

# Interactive Multimodal Recognition Of Household Objects by a Robot

Taylor Bergquist, Ugonna Ohiri, Conner Schenck,  
Mentors: Jivko Sinapov, Shane Griffith, and Alex Stoytchev

Developmental Robotics Laboratory  
Iowa State University  
{knexer, uohiri, cschenck, jsinapov, griffith, alexs}@iastate.edu

**Abstract**—This paper proposes a system for interactive multimodal recognition of household objects by a robot. Robots will need as many senses as possible in order to interact with these objects in our humanoid environments. A robot was equipped to observe both proprioceptive and auditory sensory input while performing five different behaviors on the objects in its environment. The robot learned from the sound produced by the object and from the change in joint torque during each interaction. The results show that an interactive approach to object recognition using both auditory and proprioceptive feedback performs better than auditory feedback alone. We conclude that a multimodal approach to object recognition is advantageous because one mode of sensory input may uncover important features that other modes cannot sense.

## I. INTRODUCTION

Sound has been shown to be very useful for interaction-based object recognition tasks, but not all objects can be identifiable with that sensor alone. For example, when a robot interacts with a sponge, the sound of the sponge is so faint that the sound of the robot drowns it out. According to *Sapp et al.* [8], without multimodal input, even humans can be easily deceived (i.e. the appearance and reality distinction). For example, people are often fooled by a bowl of fake fruit. This is especially true when people use only visual information. Only when a person receives proprioceptive feedback by touching or picking up the table ornament or by trying to take a bite do they realize that it's fake. Presumably, a robot that is also



Fig. 1. The upper-torso robot used in the experiments.

equipped with proprioceptive feedback would perform better at object recognition tasks than a robot that collected a single mode of information alone.

In a study done by *Sapp et al.*, [8], toddlers were presented with a sponge that was deceptively painted as a rock. All of the toddlers believed that the object was a rock until the moment they touched it or picked it up. This shows that input from multiple modalities is very important for object recognition, and some are more important than others (in the case of the rock-looking sponge, visual information could recognize the object faster, but proprioceptive feedback was found to be more reliable). Hence, robots that interactively learn to recognize objects using multimodal sensory input may be better equipped for the task.

This paper builds on previous work on acoustic object recognition by an upper-torso humanoid robot by adding proprioceptive feedback. The hypothesis is that although object recognition is significantly better than chance with auditory information alone,

performance will increase by adding proprioceptive data. The robot interacted with 50 objects using five different interactions (lift, shake, drop, crush, and push). The robot extracted and learned to recognize objects using a Self-Organizing Map. Using either modality alone, the robot was able to recognize objects significantly better than chance, but the robot performed best when it recognized objects using both modalities. The results support the conclusion that interaction-based object recognition increases in reliability not only as the number of interactions are increased, but also as robots are given more modes for perceiving the environment.

## II. RELATED WORK

Relatively few studies have equipped robots with multiple modalities for the purpose of interactively learning to recognize objects. Most previous studies have relied visual or aural information alone. Furthermore, only a small number have included proprioceptive feedback.

Although there has not been much work using multiple modalities for object recognition, there has been work where robots have used multiple modalities to interact with the environment in other ways. The work of *Arsenio and Fitzpatrick* [2] explored associating the periods of motion of objects in both the auditory and visual modalities. Similarly, the work of *Nakamura et al.* [4] explored using one modality to infer the properties of an object in another modality (e.g. whether it would make noise when picked up after only looking at it). They found a much higher correlation between visual and haptic information than visual and auditory information. Additionally, *Fitzpatrick et al.* [1] also demonstrated that a robot could more easily delineate objects when using both visual and acoustic properties. Their approach developed periods in the visual field and tried to match it with the object's acoustic field.

*Modayil and Kuipers* [5] used visual and auditory dynamic readings to track amorphous objects, as well as those that had a well-configured shape. Essentially implementing two modalities increased the robot's ability to track moving objects. The more senses present on a robot, the more informational pieces of an object.

