Team eR@sers[DSPL] 2019 Team Description Paper

Okada,H., Yokoyama,H., Contreras,L., Inamura,T., and Sugiura,K.

Tamagawa Univ., NII, The Univ. of Electro-Communications, NICT
https://sites.google.com/site/erasers2050/home/

Abstract. Team eR@sers has taken part in RoboCup@Home since 2008. The eR@sers achieved a first place at RoboCup 2008 ,2010 and second place RoboCup 2009, 2012, 2017 and its social robot HSR obtained the @Home Innovation Award in 2016. We are one of the oldest teams in the RoboCup tournament. We have improved the ability of robots with various techniques, which are going to be applied to other robot systems or social IT systems. We introduce them and our latest research briefly in this description paper.

1 Team Summary

Team eR@sers was formed around 2000 to participate in RoboCup 4 legged league. The eR@sers achieved a first place at RoboCup 2008 ,2010 and second place RoboCup 2009, 2012, 2017. The Japanese Robot Team eR@sers(erasers) is the result of a joint effort of four Japanese research groups.

We mainly focus on the adaptability to the environmental changes, and on the integration between the sensory-motor data and symbolic representation, utilizing only the neuro-dynamical model.

All developed functions could be packed in ROS modules.

Almost all training data would be real data and the system is performed and evaluated in the real environment.

2 Innovative technology and scientific contribution

2.1 Symbol emergence in robotics

Symbol emergence in robotics (SER) [1] is an active research area, which we started around 2010. In the idea behind the SER, each robot (agent) acquires concepts and symbols (language) through interaction with its surrounding physical world and other agents. Learning is considered as a bottom-up process, therefore unsupervised learning methods are involved in this approach, where Multi-modalities is another important aspect.

The framework of the SER can be realized using hierarchical Bayesian modeling and/or deep learning methods. The SER achieves true understanding of words meanings by robots, which means the symbol grounding problem can be solved. This also suggests that the general purpose service robot (GPSR) task can be completed in a real sense. Actual methodologies will be explained later.

GPSR [**4**] We developed a method through which domestic service robots can comprehend natural language instructions. The proposed method combines action-type classification, which is based on a support vector machine, and slot extraction, which is based on conditional random fields, both of which are required in order for a robot to execute an action.

Further, by considering the co-occurrence relationship between the action type and the slots along with the speech recognition score, the proposed method can avoid degradation of the robot's comprehension accuracy in noisy environments, where inaccurate speech recognition can be problematic.

2.2 Applicability of the approach in the real world

We have already implemented above mentioned methods and, currently, we are working on integrating everything as a whole system. It is very important to consider the applicability of the proposed approach in the real world. We have tested the idea of concept formation and language acquisition for about one month [5].

We are planning to test the proposal (SER + web-enable concepts) in the real home environment. Fortunately, we have several experimental homes, which can be used for the long-term test. We also think that participation at the RoboCup@Home is a good opportunity to evaluate the proposed framework.

2.3 RoboCup@Home Simulation

Not only RoboCup@Home but also RoboCup Soccer, Robocop Rescue, and other leagues tend to evaluate physical actions such as grasping, navigation, object tracking, and object/speech recognition because it is easy to evaluate them with objective sensor signals or ground truth data. However, evaluating the quality of human-robot interaction, such as the impressions of an individual user and whether a robot utterance is easy to understand, involve dealing with cognitive events, which are difficult to observe as objective sensor signals. One of the ultimate aims of RoboCup@Home is to realize intelligent personal robots that operate in daily life. While evaluating such social and cognitive functions is important, tasks for real competitions with time and space limitations cannot be designed for such evaluation. For example, using questionnaires is one conventional method for evaluating the social and cognitive functions of robots; however, this is difficult in real competitions because the number of samples that can be obtained is quite small.

Therefore, we propose a novel platform for competition design that can be used to evaluate the social and cognitive functions of intelligent robots through VR simulation. First, we have proposed a novel software platform that integrates ROS and Unity middleware to realize a seamless development environment for VR interaction between humans and robots[6]. Second, we have proposed two tasks as examples of task design for evaluating social and cognitive functions, which is 'Interactive Cleanup' and 'Human Navigation' tasks. They aim statistical evaluation of human-robot interaction in VR. Especially, the Human

Fig. 1. RoboCup@Home Simulation

Navigation task is proposed to observe and evaluate human behavior in terms of the quality of robot utterances.

