UChile Robotics Team
Team Description for RoboCup 2018

G. Azócar, N. Cruz, R. Pérez, F. Leiva, K. Lobos-Tsunekawa, I. Bugueño, N. Aguilar, and J. Ruiz-del-Solar

Advanced Mining Technology Center (AMTC),
Department of Electrical Engineering,
Universidad de Chile
Av. Tupper 2007, Santiago, Chile
uchilert@amtc.uchile.cl
http://uchilert.amtc.cl

Abstract. This Team Description Paper describes the organization, publications, and new developments of the UChile Robotics Team for the RoboCup Standard Platform League 2018 in Montreal, Canada.

Keywords: RoboCup, SPL, Standard Platform League, UChile Robotics Team, UChileRT.

1 Introduction

UChile Robotics Team (UChileRT) is a joint effort of the Advanced Mining Technology Center (AMTC) and the Department of Electrical Engineering of the Universidad de Chile in order to foster research in mobile robotics, computer vision and learning algorithms.

Our team was created in 2002 under the name of UChile1 and first participated in the RoboCup, joining the Four-Legged Standard Platform League in 2003. In 2007 we changed our name to UChile Kiltros, and in 2010 we collaborated with the SPQR Italian team. In 2012 our current team name, UChile Robotics Team, was adopted. We reached the fourth place in RoboCup 2014 (Brazil), RoboCup 2015 (China) and RoboCup 2016 (Germany). After performing a team members renewal, we were within the top twelve teams in RoboCup 2017 (Japan).

For the RoboCup 2018 we have made several changes, mainly regarding perception, which include a new ball perceptor, an improved player perceptor and a new robot orientation detector. In addition, several developments in robot behaviors have been made in order to enhance our game and satisfy the new rules, such as the free kicks.

This paper is organized as follows: In Section 2, we introduce past contributions from our team, then in section 3 we outline the developments and changes for the 2018 competition. Finally, in Section 4, we present the current research lines of our team.

2 Past Relevant Work and Scientific Publications

UChileRT has been involved in RoboCup competitions since 2003 in different leagues: Four-legged 2003-2007, @Home in 2007-2015 and 2017, Humanoid in 2007-2009, and Standard Platform League (SPL) in 2008-2017. UChileRT's team members have served RoboCup organization in many ways: Javier Ruiz-del-Solar was the organizing chair of the Four-Legged competition in 2007, TC member of the Four-Legged league in 2007, TC member of the @Home league in 2009, Exec Member of the @Home league between 2009 and 2015, and co-chair for the RoboCup 2010 Symposium.

Among the main scientific achievements of the group are the obtaining of five important RoboCup awards: RoboCup 2004 Engineering Challenge Award, RoboCup 2007 and 2008 @Home Innovation Award, Best Science Paper Award in RoboCup 2015 [5] and the Best Paper Award for Engineering Contribution in 2017 [7]. UChile's team members have published a total of 42 papers in RoboCup Symposia (see Table 1), 31 of them directly related with robotic soccer, in addition to many papers in international journals and conferences. Finally, this year 4 papers were submitted to the RoboCup symposium.

Table 1. Presented papers in the RoboCup Symposia by year, since 2003.

Articles	2003 - 2010	2011	2012	2013	2014	2015	2016	2017
Oral	14	-	-	1	1	1	-	2
Poster	8	2	1	2	1	4	2	3

3 Developments and Changes for RoboCup 2018

3.1 CNN Based Ball Perceptor

Being able to accurately detect the ball is critical to play properly in any SPL soccer match. In order to further improve the ball detection performance of our previous vision framework, we developed a new ball perceptor, which follows the paradigm of proposal generation and subsequent classification using convolutional neural networks (CNNs).

Our new perceptor uses a proposal generator inspired on the ball hypotheses provider created by HTWK team [10], however, our proposal generator does not use any color information since it processes only raw grayscale images.

To perform the ball detection, the proposals are fed to a cascade of two CNNs which classifies them as ball or non-ball. Both CNNs have architectures based on [7] but they receive graycale images as inputs. Since most of the computational cost of the networks is associated to the first convolutional layer computation, this translates in sharply reduced inference times. The first CNN performs boosting in order to both limit the proposals' number to a maximum of five, and sort them based on their confidence. The second CNN performs the binary classification task, meaning that it processes the filtered hypotheses to detect the ball. Both networks are extremely fast and accurate, having execution times of 0.043 ms and 0.343 ms, and accuracy rates of 0.965 and 0.984, respectively.

