

# UvA@Home Team Description paper 2017

Jonathan Gerbscheid, Thomas Groot, and Arnoud Visser

University of Amsterdam  
Faculty of Science  
The Netherlands  
<http://www.uvahome.nl/>

**Abstract.** This team description paper describes the approaches that will be taken by the UvA@Home team to compete in Standard Platform League with the Softbank Robotics Pepper. The research challenges concern person recognition, object recognition, natural language processing and navigation. Modules implemented so far include 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 of two bachelor Artificial Intelligence students supported by a senior university staff member. The team was founded as a part of the Intelligent Robotics Lab at the beginning of the 2016-2017 academic year. The IRL acts as a governing body for all the University of Amsterdam's robotics teams, including the Dutch NAO Team and the UvA@Home team (both active in a RoboCup Standard Platform League). It encourages the sharing of experience between these teams to be successful in both leagues, which is possible because the Nao and the Pepper robot share the same NaoQi basis (although a slightly different version).

## 2 Background

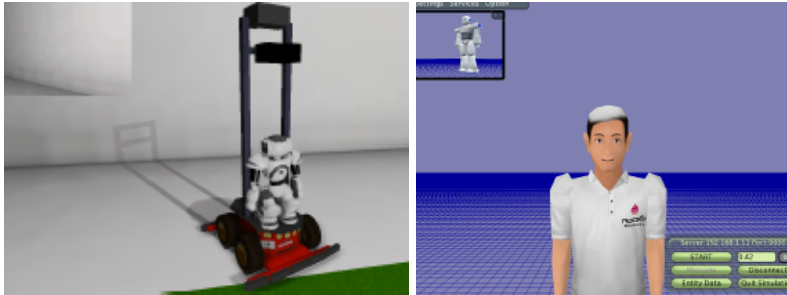
The Universiteit van Amsterdam has a very long history in RoboCup [1]. The university has been active in the Soccer Simulation League [2], the MidSize League [3], the Rescue League [4], the 4-Legged League [5], the Rescue Simulation League [6] and the Standard Platform League [7]. The teams have won several prizes<sup>1</sup>, both in the competition as with the technical challenges.

The focus of the research of the university is on perception, world modeling and decision making. The @Home competition nicely fits in our research; the lack of a standard platform withheld us from entering the competition. Instead, we have initiated studies towards the simulation of the @Home competition [8, 9].

When qualified, the Intelligent Robotics Lab has the intention to buy a Pepper robot under the conditions of Softbank Robotics <sup>2</sup>, if the investment budget

<sup>1</sup> <http://www.robocup.nl/achievements.html>

<sup>2</sup> <http://www.robocupathome.org/athome-spl/pepper>



**Fig. 1.** An @Home robot in the UsarSim [10] and an @Home scenario in the SigVerse [11] simulation environment [8, 9].

of 2017 is approved. Otherwise, the university has good contact with two Dutch companies in the possession of a Pepper robot.

### 3 Facial Recognition

The approach chosen for person recognition is a deep neural network implementation called Openface [12], this approach was chosen because of its state of the art performance and ease of use. This method detects faces and returns the most probable person and a confidence score for each face. The recognition model is trained using 10 images for each person, these images can be taken on the fly to allow for immediate retraining of the model after learning a new face. Using these confidence scores and a threshold also allows the classification of unknown faces. Using this approach an approximately 90% accuracy was achieved depending on lighting conditions. A lower accuracy of 70% was initially achieved for classifying unknown faces, using averaging over multiple pictures this was improved to 80%.

### 4 Speech Recognition

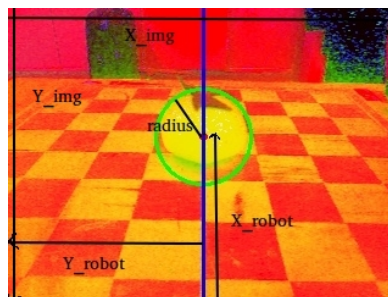
For speech recognition the Uva@Home team uses the google speech API <sup>3</sup>, which takes audio we record from the Pepper microphone and returns the recognized text. The disadvantage of this approach is that we can only use this part of our system as a black box. It would for example be preferable to limit the search space of possible sentences, yet this is not possible using this approach. However, we have found the google speech API to be much more accurate than other implementations and we believe this outweighs this inconvenience and we will therefore continue using it.