2.4 Tracking and re-identification using self-supervised learning

These days, many techniques using CNN (convolutional neural network) such as YOLO (You Only Look Once), Mask R-CNN, and OpenPose are proposed, which enables to detect object or people on line. However, their training process requires long time and huge amount of labelled images. In some situations, a service robot may need to memorize a person, furniture, etc. that are new to it. Retraining such kind of classifiers cannot achieve this via human-robot interaction.

Our idea is, instead of using a detector trained for specific classes, to construct detectors, trackers, etc. based on an image feature extractor. Although SIFT and SURF are widely used for image processing, they are not sufficiently robust because they extract only local features. Instead, we use CNN with self-supervised learning[7] as an image feature extractor.

In training process, this model attempts to colorize video frames. Instead of directly select a color for a pixel, it embeds the target pixel into a feature space, selects a source pixel in a previous frame that is similar to the target pixel in terms of a metric in the feature space, and copy the color from the source pixel. By minimizing the error of color prediction, the embedding vectors corresponding to the same part in a video get close to one another in the feature space.

2.5 Learning non-parametric policies as random variable transformations [8]

Learning how to act under uncertainty is a central problem in the field of machine learning. Particularly, it is desired for robots to learn to generate continuous control signals. One of major approaches to attack this problem is policy gradient, which explicitly represents the policy and attempts to find the optimal one. Although this makes it straightforward to generate continuous action

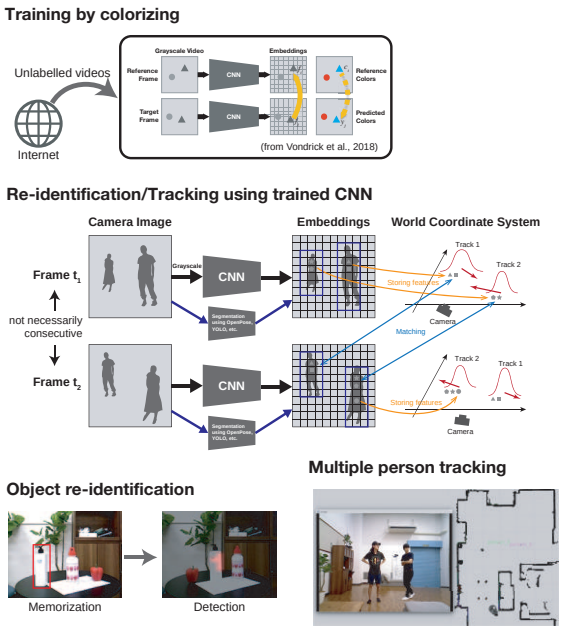

Fig. 2. Tracking/re-identification system using self-supervised learning

value, the policy is confined to a particular family of distributions such as normal distributions due to the difficulty of drawing action values from arbitrary probability distributions. To get rid of this limitation, we propose a method to learn non-parametric policy without any concern about generating action values.

Our approach, shown in Fig.3, is to find a transformation from a random variable $\boldsymbol{n}_t$ (following a normal distribution for simplicity), whose distribution is known, to one that follows desired distribution.

Consider a random variable transformation from $\boldsymbol{n}_t$ to the action value $\boldsymbol{a}_t$ depending on the state value $\boldsymbol{s}_t$, namely,

$$\boldsymbol{a}_t = f(\boldsymbol{s}_t, \boldsymbol{n}_t; \boldsymbol{\theta}), \tag{1}$$

where $\boldsymbol{\theta}$ is the parameter vector of the function approximator.

In this case, the form of the distribution of $\boldsymbol{a}_t$ can vary depending on $\boldsymbol{s}_t$ and $\boldsymbol{\theta}$. In particular, the distribution can be multimodal if the function approximator is sufficiently expressive. Once one find f such that $\boldsymbol{a}_t$ follows desired distribution, samples of the distribution can easily be drawn by calculating f.

In the field of reinforcement learning, the goal is to maximize the expected reward w.r.t. the stochastic policy. To deal with the random variable transformation, we used Dirac's delta to represent the action distribution:

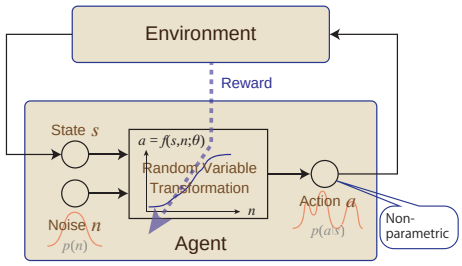

Fig. 3. A schematic illustration of our approach. Action value $\boldsymbol{a}$ is calculated from the state value $\boldsymbol{s}$ and noise $\boldsymbol{n}$. The function f of them is learned to acquire a desired distribution of $\boldsymbol{a}$. The optimization is performed with respect not to the value of $\boldsymbol{n}$ but to the distribution of $\boldsymbol{n}$.

