



Rhoban Football Club: RoboCup Humanoid KidSize 2019 Champion Team Paper

Loic Gondry, Ludovic Hofer, Patxi Laborde-Zubieta, Olivier Ly, Lucie Mathé,
Grégoire Passault, Antoine Pirrone, Antun Skuric

► To cite this version:

Loic Gondry, Ludovic Hofer, Patxi Laborde-Zubieta, Olivier Ly, Lucie Mathé, et al.. Rhoban Football Club: RoboCup Humanoid KidSize 2019 Champion Team Paper. 23rd Annual RoboCup International Symposium 2019, Jul 2019, Sydney, Australia. hal-02330530

HAL Id: hal-02330530

<https://hal.archives-ouvertes.fr/hal-02330530>

Submitted on 25 Oct 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Rhoban Football Club: RoboCup Humanoid KidSize 2019 Champion Team Paper

Loïc Gondry, Ludovic Hofer, Patxi Laborde-Zubieta, Olivier Ly, Lucie Mathé
Grégoire Passault, Antoine Pirrone, and Antun Skuric

Rhoban Football Club Team,
LaBRI, University of Bordeaux, France
`team@rhoban.com`
Corresponding author: Patxi Laborde-Zubieta
`patxi.laborde-zubieta@u-bordeaux.fr`

Abstract. In 2019, Rhoban Football Club reached the first place of the KidSize soccer competition for the fourth time and performed the first in-game throw-in in the history of the Humanoid league. Building on our existing code-base, we improved some specific functionalities, introduced new behaviors and experimented with original methods for labeling videos. This paper presents and reviews our latest changes to both software and hardware, highlighting the lessons learned during RoboCup.

1 Introduction

This article presents some of the elements which led to the fourth consecutive victory of our team in the RoboCup KidSize Humanoid league. We also obtained the first place in the drop-in tournament for the third time in a row and the first place at the technical challenges competition for the first time. Our robots scored 30 goals, received 11 goals and performed the first in-game throw-in in the history of the Humanoid league.

This year we mainly pursued three objectives: moving toward more dynamic gameplay, improving our performance at the technical challenge competition and reducing the complexity of our code base which has grown each year since we first participated in 2011. While introducing several new functionalities, we still managed to reduce the total number of lines of code used for the competition from 196,000 to 163,000. All the code and configuration file we used during the competition are available¹ along with some documentation.

The structure of the paper is as follows: section 2 introduces the tools used to model our robot, recent hardware changes are presented in section 3, new motions and improvements are detailed in section 4, our ongoing work regarding perception and data acquisition is described in section 5 and finally the dynamical aspect of our strategy is presented in section 6.

¹ https://www.github.com/Rhoban/workspace/releases/tag/public_2019

2 Model

The 3D model of the robots used during Robocup 2019 can be accessed² in the online CAD software *OnShape*³.

2.1 CAD to standard model (URDF/SDF)

Several CAD tools are commonly used to design robotics parts⁴, mostly based on constraints geometry design. On the other hand, standard robot description format emerged, notably URDF and SDF⁵, driven by the ROS community[8]. They are XML files describing the robot architecture, including transformation matrices, information about dynamics (mass, center of mass, inertia), collisions and visualisation geometry. In order to manufacture them, we used to design our robot with such CAD tools. But even if the model carried all the information needed for the standard robot description, the description was produced with separate tools. As a result, we could not ensure the consistency between the CAD and the description model. For that reason, we switched to *OnShape*, an emerging CAD software that includes an API allowing to request information about the 3D model. It allowed us to develop *onshape-to-robot*⁶, a tool that seamlessly produces an URDF from a CAD model that can be used without any change for all our applications.

2.2 Using model in online code

In order to compute frame transformations online using the robot, we developed a library on top of *RBDL*[3] to load the *URDF* and request it. *onshape-to-robot* allows you to attach manually frames in your model, directly in the CAD design-time, that appear in the final robot description and allow you to compute transformation matrices using the DOFs of your robot.

