Shaking eases Object Category Acquisition: Experiments with a Robot Arm

Shinya Takamuku, Koh Hosoda, and Minoru Asada
JST ERATO, Osaka University
2-1 Yamadaoka, Suita
Osaka 565-0871 Japan
{shinya.takamuku, hosoda, asada}@ams.eng.osaka-u.ac.jp

Abstract

This paper introduces a synthetic study which investigates the role of shaking behavior in object category acquisition. By shaking objects with several holding positions and obtaining the amplitude spectrums of the auditory signal, a robot with poor control ability can acquire object categories such as rigid objects, paper materials, and PET bottles with water. The result indicates the possiblity that dynamic touch is more effective for categorization than static touch such as grasping since they provide the agent with the information of the whole object enabling them to acquire object categories independent of size, shape, and contact condition.

1. Introduction

Many attempts to develop humanoid robots which live with humans in daily environments are made today (Fujita and Kitano, 1998) (Ishiguro et al., 2001) (Sakagami et al., 2002) (Kaneko et al., 2004). Such robots are expected to recognize and handle various objects that exist in daily environments to help and entertain the users. However, most current humanoid robots are equipped with fixed recognition systems based only on visual information such as colors and textures. Therefore, it is often the case that they can handle only few predefined objects such as balls and markers, and recognition often fails under different light conditions or backgrounds. In the case of humanoid robots which handle wider variety of objects often unknown in advance, the ideal approach would be to acquire multi-modal object categories by themselves through interaction with the objects.

The process of acquiring daily object categories is what every human infant goes through (Slater and Lewis, 2007). Human children are known to be an active explorer of the world from the very first stage of their development (Hofsten, 1984) (Molina and Jouen, 2004). They actively touch the objects to form multi-modal representation of the objects (Gibson and Walker, 1984) (Rochat, 2004) (Streri and Feron, 2005), although the concrete process is not well understood. Accordingly, several robotic experiments have been performed to investigate the role of infants' typical exploring behavior in acquiring multi-modal representation of objects. Just to mention a few, Natale et al. (Natale et al., 2004) made a humanoid robot grasp various objects to obtain haptic information such as shape and hardness which are subsequently processed in a self organizing map to form the representation of the objects. The research is extended by robots with tactile sensors to study how cutaneous sense helps in recognizing the objects (Natale and Torres-Jara, 2006) (Takamuku et al., 2007). On the other hand, Ogata et al. (Ogata et al., 2005) made their humanoid robot hit the objects to obtain their dynamic properties. In this case, the representations were built by processing the resulting multi-modal sensory sequence with a recurrent neural network with parametric bias. These existing studies have succeeded in showing the importance of active tactile exploration in enriching the representation of the objects. However, we observe a critical problem which make it difficult to apply the method to object category acquisition of humanoid robots and to explain the case in infants. The exploring behavior investigated in the existing studies such as grasping and hitting can only obtain information of object parts. Therefore, existing approaches tend to fail in recognizing the object category when the size, shape, or contact condition changes.

In order to avoid such a problem, we focus on another frequently observed behavior of infants, shaking. Although the difficulty of measuring and controlling the behavior have kept it from being a hot topic in the field of developmental psychology, there are enough convincing reasons to do so. Firstly, it is pointed out by Turvey (Turvey, 1996) and his colleagues that shaking behavior gives rich information of the whole object. This effect eases the acquisition of object categories which can be generalized to objects with different sizes, shapes, and contact conditions. Secondly, the rhythmic actuation in shaking behaviors realizes entrainment (Williamson, 1998) which are expected to enable stable recognition under rough control. This characteristic is important to considering the case in infants who also have difficulty in precise control.

In this paper, we introduce preliminary results of a robotic experiment which investigates the effectiveness of shaking behaviors in object category acquisition. Although several exisiting studies are found which shows that shaking behavior helps object recognition of rigid objects by detecting the momentum of inertia (Atkeson et al., 1985), we show for the first time that shaking behaviors are also effective for acquiring humanlike daily object categories not limited to rigid ones. The paper consists as follows. First, we introduce some related work and describe our system of categorization. Next, we explain the experiment design and show the results with some analysis and discussions. Finally, conclusion is given.

