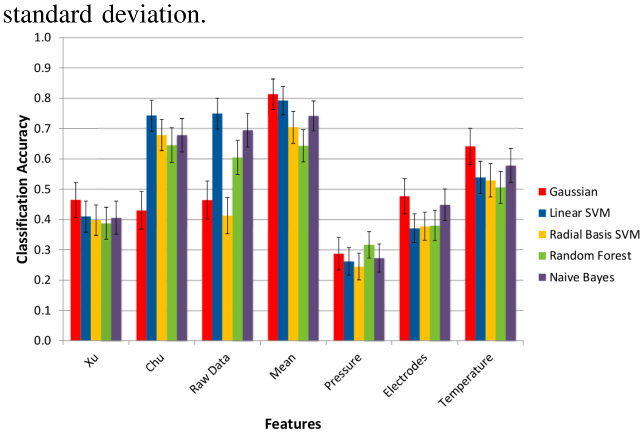


Fig. 7: Classification accuracy for each feature method and classification method, averaged over all movements, with 9 training samples per object. Error bars represent one standard deviation.

In Figure 7 we report the results per classification method, averaging over the different movements used. We see that the Gaussian classifier and linear SVM perform the best with the mean features. Rather unexpectedly in Figure 6 we see that extracting features using static contact produces the best results for four of the features and the second best results for the remaining three sets. We can attribute this to the thermal properties of the objects being highly discriminative. The features proposed by Chu et al. performed second best on average, attaining the highest accuracy using static contact and the linear SVM. The raw data and temperature features perform next best, with varying accuracy across the different motions and classifiers. Classifiers trained using our adaptation of the features proposed by Xu, as well as the pressure and electrode features, perform much worse than the top performing classifiers.

B. Movement Evaluation and Combination

We further evaluate the movements using the mean features with both the linear SVM and Gaussian classifiers, as these were the best performing methods reported above. We first examine the movements separately according to their classification accuracy. In Figure 8 we see that all methods achieve

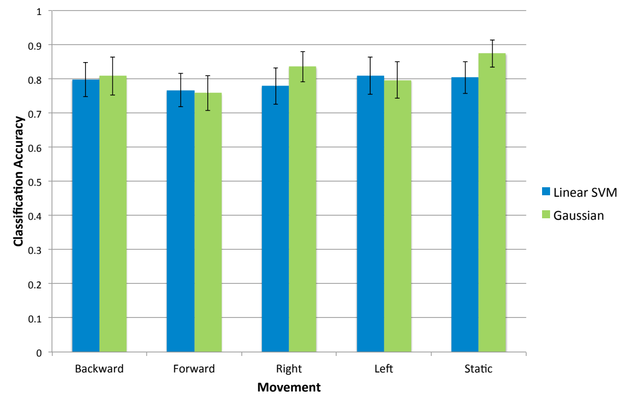


Fig. 8: Classification accuracy for each movement using mean features. Error bars represent one standard deviation.

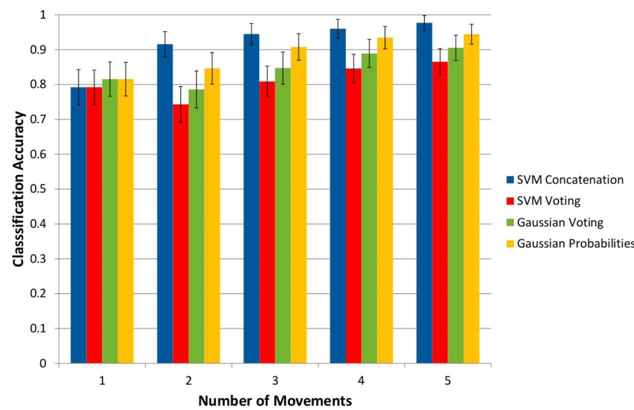


Fig. 9: Classification accuracy for different movement combination methods for an increasing number of movements using the mean features. Error bars represent one standard deviation.

above 70% classification accuracy with static contact as the best performing movement achieving an accuracy rate of $87.5 \pm 4.01\%$ with the Gaussian classifier.

