# MIMOSO@Home Team Description Paper RoboCup Brazil Vitória/ES

1st Murilo C.Bicho
*C3 - Centro de Ciências Computacionais*
*Universidade Federal do Rio Grande - FURG*
Rio Grande, Brazil
murilobicho@gmail.com

***Abstract*—This Team Description Paper (TDP) presents and details the MIMOSO robot, developed as part of an individual master's project in Computer Engineering at the Federal University of Rio Grande (FURG). The work focuses on creating a low-cost humanoid robot designed to reproduce human behaviors and expressions in a simplified manner, aiming to facilitate interaction and promote empathy with users. The focus is placed on the mechanical architecture, perception and control systems, as well as locomotion strategies.**

## I. Introduction

General-purpose domestic and social robots have been an ambition of engineers for several decades. The growing demands and complexities of modern life point to a future in which robots will play—or already play—an important role in people's daily lives. Recent technological advances have made this aspiration more feasible than ever.

In this context, this project presents an independent initiative developed as part of a master's thesis, focused on creating a small, low-cost humanoid robot named MIMOSO. Unlike many existing robots, which generally require high financial investment and complex infrastructure, this project aims to develop an accessible platform capable of reproducing human behaviors and expressions in a simplified yet effective manner.

The robot integrates affordable hardware components, including 3D-printed parts, low-cost sensors, and employs reinforcement learning techniques for autonomous locomotion. For the robot's appearance, a simple and cute design was adopted, with special attention to facial details, seeking to avoid the uncanny valley effect [3] and promote natural and empathetic interaction with users.

This document is organized as follows: Section II describes the research conducted for the robot, including its mechanical structure, hardware components, perception and control systems, and also discusses the locomotion strategy based on simulation and reinforcement learning; Section III presents the expected contributions and future work; finally, the conclusion summarizes the project's advances and perspectives.

## II. Research

For the development of a low-cost humanoid robot aimed at the mimetic reproduction of human behaviors and characteristics, with the purpose of fostering empathy and mitigating the uncanny valley effect, it is essential to conduct comprehensive research and to develop technologically appropriate methods and devices that enable the achievement of this objective.

### A. Mechanical Structure

Considering the objective of keeping the robot low-cost, it was necessary to establish dimensional constraints. Thus, the robot has a total height of 75 cm, measured from the soles of the feet to the top of the head, with 35 cm corresponding to the head region. The robot consists of two legs, a torso, and two arms, each equipped with manipulators at their ends. In total, these structures comprise 33 joints, distributed as follows: 6 in each leg, 2 in the torso, 3 in the neck, 7 in each arm, and 3 in each hand.

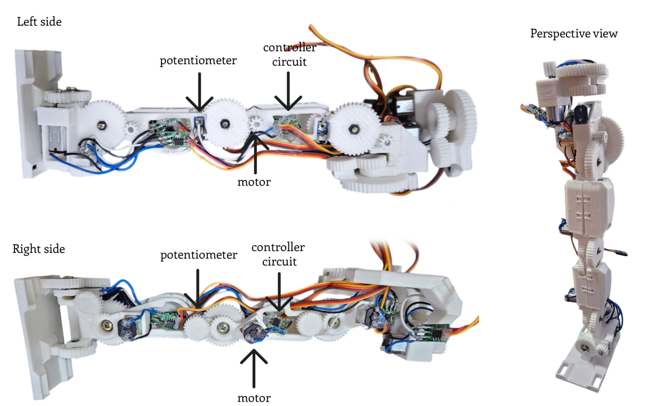


Fig. 1. Image showing the joint control and activation system, including the potentiometer, motor, and control circuit

The actuators responsible for joint movement are custom-built servomotors, assembled from N20 DC motors combined with SG90 servo motors. The output torque is increased through a gear reduction mechanism integrated into the joint assembly, which reduces rotational speed and enhances mechanical leverage. Position feedback is obtained from a potentiometer installed in the joint, which provides real-time positional data to the servomotor's internal control circuit. This closed-loop control system continuously monitors the actuator's status and adjusts motor operation to accurately maintain the commanded position. Figure 1 shows an image of the robot's knee.

