CEA Comité Español de Automática
XLVII JORNADAS DE AUTOMÁTICA CÓRDOBA 2026
UNIVERSIDADE DA CORUÑA

Jornadas de Automática

Towards a multi-robot coordination and cooperative perception for gait analysis

Diego Martinez[a,*], Alberto Brunete[a], David Álvarez[a], Miguel Hernando[a]

[a] *Dpto. de Automática, Ingeniería Electrónica e Informática Industrial, Universidad Politécnica de Madrid, C/ Ronda de Valencia 3, 28012, Madrid, España.*

To cite this article: Martinez, D, Brunete, A, Álvarez, D, Hernando, M. 2026. Towards a multi-robot coordination and cooperative perception for gait analysis. Jornadas de Automática, 47. https://doi.org/10.17979/ja-cea.2026.47.13761

Resumen

Este trabajo presenta RoboMesh, una arquitectura descentralizada de percepción para el análisis de la marcha basada en sensores RGB-D. El sistema permite que cada robot estime de forma independiente la posición del paciente, transformando los datos en variables de seguimiento. Este marco técnico sienta las bases para mitigar oclusiones en el análisis de la marcha mediante la futura integración de datos multivista. La implementación utiliza ROS2, Fast DDS y sincronización PTP sobre plataformas Rover Mini, constituyendo el pilar para futuras capacidades de mapeo compartido y localización colaborativa en entornos clínicos.

Palabras clave: Robots móviles, Control descentralizado, Percepción y sensado, Sistemas robóticos autónomos, Sistemas multiagente, Sistemas robóticos en red.

Hacia la coordinación multirrobot y percepción cooperativa para análisis de la marcha

Abstract

This work presents RoboMesh, a decentralized perception architecture for gait analysis based on RGB-D sensors. The system allows each robot to independently estimate the patient's position, transforming the data into tracking variables. This technical framework lays the foundation for mitigating occlusions through the future integration of multi-view data. The implementation utilizes ROS2, Fast DDS, and PTP synchronization on Rover Mini platforms, constituting the pillar for future shared mapping and collaborative localization capabilities in clinical environments.

Keywords: Mobile robots, Decentralized control, Perception and sensing, Autonomous robotic system, Multi-agent systems, Networked robotic systems.

1. Introduction

Human gait analysis is a fundamental tool in clinical rehabilitation and functional assessment, enabling the quantification of biomechanical parameters relevant to the diagnosis and monitoring of neurological and musculoskeletal disorders (Bonanno et al., 2025). Marker-based optical systems remain the gold standard due to their high accuracy; however, they require expensive infrastructures and controlled environments, limiting their applicability in real clinical scenarios (Pfister et al., 2014; Bravo and Álvarez, 2016; Scataglini et al., 2024).

RGB-D sensors have emerged as a non-invasive alternative for three-dimensional human pose estimation and robotic gait analysis (Álvarez et al., 2023). Despite their flexibility, these sensors remain sensitive to perceptual noise, temporary occlusions, and skeletal discontinuities (Pfister et al., 2014; Scataglini et al., 2024). As an evolution of the RoboGait platform, the RoboMesh project seeks to expand spatial coverage through multiple mobile robots capable of observing the patient from different viewpoints.

Compared with single-robot systems, multi-robot perception provides redundant observation perspectives, reduces the impact of occlusions, and improves monitoring robustness (Dahiya et al., 2023). In addition, cooperative robotic systems support different coordination paradigms, including centrali-

*Autor para correspondencia: diegoalejandro.martinez@alumnos.upm.es
Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0)

zed control, decentralized cooperation, leader–follower formations, and swarm behaviors (Hu et al., 2021), making them attractive for large clinical environments.

Nevertheless, multi-robot perception introduces challenges related to temporal synchronization, spatial coherence, distributed control, and relative localization (Chen et al., 2025; Wang et al., 2023; Spletzer et al., 2001). Moreover, the choice of coordination strategy directly affects scalability, robustness, and communication requirements (Olfati-Saber and Murray, 2004; Hu et al., 2021).

This work adopts a decentralized architecture in which each robot independently estimates the patient relative position using RGB-D perception. The patient acts as the main dynamic reference of the system, while skeletal information obtained through Nuitrack®[1] is used to maintain a consistent relative position under different observation geometries.

The proposed architecture integrates perception, relative position estimation, distributed communication, and local decision-making. These capabilities establish the foundation for future cooperative localization, shared mapping, and enhanced patient coverage strategies within the RoboMesh framework.