Adding the sense of proprioception helps to understand the physical features of an object. *Montesano et al.* [3] interactively learned about objects' affordances using vision as a single modality. They argued it was rather difficult to learn several affordances without using all of our senses. However, they created a model for learning object affordances that used only one modality. The main

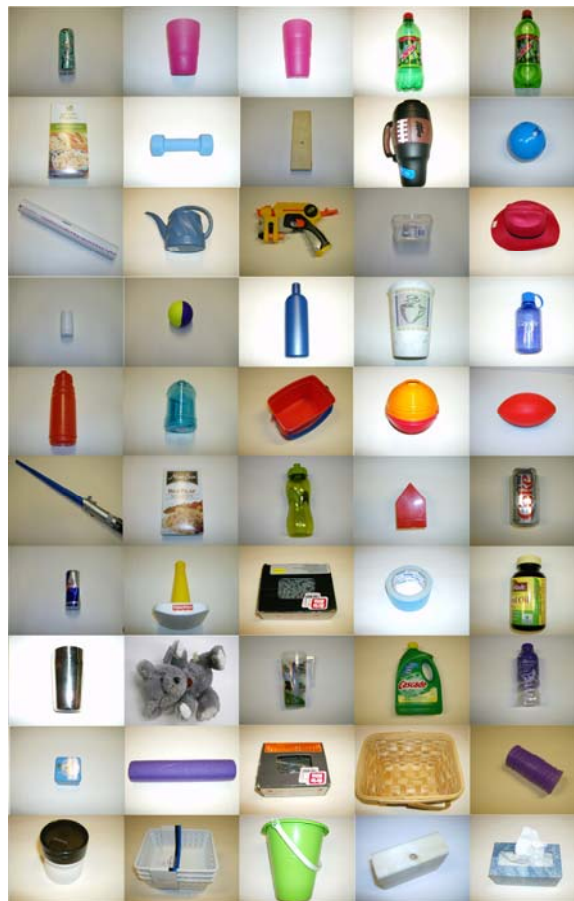


Fig. 2. The 50 objects used in the experiments (not drawn to scale). Identifying from left to right, *First row*: Arizona tea can, small plastic cup, large plastic cup, empty mountain dew bottle, full mountain dew bottle. *Second row*: macaroni shells, dumbbell, 2-by-4 wood block, football mug, blue ball. *Third row*: pbc pipe, watering can, nerf gun, closed plastic container, cowboy hat. *Fourth row*: white pills, nerf ball, shampoo bottle, Styrofoam cup, google water bottle. *Fifth row*: red water bottle, clear water jug, red container, plastic hemisphere ball, nerf football. *Sixth row*: toy light saber, rice pilaf, green water bottle, lego house, diet coke can. *Seventh row*: red bull can, peg, large screws, masking tape, brown pills. *Eighth row*: grey cup, teddy bear, clear cup, detergent bottle, clear water bottle. *Ninth row*: metal tin, water noodle, small screws, picnic basket, purple cushion. *Tenth row*: black/white mug, white container, green container, 2-by-2 wood block, tissue box.

characteristics of their model included: 1) capturing the relations between the robot's actions, objects features and the observed effects; 2) learning through examination and interaction with the world; 3) distinguishing the features that are important in each affordance; 4) providing a common model for learning. They stated that one method of improving

their approach would be to add multiple modalities of input.

This study built on previous work by *Sinapov and Stoytchev* [6]. Previously the robot used only auditory information for object recognition and categorization. The robot was able to successfully recognize and categorize objects using a single modality of input. This paper shows that by incorporating proprioception, object recognition rates greatly increase. Eventually, the goal is to incorporate many modalities such that it is possible to learn to recognize objects using very few interactions.

### III. EXPERIMENTAL SETUP

#### A. Robot

This study used an upper-torso humanoid robot, equipped with a 7-DOF Barrett Whole-Arm Manipulator and a 3-finger Barrett Hand as its end effector (as shown in Fig. 1). The robot arm is programmed and controlled from a Linux PC at 500 Hz over a CAN bus interface.

The robot is equipped with a U853AW Hanging Microphone in its right chest. Sound input was recorded at 44.1 KHz using the Java Sound API over a 16-bit channel. The microphone's output was directed through an ART Tube MP Studio Microphone pre-amplifier.

#### B. Objects

A set of 50 commonly used household objects was selected for these experiments. (see Fig. 2.) They included cups, water bottles, tissue boxes, etc. Objects were selected using three criteria: 1) it was graspable by the robot; 2) the object would not break when dropped; and 3) it would not damage the robot. Some of the objects had contents inside of them such as the box of screws and the capsule of pills, which distinguished them acoustically from other objects. The objects were also made of various substances such as metal, plastic, and wood. Each object was placed at a marked position at the center of the table within the robot's reach.

