UvA@Home Team Description paper 2018

Gerbscheid, J.; Groot, T.; Testroote, M.; Le Roy, N.; Terwijn, B.; Visser, A.

Publication date
2017

Document Version
Submitted manuscript

UvA@Home Team Description Paper 2018

Jonathan Gerbscheid, Thomas Groot, Merijn Testroote, Noah Le Roy,
Bas Terwijn, and Arnoud Visser

Intelligent Robotics Lab
Faculty of Science
University of Amsterdam
The Netherlands
http://www.uvahome.nl/

Abstract. This team description paper describes the approach that will be taken by the UvA@Home team to compete in Social Standard Platform League. The research challenges concern person recognition, object recognition, natural language processing and navigation. Modules implemented so far include people detection, face recognition, speech recognition and natural language processing. The remaining challenges will be solved using the previous research and achievements of the UvA teams in the RoboCup.

1 Introduction

The UvA@Home team consists currently of four bachelor Artificial Intelligence students, assisted by two senior university staff members. Yet, the most important team member is our Pepper robot, called Ginger. The team is part of the Intelligent Robotics Lab[1] (IRL) of the University of Amsterdam. The IRL functions as the governing body for all university robotics teams, which include the Dutch NAO Team (SPL) and the UvA@Home team (SSPL).

The focus of the university and the IRL lies in perception, world modeling, machine learning and decision making. Previous efforts done by other teams at the university also include efforts done for the simulation of the @Home competition [1, 2].

Last year the UvA@Home team competed in the RoboCup in Nagoya [3] and even though the final ranking was not very high, the team was satisfied with the demonstration of the conversational mode [4]. A lack of working localization and mapping modules were the primary reasons for not being able to continue to the final round. In preparation for RoboCup Montréal the team is planning to participate in the @Home competition at the German Open from April 25 to 29 in Magdenburg. Here the team wants to test the improved interaction modules and integration with a working localization and mapping component. The time after the German Open will be used to further train and test the neural networks for object recognition relevant to the challenges.

[1] http://www.intelligentroboticslab.nl/

2 Current Modules

For the previous RoboCup in Nagoya the UvA@Home team built a rudimentary Python framework on top of NAOqi. For the 2018 RoboCup in Montréal, the team decided to stick with the Python language due to the teams familiarity with it and due to the language being accessible to new team members. Additionally, the choice was made to rework the original framework using the qi framework instead of the NAOqi API. The qi framework is able use the Peppers processing power more efficiently than the NAOqi API for the reasons stated in the qi documentation[2].

2.1 Natural Language Processing

A rule based Natural Language Processing module was developed for 2017 Cocktail Party challenge [5] where the robot needed to be able to extract names and drinks embedded in commands. The extracting of drinks and names was done by parsing the lowest laying noun and recursively searching the tree of hypernyms for a specific word definition [6]. The hypernym tree is constructed using WordNet which is an English lexical database. In figure 1 the tree of hypernyms for the word *water* is illustrated. Because the hypernyms *Liquid* and *Fluid* occur in the tree it is likely that the word *water* is a drink. However, if *Binary Compound* and *Substance* were the only hypernyms in the tree the word would not be classified as a drink. A similar approach is used for extracting names.

This module made for the Cocktail Party challenge can be extended by adding different definitions that the system can look for. Adding rules upon rules is not a final solution for robots to work in an @Home environment but when challenges have specified objects and commands, it will be easy to add the new definitions to the database.

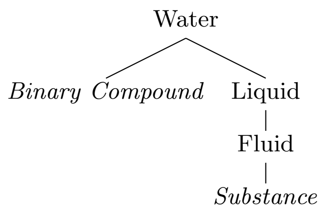

Fig. 1. Tree of hypernyms for the word: Water

2.2 Face Detection & Recognition

For the 2017 RoboCup, both the OpenCV Local Binary Patterns Histogram (LBPH) [7] and the OpenFace classifier [8] were used for face recognition. How-

[2] `http://doc.aldebaran.com/libqi/guide/py-tonaoqi2.html#qi-framework-main-features`

ever, neither were used during the competition itself as it was not needed for the challenges that the team participated in. Face detection, which was used for the speech and person recognition challenge, was done using the OpenCV Haar Feature-based Cascade Classifier. The LBPH method is still used in the current implementation due to its ease of use and low computational cost. A server-based implementation of the OpenFace Classifier can also be used but only if the LBPH method underperforms in the challenge environments.

2.3 Speech Recognition

The UvA@Home team currently relies on the Google Speech API for speech processing as it performed well in the competition, but will also implement a backup using NAOqi's built-in speech recognition in case of network connection problems. The Google Speech API works by converting an audio file containing speech to text. The main task here being able to know when to start and stop recording audio so that the Google Speech API can accurately determine the content. This is done listening to peaks in sound energy levels in the environment. Recording will start when registering a peak in sound energy levels and will stop when sound energy levels settle again. The sound energy levels are calculated using the built-in front microphone energy function in the NAOqi API [3]. A peak in sound energy levels is determined to be relevant when the difference of sound energy level passes some threshold. This threshold is increased by measuring the amount of times it is passed without any speech being properly detected and is decreased when not being passed for a predetermined amount of time.

3 Future Modules

Several modules are currently not implemented in the new framework but are planned to be implemented in the future. The team hopes to complete these modules for the 2018 competition.

