Home Robots, Learn by Themselves

Osamu Hasegawa and Daiki Kimura

Tokyo Institute of Technology, Japan
{hasegawa.o.aa,kimura.d.aa}@m.titech.ac.jp

Abstract. To build an intelligent robot, we must develop an autonomous mental development system that incrementally and speedily learns from humans, its environments, and electronic data. This paper presents an ultra-fast, multimodal, and online incremental transfer learning method using the STAR-SOINN. We conducted two experiments to evaluate our method. The results suggest that recognition accuracy is higher than the system that simply adds modalities. The proposed method can work very quickly (approximately 1.5 [s] to learn one object, and 30 [ms] for a single estimation). We implemented this method on an actual robot that could estimate attributes of "unknown" objects by transferring attribute information of known objects. We believe this method can become a base technology for future robots.

SOINN is an unsupervised online-learning method capable of incremental learning. By approximating the distribution of input data and the number of classes, a self-organized network is formed. SOINN offers the following advantages: network formation is not required to be predetermined beforehand, high robustness to noise, and reduced computational cost. In the near future, a SOINN device will accompany an individual from birth; this will allow the agent to share personal histories with its owner. In this occasion, a person's SOINN will know "everything" about its owner, lending assistance at any time and place throughout one's lifetime. Besides having a personal SOINN, an individual can install this self-enhanced agent into human-made products - making use of learned preferences to make the system more efficient. If deemed non-confidential, an individual's SOINN could also autonomously communicate another SOINN to share information.

Keywords: SOINN (Self-organizing Incremental Neural Network), Home robots, Machine learning.

1 Introduction

To compensate for a shortage of labor in the near future, it is vital to develop an intelligent robot that works for human beings in real living environments. To build such a robot, we require a system that can autonomously learn from real world interactions with humans, its environment, and electronic data. Therefore, we proposed an ultra-fast and online incremental transfer learning method [1]. This method uses the STAtistical Recognition on Self-Organizing and Incremental Neural Network (STAR-SOINN), which is an extension of the original SOINN [2]. We performed a

C. Stephanidis and M. Antona (Eds.): UAHCI/HCII 2013, Part III, LNCS 8011, pp. 48–53, 2013.

comparative experiment with previous offline [3] and online [4] methods using "Animals with attributes" dataset [3]. In this experiment, we attempted to estimate attributes of the unseen animal's image by transferring attribute information of learned animals. On the basis of experimental results, our method can maintain recognition rate equivalent to the online method. Table I shows the learning and test times and the features of these methods. The results show that our method is an ultra-fast transfer learning method, and can have potential practical applications.

Fig. 1. Framework of our research

The online incremental transfer learning method that we propose did not use or focus on certain modalities (e.g., sound, touch, and weight) because comprehensive combining of such modalities is difficult. For example, a "heavy" attribute can only be understood by a weight sensor. This implies the system cannot understand such an attribute from an image even if the system collects a huge amount of image. Therefore, the system requires a confidence value that represents the relationship between an attribute and a modality. In this paper, we propose multimodal transfer learning using the STAR-SOINN[1]. Figure 1 shows an overview of our research. This system can estimate attributes of an "unknown" object by transferring a robot's prior experimental learning of other objects. This implies that the robot can understand conceptual categories of any object in front of the robot, and can use this knowledge to act accordingly. This feature is required an intelligent robot that can work in real environments, because there is a wide variety of objects in such environments.

2 Proposed Method

Figure 2 shows a system overview of the proposed method. In the learning phase, the robot receives real data from its environment, interactions with humans, and electronic data, e.g. Internet. It then extracts features and trains each STAR-SOINN to remember the features for each attribute. After a short training, the system also uses the learning data to update the confidence value using attribute estimation. In the test phase, the robot receives real data from its environment, extracts features, and estimates attributes using the STAR-SOINN and the confidence value.

