Happy Robot 2023 Team Description

Kosei Demura, Ryo Hidano, Yuuki Ishiyama, Masaaki Ida,

Takumi Kishimoto, Ren Kawasumi

Department of Robotics, Kanazawa Institute of Technology
7-1 Ohgigaoka Nonoichi, Ishikawa / 921 – 8501, Japan
demura@neptune.kanazawa-it.ac.jp
http://demura.net/happyrobot2023/

Abstract. This paper describes our research interests and technical information of the Happy Robot team for the RoboCup@Home league of RoboCup 2023. The Happy Robot team from the Kanazawa Institute of Technology has developed an autonomous domestic service robot for little child and elderly. This paper describes the basic architecture of the robot as well as present various algorithms and research contribution. Finally, this paper concludes and outlines our future research works.

1 Introduction

The Happy Robot team, our former team name was happy mini, has been participating in the RoboCup@Home league of the RoboCup Japan Open since 2012, and participated in the RoboCup world competition since 2015. Our team got the 9th in 2015, 8th in 2016, 9th in 2017, 5th in 2018, respectively. From 2020 to 2022, we could not participate in the world competition due to the COVID-19.

Our team is a joint team of the Yumekobo project and the Demura research laboratory in department of robotics at the Kanazawa Institute of Technology (KIT). The Yumekobo, the factory for dreams and ideas, is a unique educational system in KIT. In 1993, KIT established Yumekobo to encourage students to create things and make character building. The most important activity of Yumekobo is to support student projects called Yumekobo Projects whose purpose is to improve technology and teamwork.

The mission of the Happy Robot team is to create robots that make people happy. Our robots as shown in Fig.1 is designed in the image of a little child with a lively yellow color. We think that the design of home robots should be friendly to the elderly and children. Our robots are pioneer in such design concept in the RoboCup@Home league since 2015. The software system is based on ROS, and the state of arts frame works such as Detectron2, BlenderProc, OpenCV and so on.

The rest of paper is as follows. Section 2 describes the architecture of our robots, Section 3 describes the research contribution. Finally, this paper concludes by exploring applications and denoting future work.

2 Architecture

2.1 Design Concept

Our robots have two design concepts Kawaii and Safety.

The first concept of our robots is Kawaii. Kawaii is an adjective in Japanese. The meaning is cute, lovely, or charming. Our robot is used for childcare and conversation partner for elderly, and for persons feel lonely living alone. Kawaii robot can ease those people. Thus, the exterior design is crucial, and a lovely and friendly exterior design is suitable for those people.

The second concept of our robots is Safety. Safety is the most important aspect of service robots for children and elderly. Our robots are lightweight less than 10[kg] and have minimal power actuators for their arms.

2.2 Hardware

The platform of our robot is the Kobuki base that is a low-cost mobile research base designed for education and research. The torso is the extensible, and the commercial electric extendable cane, the KODUECHAN (ITK Co.Ltd), is used. It is only 0.38[kg], rated for up to 100 [kg], extends and retracts to 250 [mm]. The robot on the right side of Fig.1 has the 4DoF (shoulder 1, elbow 1, wrist 1, hand 1) arm. The maximum load of this arm is 0.5[kg]. The robot on the left side of Fig.1 has the 5DoF (shoulder 1, elbow 1, wrist 2, hand 1) arm. The maximum load of this arm is 0.78 [kg]. Each arm of two robots were fully developed by our team.

Fig.1. Happy Robots

The hand is also designed to have a capability of grasping an object, from ground up to 110 [mm] on the ground. Our robot has various sensors, such as the RGB-D sensor (Realsense, Intel), the LiDAR (UTM-30LX, Hokuyo), and the 360-degree camera (4KVR360, Kodak).

We have been developing a second-generation robot called Happy Bird. The details of the robot are described in Annex.

2.3 Software

Speech Recognition & Speech Localization

We have developed a speech recognition and a sound source localization system. The speech recognition system uses the Kaldi [1] gstream server. It is a real-time full-duplex speech recognition server and uses a DNN-based model for English trained on the TEDLIUM speech corpus. The sound source localization system is implemented using the robot auditory library HARK [2]. The system takes multi-channel speech waveform data from the 8ch microphone array, calculates FFT and estimates the sound source direction by the MUSIC method.