2. Related Work

The role of shaking behavior in exploring object properties was initialy pointed out by the research group of University of Connecticut headed by Turvey (Turvey, 1996). They gave experimental results which imply that shaking behavior, also refered to as dynamic touch, gives information of object lengths, shapes, and contact conditions. Their work still remains as one of the largest efforts on this topic. However, the work dealt with only rigid objects such as rods with different length. In the field of developmental psychology, very few studies on shaking behavior are found (Shimizu and Norimatsu, 2005) (Kloos and Amazeen, 2002). The lack of such studies comes from the difficulty of objective measurements. If the object to shake is complex as in the case of daily environments, it is also not plausible to study the phenomena with simulation. Recently, robotic experiments are attracting interest as an alternative approach to investigate the experience of infants in a objective manner and discuss the information process. Two studies are found for shaking where Williamson (Williamson, 1998) investigate the arm control with neural oscillators, and Suzuki et al. (Suzuki et al., 2006) attacks the problem of object recognition for two cylinders with different length. However, the role of shaking behavior in object categorization is not well investigated.

3. System

We propose a system of acquiring object categories from sensory sequence obtained through shaking behavior. Fig. 1 shows a sketch of the information flow in the system for the categorization. First a rhythmic actuation on the arm produces stable cyclic behavior under rough control by virtue of entrainment. Then, the obtained sensory sequence are subsequently processed by a Fourier transform circuit. Such circuits are also found in human ears known as the cochlea (Purves et al., 2004). Finally, the Fourier components such as amplitude spectrums are utilized as feature vectors to categorize the objects with a pattern recognition system. In order to obtain a object category which can be generalized to objects with different size, shape, and contact condition, the robot shakes each example object with several different contact conditions to obtain enough representative vectors. We consider it feasable to include such a process in the model since infants also change the grasping condition through bimanual control.

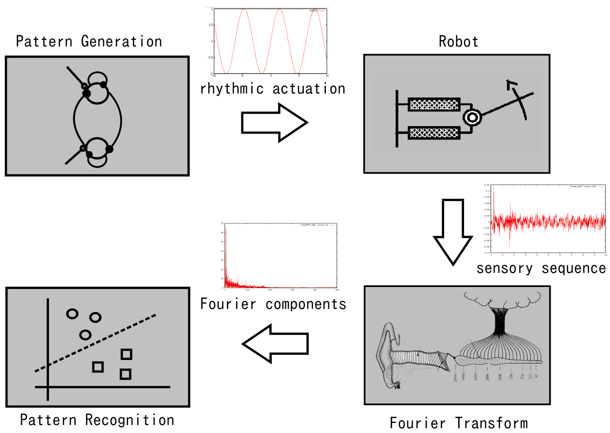

Figure 1: Information flow in proposed approach.

4. Experiment Design

In the experiment, we utilize a robotic arm with McKibben pneumatic actuators. The robot is equipped with a microphone to obtain the auditory data and a potentiometer to obtain joint angle data. Fig. 2 shows a photograph of the robot.

The arm shakes the objects in the horizontal plane to simplify the results by reducing the effect of gravity. Stable limit cycle behavior was realized by controlling the pressure of the pneumatic actuators as antiphase sinusoid curves with feedback control on the valves. We utilized a simple nearest neighbor method for pattern recognition. The distance measure d is given in the following formula where p_{1i} and p_{2i} are components of the feature vectors and N is the number of these components. In the case of amplitude spectrums or phase spectrums, $N = 10000$.

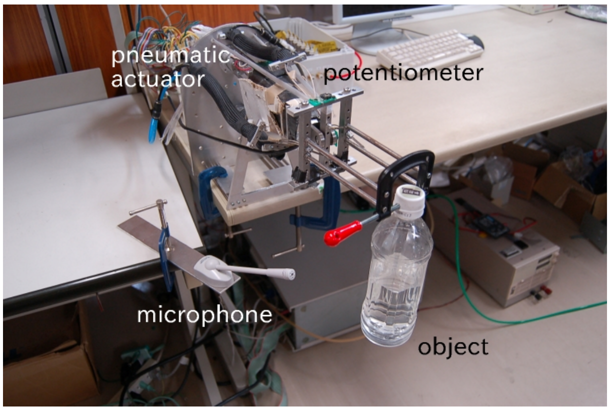

Figure 2: Robot used in the experiment.