3.1 Image Recognition

A full body people detection method was developed for the 2017 competition [9], however, integration with the full system was not completed for the competition itself, as the team decided to focus on other modules that were more important for the challenges. The detection works by combining data from Peppers 3D sensor with a CNN made in Tensorflow [10]. A color image dataset was made to test and train this method on. This dataset will be expanded with images of common objects from the @Home competition in order to train a general detection network, this network will be integrated into the new framework. This method relies on offloading the computations of the network to an external computer

[3] `http://doc.aldebaran.com/2-1/naoqi/audio/alaudiodevice-api.html#ALAudioDeviceProxy::getFrontMicEnergy`

with a dedicated GPU, making the system dependant on internet connectivity and thus more likely to have difficulties in the competition. Detection on the robot itself, however, is very limited due to its relatively low processing power. It was therefore decided that offloading detection was the most ideal solution.

3.2 Navigation & Localization

Navigation & localization are currently done using NAOqi's built-in navigation module. The performance of this module does not seem to be at the desired level required for the competition therefore the team will conduct additional tests on the navigation module in a variety of indoor environments to asses its effectiveness. If it is found to consistently underperform, the team will have to consider switching to the ROS framework in order to use the ROS SLAM gmapping implementation.

4 Contribution

The team plans to contribute to the competition by making all code public and documented[4]. Modules and classes in general are made to be straight forward and intuitive as to help other teams integrate the source code. Because the code is built upon the qi framework which is already present on all Peppers, installing the code on a new robot is relatively easy compared to ROS based frameworks, allowing inexperienced teams to quickly have a working basis for competition. The team also believes that because the competition is based around a standard platform, development for the robot should strive to make it as autonomous as possible. Where this is not possible external computing should only be used sparingly and the robot should still be able to function without it.

A module has also been developed allowing integration with the Slack[5] chatting program. Updating the robot with the latest code from the GitHub repository and running a specific program can be done through Slack to make interaction with the robot from a software perspective easier.

5 Conclusion

The UvA@Home team has learned a lot from the previous competition and is looking forward to demonstrating the progress made for the 2018 RoboCup. The team believes a better understanding of the challenges and the @Home league in general will allow it to improve upon previous performances and compete in more challenges. Yet, note that this team description paper is written on a moment that the rules for the 2018 competition are not finalized [11][6]. If new challenges are designed for the Social Standard Platform League, then it would

[4] `https://github.com/SpinazieSin/UvA-Home`
[5] `https://slack.com/`
[6] Rules Last Build Date: September 5, 2017

be preferable if they are announced well in advance to allow the teams to initiate the new research projects to solve these challenges.

In summary, our aim for the 2018 RoboCup is to offload image processing to neural networks on external computing devices and to improve existing modules such as speech recognition, face detection, natural language processing and navigation & localization.

Software List

Main software

- Operating System/Robot Control:
  NAOqi qi API `http://doc.aldebaran.com/libqi/guide/py-tonaoqi2.html#qi-framework-main-features`.
- Face recognition:
  LBPH (`https://docs.opencv.org/2.4/modules/contrib/doc/facerec/`),
  OpenFace (`https://cmusatyalab.github.io/openface/`),
  Cascade classifier (`https://docs.opencv.org/2.4/modules/objdetect/doc/cascade_classification.html`)
- Navigation:
  Naoqi (`http://doc.aldebaran.com/2-5/naoqi/motion/alnavigation.html`)
  ROS gmapping (`http://wiki.ros.org/gmapping`) .
- Natural Language Processing:
  Wordnet (`https://wordnet.princeton.edu/`),
  NLTK (`http://www.nltk.org/`)

Used Cloud service:

- Speech recognition: Google speech to text API `https://pypi.python.org/pypi/SpeechRecognition/`
- Object & People Detection: Tensorflow `https://www.tensorflow.org/`

Bibliography

[1] Sander van Noort and Arnoud Visser. *Extending Virtual Robots towards RoboCup Soccer Simulation and @Home*, pages 332–343. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.

[2] Victor I.C. Hofstede. The importance and purpose of simulation in robotics. Bachelor thesis, Universiteit van Amsterdam, June 2015.

[3] Jonathan Gerbscheid, Thomas Groot, and A. Visser. Uva@home team description paper 2017. RoboCup Social Standard Platform League, July 2017.

[4] Jonathan Gerbscheid, Thomas Groot, and Arnoud Visser. Intelligent news conversation with the pepper robot. In *Proceedings of the 29th Belgian-Netherlands Conference on Artificial Intelligence (BNAIC 2017)*, Groningen, The Netherlands, November 2017.

[5] Arnoud Visser. A new robocup@ home challenge. *Benelux A.I. Newsletter*, 31(1):3–6, 2017.

[6] Tirza F.E. Soute. Discovering available drinks through natural, robot-led, human-robot interaction between a waiter and a bartender. Bachelor thesis, Universiteit van Amsterdam, July 2017.

[7] Timo Ahonen, Abdenour Hadid, and Matti Pietikäinen. Face recognition with local binary patterns. In *Computer Vision - ECCV 2004: 8th European Conference on Computer Vision, Prague, Czech Republic, May 11-14, 2004. Proceedings, Part I*, pages 469–481, Berlin, Heidelberg, 2004. Springer Berlin Heidelberg.

[8] Brandon Amos, Bartosz Ludwiczuk, and Mahadev Satyanarayanan. Openface: A general-purpose face recognition library with mobile applications. Technical report, Technical report, CMU-CS-16-118, CMU School of Computer Science, 2016.

[9] Jonathan R. Gerbscheid. People detection on the pepper robot using convolutional neural networks and 3d blob detection. Bachelor thesis, Universiteit van Amsterdam, July 2017.

[10] Martín Abadi et al. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. Software available from tensorflow.org.

[11] Loy van Beek, Dirk Holz, Mauricio Matamoros, Caleb Rascon, and Sven Wachsmuth. Robocup@home 2018: Rules and regulations. `http://www.robocupathome.org/rules/2018_rulebook.pdf`, 2018.