Object Recognition

Object recognition was performed using Mask R-CNN [3] trained with automatically generated training data. The system uses Blender, BlenderProc [4], OpenCV, and Detectron2. BlenderProc is a software that can generate training data using Blender and 3D models. Fig.2 shows examples of generated data. The generalization performance of Mask R-CNN is improved by data augmentation, i.e., adding noise, random rotating, grayscaling, darkening and brightening and color modification using OpenCV.

Fig.2. Examples of generated data

Fig.3. Recognition results

3 Research Contribution

3.1 The Fluorescent AR Marker

We have developed a 6DoF pose estimation network using a dataset that is automatically annotated with the Fluorescent Augmented Reality (AR) Marker [5] as shown in Fig.4. The marker is a transparent film coated with fluorescent paint that glows under ultraviolet (UV) light and is cut to match the texture of the object it is being used on. By switching between visible and UV light, the marker can automatically annotate RGB images, mask images for semantic segmentation, and provide 6DoF pose information. The Fluorescent AR Marker is transparent to visible light and preserves the texture of the target object, making it possible to automatically annotate 6DoF pose information that is difficult to obtain using other methods. The 6DoF pose estimation network uses the Fluorescent AR Marker to learn the 6DoF pose of the target object and can output the 6DoF information without the marker during inference. 6DoF pose information is crucial for manipulation tasks and has potential future applications.

The Fluorescent AR marker makes it easy to create a dataset containing 6DoF pose information for Deep Learning, which was previously difficult to obtain. This research makes a significant contribution to robotics and Deep Learning.

We also proposed a 6DoF pose estimation network that has a capability of learning the 6DoF pose of an object with a Fluorescent AR marker as shown in Fig.5. During inference, the network can output the 6DoF pose of the object without the marker. The average error between the inferred pose and the supervised data is 7.2 [mm], which is similar to that of a traditional AR marker. The ability to estimate 6DoF pose is critical for tasks such as robot manipulation, and future applications are anticipated.

In the future, we plan to apply our proposed method to transparent objects such as plastic bottles and glasses, as well as shiny metals, which are extremely difficult to recognize and estimate their 6DoF pose using existing methods.

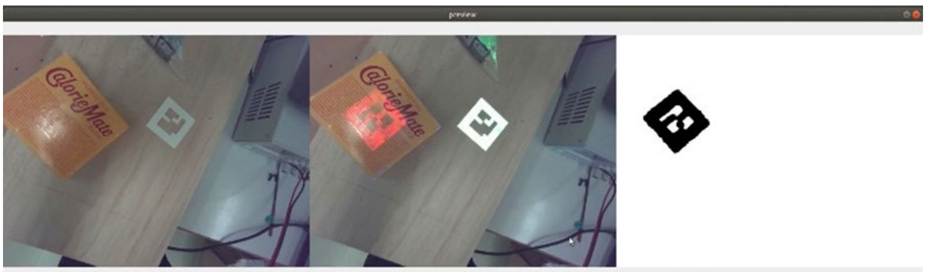

Fig.4. The Fluorescent AR Marker: The left image is not irradiated with UV light. The center image is irradiated with UV light. The right image is difference between left image and center image.

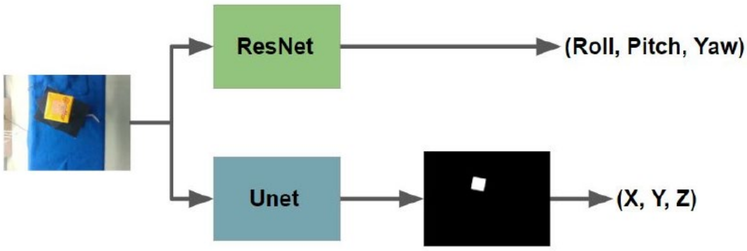

Fig.5. A proposed 6DoF network

3.2 The jamming gripper with an adsorption mechanism

Japan is facing a serious labor shortage due to the declining birthrate, aging population, and shortage of workers. This has led many companies to consider the introduction of robots. However, with the recent development of soft robotics, soft grippers that can grasp objects with complex shapes and soft objects such as food, which have been difficult with conventional grippers, are attracting attention in the industrial world. This research aims to develop a soft gripper that can grip a wide variety of objects.