<sup>3</sup> <https://cloud.google.com/speech/>

## 5 Object Recognition

Recognizing and localizing object is difficult to execute on robots who generally have low-end CPUs. Some of the previous approaches done by the Intelligent Robotics Lab and its members/collaborators include:

- The ROS object detector which relies on the OpenCV2 library used for the UvA@Home league [13].
- A series of color, contour, size and blob detection based approaches used for the Roasted Tomato Challenge [14].



**Fig. 2.** Detecting of a tomato with a Nao robot using combined contour and color detection[14].

- A color invariant cognitive image pro-cessing module (CIP-module) based on the Recognition-By-Components (RBC) [15] used for ball and goal recognition in the Robocup SPL soccer competition.
- Optimizing the amount of perspectives that are required to correctly identify an object from the RoCKIn@Work competitions [16].
- In [17] categorization was accomplished by a Bag of Key-points approach, inspired by the method of Csurka et al. [18], to distinguish 10 different objects from the RoCKIn@Work competitions.
- In [19] categorization was accomplished using a decision tree approach based on shape and color, inspired by Alers et al. [20], to distinguish 10 different objects from the Ikea Duktig fruit and vegetable set.

The later two approaches make use of datasets recorded with an ASUS Xtion 3D sensor, which make their algorithms direct applicable to the Pepper robot.

## 6 Object Manipulation

In most cases the goal of object recognition in the setting of the @Home league is the manipulation of said objects. Earlier work on the UvA@Work League [13] and the Roasted Tomato Challenge [14] have required the manipulation of

objects. In both cases the MoveIt library [21] from the ROS framework was used. In this approach the MoveIt ROS Node receives a 3D world position from the object detection Node, the MoveIt Node contains a representation of the limbs and joints of the robot and uses those to solve the inverse kinematics. The joints are then moved to position the grabbing actuator to the object location.

## 7 Localisation

The Pepper robot has an extensive set of sensors, but only the HR cameras provide long range measurements, meaning that for localisation we have to rely mainly on visual SLAM (or adjust our navigation strategy).

Although for long range measurements the baseline between the two HR cameras is too short, much can be learned from the motion of the robot. To verify the applicability of visual SLAM, tests with the available ros-packages<sup>4</sup> will be made.

## 8 Navigation

The Intelligent Robotics Lab has extensive experience with the application of laser-based simultaneous localisation and mapping algorithms for robots in a natural environment<sup>5</sup>, yet the point-clouds of the 6 laser scanners with their limited range of 5 meters will force us to fall back to the coastal navigation algorithms developed for the Minerva museum tour-guide robot [22].

## 9 Open Challenge

The Uva@Home team has started the *Genuine Conversation* project that aims to enable the Pepper robot to engage in an opinionated conversation about current news topics, this project combines research from facial recognition, speech recognition and, natural language processing to construct what we call the *Conversation Engine*. Work on this project is progressing steadily and the results so far indicate that we will be able to demonstrate it during the 2017 Robocup. A flowchart describing the workings of the system can be seen in figure 3.

## 10 Language Processing

So far, a conversation engine using a rule-based approach with the *Stanford POS tagger* [23] has been implemented for queries about the news for the Open Challenge. The system uses the *Stanford POS tagger* to turn sentences into syntax trees and parse the lowest laying *noun phrase* (NP) in the tree. Studying

<sup>4</sup> [http://wiki.ros.org/vslam\\_system/Tutorials/RunningVslamOnStereoData](http://wiki.ros.org/vslam_system/Tutorials/RunningVslamOnStereoData)

<sup>5</sup> the Care-O-Bot of the European ACCOMPANY project, the Giraff robot of the European TERESA project, the MBot robot of the European MOnarCH project

