# Birmingham Autonomous Robotic Club (BARC)
## Motto: *Learning by doing*

Members: Lenka Mudrova, Manolis Chiou, Marco Becerra,
Sean Bastable, Joshua Smith, Edoardo Bacci, Tanya Daskova and Anya Hristova
Email: robot-club@cs.bham.ac.uk
Affiliation: School of Computer Science, University of Birmingham, UK
Website: *http://lenka-robotics.eu/index.php/robotic-club*

***Abstract*—Birmingham Autonomous Robot Club (BARC) connects students from the University of Birmingham, with a strong interest in robotic applications and competitions. This paper is part of our qualification for the RoCKIn@Home 2015 competition. Therefore, it provides an overview of our robot Dora and the developed software structure based on ROS middleware.**

**It describes how this challenge relates to our interests and experiences and how we can achieve high reusability of our system by integrating different subsystems from other projects. Moreover, team members, their experiences and research interests are described in detail.**

**Finally, the conclusion summarises our motivation and relevance for this competition.**

## I. Introduction

Birmingham Autonomous Robotic Club (BARC) was established in 2011 in the School of Computer Science at the University of Birmingham. The main purpose was to provide an extra opportunity for students to gain additional knowledge about robotics and to work on real robotic platforms and projects. Several students involved in the team contributed to projects which were mainly used to promote robotics during the school's open days. For example a robotic waitress which accepted orders for drinks and delivered them.

In 2014, the structure of BARC was changed in order to incorporate the lessons learned and to allow the team to take part in robotics competitions. We participated successfully in Sick Days 2014 [1] and RoCKIn@Home 2014. In the latter we won two of the competition challenges. This year, we would like to participate again in RoCKIn@Home as we still have many things to learn, improve and most importantly contribute.

The team has the support of the Intelligent Robotics Lab [2] in the School of Computer Science. The lab conducts research and has expertise in a variety of fields including but not limited to computer vision, manipulation, planning, architectures, reasoning and mobile robots. Furthermore, the lab has strong links with the industry. We are also starting to cooperate with other departments within our university, such as electrical and mechanical engineering. We hope to create a more interdisciplinary team than we had last year.

This paper provides an overview of our system from last year and analyses the drawbacks of our solutions. Moreover, it extensively discusses what we are currently working on.

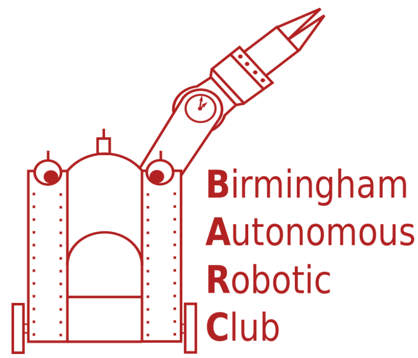


Fig. 1. Our logo represents our university campus with its famous clock tower Old Joe.

## II. Hardware

Our robot Dora is essentially a Pioneer 3D-X differential robot platform with an in-house central supporting construction attached to it. This construction serves as support for the various sensors and the laptop which is the brain of the robot (see Fig 2). In terms of sensing Dora is equipped with 2 laser rangefinders and 1 depth camera. These sensors are primarily used to ensure safe navigation and people detection. This year, we are going to significantly improve Dora's hardware itself as well as how the software handles this hardware. This will provide more support for:

- *Human-Robot-Interaction (HRI)* - Last year, speech understanding and reproduction were strongly limited due to the use of the laptop's build-in audio equipment. We are going to extend audio recording and reproduction capabilities by mounting a microphone and a speaker. Additionally, we are going to improve Dora's appearance to be more user-friendly which in turn will improve HRI.
- *Safe and robust navigation* - We are investigating usage of Pioneer's sonar to detect small objects very close to the ground level. Moreover, Pioneer's bumpers will be used to stop the robot when an object is hit.
- *Object recognition* - another depth camera with short range will be added to allow detection and recognition of small household objects.