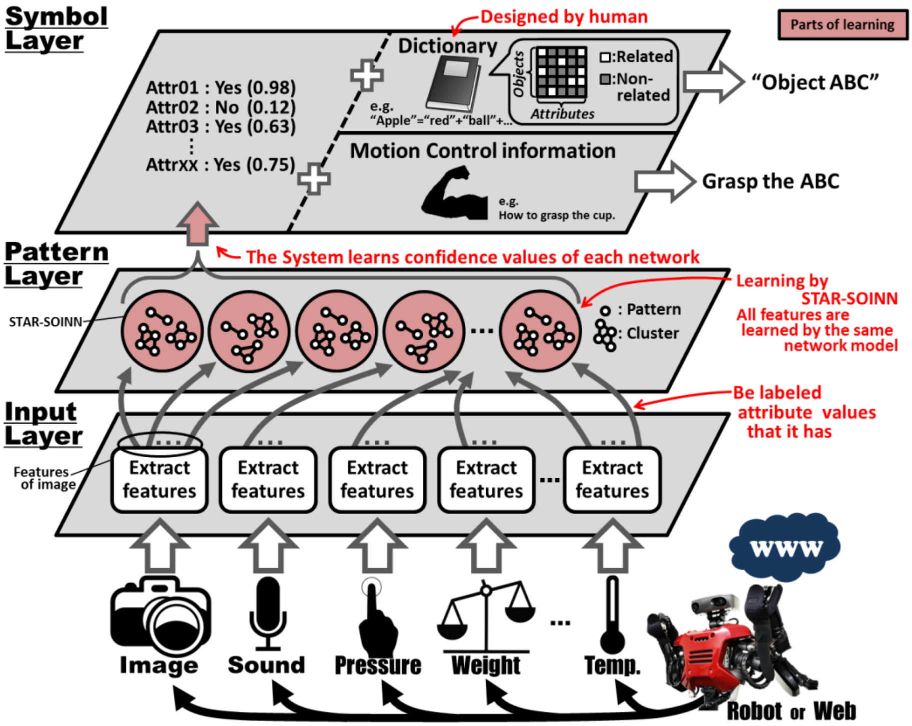

Fig. 2. Outline of our system

Table 1. Comparison of previous research

Method	**Proposed[1]**	[4]	[3]
Online learning	○	○	×
Learning time	**7 mins**	6 hours	70 days
Test time	**1.5 mins**	4 hours	2 days
Use real-valued attr.	○	×	×
Easy to update attr.	○	×	×

Fig. 3. A multi-modal robot used in our experiment

3 Experiment and Result

We conducted two experiments to evaluate the proposed method. One experiment was to verify the effects of the confidence value, and the other was to estimate the "unknown" object using a humanoid robot. In these experiments, we defined 23 attributes for each object used.

1. Verify the effects of the confidence value

We developed two systems; the proposed system, and a system that simply added modalities (i.e. it did not use the confidence value). We used these objects shown in Figure 3, and 3 modalities (image, sound, and depth). Table II shows the accuracy of these methods. The proposed method is more accurate than the method that simply added these modalities.

2. Estimate "unknown" objects using the robot

Figure 3 shows the robot, sensors, and learning objects used in this experiment. The robot learned these 20 objects from the 5 modalities using the proposed method. The system took approximately 1.5 [s] to learn one object. We then gave the robot four unknown objects, and checked the results of the attribute estimation for these objects. Figure 4 shows the result for the "mate tea cup"(=unknown object). The system correctly determined the attributes such as cylinder and wooden-made. Approximately 90% of estimations succeeded in determining the correct attributes. The system took approximately 30 [ms] for a single estimation, excluding action time.