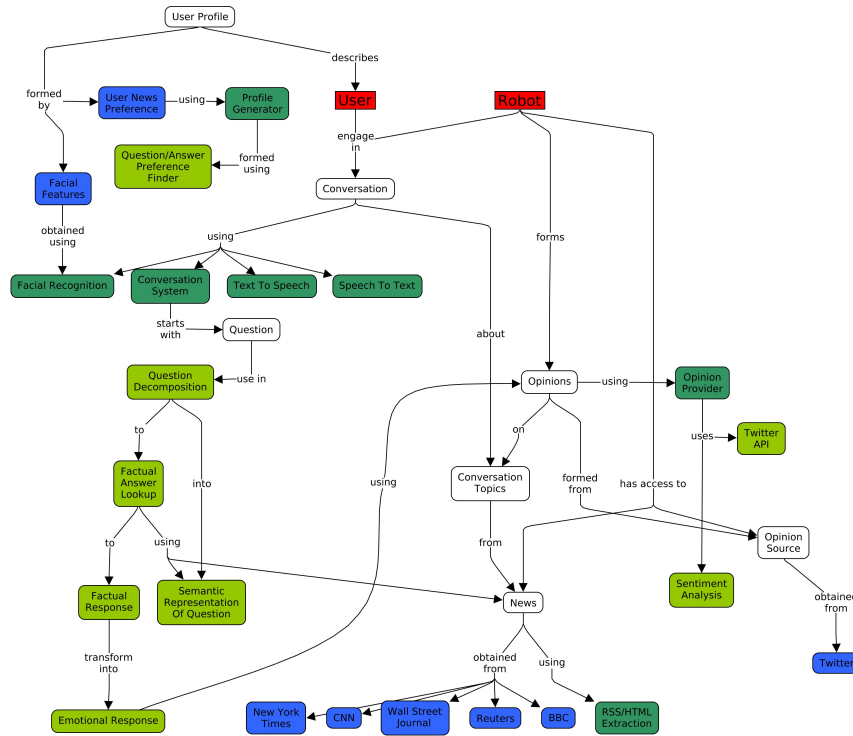


Fig. 3. News Conversation System

leafs of other NPs the system is able to derive meaning from questions given by the user. The conversation domain is generally limited, so only a few interpretations of sensible trees (that is, relevant for the conversation) are possible. While rule-based systems like this are limited to certain types of sentences, it works well within a closed domain conversational environment. The system can likely also extended to different uses perhaps by applying modern machine learning methods.

## 11 Conclusions and future work

We are looking forward to demonstrate our research for the Softbank Robotics Pepper robot and our progress on the challenges imposed by the RoboCup@Home competition. The current working modules have all been tested on the Nao robot and should also work on the Pepper robot as they share the same operating system. A large benefit of this league is that the achievements made are directly applicable to relevant scenarios in a social environment, something that can directly be communicated and disseminated to interested companies and the community.