It is important to note that because objects with distinct shapes and substances were used, the robot was able to learn to find similarities among them by using both aural and proprioceptive data. Still even with a limited number of substances, the actual number of differences present is much larger – for example the plastic used to make a cup is much different than the plastic used to make a water bottle.

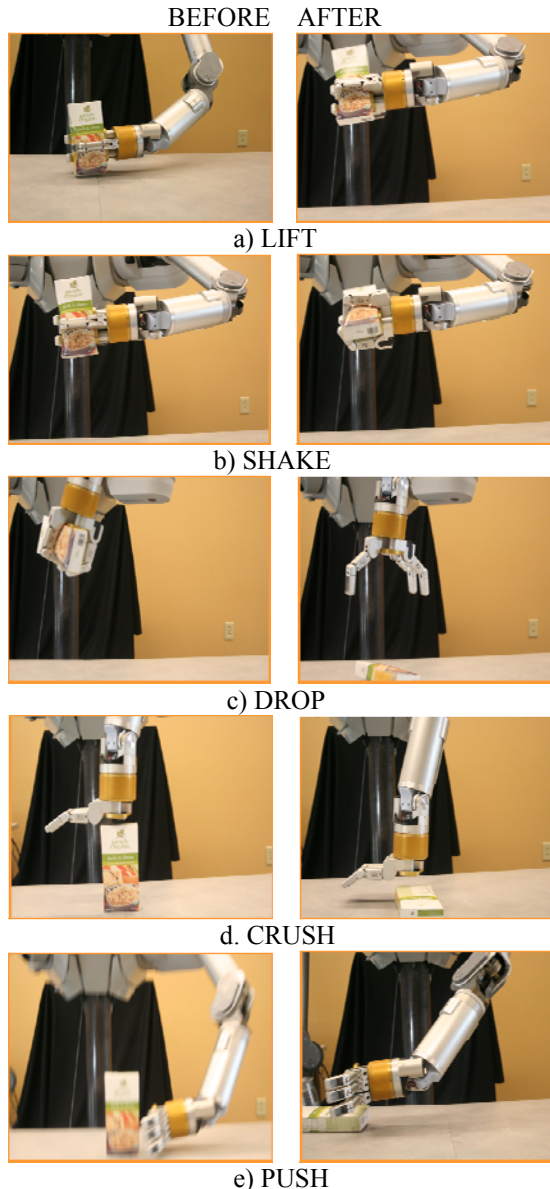


Fig. 3. The upper-torso robot, as it performed five different behaviors in a single trial with the box of macaroni. The robot a) lifted and shook it, b) dropped it, c) crushed it, and d) pushed it.

Such differences were expected to make object recognition possible.

#### C. Behaviors

The robot performed five different behaviors on each of the 50 objects: lift, shake, drop, crush, and push during each trial (see Fig. 3.). After grasping the object, the robot lifted it, shook it, dropped it, and pushed it, respectively. The robot completed these interactions using the Barrett WAM API. The crush interaction did not actually crush the objects. Usually,

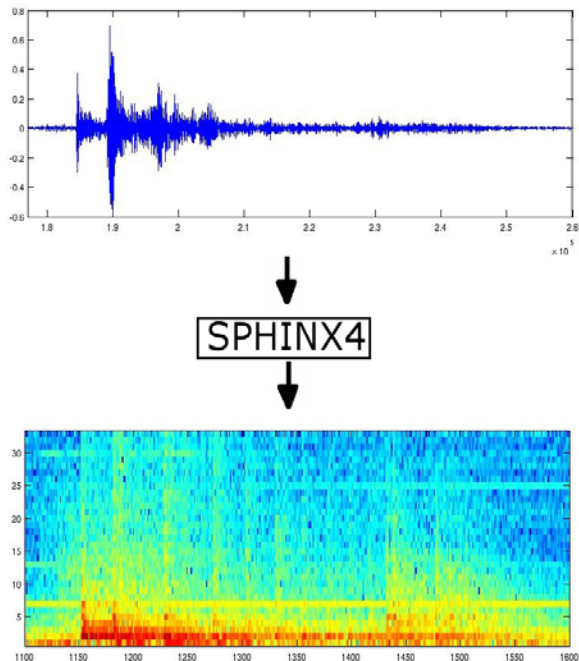


Fig. 4. Audio signal processing. The top graph is the raw sound recorded for object 16 (watering can) when dropped. The bottom graph is the spectrogram of the discrete fourier transform. The horizontal axis denotes time and the vertical axis denotes the 33 frequency bins.

the downward force made the object fall over. Also, the push was relatively swift, which frequently caused the object to topple.