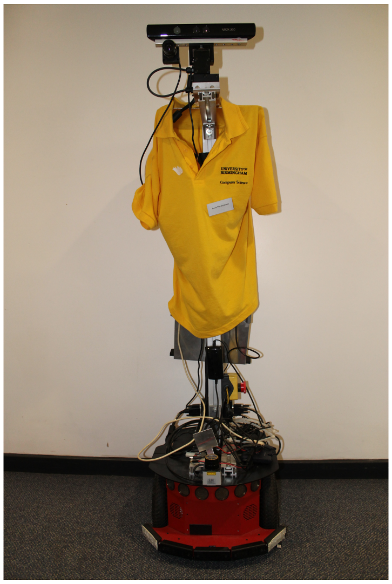

Fig. 2. Dora is an extended Pioneer 3D-X robot with sensors such as laser range finders, depth cameras and a laptop mounted on top.

## III. Software architecture

As our goal is to achieve high re-usability, modularity and openness, we built our software architecture on the Robot Operating System (ROS) middleware, using version Indigo [3] running on Ubuntu Trusty Tahr. Additionally, our system this year uses some of the publicly released packages from the STRANDS project [4]. This allows us to not only integrate the cutting edge research to our robot, but also to contribute to the evaluation of this research. Moreover, we integrate our own research where our interest overlaps with the focus of the competition. Our system architecture contains the following modules.

### A. State machine

Even though AI planning is in our expertise, we use *finite state machines* to control and monitor the robot's states and actions. We have two main reasons for this decision. First, the competition defines all *task benchmarks* as short scripts, so a robot does not have too much freedom in decisions on how to fulfil a task. Second, a state machine provides us with *repeatability* during testing.

For each of the benchmarks, we develop the state machine using ROS's SMACH package, that is "a task-level architecture for rapidly creating complex robot behaviour" [5]. All of the task's state machines are linked with the central state machine which communicates with the referee box and based on the accepted benchmarking test triggers specific state machines.

### B. Database for robot knowledge

The robot's knowledge is managed in a database, which allows us not only to record all data as rosbags, but also to extract and separate data offline in order to evaluate our system in different scenarios. We are using Robomongo [6].

### C. Navigation

We use standard ROS packages for localisation, mapping and low-level navigation (move-base), e.g. navigation planning path from initial coordinates to the goal's ones. These systems proved great for robustness in last year's competition and we are reusing them this year. However, we observed that at some cases, such as passing narrow doors and overcoming a doorstep, a special robot behaviour would be better. Therefore, we are extending our system using high-level topological navigation [7] and we can strongly benefit from three main features:

- Waypoints and edges are managed in a database which allows even online modification.
- A special robot behaviour can be specified on the edges overriding standard move-base.
- A navigation policy [8] provides paths on the top of the topological map containing waypoints. Thus, the robot must pass the surrounding area of the waypoint on the path. This can be mainly used to demand certain robot movements and keep the robot away from obstacles in an environment.

  Moreover, this year the system is extended by an extra monitoring level, which overrides the standard move-base behaviour when the robot needs to recover from being stuck, etc.

### D. Mapping

is done by *OpenSlam's GMapping* algorithm [9] through the ROS wrapper package called slam gmapping. This approach uses a Rao-Blackwellized particle filter in which each particle carries an individual map of the environment. The particles are updated by taking into account both odometry and the latest observations from a laser range finder.

### E. Localisation

in a known map uses an *Adaptive Monte Carlo Localization* (AMCL) [10] algorithm. This node is part of the ROS navigation stack package. It uses laser range finder readings to update a particle filter. Every particle represents a specific discrete state and stores the uncertainty of the robot being in that state/position within the known map. Also, every time the robot moves it uses the motor information to shift the particles to the corresponding direction.