3.2 Efficient CNN Based Robot Detector

Robot detection is a critical capability of robotic soccer players as it enables the players to avoid obstacles and implement complex strategies. Traditional heuristic robot perceptors such as [16] achieves excellent results in controlled environments. However, the static color segmentation used by this kind of perceptors performs poorly under challenging light conditions, detecting players on shadows and spotlights.

Last year we developed a robot detector algorithm based on proposal generation and subsequent classification using CNNs [7]. This work won the 2017 Best Engineering Paper Award. Our improvements to this algorithm include a new contrast based robot proposal generator as well as a grayscale proposal CNN, which results in better performance and faster inference times.

3.3 Robot Orientation Determination

Inspired on [14], we propose an improved Vision-Based Orientation Detection for SPL League, which makes use of CNNs in order to achieve much better prediction accuracy than the original system.

The system relies on determining lines that fit the contour of the robots' feet which are then classified as aligned to the front, back or side of the NAO. This information is then used to determine the final orientation.

The orientation determination system uses the bounding boxes of the detected robots as inputs. Over these windows, the set of points that compose the robots' lower silhouette [14] is calculated. Then the subset of points that make up a closed convex region is obtained using Andrew's convex hulls algorithm [1]. For each pair of points of the convex set we calculate a line model in field coordinates. Each line model is then validated using heuristic algorithms. To classify the lines, a region that includes the robot's feet and legs is constructed around each line. For each of these regions a CNN that measures its quality is first applied. Regions with too much motion blur or regions that were incorrectly estimated are discarded. If a region for a line is accepted, it is then fed to a second CNN that classifies it as side, front or back.

Finally, the orientation determination is performed by combining the rotation given by the inverse tangent from two points belonging to the analyzed line with the direction of the line determined by its class. The resulting orientation is then added to a buffer that stores the last 11 measurements. Then, a circular median filter is applied over the buffer to obtain better results.

The deployed system presented high performance when both observing static and moving robots, with a correct estimation rate of 0.9988 and 0.9552, respectively. The average execution time of this module is $\sim$1.366 ms.

3.4 Interactive Machine Learning for Dribbling

One of the key features of the UChile Robotics Team strategy consists on having a good dribbling mechanism. Our approach is based on a Fuzzy Logic Controller with adaptive gains that was first proposed to be tuned using Reinforcement Learning [12]. Then, the tuning strategy was replaced by an Interactive Machine Learning algorithm called COACH [5]. Under this paradigm, while a robot dribbles, a human trainer is able to give corrective feedback to the actions taken by the dribbling mechanism, which modifies the values of the controller gains in real time [4].

In the recent years, the dribbling controller gains have been obtained using COACH with a simulated robot in the B-Human simulation engine. But, recent changes in the robots' environment and actuation (synthetic grass and rUNSWift walk2014 [9]) have distanced the physics of the simulator from the ones of the real world. In consequence, a dribbling mechanism trained in the simulator does not perform at its best in the real word. To tackle this shortcoming, an interface for fine-tuning the dribbling mechanism outside the simulator using COACH has been created. With this, it is possible for a human trainer to give feedback using a remote control while a robot dribbles in the real world, correcting the mismatches that are generated when first tuning the gains in the simulator.

Finally, given that now we have estimations of the opponent robots' rotation (see Section 3.3), this information is added to the dribbling direction computation. This endows our robots with the ability to dribble more strategically, taking into account the orientation of the opponents.

The direction computation [4] is inspired by the potential fields methods for robot navigation [8, 17] in which there are attractive and repulsive forces. These forces shape the trajectory that the ball should follow while the robot dribbles, having as target the opponents' team goal. The way in which the opponent robots' rotation estimation is added to the potential field is that it generates a new force for each opponent robot. Each one of these forces point perpendicularly to each ball-opponent vector. The sign of these vectors is chosen such that they point in the direction that increments the absolute value of the angle between the orientation of the opponents and the ball. The magnitude for each of these vectors is defined in the following equation:

$$M_F = \frac{A \sin(\theta_{\text{opponent-ball}})}{d^2_{\text{opponent-ball}}}$$

A is a constant, $\theta_{\text{opponent-ball}}$ is the angle between the opponent orientation and the ball and $d_{\text{opponent-ball}}$ is the distance between the opponent and the ball.