The five behaviors were chosen to be able to examine the effects of object recognition using multiple modalities. An inspiration arose from the way humans interact with objects. For example, behaviors such as shake and drop would help them be able to acoustically tell the difference between a box of macaroni and a box of rice. Other behaviors such as lift and crush could tell whether they had an empty vs. a non-empty bottle of milk.

## IV. LEARNING METHODOLOGY

### A. Data Collection

The robot performed 10 trials with each object for a total of 50 objects \* 10 trials \* 5 interactions = 2500 data segments. For each interaction the object was placed in the same position and orientation in order to maintain consistency.

Data collection began when the trajectory for each interaction started and continued for a fixed amount of time, which ended after the interaction was over. After determining the length of each interaction, both the audio and proprioceptive data

were automatically cropped in order to isolate the relevant data. This resulted in excess data for each interaction. The proprioceptive data was recorded at 500 Hz. The joint positions, torques, and hand positions were all recorded. No filtering was used on the audio.

### B. Feature Extraction

The log-normalized Discrete Fourier Transform (DFT) was run on the audio using the SPHINX4 natural language processing library with default parameters in order to get a good representation of the data. It split the audio into thirty-three frequency bins at each time slice, each bin containing the intensity of the corresponding frequency range. An example of a sound wave and the resulting DFT are shown in Fig. 4.

The feature extraction algorithm used in this study was a Self-Organizing Map (SOM). It creates a predefined number of states, and then distributes them throughout the feature space to represent the data given a small portion of the data to train on. Then each data point in the set is converted to a sequence of states. A different instance of the same SOM was used for each data set (one for proprioception and one for audio). Due to limits on computing resources and time, only twenty percent of the data, sampled at random, was used to train each SOM.

The SOM's used were of size 6 x 6, or contained a total of 36 nodes. The SOM for the audio data was trained with 33-dimensional input data (all the frequency bins at a given time slice). The SOM for the proprioceptive data was trained with 7-dimensional input data (all the torque values at a given time slice). An example of the unfiltered and the filtered torque values can be seen in Fig. 5. The Growing Hierarchical SOM toolbox for Java was used to train the SOM's [8]. The SOM was trained using the default parameters for a non-growing 2-D single layer map. Fig. 6. gives a visual overview of the training procedure.

After training, each vector (time slice) of both the proprioceptive and the audio data were mapped to the state with the highest activation in their corresponding SOM's. This created a sequence of states for each data point. The state sequences were used for the rest of the algorithms.

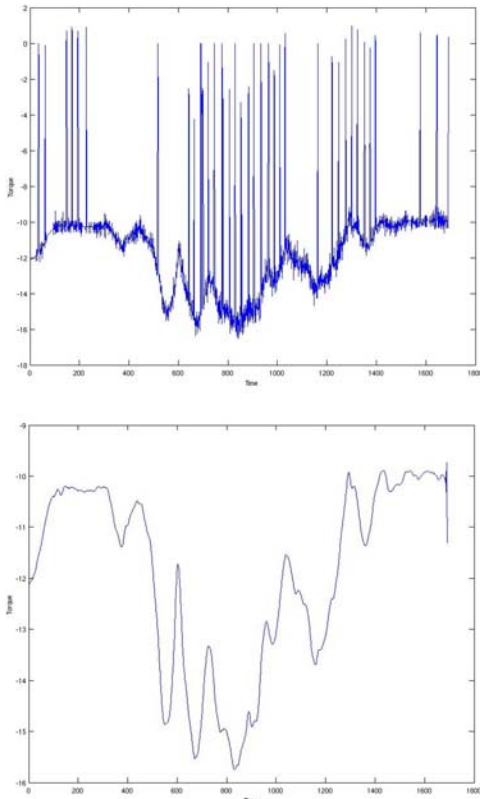


Fig. 5. The top graph shows the torque value for joint two during lift before filtering. First an outlier filter was applied to remove the random spikes in the data. All values outside of three standard deviations of a moving window were filtered. After a running average operation with a window size of ten was used to smooth the curve. The bottom graph shows the results of the filters.