### F. Low-level navigation using move-base

was used last year as the only navigation system. It is a proven robust solution for domestic environments [11]. More specifically this node reads the odometry, the current pose estimate and the laser range finder scans and 3-D points clouds from a depth camera and safely drives the robot in the environment to a predefined goal. In order to navigate smoothly, it uses a combination of a global and a local planner. The global planner creates an optimal global path based on the robot's pose and a global cost-map. Then the local planner, which uses the *Dynamic Window Approach* algorithm [12], is responsible for following the global path and reactively avoiding obstacles.

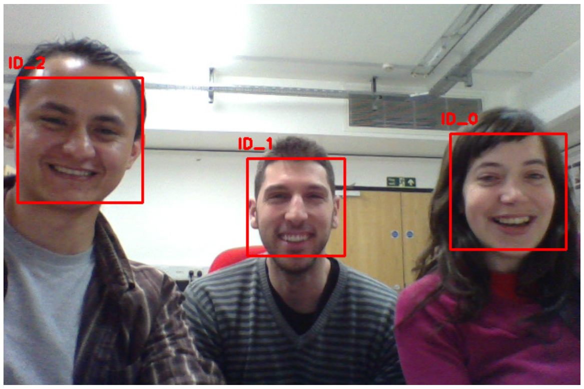


Fig. 3. An example of the face detection and face recognition algorithms. The red bounding boxes surround the successfully detected faces, while each of them is given a corresponding identification code.

### G. Human detection and recognition

Human detection can be split to indirect, when a robot must use only an RGB camera of the flat, and direct, when a robot can use any sensors.

*1) Indirect Face detection and recognition:* Face detection, see Fig. 3, is performed using the *Viola-Jones* algorithm [13]. The algorithm looks for faces by incrementally applying many simple Haar classifiers. The composition is performed by a cascade function, which needs to be trained *a priori* with many positive and negative images. The resulting classifier can find faces efficiently with independence of the size of the faces and light conditions.

Face recognition is performed by applying a *Local Binary Pattern Histogram* (LBPH) algorithm [14]. The principle of the algorithm is to build local binary patterns (LBP) for each pixel depending on a variable neighbourhood scheme. Then, it divides the formed LBP image into $m$ local regions and computes the LBP distribution histogram from each region. Finally, classification is performed by comparison between the LBP histograms of a new face with the ones from the dataset.

*2) Indirect uniform recognition:* Our uniform recognition was based on a simple segmentation of an input image to two regions. A positive region was corresponding to a colour within previously hand calibrated lower and upper bounds. In contrast, a negative region correlated to the colours outside of these limits. Based on the size of the positive region, the person was clarified as a Postman, Deliman or other. However, this detection is infeasible for robust performance due to high sensitivity on hand calibration and no knowledge about a detected area (for example any significantly big object in a background can create fall positive). Therefore, we are currently working on an extension using an upper body detector. The colour of the uniform will be then recognise only on this limited area, minimising the fall positives.

## IV. Our focus and plan

This section summarises our intentions for our system this year. We would like to aim for high re-usability of code in our robotic system and to demonstrate that the current state of the art in AI algorithms mixed with our extensions can be successfully used to produce a robust, effective and complete robotic system with domestic applications.

### A. Catering for Granny Annie's Comfort

A robot will receive commands from Granny Annie by natural speech. These commands may contain interaction with an intelligent flat or require the robot to bring her a specific object. We did not participate in this task last year because we did not have robust speech recognition as neither of us is an expert in the field. However, we have started a cooperation with a research group in this field and we hope to be able to recognise speech more robustly than using CMU Sphinx last year, see Sec. IV-F.

We would like to prepare our robot so that it can perform the first part - cooperation/interaction with an intelligent flat. This subtask still contains interesting challenges, such as robust speech recognition and robot navigation in a way which is comfortable and natural for a human.