For the construction of the robot's external structure, a white canvas was used to represent a covering similar to human skin. This canvas is shaped according to the robot's form and inflated with air, resulting in a balloon-like configuration. In addition to its aesthetic aspect, this external layer serves to protect the internal components from moisture and from potential impacts caused by falls during task execution. The figure 2 shows the full-body robot

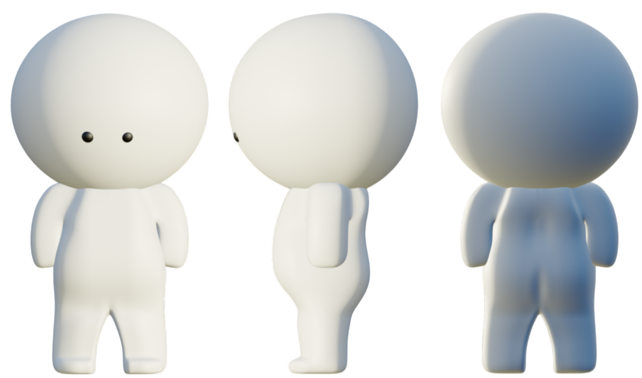

Fig. 2. Front, side, and rear views of Mimoso

All parts of the robot were designed using the Onshape CAD modeling software[1] and manufactured through 3D printing. The components that require greater strength, such as those of the legs, arms, and abdomen, were produced using FDM (Fused Deposition Modeling) technology. The parts that demand higher detail, due to their reduced size and complexity—such as the face and hands—were printed in resin.

The robot's torso houses the battery and the microcontrollers responsible for sending position commands to the joints. Figure 3 presents a schematic drawing of the robot's lower section and torso, where the battery is located.

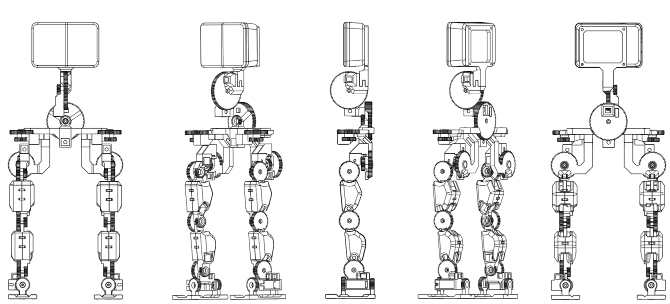

Fig. 3. Schematic drawings of the lower part of the robot in front, side, and rear views

## B. Head

The most important element for conveying friendliness in a robotic agent is the face. Designing a face with features too close to those of a human can cause discomfort (the uncanny valley effect), which is why a more simplified design was chosen: a completely spherical head composed of only two eyes and a mouth. The back of the head is made of canvas, as is the rest of the body, while the front is molded from white silicone. Silicone was chosen due to its flexibility, an essential characteristic for representing facial expressions.

Facial movements are carried out by a mechatronic system composed of 12 servomotors. Two of them control eyelid movement, while four actuators deform a cavity representing the mouth, enabling actions such as opening, closing, and even smiling. The remaining six motors deform the silicone positioned just above the eyes, creating a shadow that simulates an eyebrow. This indirect "eyebrow" allows the robot to convey different emotional expressions, such as anger, happiness, doubt, fear, and sadness, as illustrated in the figure.4.