3.5 Visual Navegation using Deep Reinforcement Learning

We developed a map-less visual navigation system for NAO robots, which uses color images to derive motion commands using Deep Reinforcement Learning (DRL). The map-less visual navigation policy is trained using the DDPG algorithm [13]. In addition to convolutional and fully connected layers, LSTM layers are included to address the limited observability present in the problem. The use of DRL allowed to obtain a complex and high performant policy from scratch, without any prior knowledge of the domain, or the dynamics involved. The visual navigation policy is trained in SimRobot [15] and then successfully transferred to the physical robot where it is able to run in real-time with no external processing.

3.6 Active Ball Search Behavior

Our ball search behavior has been updated to improve its performance. We developed a new behavior to search the ball which follows active vision principles.

The method initially conducts a search in the regions where the ball is most likely to be. For example, in the position predicted by the ball model, or around the feet of the robot that had it before it was lost. If the initial search does not return any positive result, then a new search process is undertaken by using all the active players. Together, the robots maximize the field coverage by visually exploring it in a coordinated manner, giving priority to the areas that have not been seen in a while. In order to do this, the field is virtually divided in four areas which are assigned to each player (excluding the goalie). Each robot calculates its euclidean distance to the four field areas, and then a joint decision is performed in order to minimize the global sum of the distances between robots and areas.

3.7 Coordinated Defense Decision Making

In a cooperative game, the coordination of players allows the adoption of more flexible strategies that might be better for the team, even if each robot performs individually sub-optimal actions. In the past, each defense player's decision was independent from the others, depending only on the ball position and the side of the field they were assigned to. Therefore, based on [19], we developed an improved decision making process based on coordination graphs and Nash Equilibrium to create a coordinated defense.

In this process, each of the robots involved in defense have a set of rules that determine an action of the respective robot and may include actions by other players. Each of these rules has a value assigned to it, that may be either a reward or a penalty depending on the expected outcome of the set of actions. The reward function is hand-crafted and depends on the change in position, distance to the ball, relative position between players and to the goal. After the value of each rule is calculated, the rule sets of all involved robots are brought together and then combined, creating a new set of rules that is then ordered by their respective value. The rule with the best combined reward is then selected, determining the next action of each involved player. This way, a better defensive positioning involving more robots is achieved and actions that together would be detrimental are avoided.

4 Current Research Lines

4.1 Generative Models for Simulated Environments

This line of work is being followed by a Master's student. A methodology for implementing a visually realistic simulator for the SPL is being studied. The approach uses Deep Neural Networks to transform semantic maps generated by the simulator into realistic images. Images generated by this method have an almost non-existing reality gap. By using these realistic simulated images the algorithms of our vision system can be tuned for real world operation without the need to test them intensively on the robot. This is specially important for the CNNs used for object classification since a realistic simulated environment allows to train directly on the simulator, without the need of using the robot to obtain large databases. Overall, the approach will allow for faster and better development of vision algorithms.

4.2 Reinforcement Learning

This research line is being followed by a Master's student. The main objective is to develop systems to control robot behavior during games by learning policies using Reinforcement Learning.

Some challenges that naturally arise when trying to implement and deploy such systems are the problem modeling (e.g. reward shaping), the processing constraints that NAO robots inherently have, and the reality gap between the training environment (B-Human's simulator) and the test environment (a real SPL soccer field).

4.3 Interactive Machine Learning

This line of work is part of the Master's thesis of one of the team members. It is proposed to develop strategies for maximizing the information obtained from users who participate in the learning process of decision making systems. Interactive Machine Learning frameworks allow learning agents to be trained by human teachers who provide different kinds of information for supporting the learning process. There are paradigms like Learning from Demonstrations (LfD) [2, 3], in which the human feedback is in the actions domain, or approaches of Interactive Reinforcement Learning [11, 18], in which the human feedback is in the evaluative domain.

This line of research is specially focused on methods for learning tasks of continuous actions with corrective human feedback in the actions domain. The Ball Dribbling problem has been approached with these learning methods, and currently, the training of the associated controller is based on a method proposed in the work of a former teammate [5, 6, 4]. Currently, new developments are being made by combining this work with Deep Neural Networks.

Acknowledgments

This research is partially supported by Fondecyt project 1161500.