### B. Welcoming visitors

In this task, a robot should welcome visitors, recognise them and accompany them to a specific location in Granny's flat. Several components are needed in this task: computer vision for face recognition and for uniform detection. Moreover, machine learning techniques are needed to be able to learn these two patterns in order to recognise them. A robot can also benefit from speech recognition if face/uniform recognition does not provide a certain decision. Additionally, a robot must have robust navigation in an environment in order to accompany a person to a specific place. Finally, human-robot interaction is needed while accompanying a person.

Last year, we successfully completed this task. Our system was built on several components which are mentioned in Sec. III. Here, we would like to discuss their properties, advantages and disadvantages and their future extensions. Our core state machine is built for this task calling several components. We do not plan changes in our face recognition system as neither of us is an expert in this field. However, we will create a larger database of faces in order to support a machine learning algorithm.

Uniform recognition needs to be improved due to the following reasons. First, recognition was based on manual sensitive calibration of colours last year, which slowed down preparation and the calibration was sometimes impossible. Second, uniform detection was based only on segmentation of all of the picture. This means, if there is a significantly large object of a certain colour belonging to a uniform (white/yellow) in the background, this object will be classified as the person. Both issues were limitations last year and we are improving them.

The last issue of our system last year was insufficient detection of a human. Thus, our human-robot interaction was strongly limited. We were using only a leg detector based on the readings from the laser scanner. This system was not robust enough. Hence we are testing how to combine this system with an upper-body detector based on a depth camera or an IR camera.

### *C. Getting to know my home*

In this task, a robot should recognise changes in a flat, such as open/closed doors, furniture which has been moved to another room or everyday objects which were moved. The robot can detect these changes autonomously or by cooperation with a human who uses speech or gestures. After detection, the robot is asked to use this new knowledge to execute some commands, such as to bring a mug on the dining table which is now in living room.

The state machine for this task from last year is in Fig. 4. Our system was built on two parts - autonomous detection of doors using the laser rangefinder and cooperation with a human in order to detect where the furniture is. However, the robot was unable to detect the exact position of the furniture which is important to be able to use the knowledge later. Therefore, we would like to integrate one of the STRANDS packages this year, which is able to detect dynamic clusters from 3D point clouds and recognise which object they belong to. Additionally, we would like to detect objects as well.

Finally, we used CMU Sphinx last year for speech recognition. However, this system is not proper for this usage due to its limited dictionary and usage. Therefore, we are going to change our speech recognition system, see Sec. IV-F.

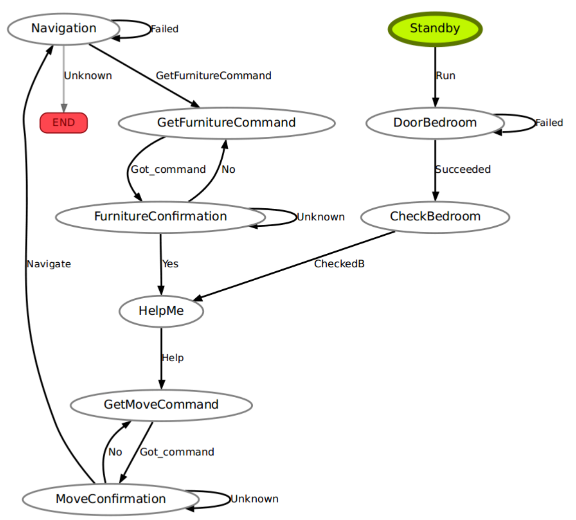


Fig. 4. The state machine for the task Getting to know my home

### *D. Object perception*

We are currently comparing two different approaches for object detection and recognition. The first approach uses manually created point-cloud models with rgb information. The advantage is that this algorithm detects the exact objects which were used for training. However, if the object is occluded then the detection is significantly worse. Additionally, this approach is possible for a small set of objects because each object needs a manually created model.