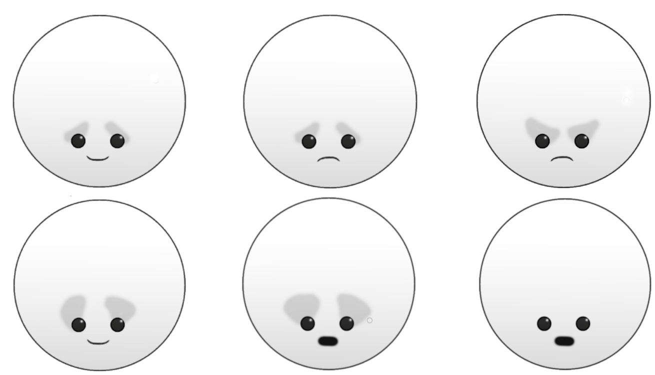

Fig. 4. Facial expressions

On the sides of the head, two INMP441 omnidirectional microphones are installed for voice recognition. Inside the head, just behind the mouth, there is an Altomex AL1115 active mini speaker responsible for reproducing the robot's voice. The robot's eyes are composed of two HQ-USB-1080HD USB camera modules.

## C. Vision

Considering that the goal of this device is to be a low-cost humanoid robot, the only sensors used to estimate its location in the environment are the two camera modules positioned on the robot's face, together with an IMU. Mimoso's vision package integrates several advanced techniques, including depth estimation such as Depth Anything V2 [2], stereo vision, and back projection, with the purpose of reconstructing the three-dimensional environment in which the robot operates.

These methods work together to extract detailed spatial information from the images captured by the cameras, enabling the robot to understand the geometry of its surroundings without the need for additional, more complex, and costly sensors such as LiDAR or ultrasonic sensors.

Depth estimation enables the generation of depth maps from two-dimensional images, while stereo vision allows for the precise calculation of object distances through triangulation

[1] https://www.onshape.com/.

of images captured from different viewpoints. Back projection complements this process by transforming the obtained information back into three-dimensional space, facilitating environmental modeling and navigation. The entire process can be seen in the figure 5

In this way, Mimoso's vision system provides a cost-effective and efficient solution for spatial perception, which is essential for the humanoid robot's autonomy and interaction within its operational environment.

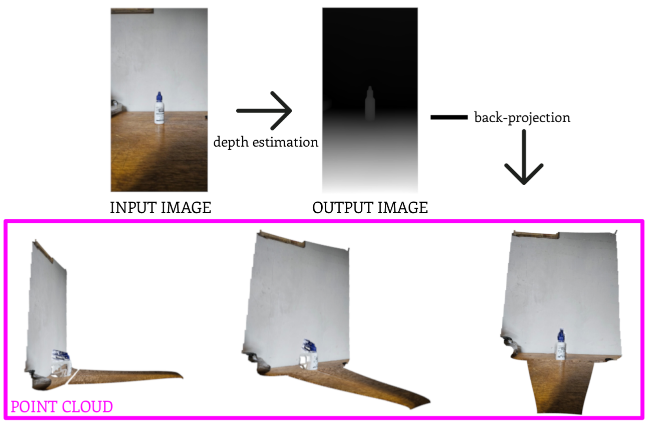


Fig. 5. Process of converting an image into a point cloud

### D. Locomotion

To enable the robot to learn to walk, the Isaac Sim simulation environment[2], developed by NVIDIA, was used. Isaac Sim offers Isaac Lab — a comprehensive platform for training robotic agents through reinforcement learning [1]. MIMOSO can be seen in Isaac Lab in the figure 6.

Within this environment, the robot underwent training cycles where, through trial and error, it learned to perform stable and efficient walking movements. The use of reinforcement learning allowed the robot to develop control policies directly from interaction with the simulated environment, without the need for manual modeling of movements or balance.

Isaac Lab provides advanced tools for simulating realistic physics, sensors, and actuators, as well as support for deep learning frameworks, enabling the creation of a robust and adaptable control model for the robot's bipedal locomotion.

Thus, the training performed in Isaac Sim enabled the autonomous development of the robot's locomotion capabilities, accelerating the process in parallel with the construction of the physical robot and reducing the costs and limitations inherent to tests conducted directly on hardware.