2.3 Physics simulation

The robot description model that is produced this way can also be used for physics simulation, like *Bullet*⁷. However, simulating the collisions of the actual parts is computationally expensive, first because the exported parts are represented by unstructured triangular surfaces, but also because of all the small details (like screw holes) which are not relevant for our use case.

To tackle this issue, we introduced a semi-automatic system that allows to approximate those complex 3D shapes with pure geometry: union of cubes, spheres

² <https://cad.onshape.com/documents/f3bdef32bffd81536fce83d1/v/779c691df8f135bba01eead1/e/a530b1889ee09acb5e1d7ff9>

³ <https://www.onshape.com/>

⁴ Famous examples are *Dassault Solidworks*, *Autodesk Inventor* and *Catia*

⁵ <http://sdformat.org/spec>

⁶ <https://github.com/rhoban/onshape-to-robot/>

⁷ <https://github.com/bulletphysics/bullet3>

and cylinders (Fig. 1). This also allows to ignore some small parts that are not useful in the collision world. We are then able to simulate a physics model of our robot (Fig. 2). Even if the discrepancy between simulation and real world is high, motions like walking, kicking and standing up can be reproduced, allowing to do some motor test before porting it to the real robot.

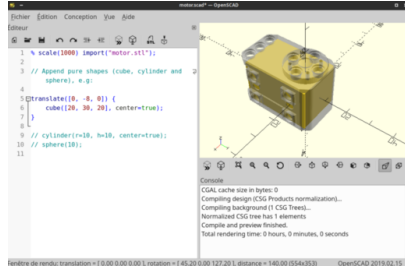


Fig. 1: Semi-automatic system to approximate 3D model with shape representation.



Fig. 2: Sigmaban approximated with 3D shapes in PyBullet physics simulator.

3 Hardware

We made only few changes on the hardware. We added some piano wire arcs to protect from falls, we switched to a four cells battery and we changed the shape of the feet and the hands. The last two improvements are described in section 4.

3.1 Protections

One of the challenges in our league is that robots should be able to withstand falls. For example, our robots can fall up to 20 times during a game. It regularly resulted in the breaking of a motor of the neck. After several attempts, it was clear to us that a software safety on its own was not enough. Hence, to absorb a part of the impact we added 3mm thick piano wire arcs at the front and at the back of the robot. By doing so, in 2018 and 2019 we had a significant decrease in the number of broken motors in the neck. But it is still not reliable enough as the many shoulder motors that we broke can tell.

To improve the protection of our robots, we would like to try other materials such as spring steel strips, which are less likely to bend perpendicularly to the impact, or dense foam. In order to assess rigorously the effectiveness of different solutions, the impacts when falling should be measured. For example, using motion capture or force platforms. Another way to protect the motors is to use clutches, but for the moment we did not find any solution fulfilling the requirements in term of weight and space used.

As mentioned during the second edition of the workshop “Humanoid Robot Falling: Fall Detection, Damage Prevention and Recovery Actions”⁸, this problem is a research topic of growing importance. And the robots of the KidSize league happens to be an interesting benchmarking environment for this subject: solutions can be tested under realistic conditions during matches and it is much easier to safely experiment with smaller humanoid robots.

3.2 Battery

In our design, the 6 motors of a leg are connected in series. We observed voltage drop up to 4 volts between the first and the last motor of the leg during dynamic motions. When using three cells batteries, this leads to a voltage around 8 Volts in the ankle. While the four cells batteries are out of the specified range for the dynamixel motors since they deliver 16V, they strongly increase the available torque and reduces the ohmic power loss. In the the past, we were using MX-64 motors in the legs and a lower position of the center of mass during the walk motion, this led to frequent overheating when using four cells batteries. Now that we have MX-106 motors and a smoother walking engine, we can safely use four cells batteries without risking overheating. Increasing the voltage was one of the key elements to obtain a more powerful kick. When the ball was properly positioned, our robot managed to kick more than seven meters.