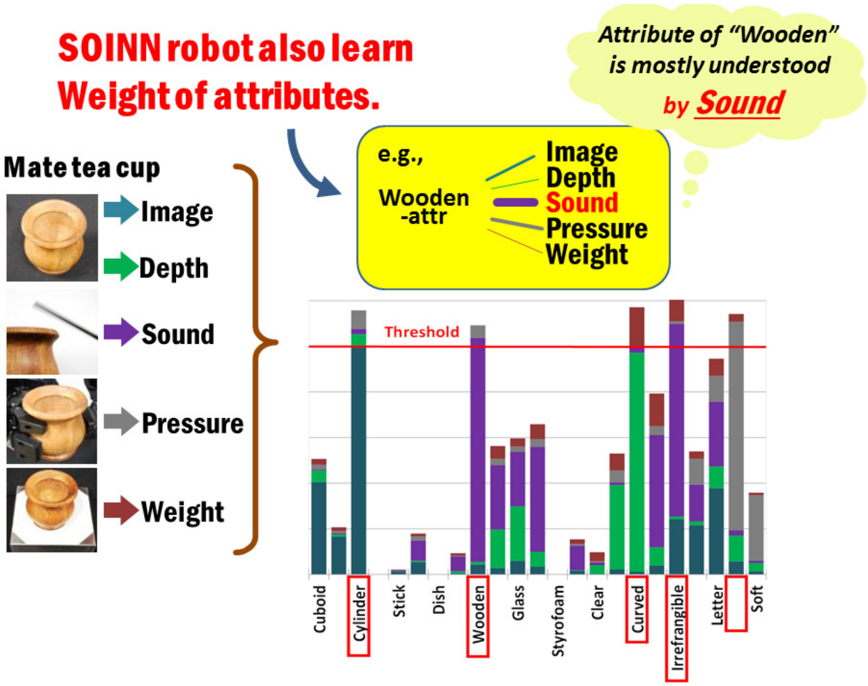

Fig. 4. Prediction results of "unknown" object

Table 2. Accuracy of these methods

Methods and modalities	Accuracy
Proposed method (Image, Sound, and Depth)	**94.3%**
Simply added modalities (Image, Sound, and Depth)	80.2%
Using only Image sensor	40.7%
Using only Sound sensor	62.3%
Using only Depth sensor	67.6%

4 Future Work

In the future, we think the Internet has a huge amount of information about many objects by such modalities using robots all over the world. Therefore, the robot will ask to other robots when it get information. We hope this system has the potential to become as the base technology for future robots.

Acknowledgment. This work was sponsored by the Japan Science and Technology Agency's CREST project.

References

1. Kimura, D., Pichai, K., Kawewong, A., Hasegawa, O.: Ultra-fast and online incremental transfer learning. In: The 17th Symposium on Sensing via Image Information, Kanagawa, Japan (June 2011) (in Japanese)
2. Shen, F., Hasegawa, O.: A fast nearest neighbor classifier based on self-organizing incremental neural network. Neural Netw. 21(10), 1537–1547 (2006)
3. Lampert, C.H., Nickisch, H., Harmeling, S.: Learning to detect unseen object classes by between-class attribute transfer. In: IEEE Conference on Computer Vision and Pattern Recognition (2009)
4. Kawewong, A., Hasegawa, O.: Fast and incremental attribute transferring and classifying system for detecting unseen object classes. In: Diamantaras, K., Duch, W., Iliadis, L.S. (eds.) ICANN 2010, Part III. LNCS, vol. 6354, pp. 563–568. Springer, Heidelberg (2010)
5. Araki, T., Nakamura, T., Nagai, T., Funakoshi, K., Nakano, M., Iwahashi, N.: Autonomous acquisition of multimodal information for online object concept formation by robots. In: 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 1540–1547 (September 2011)
6. Bosch, A., Zisserman, A., Munoz, X.: Representing shape with a spatial pyramid kernel. In: Proceedings of the 6th ACM International Conference on Image and Video Retrieval, CIVR 2007, pp. 401–408. ACM, New York (2007)
7. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 886–893 (June 2005)
8. Haasch, A., Hofemann, N., Fritsch, J., Sagerer, G.: A multi-modal object attention system for a mobile robot. In: Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems, Edmonton, Alberta, Canada, pp. 1499–1504. IEEE (August 2005)
9. Kimura, D., Pichai, K., Hasegawa, O.: Ultra-fast and online incremental transfer learning with the internet. In: The 18th Symposium on Sensing via Image Information, Kanagawa, Japan (June 2012) (in Japanese)
10. Kimura, D., Pichai, K., Hasegawa, O.: Unknown object recognition by images on the internet using ultra-fast and online incremental transfer learning. In: Meeting on Image Recognition and Understanding, Fukuoka, Japan (August 2012) (in Japanese)
11. Kumar, N., Berg, A.C., Belhumeur, P.N.: Shree K. Nayar. Attribute and simile classifiers for face verification. In: IEEE International Conference on Computer Vision (ICCV) (October 2009)
12. Nakamura, T., Nagai, T., Iwahashi, N.: Multimodal object categoriza-tion by a robot. In: Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 2415–2420 (2007)
13. Ozasa, Y., Iwahashi, N., Takiguchi, T., Ariki, Y., Nakano, M.: Integrated multimodal information for detection of unknown objects and unknown names. In: 2012 RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing (March 2012)
14. Pan, S.J., Yang, Q.: A survey on transfer learning. IEEE Trans. on Knowl. And Data Eng. 22(10), 1345–1359 (2010)
15. Rohrbach, M., Stark, M., Szarvas, G., Gurevych, I., Schiele, B.: What helps where – and why? semantic relatedness for knowledge transfer. In: CVPR (2010)