2. Cooperative perception and relative localization challenges

Multi-robot architectures can be classified into centralized and decentralized approaches. In centralized systems, a master node concentrates information and is responsible for global coordination and information fusion. Although this approach simplifies coordination, it introduces a single point of failure and limits scalability as the number of agents increases Dahiya et al. (2023). In contrast, decentralized architectures distribute perception and decision-making among all robots, improving robustness and fault tolerance (Olfati-Saber and Murray, 2004; Hu et al., 2021).

Coordination schemes include leader–follower formations, cooperative encirclement, distributed consensus algorithms, and swarm-based self-organized behaviors (Fareh et al., 2023). These approaches have demonstrated effectiveness in applications such as formation control, cooperative navigation, target tracking, and distributed localization. In environments shared with humans, decentralized leader–follower strategies are particularly attractive because they allow each robot to react directly to local observations while preserving group coherence (Talebpour and Martinoli, 2018; Dahiya et al., 2023).

This work adopts a decentralized leader-follower strategy. The patient act as the leader and the robots move to keep a relative view-point. The relative position of the patient is determined using an RGB-D sensor and Nuitrack software, an external middleware specialized in real-time skeletal estimation. The information provided is the joint position, such as shoulders, hips, and ankles, in the camera frame. However, individual estimations of each robot remain sensitive to perceptual noise, partial occlusions, and temporary skeletal losses (Pfister et al., 2014; Scataglini et al., 2024).

[1] https://nuitrack.com/

Relative localization also presents different challenges depending on the availability of a global map. In mapless environments, robots rely exclusively on local observations and inter-agent communication to verify spatial consistency. In contrast, shared-map configurations enable cooperative triangulation strategies that compare the estimated human orientation and distance with the global relative position of each robot (Wang et al., 2023; Spletzer et al., 2001). This cooperative mechanism allows the construction of a virtual geometric triangle between robots and patient to validate inconsistencies in local orientation estimations.

To provide a common temporal framework, the RoboMesh architecture employs ROS2 and Fast DDS as distributed communication infrastructure for information exchange between robotic agents (Chovet et al., 2025). Besides, it uses temporal synchronization through PTP (Precision Time Protocol) to temporary align skeletal observations obtained from multiple RGB-D cameras.

In this work, the focus is on validating the robustness of the decentralized communication infrastructure, as it is an essential prerequisite for the subsequent implementation of the described cooperative gait analysis system.

3. System architecture and coordination

The RoboMesh multi-robot frameworks is implemented using two Rover Mini 4WD mobile platforms equipped, each one, with an Orbbec Femto Mega RGB-D sensor. Local processing is handled by an onboard NUC-computer running ROS2 Humble. A single-degree-of-freedom Dynamixel actuator permits to adjust the camera orientation to maintain the patient within the field of view, if necessary.

3.1. Communication and synchronization infrastructure

The RoboMesh architecture relies on the DDS standard for decentralized and efficient data exchange. Within the ROS2 ecosystem, several middleware implementations are available with distinct performance profiles, such as Cyclone DDS, Zenoh, Micro XRCE-DDS, and Fast DDS. While Zenoh targets large-scale distributed systems and Micro XRCE-DDS specializes in resource-constrained devices, this work utilizes Fast DDS due to its maturity and high performance in low-latency environments (Chovet et al., 2025).

This selection is justified by the ability of the middleware to robustly handle the exchange of skeletal information obtained through NUITRACK. Its status as the default for ROS2 Humble ensures native integration with discovery tools. Each robot operates within an independent ROS2 namespace to prevent topic conflicts.

To maintain temporal consistency, the system incorporates the Precision Time Protocol (PTP) through the ptpd Linux-based implementation. Although currently utilized for observation alignment, future iterations will employ PTP synchronization to match skeletal joints detected by different cameras, which establishes a coherent and unified perceptual frame using sensor fusion techniques.

3.2. Coordination and behaviour logic

The decentralized approach allows each agent to execute movement decisions using only local perceptual data, eliminating the need for global shared maps during patient tracking. Figure 1 illustrates the functional organization of the perception and control modules implemented on ROS2.

This architecture is designed to handle relative localization challenges with and without a map. In mapless scenarios, robots rely on their local information. However, if a map is available, robots can utilize cooperative triangulation to cross-verify if the estimated distance and angle relative to the patient and align with their own global position and orientation. This process would help to detect spatial inconsistencies and improves the accuracy of the tracking.