4 Motions

4.1 New walk engine

We designed a new walk engine, still based on single support, with the goal of reducing the complexity of the former one that included too many unused parameters, making it less maintainable. We now use cubic splines to represent the trajectories of the feet (see Fig. 3). They allow us to control both position and speed at specific knots. Trajectories are updated only at each support foot swap and described in the trunk frame.

We wanted the foot to reach its nominal speed before touching the ground and to decelerate after leaving the ground. This way, the foot in contact with the ground would have a constant speed which should lead to a steady and continuous speed of the trunk. However, the robot is performing much better when the foot touches and leaves the ground with vanishing speed. One hypothesis is that it is better to touch the ground with no speed for stability reasons. Moreover, the exact time when the foot touches the ground can vary from one step to one another.

The walk engine code can be found on our open source repository under `Motion/engines/` directory.

⁸ <https://iros2018wsfallingrobots.wordpress.com/>

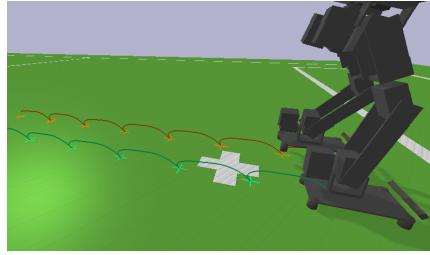


Fig. 3: Trajectory of the feet during walk engine (physics simulation).

4.2 Throw-in

This year, the throw-in rule has been introduced. While it is currently allowed for robots to perform it by kicking, we decided to move directly to human-like throw-in performed with the hands. The motion for the throw-in is created the same way as other kicks: we define splines associating time points with angular targets for the motors. The target between specified time points is obtained through linear interpolation.

We designed the throw-in motion using splines with several keypoints, see Fig. 4: initialization, bending the knees, unfolding the arms, leaning forward, holding the ball, straightening, lifting the arms and unfolding the knees, moving the arms behind the head, throwing the ball and back at initialization.

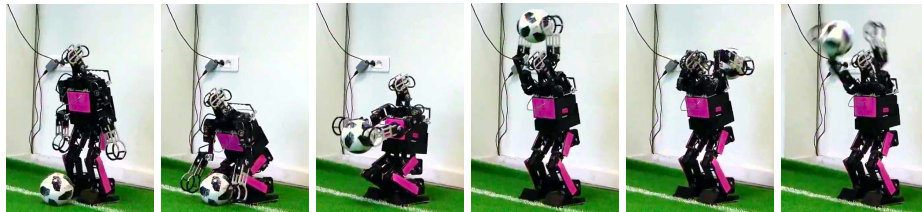


Fig. 4: Different steps of the Throw-in.

While simulation using *PyBullet* was not accurate, it allowed to design a coarse approximation of the motion before fine tuning the targets on the robot.

One of the main challenge to design the throw-in was to handle the fact that our arms have only three degrees of freedom. While the mechanical design is entirely sufficient for standing-up, it does not allow to control the distance between the hands when the elbow are bent. Therefore, amplitude of the motion on the elbow was limited in order to maintain grasp on the ball.

In order to lift the ball more easily, we designed new hands for the robot, see Fig. 5. The main idea was to create two metal plates in each side of the elbow motor and join them with spacer screws. On one side, the plate has a notch for the motor and, on the other side, there is large hole to surround the ball.

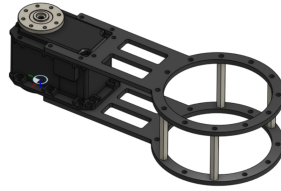


Fig. 5: Assembly of one hand of the robot in *OnShape*.

While designing the motion, we managed to have the ball bouncing and rolling up to 4.5 meters. Since the field is only 6 meters wide, we had to slow down the speed during the throw-in phase in order to have a more appropriate length for passes.

Next year, we plan to improve the throw-in so that the robot can adapt the direction by changing the orientation of the torso while the ball is in the air.

4.3 High Kick

One of the key aspect to lift the ball while kicking is the point in contact with the ball at impact. We refer to this part by the name of *kicker*. Previously, we used a *kicker* designed to give a rotational impulsion to ball so that it could roll on the grass. This year, in order to perform in the High Kick technical challenge, we designed new *kickers*.