We use three different methods to combine the features extracted from different movements executed on the same object: concatenation, voting, and joint probability. For concatenation, we concatenate the features from the different movements into a single feature vector for classification. Voting consists of executing each movement classifier independently and choosing the classification result that occurs most often. For joint probability we classify each movement separately and sum up the probabilities for each class. We combine the linear SVM classification using the first two methods. As SVMs produce no probability distributions, the third method is not possible. We use the second two methods with the Gaussian classifier approach, since we found the dimensionality of the feature vectors produced by concatenation to be too large and produce numerical instability in computing the feature covariances.

We train classifiers with an increasing number of movements and report the accuracy in Figure 9. Motions are added from best to worst performing based on the results shown in Figure 8. The SVM concatenation approach obtains the highest classification accuracy as soon as two motions are used. With all five motions it results in 97.6$\pm$2.12% classification accuracy. We see that average performance with voting actually decreases with two motions, since disagreements must be resolved randomly. However with three or more motions all methods increase their classification accuracy over that of using one motion alone.

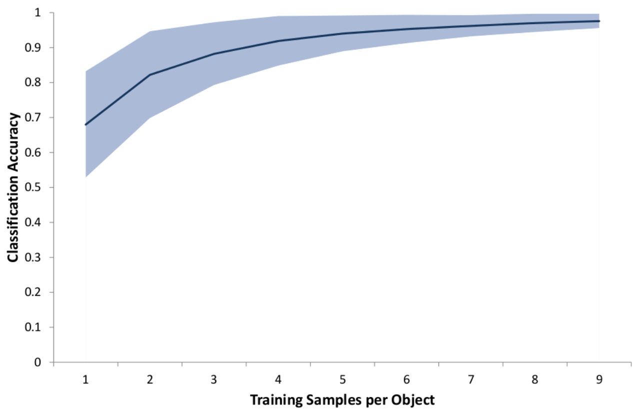


Fig. 10: Classification accuracy for the concatenation of all movements with mean features, linear SVM, and an increasing number of training samples per object. Shaded region represents one standard deviation of error.

Figure 10 shows the classification accuracy of this approach depending on the number of training samples. With four training samples per object the classification accuracy already exceeds 90%. In the next sections we focus on the results of the linear SVM concatenation combination.

C. Material Classification

We take a closer look at the accuracy rates for each object in order to determine the possible reasons for missclassifications. The confusion matrix (Figure 14) shows that the majority of missclassifications are between objects of the same material class. This is especially true for the ceramic, fabric, and wood material classes which account for $\approx 85\%$ of errors. To further investigate this we trained classifiers to only distinguish objects coming from the same material class. For example we take only the objects from the fabric class and train a classifier to distinguish the five objects from one another. We do this as with fewer objects to distinguish the classifier may achieve better performance. The results are displayed in Figure 11. We see that the above mentioned materials (wood, ceramic, fabric) obtain the lowest classification accuracies. This observation suggests, that the objects belonging to these classes have a low in-class variance. The objects made of paper, sponge, or plastic are perfectly distinguishable with nine training samples per object.

This leads us to examine classifying objects not by instance identify, but by material class. We train a classifier on the data belonging to the 40 objects from the material classes plastic, metal, paper, fabric, ceramic, stone, wood, and sponge. The nine objects that do not belong to any material class are left out. We again use a linear SVM with

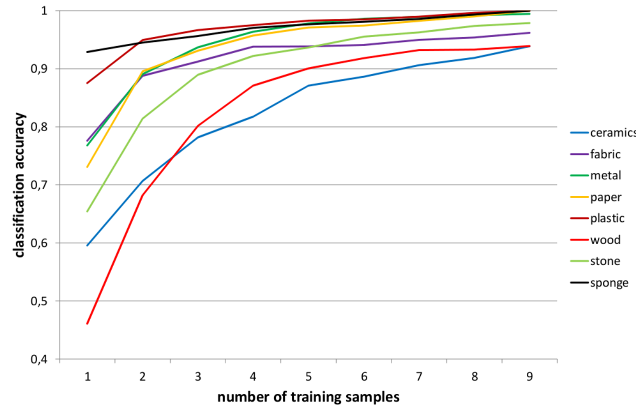


Fig. 11: Classification accuracy for objects of each material class separately with concatenation of all movements, feature calculation mean features, linear SVM and an increasing number of training samples per object.