$P(\boldsymbol{a}|\boldsymbol{s},\boldsymbol{n}) = \delta(\boldsymbol{a} - f(\boldsymbol{s},\boldsymbol{n};\boldsymbol{\theta}))$, which yields the expected reward to be maximized:

$$\mathrm{E}[r|\boldsymbol{\theta}] = \iint R(\boldsymbol{s},\boldsymbol{a})\, P(\boldsymbol{a}|\boldsymbol{s})\mathrm{d}\boldsymbol{a}\, P(\boldsymbol{s})\mathrm{d}\boldsymbol{s} \tag{2}$$

$$= \iiint R(\boldsymbol{s},\boldsymbol{a})\delta(\boldsymbol{a} - f(\boldsymbol{s},\boldsymbol{n};\boldsymbol{\theta}))P(\boldsymbol{n})\mathrm{d}\boldsymbol{n}\mathrm{d}\boldsymbol{a}P(\boldsymbol{s})\mathrm{d}\boldsymbol{s}, \tag{3}$$

where $R(\boldsymbol{s},\boldsymbol{a}) = \int rP(r|\boldsymbol{s},\boldsymbol{a})\mathrm{d}r$ with the reward r.

Using generalized stoke's theory, we derived the gradient of $\mathrm{E}[r|\boldsymbol{\theta}]$ w.r.t. elements of $\boldsymbol{\theta}$, which yields an update rule

$$\theta_i \leftarrow \theta_i - \alpha r \left(D_{\boldsymbol{v}_i} \log P_n(\boldsymbol{n}) + \operatorname{div} \boldsymbol{v}_i\right), \tag{4}$$

$$\boldsymbol{v}_i = \left(\frac{\partial}{\partial \boldsymbol{n}^\top} f(\boldsymbol{s},\boldsymbol{n};\boldsymbol{\theta})\right)^{-1} \frac{\partial}{\partial \theta_i} f(\boldsymbol{s},\boldsymbol{n};\boldsymbol{\theta}) \tag{5}$$

where α denotes the learning late.

Note that this update rule has a differential operator ($\operatorname{div}\boldsymbol{x} = \sum_j \frac{\partial x_j}{\partial n_j}$), and accordingly the second order derivative of f is to be calculated. For one-dimensional case, the update rule can be written as

$$\theta_i \leftarrow \theta_i - \alpha r \frac{((\log p(n))' f_n - f_{nn}) f_{\theta_i} + f_n f_{n\theta_i}}{f_n^2} \tag{6}$$

where f_ξ (resp. $f_{\xi\zeta}$) denotes partial derivative $\frac{\partial}{\partial \xi} f$ (resp. $\frac{\partial^2}{\partial \xi \partial \zeta} f$).

Although its implementation can be involved according to the complexity of the function f, the automatic differentiation technique (e.g. provided in TensorFlow) can be utilized to achieve simple codes.

We trained a fully-connected four-layer neural network, whose input and output are n and a respectively, and whose weight parameters are constrained

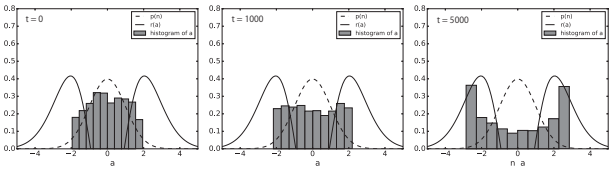

Fig. 4. A numerical experiment. The input noise n was drawn from $N(0, 1)$, which is indicated by dashed curves (the horizontal axes are shared by n and a). The histograms show the distribution of $a = f(n)$. The solid curves show the reward function $r = 0.7e^{-(a+1)^2/4} + 0.7e^{-(a-1)^2/4} - 1.5e^{-a^2/2}$.

to be non-negative. The frequency of a around zero decreased to avoid negative reward, whereas that around ± 2 increased pursuing the reward. This shows that a bimodal distribution can be acquired with our algorithm, and also suggests its potential to achieve a variety of distributions depending on the problem. We are currently applying our algorithm to neural networks with state input, to deal with control problems and to achieve social behaviors through learning with our robot.

2.6 Active robot-object interaction [9]

We use a multimodal system for active robot-object interaction using laser-based SLAM, RGBD images, and contact sensors. In the object manipulation task, the robot adjusts its initial pose with respect to obstacles and target objects through RGBD data so it can perform object grasping in different configuration spaces while avoiding collisions, and updates the information related to the last steps of the manipulation process using the contact sensors in its hand. We compare our approach with a number of baselines, namely a no-feedback method and visual-only and tactile-only feedback methods, where our proposed visual-and-tactile feedback method performs best.