## Bibliography

- [1] Emiel Corten and Erik Rondema. Team description of the windmill wanderers. In *Proceedings on the second Robocup Workshop*, pages 347–352, July 1998.
- [2] Jelle R. Kok and Nikos Vlassis. Uva trilearn 2005 team description. In *Proceedings CD RoboCup 2005*, Osaka, Japan, July 2005.
- [3] Matthijs Spaan, Marco Wiering, Robert Bartelds, Raymond Donkervoort, Pieter Jonker, and Frans Groen. Clockwork orange: The dutch robosoccer team. In *RoboCup 2001: Robot Soccer World Cup V*, pages 627–630, Berlin, Heidelberg, 2002. Springer Berlin Heidelberg.
- [4] A. Visser and S. Oomes. Uva flying rescue 2003. Participant Information and Team Description - RoboCup Rescue Robot League, July 2003.
- [5] Arnoud Visser, Paul Van Rossum, Joost Westra, Jürgen Sturm, Dave Van Soest, and Mark De Greef. Dutch aibo team at robocup 2006. *Proceedings CD RoboCup*, June 2006.
- [6] Raymond Sheh, Sören Schwertfeger, and Arnoud Visser. 16 years of robocup rescue. *KI-Künstliche Intelligenz*, 30(3-4):267–277, 2016.
- [7] Jonathan Gerbscheid Thomas Groot Sebastien Negrijn Patrick de Kok Caitlin Lagrand, Michiel van der Meer. Dutch nao team - technical report. Technical report, Universiteit van Amsterdam, FNWI, 2016.
- [8] Sander van Noort and Arnoud Visser. *Extending Virtual Robots towards RoboCup Soccer Simulation and @Home*, pages 332–343. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.
- [9] Victor I.C. Hofstede. The importance and purpose of simulation in robotics. Bachelor thesis, Universiteit van Amsterdam, June 2015.
- [10] Stefano Carpin, Mike Lewis, Jijun Wang, Stephen Balakirsky, and Chris Scrapper. Usarsim: a robot simulator for research and education. In *Proceedings 2007 IEEE International Conference on Robotics and Automation*, pages 1400–1405. IEEE, 2007.
- [11] Tetsunari Inamura, Tomohiro Shibata, Hideaki Sena, Takashi Hashimoto, Nobuyuki Kawai, Takahiro Miyashita, Yoshiki Sakurai, Masahiro Shimizu, Mihoko Otake, Koh Hosoda, et al. Simulator platform that enables social interaction simulationsigverse: Sociointelligence simulator. In *System Integration (SII), 2010 IEEE/SICE International Symposium on*, pages 212–217. IEEE, 2010.
- [12] Brandon Amos, Bartosz Ludwiczuk, and Mahadev Satyanarayanan. Open-face: A general-purpose face recognition library with mobile applications. Technical report, Technical report, CMU-CS-16-118, CMU School of Computer Science, 2016.
- [13] Valerie Scholten, Victor Milewski, Tessa Bouzidi, Celeste Kettler, and Arnoud Visser. Uva@ work, team description paper rockin camp 2015-peccioli, italy. *Team Description Paper, Intelligent Robotics Lab, Universiteit van Amsterdam, The Netherlands*, 2015.

- [14] Caitlin Lagrand, Michiel van der Meer, and Arnoud Visser. The roasted tomato challenge for a humanoid robot. In *Proceedings of the IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC)*, Bragana, Portugal, 2016.
- [15] Gabrielle E.H. Ras. Cognitive image processing for humanoid soccer in dynamic environments. Bachelor thesis, Universiteit Maastricht, 2014.
- [16] Sébastien Negrijn. Rockin@ work visual servoing - active vision using template matching on rgb-d sensor images. Bachelor thesis, Universiteit van Amsterdam, June 2015.
- [17] Areg Shahbazian. Taking up the rockin@work object recognition challenge with the bag of keypoints approach. Bachelor thesis, Universiteit van Amsterdam, June 2015.
- [18] Gabriella Csurka, Christopher Dance, Lixin Fan, Jutta Willamowski, and Cédric Bray. Visual categorization with bags of keypoints. In *Workshop on statistical learning in computer vision, ECCV*, volume 1, pages 1–2. Prague, 2004.
- [19] Robin Bakker. A comparison of decision trees for ingredient classification. Bachelor thesis, Universiteit van Amsterdam, June 2016.
- [20] Sjriek Alers, Daniel Claes, Joscha Fossel, Daniel Hennes, Karl Tuyls, and Gerhard Weiss. How to win robocup@ work? In *Robot Soccer World Cup*, pages 147–158. Springer, 2013.
- [21] Sachin Chitta, Ioan Sucan, and Steve Cousins. Ros topics. *IEEE robotics and automation magazine*, 2012.
- [22] Nicholas Roy, Wolfram Burgard, Dieter Fox, and Sebastian Thrun. Coastal navigation-mobile robot navigation with uncertainty in dynamic environments. In *Robotics and Automation, 1999. Proceedings. 1999 IEEE International Conference on*, volume 1, pages 35–40. IEEE, 1999.
- [23] Kristina Toutanova and Christopher D Manning. Enriching the knowledge sources used in a maximum entropy part-of-speech tagger. In *Proceedings of the 2000 Joint SIGDAT conference on Empirical methods in natural language processing and very large corpora: held in conjunction with the 38th Annual Meeting of the Association for Computational Linguistics-Volume 13*, pages 63–70. Association for Computational Linguistics, 2000.