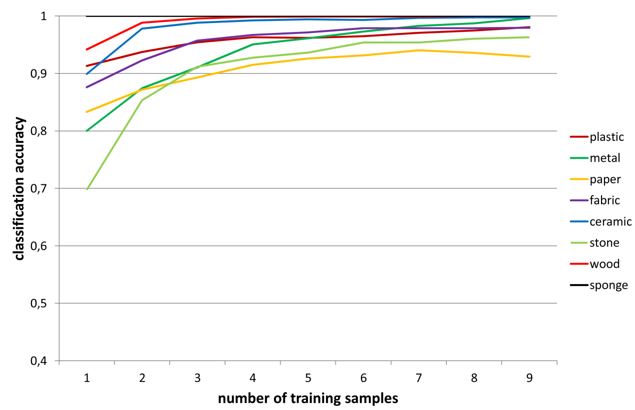


Fig. 12: Classification accuracy of material class classification. Classifiers trained using concatenation of all movements, mean features, linear SVM, and an increasing number of training samples per class.

mean features and all motions concatenated. We produce training and validation data sets as before, but label all objects made from the same material identically. The results shown in Figure 12 correspond to our previous observations. Those materials, whose objects are the most difficult to distinguish, such as wood or ceramic, achieve the highest material classification accuracies. However a low in-class variance is no requirement for high material accuracy as the material class sponge shows.

To further examine the utility of material class classification, we examine how the robot would classify objects it has not previously encountered. We do this by training our eight material class classifier, holding out all examples from one specific object (e.g. "DVD case"). We then take the held out object as a test case and perform classification. The results can be seen in Figure 13. The majority of objects are classified into the correct material class, even though the robot has never encountered them before. We see that the highly confused wood objects are all correctly classified even when not being seen before. However, the ceramics class does not perform as well with the emaille

	Blue Plastic Box	Yellow Plastic Box	DVD Case	CD Case	Blue Board	Metal Box	Smooth Aluminium Case	Brushed Aluminium Case	Can	Saucepan	Cardboard	Notepad	Paper Box	Book	Newspaper	Velvet Pillow	Pencilcase	Linen Towel	Towel	Fleece Blanket
Plastic	10	10	10	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Metal	0	0	0	0	0	10	10	7	0	0	0	10	0	5	0	0	0	0	0	0
Paper	0	0	0	5	4	0	0	0	10	0	10	0	5	5	0	3	0	10	0	0
Fabric	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	7	10	0	10	10
Ceramic	0	0	0	0	0	0	0	3	0	10	0	0	0	0	0	0	0	0	0	0
Stone	0	0	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Sponge	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0
Wood	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

	Beige Tile	Brown Tile	White Tile	Ceramic Baking Pan	Emaille Pot	Brick	Pavement	Granite	Pebble	Slate	Brown Sponge	Beige Sponge	Yellow Sponge	Red Sponge	White Sponge	Birch	Beech	Plywood	Maple	Oak	Cork	Styrofoam	Acryl	Raffia Mat	Metal Sieve	Leather	Raffia Basket	Neoprene	Mirror
Plastic	0	0	0	2	0	0	0	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Metal	1	0	0	1	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Paper	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Fabric	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ceramic	9	10	10	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	9	3	10	0	0	0	0	0	9
Stone	0	0	0	7	0	8	8	10	9	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Sponge	0	0	0	0	0	2	0	0	0	0	10	10	10	10	9	0	0	0	0	0	0	2	0	0	0	0	0	0	0
Wood	0	0	0	0	10	0	0	0	0	0	0	0	0	0	0	10	10	10	10	10	1	5	0	10	10	10	10	10	0

Fig. 13: Confusion matrix for each material with leave one out classification per object. Classifier trained on concatenation of all movements, mean features, and linear SVM. Left column is classified class label. Top label is held out test object.

pot always classified as wood and the ceramic baking dish being confused as stone, metal, or plastic. The resulting class sponge contains all objects we assigned to this class and few other objects. In contrast, the paper class has more false positives than objects actually made of paper. The ceramics class that achieved good results in the last section now achieves a relatively low accuracy of 58%. We additionally classified the nine objects for which we did not name a specific material class. These were mostly assigned to the wood class. The mirror was assigned to the ceramic class for 9 of its 10 cases, which is the most intuitive class for glass.