(a) rigid bodyobjects

(b) papermaterials

(c) PET bottles with water

Figure 3: Objects used in experiment.

$$d(\boldsymbol{p}_1, \boldsymbol{p}_2) = \sum_{i=1}^{N} |p_{1i} - p_{2i}|$$

The task was to acquire and recognize three object categories, namely rigid objects, paper materials, and PET bottles with water. The objects utilized in the experiment are shown in Fig. 3. We intended to utilize daily objects instead of controlled objects to avoid being impractical. The task of acquiring daily object categories is difficult to formalize and cannot be investigated by utilizing controlled objects. The object in the left was used to teach the category by shaking them in three different grasping conditions. Then, the other two objects were shook with five different grasping conditions to test the success rate of categorization. The duration of shaking was approximately 30 seconds. The rigid objects differ in size and shape, papers differ in size and thickness, and PET bottles differ in size, shape and the amount of water inside. Since the objects have different appearances including transparent cases, it would be difficult to categorize them with only the visual information considering the change in light condition and backgrounds.

5. Results and Discussions

Before investigating the effectiveness of the proposed categorization system, we performed some analysis on which sensory data and which Fourier components would give feature vectors suitable for categorization. In order to visualize the differences within/between different object categories that would be obtained during the learning phase, we fed the sensory sequences and the Fourier components of them into self-organizing maps (Kohonen, 1995). The self-organizing map reduces the dimension while keeping the phase relation. The result of the clusterings for raw auditory data, its amplitude spectrum and its phase spectrum are shown in Figs. 7, 8, and 9, respectively. The network had two layers with 32 units each. The best matching unit for each input pattern is plotted on the 32×32 grid. The figures show that although the raw auditory data does not form a feature vector suitable for categorization, the amplitude spectrum of the auditory data turns out to be a suitable one, independent of size, shape, and contact condition. Figs. 10, 11, and 12 show the results of clusterings for SOM with raw joint angle sensory data, its amplitude spectrum, and its phase spectrum. A network with the same structure and size as the case of auditory data was utilized for the analysis. In either cases, the feature vectors were dependent to the sizes, shapes, and contact conditions and thus not feasible for the categorization. Joint angle data seemed dependent on the macroscopic feature such as mass rather than the microscopic feature such as material.

The amplitude spectrums of auditory data from the paper materials, the bottles with water, and the rigid objects used in the experiment are shown in Figs. 4, 5, and 6, respectively. The amplitude spectrum of auditory data from an arm shaking without any objects are also included in the graph to show which component is produced from the robot itself. All paper materials showed spiky curves with equally spaced decaying peaks, whereas all PET bottles with water produced relatively smooth curves with a peak in a low frequency region. The amplitude spectrums from rigid objects used in the experiment were almost the same as those obtained in no object condition. The shapes of the amplitude spectrums of auditory data was qualitatively different for objects from different categories, but qualitatively similar for objects within the same category. The result indicates the possibility to improve the robustness of categorization by introducing much more representative data for each object category. Shaking the objects with multiple different contact conditions is one way to improve the performance.

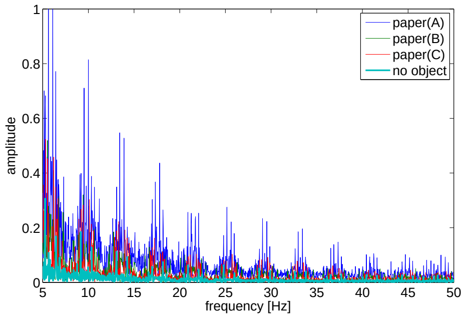

Figure 4: Amplitude spectrum of the paper materials.