In contrast, the second approach uses a database of 3D models of objects. These objects do not represent any specific object, they are created in software for 3D visualisation to represent a general object. Hence, the second approach uses such a database to train its knowledge of what a general chair, mug, banana, etc. looks like.

We are additionally evaluating an approach which uses transfer learning to recognise objects using a convolutional neural network (CNN) which has been pre-trained for another task, in our case, image classification [15]. An RGB-D pair of images is first pre-processed as shown in Fig. 5. Next, the images are processed by the CNN in order to extract a vector of image features for each RGB-D frame. This feature vector is used with a tree of classifiers and regressors to determine first the category of the object, then the object instance, and finally the pose of the object.

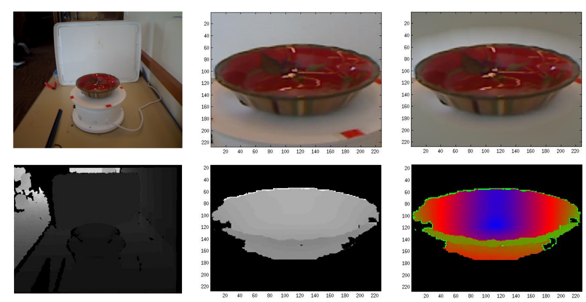

Fig. 5. The image pre-processing pipeline for the described algorithm. Top: The RGB image is first cropped down to size and then a fading algorithm is used to reduce the chance of unwanted features being detected in the background of the image. Bottom: The depth image is cropped in the same way and then coloured based on the depth data. This is required because the CNN is trained to detect features in RGB images and so we want the new image to describe the depth information in RGB.

### *E. Navigation*

We had robust configuration of navigation stack from ROS last year, which combined benefits from laser rangefinder and a depth camera. Additionally, we have added monitored navigation to recover from failure and topological navigation in order to control behaviour of the robot in different rooms. Details are explained in Sec. III.

### *F. Speech understanding*

Last year, we used CMU Sphinx for speech recognition with no added microphone. CMU Sphinx is not good enough for this benchmark. Therefore, we have started to cooperate with the speech recognition research group in our faculty in order to deploy a more sophisticated algorithm on our robot. Unfortunately, no details are known yet, as the work on this benchmark will start in September.

### *G. Timetable*

We are doing our best in order to participate in all tests. However, our priorities are the following. First, we will perform the complete *Welcoming visitors* task. Moreover, we will detect not only doors changes, but also precise locations of furniture in the *Getting to know my home* task. We will participate in the *Navigation* benchmark. We hope to participate in the rest of the tasks, but it depends on whether we will be able to recruit new team members.

## V. Evaluation and benchmarking

In order to develop a robust system, we have decided to focus on intensive testing and evaluation. Therefore, we tried to find possible ways to build a test arena and equip it with real objects. However, this is impossible for us due to a strongly limited budget and lack of space. However, we came up with a better idea - we have decided to use one of our flats, see Fig. 6. Hence, our robot is deployed permanently in this flat now and we record datasets during regular testing which we plan to publish later.

We have found such an environment really close to the arena with all of the tricky situations such as narrow corridors, door steps, a large mirror, etc. Therefore, we are able to test our algorithms in a realistic environment and improve their robustness. We created a video from our testing, which can be found on our website [16]. A more detailed description of our evaluation will be published there before the competition in November.

Fig. 6. Long-term deployment of Dora in a flat

## VI. Team members

Currently, the team has eight active members - three bachelor's, one master's and three PhD students. All team members are students in the School of Computer Science, University of Birmingham. Bachelor students are getting familiar with ROS and robotics. Hence, they do not have any specific background, research interests or position in the team yet. Therefore, the members description and details are only given for the master's student and the PhD students. The final team line-up is likely to change before the competition as more members will contribute.