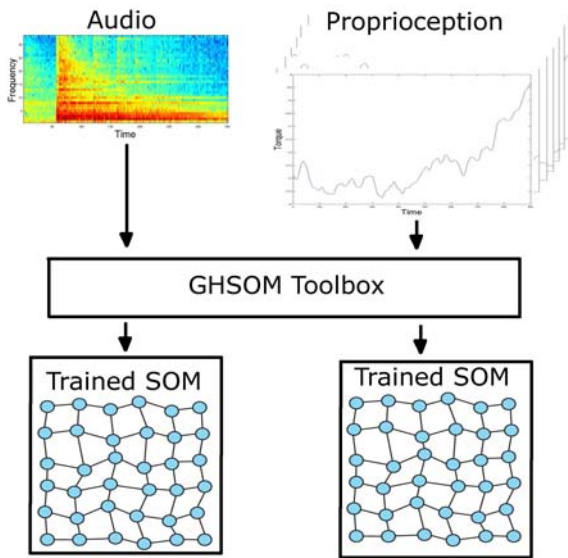


Fig. 6. Illustration of the procedure used to train the SOM's. Given a set of spectrograms for the audio data and a set of vectors for the proprioceptive data, column vectors were sampled at random from each and used to train each SOM.

### C. Learning Algorithms

Two learning algorithms were used to solve the task of object recognition: k-Nearest Neighbor (a distance-based learning algorithm) and Multinomial Naïve Bayes (a Bayesian probabilistic model).

#### 1) K-Nearest Neighbor

K-Nearest Neighbor (k-NN) is a distance-based algorithm [12], which does not build an explicit model of the training data. Instead, given a test data point, it simply finds the  $k$  closest neighbors and output a prediction, which is a smoothed average over those neighbors. In this study,  $k$  was set to 3.

The k-NN algorithm requires a distance measure, which can be used to compare the test data point to the training data points. Since each data point in this study is represented as a sequence over a finite alphabet, we used the Needleman-Wunsch global alignment algorithm [9], [10], which can estimate how similar two sequences are. While normally used for comparing biological or text sequences, the algorithm is applicable to other situations that require a distance measure between two strings. The algorithm requires a substitution cost to be defined over each pair of possible sequence tokens (i.e., letters): e.g., the cost of substituting 'a' with 'b'. Since each token represents a state on a Self-Organizing Map, the cost for each pair of tokens was set to the Euclidean distance between their corresponding SOM states in the 2-D plane.

#### 2) Multinomial Naïve Bayes

The second learning algorithm used in this study was Multinomial Naïve Bayes (MNB), which falls into the family of probabilistic models. MNB is commonly used for sequence classification tasks and has found wide applicability in natural language processing, bioinformatics, natural language processing, and more. [13]

Under the MNB model, each sequence  $S_i$  is represented as a vector  $d_i = (x_{i1}, \dots, x_{i|V|})$  of counts where  $V$  is the vocabulary and each  $x_{it} \in \{0, 1, 2, \dots\}$ . Each  $x_{it}$  indicates the number of times word  $w_t$  occurs in  $S_i$ . For example, if the sub-sequence "ab" appears 50 times in the sequence, then  $x_{i'ab} = 50$ . Given this representation, the task of the MNB model is to assign the correct object class label  $c_i$  given an audio (or a proprioceptive) sequence  $S_i$ . Given model parameters  $p(w_t | c_j)$  and class prior probabilities  $p(c_j)$ , MNB computes the most likely class for a data point  $d_i$  in the following way:

$$\begin{aligned}
c^*(d_i) &= \operatorname{argmax}_j p(c_j)p(d|c_j) \\
&= \operatorname{argmax}_j p(c_j) \prod_{t=1}^{|V|} p(w_t|c_j)^{n(w_t,d_i)}
\end{aligned}$$

where  $n(w_t, d_i)$  is the number of occurrences of word  $w_t$  in sequence  $S_i$  as specified in the feature vector  $d_i$ . The probabilities  $p(w_t|c_j)$  and  $p(c_j)$  are estimated from the available training data using maximum likelihood with a Laplacian prior (see [13] for details).