To realize a gripper capable of gripping various objects, we have developed a gripper with a suction pad attached to the center of a jamming gripper to compensate for the respective problems without losing the advantages of a jamming gripper and a suction gripper as shown in Fig.6 [6]. In addition, to enable grasping of various objects, we made it possible to switch between the suction pad and the jamming gripper, and to select the grasping method between the suction pad and the jamming gripper according to the object to be grasped. Fig.7 and Fig.8 show the results of grasping experiments. The gripper can grasp various objects such as a small, middle, large screwdriver, a plastic bottle, a can, a mug cup and an aluminum plate.

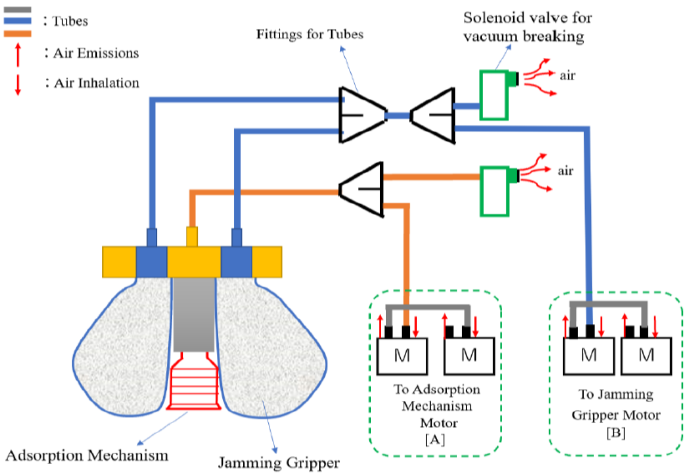

Fig.6. Diagram of a jamming gripper with an adsorption mechanism

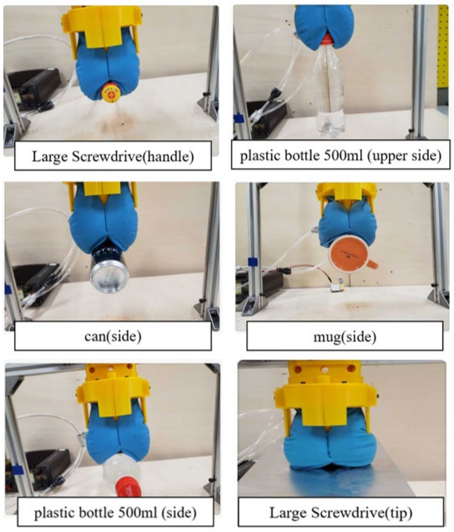

Fig.7. Grasping with an adsorption pad

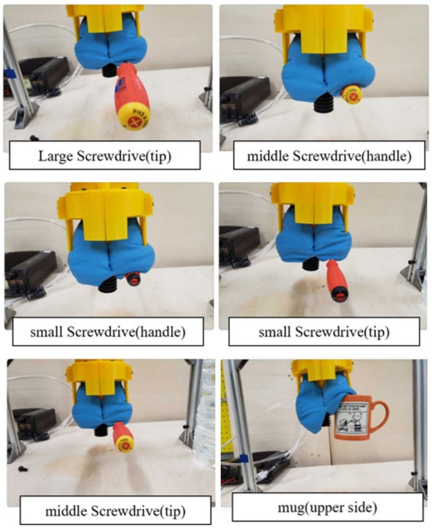

Fig.8. Grasping with a jamming gripper

4 Conclusions and future work

This paper has described the main features of the Happy Robot team that is designed with the goal of taking care of children, elderly, and for persons feel lonely living alone. The design concepts are kawaii, simplicity, safety, and usability. Thus, the robot is suitable for not only research, but also for education.

Our mission is to create robots that make people happy. We are using the RoboCup@Home challenge as a basis for the robot and working toward completing more important tasks for those persons in real-life situations. In the near future, we are planning to test the picture book reading application using Happy Mini in a real kindergarten.

