Quantitative Evaluation of Humanoid Robots

Enrico Piazza, Giulio Fontana, Matteo Matteucci
Politecnico di Milano - DEIB
Milan, Italy
{enrico.piazza;giulio.fontana;matteo.matteucci}@polimi.it

Daniele Nardi
Università di Roma La Sapienza
Roma, Italy
daniele.nardi@uniroma1.it

Martino Migliavacca
Nova Labs
Roverto, Italy
m.migliavacca@novalabs.io

Abstract*—Evaluation of the performance of autonomous robots, though important both to research and to industry, is a challenging task. This is especially true for humanoid robots designed to operate in human environments. One effective approach to this problem is to define constrained versions of everyday human activities, trying to strike the right balance between enabling the definition of suitable performance metrics and making the whole task too artificial for such metrics to be meaningful. This approach is the one chosen by the European project EUROBENCH to assess the performance of humanoid robots, exoskeletons and prostheses. In this paper we present we present two projects focuses on specific "benchmarking scenarios": MADROB for scenario *Opening/Closing Doors* and BEAST for scenario *Pushing a Shopping Trolley".

***Index Terms*—Robotics benchmarking, performance evaluation, humanoid robots**

I. Introduction

Evaluation of the performance of autonomous robots, though important both to research and to industry, is a challenging task. This is especially true for humanoid robots designed to operate in human environments, where the loosely structured nature of most tasks makes quantitative evaluation difficult. One effective approach to this problem is to define constrained versions of everyday human activities, trying to strike the right balance between enabling the definition of suitable performance metrics and making the whole task too artificial for such metrics to be meaningful.

This approach is the one chosen by the European project EUROBENCH [1], [2] to assess the performance of humanoid robots, exoskeletons and prostheses. In its first phase the project selected, via an open call, a set of *EUROBENCH subprojects* devoted to developing benchmarks for such machines. Each of the chosen subprojects focuses on a specific "benchmarking scenarios" and provides a benchmark for it. This paper describes two such subprojects: **MADROB**, for scenario "Opening/Closing Doors"; and **BEAST**, for scenario "Pushing a Shopping Trolley".

Both (sub)projects involve the same consortium of partners, and both are applications of the same basic idea, i.e., transforming a common task into a benchmark by transforming the "passive" device that the task revolves around into a robot. In order to maintain the significance of the tasks, MADROB's and BEAST's robots retain the appearance, function and behaviour of the passive device they mimic, while endowing it with two abilities that are key to benchmarking: (i) sensing (to quantitatively measure the actions of the agent –robot, person with exoskeleton, person with prostheses– and their effects), and (ii) actuation (to apply controlled disturbances to the evolution task).

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 779963.

Fig. 1. Robotized door used for MADROB benchmark. The fully assembled door (left) and a detail of the lock assembly (right) internal structure.

Fig. 2. MADROB at the SciRoc Smart Cities Challenge benchmarking competition. The testbed comprising the door mimicked a domestic setting.

II. The MADROB benchmark

In MADROB (Modular Active Door for RObot Benchmarking) the transformed device is a (hinged) door, as shown in Figure 1. While closely matching a standard door in all its external features (including mass), MADROB's robotized door perceives the opening angle of the door panel and is capable

Fig. 3. Robotized shopping trolley used for the BEAST benchmark. The box houses motors, electronics and communication gear.

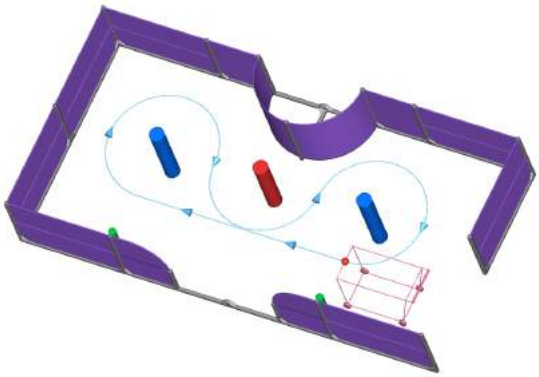

Fig. 4. Testbed of the BEAST benchmark.

of applying controlled torque to the panel. This is used to simulate a range of phenomena, such as hydraulic damping, the presence of an object behind the door (so that the user has to overcome the object's friction against the floor), the effect of wind or of a pebble stuck under the door, and so on. Additionally, the robot is able to measure the forces applied to the handle, using load cells and custom signal acquisition electronics. Finally, two arrays of infrared distance sensors on both sides of the threshold are used to perceive the passage of the user through the door.