V. Conclusion and Future Work

We have presented a comprehensive examination of tactile object recognition using the BioTac sensor. Our robot collected data covering 49 objects with five different motions creating a dataset of 2450 total interactions. We analyzed a diverse set of seven feature extraction methods using a number of different classification approaches. Our results suggest that rather simple features, dimensionality-reduced mean values of filtered data signals performed better than more outperform more elaborate and physically motivated feature extraction. This best performing method achieves a very high average classification accuracy above 97%. Most important was extracting these features from a varying set of motions to gather more discriminative information across the diverse set of objects. Our material classification results suggest that some objects are more naturally categorized through touch at the material class level than the instance level. However, our results in classifying held out objects by material type suggest that the boundaries between some materials are difficult to correctly set without seeing all objects of interest from those categories.

We wish to see how our results perform in the context of object manipulation, such as grasping or searching for objects in clutter. In such contexts the robot will be more constrained in selecting interaction motions and the location of object boundaries may not be known. Additionally we believe that it should be possible

References

[1] P. Dallaire, P. Giguère, D. Émond, and B. Chaib-Draa, "Autonomous tactile perception: A combined improved sensing and bayesian nonparametric approach," *Robotics and Autonomous Systems*, vol. 62, no. 4, pp. 422–435, 2014.

[2] J. Sinapov, V. Sukhoy, R. Sahai, and A. Stoytchev, "Vibrotactile recognition and categorization of surfaces by a humanoid robot," *Robotics, IEEE Transactions on*, vol. 27, no. 3, pp. 488–497, 2011.

[3] J. A. Fishel and G. E. Loeb, "Bayesian exploration for intelligent identification of textures," *Frontiers in neurorobotics*, vol. 6, 2012.

[4] V. Chu, I. McMahon, L. Riano, C. McDonald, Q. He, J. Martinez Perez-Tejada, M. Arrigo, N. Fitter, J. Nappo, T. Darrell, and K. Kuchenbecker, "Using robotic exploratory procedures to learn the meaning of haptic adjectives," in *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)*, May 2013, pp. 3048–3055.

[5] D. Xu, G. E. Loeb, and J. A. Fishel, "Tactile identification of objects using bayesian exploration," in *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2013, pp. 3056–3061.

[6] D. Tanaka, T. Matsubara, K. Ichien, and K. Sugimoto, "Object manifold learning with action features for active tactile object recognition," in *Proc. of the IEEE/RSJ International Conference on Intelligent Robotics and Systems (IROS)*, 2014.

[7] N. F. Lepora, U. Martinez-Hernandez, and T. J. Prescott, "Active bayesian perception for simultaneous object localization and identification," in *Robotics Science and Systems (RSS)*, 2009.

[8] P. Payeur, C. Pasca, A.-M. Cretu, and E. M. Petriu, "Intelligent haptic sensor system for robotic manipulation," *Instrumentation and Measurement, IEEE Transactions on*, vol. 54, no. 4, pp. 1583–1592, 2005.

[9] Z. Pezzementi, C. Reyda, and G. D. Hager, "Object Mapping, Recognition, and Localization from Tactile Geometry," in *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)*, May 2011.

[10] S. J. Lederman and R. L. Klatzky, "Hand movements: A window into haptic object recognition," *Cognitive psychology*, vol. 19, no. 3, pp. 342–368, 1987.

[11] S. Chitta, J. Sturm, M. Piccoli, and W. Burgard, "Tactile sensing for mobile manipulation," *Robotics, IEEE Transactions on*, vol. 27, no. 3, pp. 558–568, June 2011.

[12] T. Bhattacharjee, J. M. Rehg, and C. C. Kemp, "Haptic Classification and Recognition of Objects Using a Tactile Sensing Forearm," in *Proc. of the IEEE/RSJ International Conference on Intelligent Robotics and Systems (IROS)*, 2012.

[13] SynTouch, LLC, "Biotac product manual," 2014.

[14] K. P. Murphy, *Machine learning: a probabilistic perspective*. MIT press, 2012.

[15] C.-W. Hsu, C.-C. Chang, C.-J. Lin, *et al.*, "A practical guide to support vector classification," 2003.

[16] A. Criminisi, J. Shotton, and E. Konukoglu, "Decision forests for classification, regression, density estimation, manifold learning and semi-supervised learning," *Microsoft Research Cambridge, Tech. Rep. MSRTR-2011-114*, vol. 5, no. 6, p. 12, 2011.