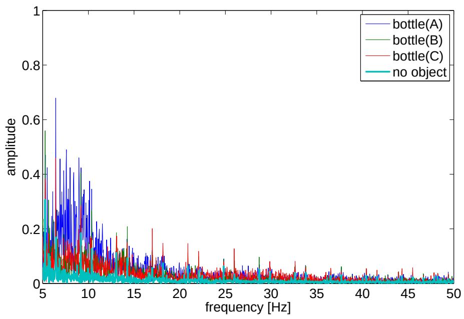

Figure 5: Amplitude spectrum of the bottles with water.

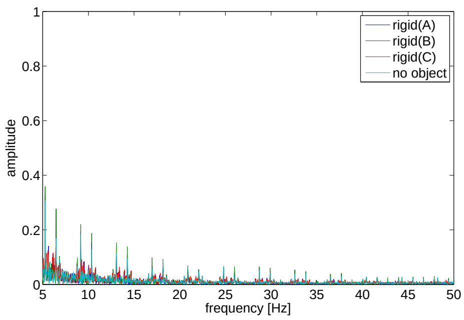

Figure 6: Amplitude spectrum of the rigid objects.

Finally, Table 1 shows the success rate of categorization by the nearest neighbor method when amplitude spectrums of different sensory data was given. The number of different contact conditions when the example object was taught is also varied to observe the effect of varying the contact condition to obtain the generalization ability. We could observe that by utilizing the auditory data, the system is able to acquire object categories which can be generalized to objects with different size, shape, and contact condition. The robot needed to shake the example objects with multiple contact conditions to obtain categories that can be generalized. However, the number of contact condition could be very few. Categorization failed with joint angle data as expected from the results shown in Fig. 11.

Table 1: Success rate of categorization.

sensor	# of contact for teaching	rigid	paper	bottle
potentiometer	3	100%	0%	0%
microphone	1	70%	10%	70%
microphone	2	100%	100%	60%
microphone	3	100%	100%	80%

6. Conclusion

The results of the experiment show that an agent with poor control ability can acquire object categories independent of size, shape, and contact condition by shaking objects in different contact conditions and utilizing the amplitude spectrum of audio sensory data as feature vectors. The difficulties of categorization faced in the current experiment are similar to those that real human infants face. The fact that extraction of amplitude spectrums of auditory data, the information processing also found in infants, proved to be effective for the task, indicates the possible role of infants' shaking behavior on object category acquisition. Observation experiments of infants are required to investigate the possibility. The acquisition of humanlike daily object categories through active exploration is a fundamental task of communication since it forms the basis for lexicon acquisition. However, the task is difficult to formalize and no standard approach is found to explain the mechanism behind this process. Our work shows that robotic experiments might turn out to be the standard approach for this issue. At the current experiment, all the objects were shook with the same actuation. As for future work, we plan to conduct observation experiments to investigate how the shaking motion should vary according to the sensory feedback. Investigating haptic sensory feedbacks of shaking behavior is another topic we plan to work on in the future.

Acknowledgement

This research was supported by the JST ERATO project: "Asada Synergistic Intelligence Project".

References

Atkeson, C. G., An, C. H., and Hollerbach, J. M. (1985). Rigid body load identification for manipulators. *Proceedings of 24th Conference on Decision and Control*, pages 996–1002.

Fujita, M. and Kitano, H. (1998). Development of a quadruped robot for robot entertainment. *Autonomous Robots*, 5:7–18.

Gibson, E. and Walker, A. (1984). Development of knowledge of visual-tactual affordances of substance. *Child Development*, 55:453–460.

Hofsten, C. (1984). Eye-hand coordination in newborns. *Developmental Psychology*, 18:498–503.

Ishiguro, H., Ono, T., Imai, M., Maeda, T., Kanda, T., and Nakatsu, R. (2001). Robovie: an interactive humanoid robot. *Industrial Robot*, 28(6):498–503.

Kaneko, K., Kanehiro, F., Kajita, S., Hirukawa, H., Kawasaki, T., Hirata, M., Akachi, K., and Isozumi, T. (2004). Humanoid robot hrp-2. *Proceedings of IEEE International Conference on Robotics and Automation*, pages 1083–1090.

Kloos, H. and Amazeen, E. (2002). Perceiving heaviness by dynamic touch: An investigation of the size-weight illusion in preschoolers. *British journal of developmental psychology*, 20(2):171–183.