### A. Lenka Mudrova

She is a PhD student with research interests in AI planning and scheduling. These aspects are important techniques for a robot to make decisions on when, how and what needs to be performed. Such decisions are necessary in service robotics when a robot needs to complete tasks assigned by a human. The robot needs to have a control framework that makes decisions concerning which particular task should be executed. The quality of the decisions influence the overall performance of the robot and of course, the robot's goal must satisfy as many of the requirements assigned by the humans as possible.

In the team, she has two roles. First, she is the team leader, which mainly includes representation of the team when formal communication outside of the team is necessary. Also, she makes sure that every member of the team knows what is happening, how their modules will be used within the system and what is required from them. The second role is that she is working also on the robot's subsystems. Her research interest in AI planning will be useful for creating the robot's overall behaviour. Moreover, she is working in computer vision and speech recognition. Lastly she contributes with her experience as she was involved as a team leader of the student robotics team FELaaCZech that took part in the international competition Eurobot for four years.

### B. Marco Antonio Becerra Pedraza

He is a PhD student in the School of Computer Science. His research interests are 3D perception, human sensing, knowledge representation and reasoning. More specifically his PhD research is about semantic mapping of human events. The topic has two main components: a) Semantic mapping, which can be conceived as an extension of the mapping problem [17]. The environment has additional spatial information that needs to be handled. Semantic maps extend the concept of maps to handle more features from the environment (e.g. structure, functionalities, events); b) Activity recognition is about using observations from the world to build representations of the ongoing actions. Later, these representations can be used to find associated structures.

Inside the team he has worked with the human sensing capabilities of the robot. He is also investigating how his PhD research can contribute to our current system. Lastly, he is a former member of the PUMAS RoboCup@Home team and as such, brings useful experience to the team.

### C. Manolis Chiou

He is a PhD student with a multidisciplinary background in control engineering, robotics and AI. His PhD work addresses the problem of variable autonomy in teleoperated mobile robots. Variable autonomy refers to the different levels of autonomous capabilities that are implemented on a robot. This can be potentially useful for robots used on demanding and safety critical tasks (e.g. search and rescue, hazardous environments inspection, bomb disposal), which are currently teleoperated and could soon start to benefit from autonomous capabilities. Robots could usefully use AI control algorithms to autonomously take control of certain functions when the human operator is suffering a high workload, high cognitive load, anxiety, or other distractions and stresses. In contrast, some circumstances may still necessitate direct human control of the robot. His research aims to tackle the problem by designing control algorithms for switching between the different autonomy levels in an optimal way.

He co-leads BARC and he shares responsibilities in administration and organization of the team with Lenka. His hands-on work on the system involves but is not limited to control, navigation, localization, data logging and Human-Robot-Interaction (HRI).

### D. Sean Bastable

He is a master's student in Robotics and is expected to follow up with a PhD in computer vision. He has worked on a number of student projects related to robot vision such as the task of visual localization of a mobile robot. This involves the use of a ceiling facing omnidirectional camera and PCA in order to determine an estimate of the robot's location relative to a set of training data. This approach is particularly useful in situations when other localization methods may be unavailable or unreliable. For example, laser based localization methods may produce incorrect location estimates in crowded environments where key features of the known environment may not be visible to the robot. Recently, he has been working on object recognition using convolutional neural networks (CNNs).

As the oldest BARC member, he has solid experience with ROS and various robotic platforms. His work on Dora includes face detection and recognition, object perception and overall system integration.

## VII. Conclusion

Our team has strong relevance to domestic service robotics, as Lenka's, Marco's and Manolis's research interests involved cooperation with humans. We would like to use our expertise to combine the state of the art in AI techniques with our own contributions. As a result, a robotic system will be created with high reusability, as we use ROS middleware.

We would like to achieve high robustness of our robotic system so that it can perform tasks repeatedly. In order to achieve this goal, we would like to put attention to verification and evaluation of our system. We expect that our robot Dora will be able to perform completely the *Welcoming visitors* task, *Getting to know my home* task and *Navigation* task. In order of fulfil these tasks, we plan to merge techniques for face and uniform detection, speech recognition and synthesis, navigation, people detection, HRI behaviour that will ensure that a person is following the robot, door detection, following a person and gesture recognition.