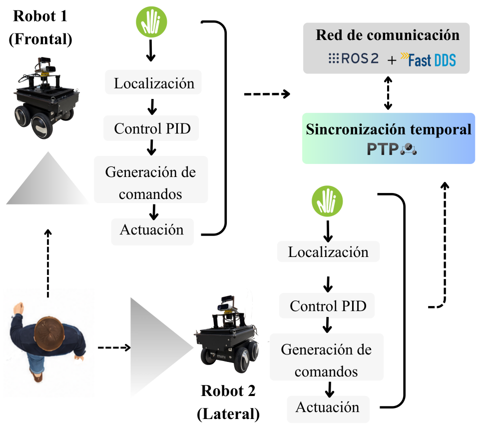

Figura 1: RoboMesh distributed coordination and local decision-making architecture.

The behavior of the system is governed by a state machine with three modes: Idle (no patient detected), Tracking (active follow-up) and Recovery. The Recovery state is crucial for clinical environments; if the system loses skeletal data due to an occlusion, the robot maintains its last valid state while attempting to reacquire the patient, avoiding unexpected behaviors.

4. Geometric proposal of the robot formation framework

The primary objective clinical gait assessment is to help clinicians by providing real-time estimates of human joint positions and trajectories during gait analysis. When using robots equipped with cameras for the analysis, depending on the relative orientation between the robot and the patient, different anatomical planes can be evaluated.

Depending on the relative orientation between the robot and the patient, different anatomical planes can be observed. Lateral viewpoints 90° or 270° provide sagittal-plane information, frontal viewpoints 0°. enable coronal-plane analysis, and diagonal configurations (45° and 315°) offer complementary views while reducing joint occlusions. The multi-robot approach enables the simultaneous acquisition of complementary viewpoints during a single clinical trial, reducing occlusions and avoiding repeated tests.

One of the main challenges in robotic person follow-up is representing the relative position of the patient consistently under different observation geometries. In frontal, lateral, or diagonal configurations, the variables observed by the RGB-D sensor present different behaviors, which complicates the use of a unified control strategy.

Figure 2 illustrates the relationship between the human, robots, and cameras coordinate systems. In the left side, the Y_p axis is defined as the positive human movement. This orientation is defined as the zero-angle in robot positioning in the formation. This angle increases in the clockwise direction. Then, each robot must advance along its own positive (X_r) axis due to its kinematic configuration.

The robot lateral axis (Y_r) corresponds only to a rotational adjustment because the platform is not equipped with omnidirectional wheels. The robot location is defined by a distance to from the human to the camera, and an angle θ defined with respect to the human positive movement (Y_p). The right side of Fig. 2 represents the RGB-D camera coordinate system under different observation positions. This representation illustrates how camera field of view, which is shown in gray, varies according to the relative angle θ between the human and the sensor.

The relative position of the patient with respect to the camera is obtained using the Nuitrack software. The position vector is expressed in the reference frame of the camera as: $P_c = \left[X, Y, Z\right]'$. Here, X represents the lateral displacement of the human with respect to the camera, and Z represents the depth measured by the RGB-D sensor. Nuitrack computes a full skeleton with more than 15 keypoints. In this study, the position of the trunk is used as the reference positions of the patient.

The architecture follows a decentralized scheme in which each robot independently estimates the patient position and executes the control strategy, while the patient acts as the dynamic reference of the system.

Future developments will extend the framework to define each robot position through a distance-angle representation around the patient, allowing a unified controller for the complete formation.

5. Experimental Validation

The following results correspond to single representative trials selected to illustrate the functional integration of the architecture and the consistency of the geometrical model. As a preliminary technical demonstration, the focus is on verifying the qualitative stability, communication robustness via Fast DDS, and temporal alignment through PTP, rather than providing a multi-subject statistical characterization.

This section describes the implementation of the proposed architecture and the experimental conditions used to assess the integration between perception, orientation estimation, and tracking control. Since the system targets clinical environments, the current experimental configuration considers straight-line trajectories to reproduce clinical functional assessments such as the 6-Meter Walk Test (6MWT) and the Timed Up and Go (TUG) test (Podsiadlo and Richardson, 1991).