[17] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[18] C.-C. Chang and C.-J. Lin, "Libsvm: a library for support vector machines," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 2, no. 3, p. 27, 2011.

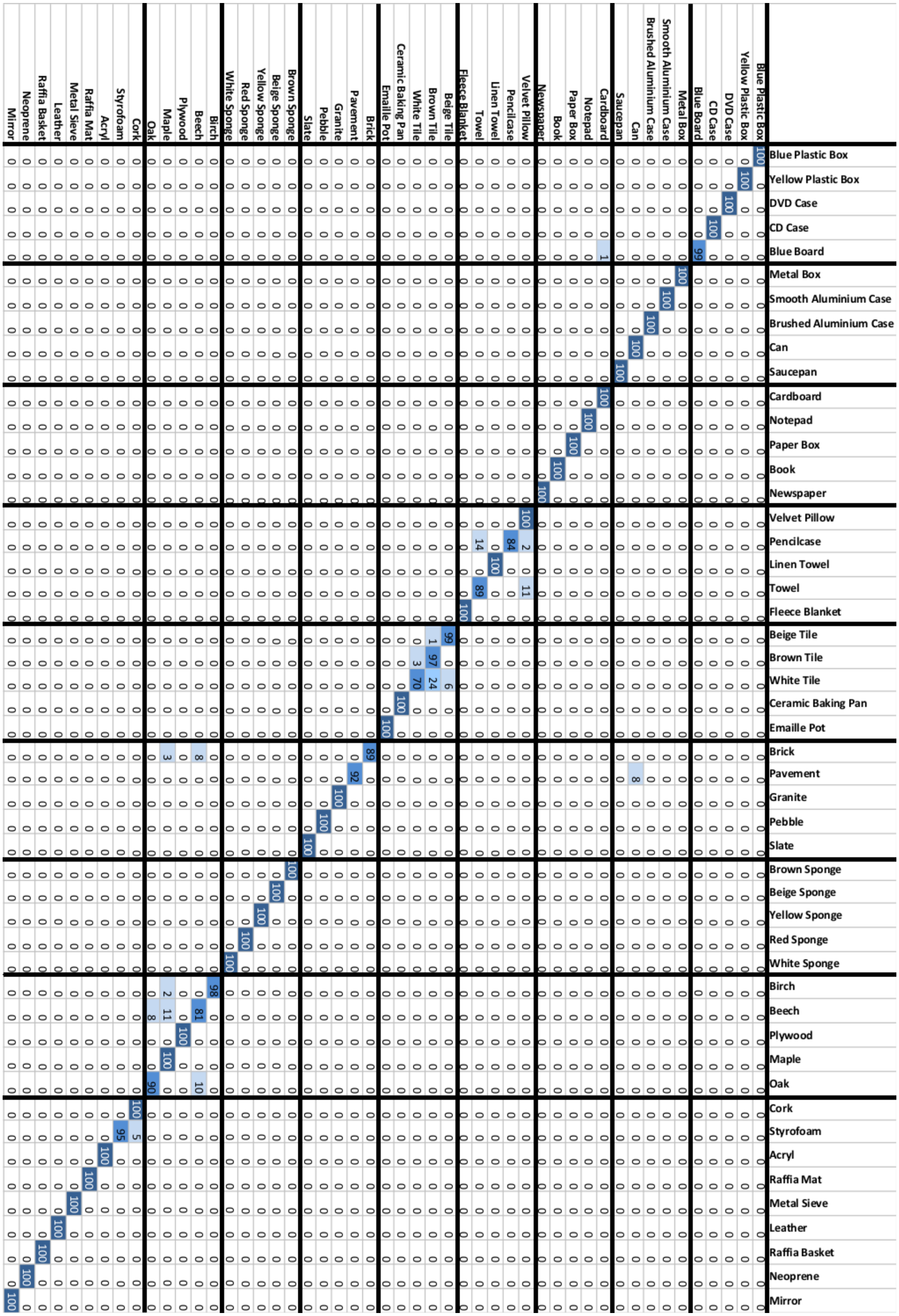


Fig. 14: Confusion matrix for all 49 objects, with concatenation of all movements, mean feature and linear SVM. Labels on page right are ground truth. Top of page labels are classifier output. Thicker lines indicate material classes.