## III. Contributions

Although the project is still in its early stages, it is expected to contribute significantly to the advancement of low-cost humanoid robot development, especially those focused on mimetic reproduction of behaviors and empathetic interaction with humans.

[2]https://developer.nvidia.com/isaac/sim.

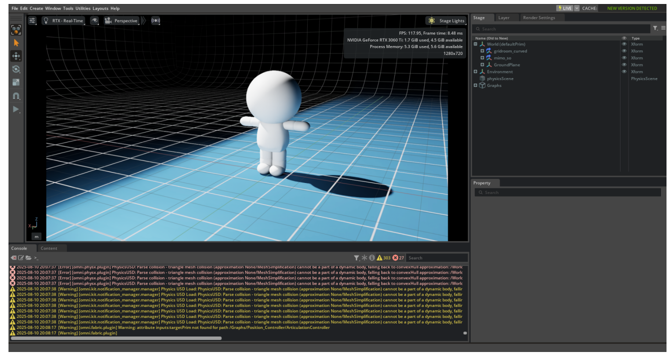

Fig. 6. MIMOSO in Isaac Lab

Among the main anticipated contributions are:

- **Development of an accessible platform**: The proposal to use economical materials and components, combined with open-source solutions and advanced machine learning techniques, aims to create a replicable and financially viable platform for academic institutions, startups, and robotics enthusiasts.
- **Solutions in visual perception**: By integrating techniques such as stereo vision, depth estimation, and back projection into a simplified system, it seeks to provide an alternative for spatial perception in humanoid robots without the need for expensive and complex sensors.
- **Promotion of human-robot interaction**: The focus on simplified facial expression, combined with voice communication and auditory recognition, aims to foster a more natural and empathetic relationship between the robot and users, paving the way for applications in education, assistance, and entertainment.

## IV. Conclusion

This project presents an initial approach to the development of a low-cost humanoid robot, focusing on reproducing human behaviors and characteristics that promote empathy and reduce the uncanny valley effect. The combination of an optimized mechanical structure, visual perception systems based on economical cameras, and advanced reinforcement learning techniques for locomotion demonstrates the potential to create an accessible and functional platform.

Although still in the early phase, the project establishes the foundation for future developments, aiming to contribute to advances in accessible humanoid robotics and more natural human-robot interaction.

## References

[1] Ilija Radosavovic, Tete Xiao, Bike Zhang, Trevor Darrell, Jitendra Malik, Koushil Sreenath. Real-world humanoid locomotion with reinforcement learning. *Science Robotics*, 9(89): eadi9579, 2024. DOI: 10.1126/scirobotics.adi9579. Disponível em: https://www.science.org/doi/abs/10.1126/scirobotics.adi9579.

[2] Lihe Yang, Bingyi Kang, Zilong Huang, Zhen Zhao, Xiaogang Xu, Jiashi Feng, Hengshuang Zhao. Depth Anything V2. 2024. Disponível em: https://arxiv.org/abs/2406.09414.

[3] Masahiro Mori. The uncanny valley. *Energy*, 7(4), 33–35, 1970.

**Title:** MIMOSO@Home Team Description Paper RoboCup Brazil Vitória/ES
**Email:** murilobicho@gmail.com
**Team name:** MIMOSO@Home
**Members:** Murilo Castanheira Bicho

## ROBOT’S HARDWARE DESCRIPTION

Components that form the robot’s structure:

- Raspberry
- jetson Orin Nano
- HQ-USB-1080HD USB camera module
- INMP441 omnidirectional microphone
- Altomex AL1115 active mini speaker
- MPU-6050
- Battery 4400mah 36v
- Two 6 DOF leg
- Two 7 DOF arm
- 2 DOF torso
- Tow 3 DOF gripper
- 12 DOF head

## ROBOT’S SOFTWARE DESCRIPTION

Components that make up the robot’s software:

- ROS2 - Robot Operating System
- Isaac Sim
- Isaac Lab
- Depth Anything V2
- YOLOV8