To compute the feature vector  $d_i$  for each sequence  $S_i$ , we used k-Gram features with  $k = 2$ . Hence, the vocabulary  $V$  consisted of all possible single and double letter combinations. With 36 states in the SOM, this corresponds to a feature vector of length  $36+36^2=1332$ .

#### D. Evaluation

Ten-fold cross-validation was used to evaluate the learning scheme. To do this, each trial was systematically selected to be the test data set while the other nine trials were used as the training set. The object recognition accuracy was averaged over all the objects for each interaction. When combining multiple object recognition rates (e.g. using more than one interaction or modality), the probabilities that the test object is each object were averaged, resulting in an overall probability for each object between interactions and modalities. This method is scalable to an arbitrary number of modalities and interactions.

Because the k-Gram ignores much of the relationships in the sequences generated by the SOM (e.g. a temporal ordering longer than  $K$ ), it was not expected to perform as well as global alignment and k-NN. It was included in this study as a baseline.

## V. RESULTS

In the following experiments, the robot was tested on its ability to correctly predict the object in the interaction given the auditory information, the proprioceptive information, or both. That is, given novel information, the robot attempted to predict the object that generated the information. The performance was evaluated using 10-fold cross-validation: the 2500 data points are distributed evenly amongst 10 folds such that each fold contains exactly one data point for each object and interaction. During each iteration, nine of the ten folds are used for training the k-NN and Bayesian models while the remaining fold is used for testing. The performance

of the models is reported in terms of the percentage of correct predictions (the accuracy). Note, the accuracy is reported with only one interaction. It is expected that recognition rates would increase as the number of interactions used for recognition increased. Experiment 4 shows how object recognition performance increases with more interactions.

#### A. Auditory Recognition

In the first experiment, the robot is given only auditory information. This experiment is comparable to *Sinapov et al.* [11] work. Table I shows the recognition accuracy of the k-NN and Bayesian models [14] for this experiment. The k-NN model was on average 51.6% accurate, while the Bayesian model was on average 38.24% accurate. The k-NN model was also more accurate than the Bayesian model for every interaction except for 'shake'. The 'drop' interaction was the most useful in both models, although the difference was much more pronounced in the Bayesian model. The 'lift' behavior was the least useful of the five in both cases.

The relative performance of the models makes sense. The k-Gram feature extraction used by the Bayesian model discards all temporal relationships that are not between adjacent time slices. This information can be very important, as the object may bounce or fall in a specific fashion. It also makes sense that the 'drop' interaction was the most useful, as the sounds objects produced when they fell were very distinctive.

Overall, the k-NN model performed well, but not as well as it did in *Sinapov and Stoytchev's* [6] work. This is most likely due to the larger dataset in this work (50 objects instead of 36) and to the exclusion of the 'grasp' interaction in favor of the 'lift' interaction. The 'lift' interaction contains a significant amount of proprioceptive information, but very little auditory information.

#### B. Proprioceptive Recognition

In the second experiment, the robot is given only proprioceptive information. Table II shows the recognition accuracy of the k-NN model and the Bayesian model for this experiment. The k-NN model was on average 45.12% accurate, while the Bayesian model was on average 30.24% accurate. As before, the k-NN model was generally more accurate than the Bayesian model. The 'crush' interaction was the most useful in both models. The 'push' behavior was the least useful in the Bayesian model, while the 'push' and 'shake' behaviors were essentially tied for least useful in the k-NN model.

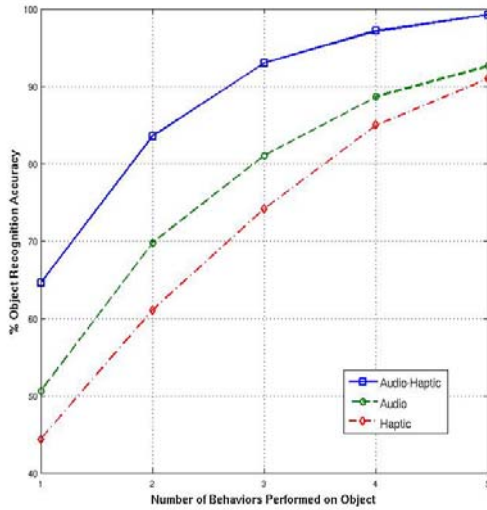


Fig. 7. Object recognition rates with k-NN as the number of interactions utilized is varied from 1-5 with each combination of modalities.