To lift the ball, the contact point at impact has to be below the center of mass of the ball. Since we use high studs to stabilize on artificial grass and the ball has a radius of only 7.5 centimeters, the margin for the motion is relatively small. We decided to separate the new *kicker* in two different parts: the first one is thin and raises the ball which rolls over it, the second one hits the ball after it has left the ground making it easier to hit below the center of gravity.

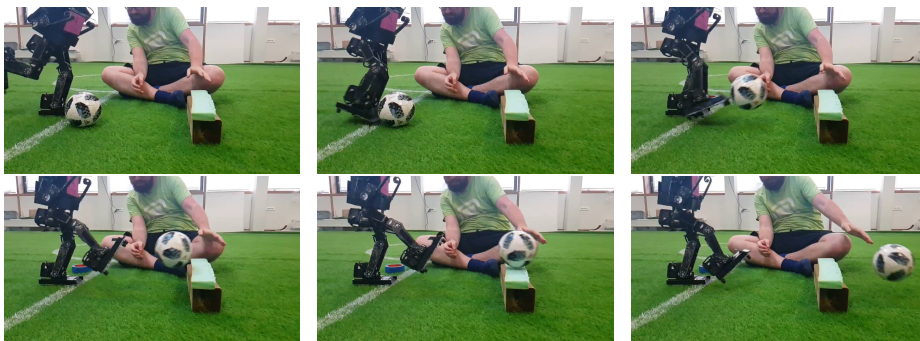


Fig. 6: Different steps of the High kick.

This new kicker allowed us to outperform the other teams during the technical challenge, our robot scored over a bar of 20cm, while the second best performance in our league was achieved by the CIT Brains who kicked the ball above 12cm. Fig. 6 shows the steps of this motion. An interesting fact about this high kick is that one of our robot did kick above another fallen robot during the quarter finals. Although not intentional, this kick is definitely a step toward the use of a third dimension in the RoboCup Humanoid league.

5 Perception and Localisation

This year, we aimed high regarding the modifications of our perception and localisation modules. Unfortunately, our system for labeling videos still lacked some robustness and the training procedure for our neural networks included a few bugs which impacted negatively its results during the competition. This sections presents our promising development of these modules along with preliminary conclusions based on their use during RoboCup 2019.

5.1 Labeling videos

Most of the teams in the RoboCup Humanoid league now uses neural network to detect or classify features in their perception module. In order for the module to work on-site, it is generally required to manually label large datasets of images acquired on-site. The labeling of images requires a significant amount of human time and adding new features to detect for the robots increases the time spent labeling. We also previously used manual tagging, with the help of a collaborative on-line tool we developed for that purpose ⁹.

This year, we decided to take a paradigm shift, moving from labeling images to labeling videos. The main idea is quite simple, if we can retrieve the pose and orientation of the robots camera inside the field referential for each frame, then the position of field landmarks inside images can easily be obtained. Moreover, by synchronizing the video streams from multiple robots, it is possible to share annotations among them.

Accurate estimation of the orientation of the cameras is a difficult problem, in order to tackle it, we experimented two different methods: using ViveTrackers on the head of the robots and combining manual labeling with odometry. Both methods share similar issues regarding time synchronization between devices and its impact on orientation estimation. We used two complementary schemes to reduce the uncertainty on camera orientation.

1. We used specific tool (`chrony`¹⁰) to synchronize all the information streams.
2. The impact of timing differences is mitigated using the following acquisition method for videos. Robots alternate between two different phases: walking to a randomly generated location and slowly scanning the environment while

⁹ <https://github.com/rhoban/tagger>

¹⁰ <https://chrony.tuxfamily.org/>

standing still. During extraction of labels, only the images obtained during the scanning phase are considered. This proved to be necessary because the head of the robot is shaking when the robot is walking, thus increasing the uncertainty on the orientation of the camera.

Access to the fields for data acquisition is a scarce resource during the setup days. In order to optimize the usage of the time we were given, we created a specific training scenario in which the field is separated in as many zones as the number of robots used for acquisition.