References

1. Kaldi official Website. http://github.com/kaldi-asr/, (accessed December 11, 2022).
2. HARK official Website. http://www.hark.jp/, (accessed December 11, 2022).
3. K. He, G. Gkioxari, P. Dollár and R. Girshick, "Mask R-CNN," 2017 IEEE International Conference on Computer Vision (ICCV), pp. 2980-2988, 2017.
4. M. Denninger, M. Sundermeyer, D. Winkelbauer, Y. Zidan, et.al.: BlenderProc, arXiv preprint arXiv:1911.01911, 2019.
5. S. Okano, T. Makino, K. Demura: Fluorescent Texture: Proposal of a 2-3D Automatic Annotation Method for Deep Learning (in Japanese), Journal of the Robotics Society of Japan, vol.40, no.1, pp.71-82, 2022.
6. R. Matsubara, K. Demura: Development of a general-purpose gripper with selectable gripping method by jamming transfer and adhesion (in Japanese), Journal of the Robotics Society of Japan, 2023. (to be published)

Annex

Robot Description

1. The first generation: Happy Mini

Fig.9. Robot1

Hardware

- Base: Kobuki base, 0.7m/s max speed.
- Torso: The commercial electric extendable cane, the KOZUECHAN (ITK Co. Ltd).
- Head: Lovely face and sensors is mounted on the head.
- Computer: Thinkpad T450. CPU: Core i7 (Intel), GPU: 940m (NVidia)

2D LIDAR: UTM-30LX (Hokuyo Automatic)

- RGB-D sensor: RealSense (Intel)
- Omni camera: PIXPRO SP360 4K (Kodak)
- Gun microphone: CS-3e (Sanken)
- Audio interface: MobilePre (M-Audio)
- Microphone array: TAMAGO-03 (SYSTEM IN FRONTIER)
- Ultrasonic Distance Sensor: PING (Parallax Inc.)

Robot 1

- Robot dimensions: height: 9.3(min)-1.16m (max), width: 0.35m depth 0.38m
- Robot weight: 9.45kg.
- Arm: Mounted on torso. 4-DOFs, Maximum load: 0.5kg.

Robot 2

- Robot dimensions: height: 0.96(min)-1.24m (max), width: 0.35m depth 0.37m
- Robot weight: 8.15kg.
- Arm: Mounted on torso. 5-DOFs, Maximum load: 0.78kg.

Fig.10. Robot2

Software

- Platform: ROS Indigo
- Navigation: ROS Navigation stack
- Localization and mapping: AMCL and Gmapping
- Object recognition: YOLO v2 and CNN(LeNet)
- Gender classi_cation: Convolution DNN (LeNet)
- Face recognition: HaarCascade and DNN
- Speech recognition: Kaldi ASR
- Sound source localization: HARK
- Speech generation: MaryTTS
- Arm control: ROS MoveIt!.

2. The second generation: Happy Bird

Fig.11 shows the second generation robot called Happy Bird.

Hardware

- Robot dimensions: height: 1.2 [m]
  width: 0.55[m], depth: 0.6[m]
- Weight: 49 [kg]
- Base: Scout Mini Omni: Max Speed 2.7[m/s]
- Arm: xArm6: 6Dof payload 5[kg], weight 9.5[kg] reach 0.7[m]
- PC: Dell Alienware R17

Software

- Platform: ROS Melodic
- Navigation: ROS Navigation stack
- Localization and mapping: AMCL and Gmapping
- Object recognition: Detectron2 Mask R-CNN
- Speech recognition: VOSK offline speech recognition API
- Speech generation: MaryTTS
- Arm control: ROS MoveIt!.

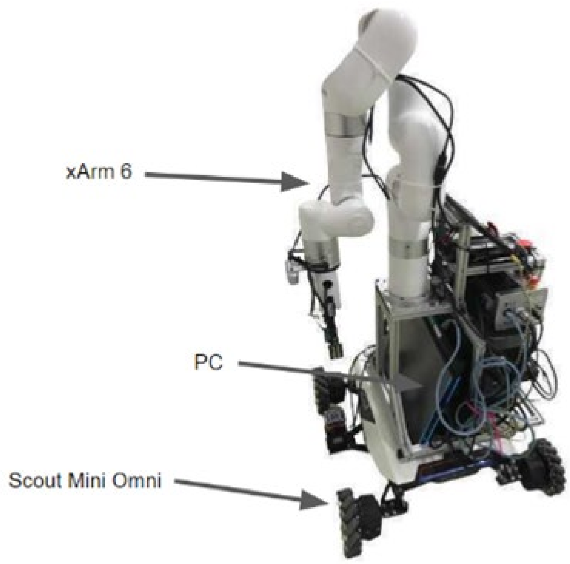

Fig.11. Happy Bird