The recognition rates in this experiment were clearly better than chance for every interaction, and for a few - notably the 'crush' interaction in the k-NN model - they were very high. It is clear that the model used in *Sinapov and Stoytchev's* work [6] is effective for proprioceptive recognition as well as auditory recognition

### C. Multimodal Recognition

In the third experiment, the robot is given both proprioceptive and auditory information. Table III shows the recognition accuracy of the k-NN model and the Bayesian model for this experiment. The k-NN model was on average 65.56% accurate, while the Bayesian model was on average 50.64% accurate. Both models benefited from the combination of multiple modalities, but the benefit was larger for the k-NN model. The benefit appeared to be the greatest to those interactions which were equally reliable in the two modalities - if one was much more accurate than the other, combining their predictions yielded no discernible improvement over the more accurate of the two. As a result, the majority of the improvement in average recognition rates comes from the robot's reliance on the more reliable modality in each case.

### D. Recognition with Multiple Interactions

In the fourth experiment, the robot's performance is evaluated when it is given multiple data points from different interactions with the same object. In this scenario, the k-NN model calculated the probability that each data point belonged to each

Table 1: Recognition Results using Audio

| Interaction                    | Accuracy     |
|--------------------------------|--------------|
| <i>Multinomial Naïve Bayes</i> |              |
| Lift                           | 17.2%        |
| Shake                          | 39.8%        |
| Drop                           | 71.4%        |
| Crush                          | 23.6%        |
| Drop                           | 39.2%        |
| <b>Average</b>                 | <b>38.2%</b> |
| <i>k-NN</i>                    |              |
| Lift                           | 17.4%        |
| Shake                          | 27.0%        |
| Drop                           | 76.4%        |
| Crush                          | 73.4%        |
| Push                           | 63.8%        |
| <b>Average</b>                 | <b>51.6%</b> |

Table 2: Recognition Results using Proprioception

| Interaction                    | Accuracy     |
|--------------------------------|--------------|
| <i>Multinomial Naïve Bayes</i> |              |
| Lift                           | 36.8%        |
| Shake                          | 17.0%        |
| Drop                           | 21.8%        |
| Crush                          | 65.2%        |
| Push                           | 10.4%        |
| <b>Average</b>                 | <b>30.2%</b> |
| <i>k-NN</i>                    |              |
| Lift                           | 64.8%        |
| Shake                          | 15.2%        |
| Drop                           | 45.6%        |
| Crush                          | 84.6%        |
| Push                           | 15.4%        |
| <b>Average</b>                 | <b>45.1%</b> |

Table 3: Recognition Results using Audio & Proprioception

| Interaction                    | Accuracy     |
|--------------------------------|--------------|
| <i>Multinomial Naïve Bayes</i> |              |
| Lift                           | 37.4%        |
| Shake                          | 40.0%        |
| Drop                           | 71.6%        |
| Crush                          | 65.2%        |
| Push                           | 39.0%        |
| <b>Average</b>                 | <b>50.6%</b> |
| <i>k-NN</i>                    |              |
| Lift                           | 66.0%        |
| Shake                          | 29.4%        |
| Drop                           | 81.2%        |
| Crush                          | 88.0%        |
| Push                           | 63.2%        |
| <b>Average</b>                 | <b>65.6%</b> |

object class. These probabilities were then summed, and the object class with the greatest resulting value was taken as the robot's prediction. The number of interactions utilized was varied from 1 (the default setting, used to generate Tables I, II, and III) to 5 (using all of the interactions).

Fig. 7. shows the recognition accuracy of the k-NN model as the robot uses information from multiple interactions. As the robot uses information from more interactions, performance improves dramatically. This figure also shows that the improvement seen when allowing the robot to use information from multiple modalities extends to the case where multiple interactions are used as well. When the robot uses only audio as a single modality [6] its recognition performance is 92.6%; but when the robot uses both the auditory and the proprioceptive information from all five interactions, its recognition performance jumps to 99.2%.