Kohonen, T. (1995). *Self-Organizing Maps*. Springer-Verlag.

Molina, M. and Jouen, F. (2004). Manula cyclical activity as an exploratory tool in neonates. *Infant Behavior and Development*, 27:42–53.

Natale, L., Metta, G., and Sandini, G. (2004). Learning haptic representation of objects. *Proc. of International Conference on Intelligent Manipulation and Grasping*.

Natale, L. and Torres-Jara, E. (2006). A sensitive approach to grasping. *Proc. of International Conference on Epigenetic Robotics*.

Ogata, T., Ohba, H., Komatani, K., Tani, J., and Okuno, H. G. (2005). Extracting multi-modal dynamics of objects using rnnpb. *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 160–165.

Purves, D., Augustine, G. J., Fitzpatrick, D., Hall, W. C., Lamantia, A.-S., McNamara, J. O., and Williams, S. M. (2004). *Neuroscience*, chapter 13. Sinauer Associates Inc.

Rochat, P. (2004). *The Infant's World*, chapter 3. Harvard Univ Press.

Sakagami, Y., Watanabe, R., and Aoyama, C. (2002). The intelligent asimo: System overview and integration. *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2478–2483.

Shimizu, T. and Norimatsu, H. (2005). Detection of invariants by haptic touch across age groups: rod-length perception. *Perceptual and Motor Skills*, 100:543–553.

Slater, A. and Lewis, M. (2007). *Introduction of Infant Development*, chapter 7. Oxford Univ. Pr.

Streri, A. and Feron, J. (2005). The development of haptic abilities in very young infants: From perception to cognition. *Infant Behavior and Development*, 28:290–304.

Suzuki, M., Noda, K., Suga, Y., Ogata, T., and Sugano, S. (2006). Dynamic perception after visually guided grasping by a human-like autonomous robot. *Advanced Robotics*, 20(2):233–254.

Takamuku, S., Gomez, G., Hosoda, K., and Pfeifer, R. (2007). Haptic discrimination of material properties by a robotic hand. *Proc. of International Conference on Development and Learning*.

Turvey, M. T. (1996). Dynamic touch. *American Psychologist*, 51:1134–1152.

Williamson, M. M. (1998). Rhythmic robot arm control using oscillators. *In Proc. of IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pages 77–83.

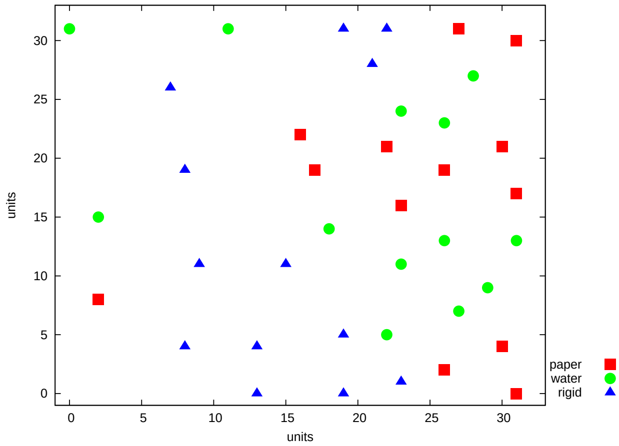

Figure 7: SOM analysis result with raw auditory data.

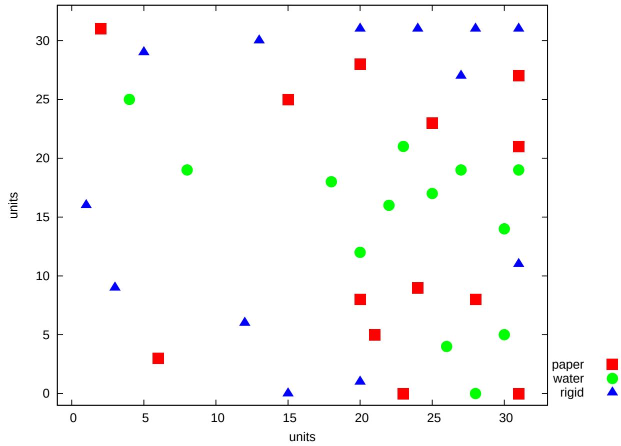

Figure 10: SOM analysis result with raw angle data.