We would like to perform other tasks as well, but it depends on our human resources. We plan to cooperate more with first year bachelor's students in order to introduce them to robotics and provide them with some interesting challenges, which can be used during the competition. We believe that participating in the RoCKIn@Home challenge will bring a lot of experience not only to the young students, but also to us. As a result, we are expecting to obtain more knowledge in the ongoing research in many fields of robotics.

## References


[1] sick, *SICK robot day*, http://www.sick.com/group/de/home/pr/events/robot_day/seiten/robot_day_overview.aspx, accessed May 2015.

[2] IRLab, *Intelligent Robotics Lab (IRLab), School of Computer Science, University of Birmingham*, http://www.cs.bham.ac.uk/research/groupings/robotics/, accessed May 2014.

[3] M. Quigley, B. Gerkey, K. Conley, J. Faust, T. Foote, J. Leibs, E. Berger, R. Wheeler, and A. Ng, "Ros: an open-source robot operating system," in *IEEE International Conference on Robotics and Automation, workshop on open source software*, vol. 3, 2009.

[4] STRANDS, *STRANDS - Spatio-Temporal Represenations and Activities for Cognitive Control in Long-Term Scenarios*, http://strands-project.eu/, accessed May 2014.

[5] "SMACH - a task-level architecture for rapidly creating complex robot behaviour," http://wiki.ros.org/smach, accessed May 2014.

[6] robomongo, *Robomongo - Shell-centric cross-platform MongoDB management tool*, http://www.http://robomongo.org/, accessed May 2015.

[7] J. Pulido Fentanes, B. Lacerda, T. Krajník, N. Hawes, and M. Hanheide, "Now or later? predicting and maximising success of navigation actions from long-term experience," in *International Conference on Robotics and Automation (ICRA)*, 2015.

[8] B. Lacerda, D. Parker, and N. Hawes, "Optimal and dynamic planning for Markov decision processes with co-safe LTL specifications," in *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Chicago, Illinois, USA, September, 14–18 2014.

[9] G. Grisetti, C. Stachniss, and W. Burgard, "Improving grid-based slam with rao-blackwellized particle filters by adaptive proposals and selective resampling," in *IEEE International Conference on Robotics and Automation*, 2005, pp. 2432 – 2437.

[10] D. Fox, "Kld-sampling: Adaptive particle filters," in *Advances in Neural Information Processing Systems (NIPS)*, 2001, pp. 713–720.

[11] E. Marder-Eppstein, E. Berger, T. Foote, B. Gerkey, and K. Konolige, "The office marathon: Robust navigation in an indoor office environment," 2010, pp. 300–307.

[12] D. Fox, W. Burgard, and S. Thrun, "The dynamic window approach to collision avoidance," *Robotics Automation Magazine, IEEE*, vol. 4, no. 1, pp. 23–33, Mar 1997.

[13] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, vol. 1, 2001, pp. I–511–I–518 vol.1.

[14] T. Ahonen, A. Hadid, and M. Pietikäinen, "Face recognition with local binary patterns," in *Computer Vision - ECCV 2004*, ser. Lecture Notes in Computer Science, T. Pajdla and J. Matas, Eds. Springer Berlin Heidelberg, 2004, vol. 3021, pp. 469–481.

[15] M. Schwarz, H. Schulz, and S. Behnke, "Rgb-d object recognition and pose estimation based on pre-trained convolutional neural network features," 2015.

[16] web, *Our homepage*, http://lenka-robotics.eu/index.php/robotic-club, accessed May 2015.

[17] A. Nüchter and J. Hertzberg, "Towards semantic maps for mobile robots," *Robot. Auton. Syst.*, vol. 56, no. 11, pp. 915–926, 2008.