- A zone is allocated to each robot.
- During 2 minutes, each robot alternates between moving to a random location inside its zone and scanning its environment.
- Zones are separated by a safety buffer of around 50cm to reduce the risk of collision between robots.
- During the training, perception is disabled and robot relies solely on odometry to estimate its position.

Automatically through Vive *Vive* is an indoor tracking system developed by *HTC*. It is based on active laser emitters base stations called *Lighthouses* that sends laser sweeps on a known frequency. Infrared receivers are used and the time when they are hit by the sweeps of lighthouses is used to compute the position of the object they are attached to. There are two generations of the lighthouses. The first generation only works by pairs and covers a maximal tracking area of $5m \times 5m$. With the second version, the area can be increased up to $10m \times 10m$ by using four lighthouses. Initially designed to track *Vive* controllers and helmets for virtual reality application, it is now also possible to buy simple trackers that you can attach to anything to track its position.

This makes ground truth possible with attaching a *Vive* tracker to the head of the robot, and capture some logs. The only thing that is required is then the transformation from the tracker frame to the camera frame, and the ability to project a known object 3D position onto image.

Vive trackers and lighthouses are easy to carry and deploy. Moreover, the calibration phase before being able to track objects is fast. This makes it a better choice than motion capture for on-site calibration, while being more affordable and with a wider working area. We used it during the *German Open 2019 Humanoid KidSize* competition to generate automatically labelled images.

A calibration phase is still needed to find the 3D transformation from the lighthouses to the field frame, and also to tag the balls position on the field every time we move them. To achieve that conveniently, we use a *Vive* controller which is itself tracked and equipped with trigger button to mark some known position on the field and find the optimal 3D transformations, or show the balls position at the beginning of a log. We developed a custom tool to do that¹¹ on the top of OpenVR SDK¹².

¹¹ https://github.com/Rhoban/vive_provider

¹² <https://github.com/ValveSoftware/openvr>

Even if most of the data obtained this way was suitable for training our neural networks, the accuracy of such technology can be discussed[6]. In order to have a better control on the calibration and the data fusion algorithm used to compute the position using the IMU and light beam datas, the authors of LibSurvive¹³ reverse engineered *Vive*. Having access to the low level data allowed better positioning results[2]. The support for the second lighthouse generation is currently being completed.

One of the drawbacks intrinsic to this method is that it cannot be used during real games.

Combining labeling and odometry Finding a camera pose from 3D-2D correspondances is a well-known problem[5], by labeling the position of keypoints of the field in an image, it is relatively easy to retrieve an accurate estimation of the pose of the camera which took the image. However, it is not realistic to apply this method for every image of a video for two majors reasons: this would require large amount of human-time to label a video and some frames do not even contain keypoints from the field.

Using odometry to extrapolate the pose of the camera before and after a labeled frame helps to strongly reduce the labeling burden. Experimentally, we noticed that labeling around 10 frames for a 2 minute session containing more than 1500 usable frames hold satisfying results.

Currently, the major flaw of this method is that it requires the robots to stop walking and reduce the scanning speed in order to improve the pose estimation. Those conditions are rarely met during real games, making it difficult to extract data from real games, a crucial point to enhance opponent detection. In the future, we hope to take leverage on visual-inertial odometry methods[4] to enhance the accuracy of pose estimation in more dynamic situations. This would allow to easily label videos of entire games quickly, thus tackling the problem of building large datasets for neural network training.

5.2 Multi-class ROI and classification

Last year, our perception system was mainly based on three specific pipelines, one for detecting the ball, a second one detecting the base of the goal posts and a third one to detect the corner of the arena field. The two first systems were roughly similar, but each system had its own neural network and some specificities to identify region of interests[7]. In order to make the perception system simpler while covering more type of features, we decided to use a single system to identify the region of interests and classify the type of feature. While less accurate for centering the features in the region of interests, it allows to incorporate additional features such as line intersections or opponents more easily.