References

1. Andrew, A.: Another efficient algorithm for convex hulls in two dimensions. Information Processing Letters 9(5), 216–219. (1979)
2. Argall, B., Browning, B., Veloso, M.: Learning robot motion control with demonstration and advice-operators. In: n Proceedings IEEE/RSJ International Conference on Intelligent Robots and Systems (September 2008)
3. Argall, B.D., Chernova, S., Veloso, M., Browning, B.: A survey of robot learning from demonstration. Robotics and Autonomous Systems 57(5), 469 – 483 (2009), `http://www.sciencedirect.com/science/article/pii/S0921889008001772`
4. Celemin, C., Pérez, R., Ruiz-del Solar, J., Veloso, M.: Interactive Machine Learning Applied to Dribble a Ball in Soccer with Biped Robots. In: RoboCup 2017: Robot World Cup XXI Lecture Notes in Computer Science. Springer (2017)
5. Celemin, C., Ruiz-del Solar, J.: Interactive learning of continuous actions from corrective advices communicated by humans. In: RoboCup Symposium 2015. Hefei, China. (2015)
6. Celemin, C., Ruiz-del Solar, J.: Teaching agents with corrective human feedback for challenging problems. In: Computational Intelligence (LA-CCI), 2016 IEEE Latin American Conference on. pp. 1–6. IEEE (2016)
7. Cruz, N., Lobos-Tsunekawa, K., Ruiz-del Solar, J.: Using Convolutional Neural Networks in Robots with Limited Computational Resources : Detecting NAO Robots while Playing Soccer. In: RoboCup 2017: Robot World Cup XXI Lecture Notes in Computer Science. Springer (2017)
8. Damas, B.D., Lima, P.U., Custódio, L.M.: A modified potential fields method for robot navigation applied to dribbling in robotic soccer. In: Kaminka, G.A., Lima, P.U., Rojas, R. (eds.) RoboCup 2002: Robot Soccer World Cup VI. pp. 65–77. Springer Berlin Heidelberg, Berlin, Heidelberg (2003)
9. Hengst, B.: rUNSWift Walk2014 report. UNSW CSE RoboCup Report. (2014)
10. HTWK: Team research report, `http://www.htwk-robots.de/documents/TRR_2017.pdf?lang=en`
11. Knox, W.B., Stone, P.: TAMER: Training an Agent Manually via Evaluative Reinforcement. In: IEEE 7th International Conference on Development and Learning (August 2008)
12. Leottau, D.L., Celemin, C., Ruiz-del Solar, J.: Ball Dribbling for Humanoid Biped Robots: A Reinforcement Learning and Fuzzy Control Approach. In: Rob. 2014 Robot Soccer World Cup XVIII Preproceedings. Joao Pessoa, Brazil. (2014), `http://fei.edu.br/rcs/2014/RegularPapers/robocupsymposium2014\_submission\_58.pdf`
13. Lillicrap, T.P., Hunt, J.J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., Silver, D., Wierstra, D.: Continuous control with deep reinforcement learning. CoRR abs/1509.02971 (2015), `http://arxiv.org/abs/1509.02971`
14. Mühlenbrock, A., L.T.: Vision-based Orientation Detection of Humanoid Soccer Robots. In: RoboCup 2017: Robot World Cup XXI Lecture Notes in Computer Science. Springer (2017)

15. Röfer, T., Laue, T., Bülter, Y., Krause, D., Kuball, J., Mühlenbrock, A., Poppinga, B., Prinzler, M., Post, L., Roehrig, E., Schröder, R., Thielke, F.: B-Human Team Report and Code Release 2017 (2017), only available online: http://www.b-human.de/downloads/publications/2017/coderelease2017.pdf
16. Röfer, T., Laue, T., Kuball, J., Lübken, A., Maaß, F., Müller, J., Post, L., Richter-Klug, J., Schulz, P., Stolpmann, A., Stöwing, A., Thielke, F.: B-Human team report and code release 2016 (2016), only available online: http://www.b-human.de/downloads/publications/2016/coderelease2016.pdf
17. Tang, L., Liu, Y., Qiu, Y., Gu, G., Feng, X.: The strategy of dribbling based on artificial potential field. In: ICACTE 2010 - 2010 3rd International Conference on Advanced Computer Theory and Engineering, Proceedings. vol. 2 (2010)
18. Vien, N., Ertel, W., Chung, T.: Learning via human feedback in continuous state and action spaces. Applied Intelligence 39(2), 267–278 (2013), http://dx.doi.org/10.1007/s10489-012-0412-6
19. Wang, J., Wang, T., Wang, X., Meng, X.: Multi-robot decision making based on coordination graphs. In: 2009 International Conference on Mechatronics and Automation. pp. 2393–2398 (Aug 2009)