## VI. CONCLUSION/DISCUSSION

This paper extended a learning framework, which had been previously applied to auditory information to another modality (proprioception) and combined the predictions of an auditory model and a proprioceptive model. A large-scale experimental study evaluated the effectiveness of this extension and combination. It was found that the methodology that had been used in previous work was effective on both auditory and proprioceptive information, and that the combination of information from the two modalities could substantially and consistently improve recognition accuracy.

The robot was evaluated on 50 household objects, using 5 exploratory behaviors: lift, shake, drop, crush, and push. Recognition accuracy was fair with only one interaction and one modality, but jumped to 99.2% if both modalities and all five interactions were used. The large number of objects indicates that both auditory and proprioceptive recognition scale well with the number of objects in the experiment.

There are several logical extensions to this work. First, the number of modalities used could be increased - vision could be added, for example. It would even make sense to consider different representations of the same information to be different modalities. Second, the proprioceptive and auditory information, even in cases where they are not sufficient to confidently predict the specific object involved in the interactions, could be very useful in determining some of the physical properties of the object, such as its material.

## VII. ACKNOWLEDGEMENTS

This research was performed at Iowa State University as a part of a Research Experience for Undergraduates (REU) internship sponsored by NSF (0851976). We would like to thank the Human Computer Interaction (HCI), Virtual Reality Applications Center (VRAC), and the Program for Women in Science and Engineering for hosting and sponsoring the REU Program for Summer 2009. We would also like to thank our mentors Jivko Sinapov, Shane Griffith, and Matthew Miller for their help with the applications development.

## VIII. REFERENCES

- [1] G. Metta, and P. Fitzpatrick. "Early integration of vision and manipulation." *Adaptive Behavior*, 11(2): 109–128, 2003.
- [2] D. Perzanowski, A. Schultz, W. Adams, E. Marsh and M. Bugajska. "Building A Multimodal Human-Robot Interface", *IEEE Intelligent Systems*, vol. 16, no. 1 pp. 16-21, 2001.
- [3] L. Montesano, M. Lopes, A. Bernardino, and J. Santos-Victor. "Learning object affordances: From sensory-motor coordination to imitation", *IEEE Transactions on Robotics*, vol. 24, no. 1, pp. 15-26, 2008.
- [4] T. Nakamura, T. Nagai, and N. Iwahashi. "Multimodal object categorization by a robot," in *Proc. of IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pp. 2415–2420, 2007.
- [5] Modayil and J. Kuipers. "Autonomous shape model learning for object localization and recognition." In: *IEEE International Conference on Robotics and Automation*, IEEE Computer Society Press, Los Alamitos, 2006.
- [6] J. Sinapov, M. Weimer, and A. Stoytchev, "Interactive learning of the acoustic properties of objects by a robot," in *Proceedings of the RSS Workshop on Robot Manipulation: Intelligence in Human Environments, Zurich, Switzerland*, 2008.
- [7] F. Sapp, K. Lee, and D. Muir. Three-year-olds' difficulty with the appearance-reality distinction: is it real or is it apparent?. *Dev. Psychol.* 36, pp. 547–560, 2000.
- [8] T. Kohonen. *Self-Organizing Maps*. Springer, 2001.
- [9] G. Navarro. "A guided tour to approximate string matching," *ACM Computing Surveys*, vol. 33, no. 1, pp. 31–88, 2001.
- [10] S. Needleman and C. Wunsch. "A general method applicable to the search for similarities in the amino acid sequence of two proteins," *J. Mol. Biol.*, vol. 48, no. 3, pp. 443–453, 1970.
- [11] J. Sinapov, M. Wiemer, and A. Stoytchev. "Interactive learning of the acoustic properties of household objects," in *Proc. of the IEEE Intl. Conf. on Robotics and Automation (ICRA)*, 2009.
- [12] K. Fukunaga and P. Narendra. "A branch and bound algorithm for computing k-Nearest Neighbors. *IEEE Trans. Comptrs.* C-24, pp. 750-753, 1975.



[13] Y. Yang and X. Liu. "A re-examination of text categorization methods." In Proceedings of SIGIR-99, 22nd ACM International Conference on Research and Development in Information Retrieval (Berkeley, CA, 1999), 42–49, 1999.

[14] K. Murphy, "Dynamic Bayesian networks: Representation, inference and learning." Ph.D. thesis, UC Berkeley, Computer Science Division, 2002.

[15] X. Huang. "On global sequence alignment." Comput. Appl. Biosci. 10, pp. 227–235, 1994.