The MADROB benchmark has been experimentally validated at the 1st SciRoc Challenge international competition [3], which took place in Milton Keynes (UK) in September 2019 (see Figure 2).

The MADROB benchmark requires that the robot reaches the door, uses the handle to unlatch the lock, opens the door, reaches the other side and closes the door again. Performance evaluation uses metrics based on execution times (both overall and per-phase) and metrics that focus on safety in door operation (based on angular acceleration of the door panel and forces applied to the handle). Finally, a "capability level" metric depends on how much of the task the robot actually accomplished.

III. The BEAST benchmark

In BEAST (Benchmark-Enabling Active Shopping Trolley), the passive device transformed into a robot is a shopping mall trolley, with two fixed-axis wheels and two swiveling wheels. Differently from MADROB, in BEAST it has been possible to build the robot by modifying a commercial trolley, without need to rebuild it from scratch.

For what concerns actuation, BEAST exchanges the fixed-axis wheels with an active drive train (with differential drive kinematics). Sensing is represented by a single-plane LiDAR (Slamtec RPlidar A3, used for localisation) and by load cells, used to measure forces applied to the handle of the trolley. Both motor control and signal acquisition are done by custom electronics. Figure 3 shows the robotized shopping trolley.

An additional, passive element of the BEAST benchmarking infrastructure is the testbed, i.e. the physical environment where the benchmark takes place. This is a 3 m x 6 m environment provided with a perimeter (deliberately interrupted) and a few colored pillars used as references to define the desired trajectory. Figure 4 is a CAD rendering of the testbed.

To execute the benchmark, the robot is required to push the trolley (in purple) as accurately as possible along the specified trajectory (light blue). Trolley sensing is used to measure the forces applied by the robot to the handle, in order to detect anomalous handling (e.g., force intensities with the potential to cause harm to nearby people) and trajectory tracking precision. Trolley actuation is used to simulate typical anomalies of shopping mall trolleys, including excess friction on wheels, unbalanced friction (i.e., tendency to deviate), abrupt disturbancies such as a pebble under one wheel or a small step in the surface of the floor, and so on. Additional variants of the benchmark involve loading the trolley with mass.

Performance metrics of the BEAST benchmark include time-focused ones (e.g., times to execute the straight and curved parts of the trajectory), accuracy-focused ones (e.g., precision in following the straight part of the trajectory, minimum distance from obstacles) and an overall "capability level" assessing how much of the task the robot actually accomplished.

A second wheeled active device, a walker, is currently being developed. This will be used to create versions of the benchmark where bipedal robots (or people wearing exoskeletons or prostheses) use the walker for support and stabilisation while moving along the prescribed trajectory.

The shopping trolley version of BEAST will be one of the robot benchmarks of the 2nd SciRoc Smart Cities Challenge, taking place at Bologna (Italy) in September 2021.

References

[1] (2020) European robotic framework for bipedal locomotion benchmarking. [Online]. Available: http://eurobench2020.eu/

[2] D. Torricelli and J. L. Pons, "Eurobench: Preparing robots for the real world," in *Wearable Robotics: Challenges and Trends*, M. C. Carrozza, S. Micera, and J. L. Pons, Eds. Cham: Springer International Publishing, 2019, pp. 375–378.

[3] (2019) Sciroc smart cities challenge 2019. [Online]. Available: https://sciroc.org/challenge-description-2019/