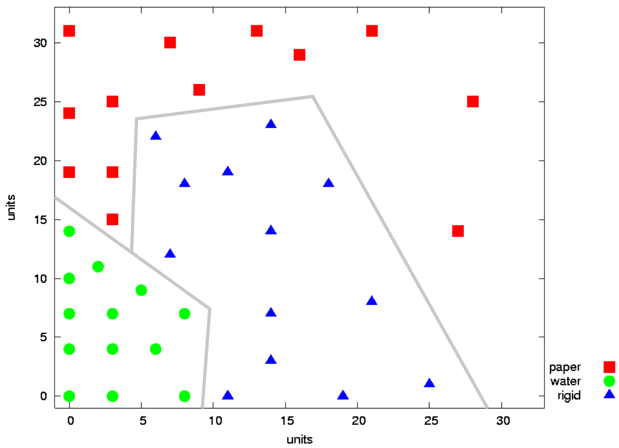

Figure 8: SOM analysis result with amplitude spectrum of auditory data.

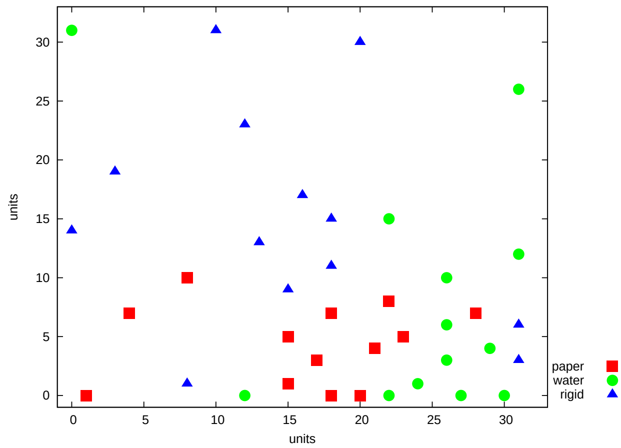

Figure 11: SOM analysis result with amplitude spectrum of angle data.

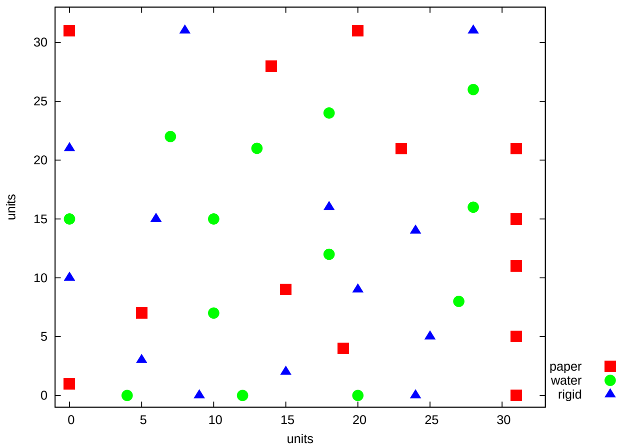

Figure 9: SOM analysis result with phase spectrum of auditory data.

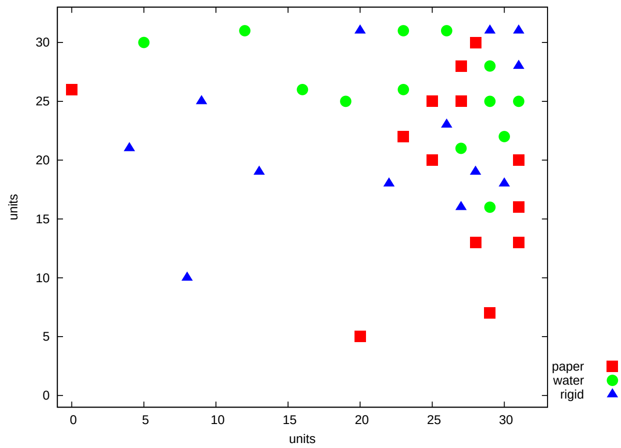

Figure 12: SOM analysis result with phase spectrum of angle data.