While this new system yielded promising results during the preparation of the RoboCup, we struggled to obtain decent results on-site and were forced to

¹³ <https://github.com/cnlohr/libsurvive>

limit the perception to two different type of features: balls and base of goal posts. We initially thought that the main problem was the fact that the posts were thinner than the posts we used in our laboratory. However, after RoboCup, we ran a thorough code review and found 3 major bugs between the training process of the neural networks and the online prediction of the class. Once we solved them, we were able to include more classes while having an accuracy rate much higher than what we had during RoboCup. Due to these bugs, it is not possible to provide meaningful results for the code we used during RoboCup. However, the final results promise a strongly improved perception system for next year.

5.3 Localisation

We use a three-dimensional particle filter for localization which includes the position and the orientation of the robot on the field. It fuses information from both: the perception module and the odometry. Due to the major issues in perception at the beginning of the competition, we decided to strongly reduce the exploration for the first games. Increasing the confidence in the odometry allowed us to stay into the competition until we improved the perception part.

The position and the orientation of the robot used to be defined respectively as the average of the positions and the orientations of the particles. The major change we introduced this year was to fit a Gaussian mixture model on the set of particles by using the Expectation-Maximization[1] algorithm. It finds iteratively the partition in k disjoint subsets that maximizes the likelihood of the corresponding Gaussian mixture model. The choice of the k is done online as follows. Let $C_k = \{c_1, \dots, c_k\}$ be the clustering into k disjoint subsets obtained by the Expectation-Maximization algorithm. Let $|c|$ and $\text{Var}_P(c)$ denote respectively the number of particles and the variance of the particles positions. Finally, we define respectively the internal variances of position and orientation of C_k as

$$\text{Var}_P(C_k) = \frac{\sum_{i=0}^k \text{Var}_P(c_i) * |c_i|}{\sum_{i=0}^k |c_i|}.$$

Starting from $k = 1$, the number of clusters is increased until we reach $k = 5$ or $\text{Var}_P(C_k) \geq 0.5\text{Var}_P(C_{k+1})$. This ensures that adding a cluster is only done if it provides a major reduction of the variance. When several clusters are considered, only the most populated one is considered for high-level decisions.

The proposed method provides two main benefits with respect to the simple solution of taking the average of all particles. First, it allows to obtain meaningful orientation while the particle filter has not converged. This situations frequently occurs when a robot comes back from penalty from the side line. During games, it allows to discard scattered particles which will be automatically attributed to a cluster containing the noise, see Fig. 7a. Second, representation of localization belief under the form of multiple clusters allows to store or broadcast the information at a much lower cost than sending the position of all the particles. Therefore, this method enables real-time monitoring under low bandwidth conditions, see Fig. 7b

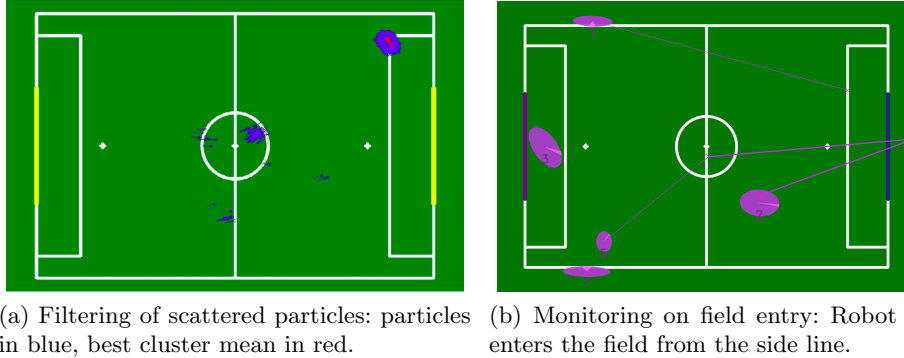


Fig. 7: In-game examples of localization.

6 Strategy

In 2017, we introduced an off-line value iteration process to compute a kick policy that chooses a type of kick as well as the orientation aiming to optimize the time to score a goal with one robot on an empty field[7]. The reward function used for the optimization is simply the time needed for the robot to reach the next ball position, a penalty if the ball is kicked out of field and 0 if a goal is scored. This process produces a value function V that gives us an estimation of the time it takes to score a goal from a given position on the field.