We propose an active object manipulation systems using a 3-DOF RGBD camera (height, pan and tilt movements) on top of a service robot and a 6-axis force sensor in the hand. Through this sensors, the robot is able to detect the obstacle's position and orientation in robot coordinates while the different states of the manipulation process take place.

In particular, the robot arrives near the dishwasher within an uncertainty given by the localisation system based on 2D laser scans, but with a localisation error big enough to affect the performance in the dishwasher's door handle grasping step using only the arm's inverse kinematics. Therefore, we propose the use of the upper RGBD camera to update the robot's relative position to the dishwasher and to locate the handle, and then we use the contact sensor in the robot manipulator to detect when the robot reaches it.

3 Contribution for RoboCup@Home

Starting from 2006, RoboCup@Home has been the largest international annual competition for autonomous service robots as part of the RoboCup initiative.

However, it is observed that the development curve of the RoboCup@Home teams have a very steep start. The amount of technical knowledge and resources (both manpower and cost) required to start a new team has made the event exclusive to only established research organizations. For instance, in domestic RoboCup Japan Open challenge, the participating teams in RoboCup@Home were merely around 10 teams, which are about the same teams for the past few years. There were actually several new team requests however the development gap was huge for them to even complete the construction of the robots.

For this reason, RoboCup@Home Education initiative had been started at RoboCup Japan in 2015. RoboCup@Home Education is an educational initiative in RoboCup@Home that promotes educational efforts to boost RoboCup@Home participation and service robot development. Under this initiative, currently there are 3 projects started in Japan:

1. RoboCup@Home Education Challenge at RoboCup AsiaPacific2017 Bangkock.
2. RoboCup@Home Education Challenge at RoboCup Japan Open sice 2014.
3. Development of an educational Open Robot Platform for RoboCup@Home
4. We host RoboCup@Home Education Workshop Roma, Italy,March 15–16,2017 `https://sites.google.com/dis.uniroma1.it/athomeedu-rome2017/home`
5. Outreach programs (domestic workshops, international academic exchanges, etc.)

(For more information, visit `http://www.robocupathomeedu.org/`)

4 The contents of the web site

Official website:
`https://sites.google.com/site/erasers2050/home/`
Photos and Videos of the robot:
`https://sites.google.com/site/erasers2050/photos-movies/`

References

1. T.Taniguchi, T.Nagai, T.Nakamura, N.Iwahashi, T.Ogata, H.Asoh, “Symbol Emergence in Robotics: A Survey,” Advanced Robotics, 30(11-12), pp.706-728, 2016
2. M.Attamimi, Y.Ando, T.Nakamura, T.Nagai, D.Mochihashi, I.Kobayashi, H.Asoh, “Learning Word Meanings and Grammar for Verbalization of Daily Life Activities Using Multilayered Multimodal Latent Dirichlet Allocation and Bayesian Hidden Markov Models,” Advanced Robotics, 30(11-12), pp.806-824, 2016
3. T.Nakamura, T.Nagai, D.Mochihashi, I.Kobayashi, H.Asoh, M.Kaneko, “Segmenting Continuous Motions with Hidden Semi-markov Models and Gaussian Processes,” frontiers in Neurorobotics, 2017

4. T.Kobori, T.Nakamura, M.Nakano, T.Nagai, N.Iwahashi, K.Funakoshi, M.Kaneko, "Robust Comprehension of Natural Language Instructions by a Domestic Service Robot," Advanced Robotics, 30(24), pp.1530–1543, 2016
5. J.Nishihara, T.Nakamura, T.Nagai, "Online Algorithm for Robots to Learn Object Concepts and Language Model," IEEE Transactions on Cognitive and Developmental Systems, Volume: PP, Issue: 99, 10.1109/TCDS.2016.2552579, 2016
6. T. Inamura and Y. Mizuchi: Competition design to evaluate cognitive functions in human-robot interaction based on immersive VR, The RoboCup Symposium 2017.
7. C. Vondrick, A. Shrivastava, A. Fathi, S. Guadarrama, and K. Murphy: Tracking Emerges by Colorizing Videos, arXiv:1806.09594v2, 2018.
8. H. Yokoyama and H. Okada: Learning non-parametric policies as random variable transformations, The 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2016), 2016.
9. L. Contreras, H. Yokoyama, H. Okada: Multimodal feedback for active robot-object interaction, In the Workshop "Towards Robots that Exhibit Manipulation Intelligence", IEEE/RSJ IROS, 2018.