This year, we also used an online value iteration that performs an optimization at depth one to include the current state of the game. It can roughly be described as follows. The discrete set of ball positions on the field is called S . An action is a type of kick together with a discrete orientation. We denote A the set of possible actions. Let $P_a(s, s')$ denotes the probability of reaching state s' from state s after performing action a kick with orientation a . The knowledge of the game is introduced with a reward function $r(s, s')$ described thereafter. The online policy is given by the formula

$$\arg \max_{a \in A} \sum_{s' \in S} P_a(s, s')(r(s, s') + V(s')).$$

To compute the reward r , we check if we are in a state where it is not allowed to score a goal. For example in the case of a throw-in, an indirect penalty kick or when we have the kick-off and the ball is still not in play (exited the center circle). In this case, a penalty score is given if a goal is scored. In that case, the robot will naturally kick the ball so that it does not score a goal, but placing it in the best situation possible to score a goal on the next kick. The reward $r(s, s')$ also includes the time for the closest robot to reach s' from which we subtract the time for the kicking robot to reach s' . Finally, we give penalties to kicks towards opponent robots if we have their location. As an example, in the situation of Fig. 8, the robot 1 is going for the ball, because he is the closest to

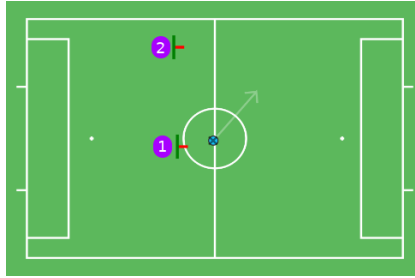


Fig. 8: An example situation using on-line Monte Carlo.

it. He is not allowed to score a goal, because of the kick-off conditions. Hence, a short kick action getting the ball out of the center circle is preferred by the on-line iteration instead of a powerful kick that would certainly score. If the robot was alone on the field, the optimal orientation would be straight forward. However, the iteration produces this left oriented kick, because by the time we estimate the kick is done, the robot 2 will be properly positioned to handle the ball and score a goal faster than if robot 1 would have kicked straight.

References

1. Bilmes, J.a.: A gentle tutorial of the EM algorithm and its application to parameter estimation for Gaussian mixture and hidden Markov models. *International Computer Science Institute* (1998). <https://doi.org/10.1.1.119.4856>
2. Borges, M., Symington, A., Coltin, B., Smith, T., Ventura, R.: Htc vive: Analysis and accuracy improvement. In: 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). pp. 2610–2615. IEEE (2018)
3. Felis, M.L.: Rbdl: an efficient rigid-body dynamics library using recursive algorithms. *Autonomous Robots* **41**(2), 495–511 (2017)
4. Gui, J., Gu, D., Wang, S., Hu, H.: A review of visual inertial odometry from filtering and optimisation perspectives. *Advanced Robotics* (2015). <https://doi.org/10.1080/01691864.2015.1057616>
5. Lepetit, V., Moreno-Noguer, F., Fua, P.: EPnP: An accurate $O(n)$ solution to the PnP problem. *International Journal of Computer Vision* (2009). <https://doi.org/10.1007/s11263-008-0152-6>
6. Niehorster, D.C., Li, L., Lappe, M.: The accuracy and precision of position and orientation tracking in the htc vive virtual reality system for scientific research. *i-Perception* **8**(3), 2041669517708205 (2017)
7. NGuyen, S., Passault, G., Pirrone, A., Rouxel, Q.: Rhoban football club: Robocup humanoid kid-size 2017 champion team paper. *RoboCup 2017: Robot World Cup XXI* **11175**, 423 (2018)
8. Quigley, M., Conley, K., Gerkey, B., Faust, J., Foote, T., Leibs, J., Wheeler, R., Ng, A.Y.: Ros: an open-source robot operating system. In: *ICRA workshop on open source software*. vol. 3, p. 5. Kobe, Japan (2009)