During these experiments, each robot moves along its longitudinal axis tracking its reference while the relative orientation angle remains constant to establish a performance baseline and isolate the longitudinal response of the control architecture. This controlled setup is a necessary step to validate

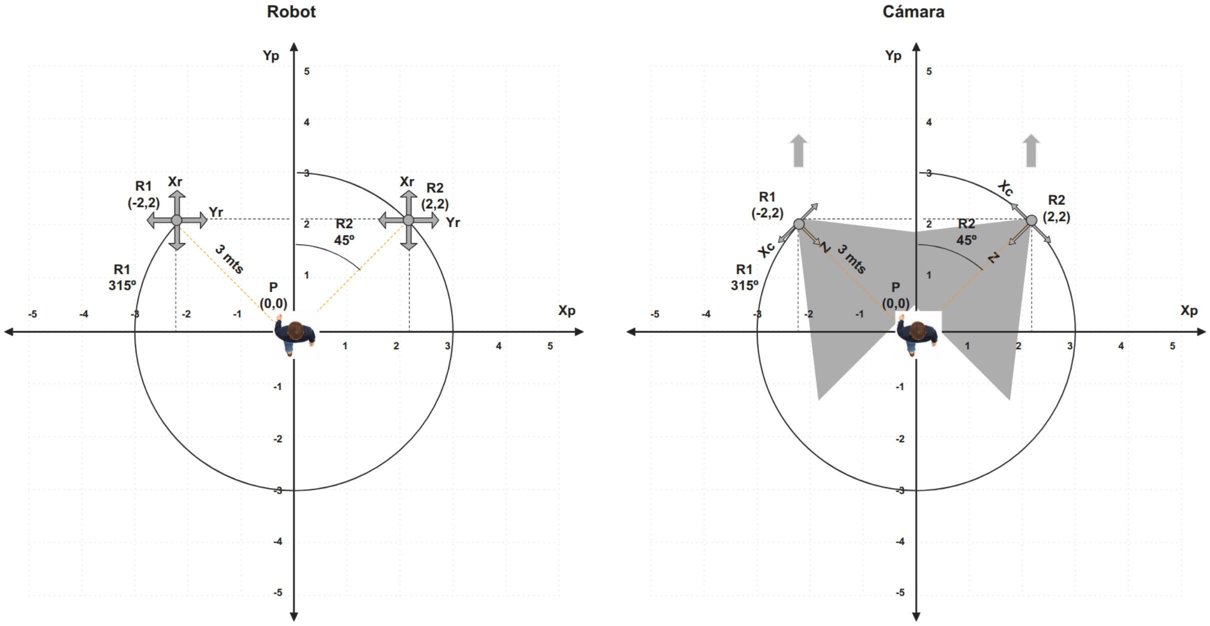

Figura 2: Relative angle estimation through skeletal joints and longitudinal projection.

tracking stability before integrating dynamic orientation estimation in future work. Therefore, each robot follows straight trajectories along its positive longitudinal axis (X_r), tracking the proper reference.

Figure 3 shows an example of the experimental configuration for frontal and lateral tracking during the validation process.

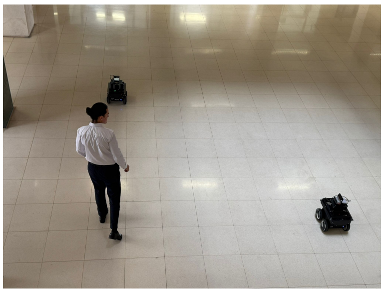

Figura 3: Experimental validation scenario.

5.1. Controller implementation

Each robot independently obtains the patient's skeletal position through Nuitrack, which then is used to generate the control reference used by the control system. Following a sequential validation approach, the robot's θ angle is maintained at a specific setpoint in these preliminary tests. Two different cases are considered in the control strategy. In the frontal and diagonal tests, the distance from the camera to the patient is used as the reference, while in the lateral case, since this distance should not change, the lateral displacement in the camera frame is considered as the reference and, therefore, the robots tries to keep the patient in the center of the field of view of the camera.

In both cases, the tracking controller employs a conventional PID structure with dead-zone and saturation mechanisms to improve stability and prevent oscillatory behavior near the reference distance. Table 1 summarizes the controller parameters used throughout the validation process.

Figure 3 illustrates the validation scenario, in which multiple robotic platforms observe the same patient from different points of view. Although several robots are present in the environment, each platform performs its perception, orientation estimation, and control processes independently according to the decentralized architecture described in Section 3.

Tabla 1: Controller parameters used during experiments

Parameter	Frontal Robot	Lateral Robot
K_p	1.8	0.3
K_i	0.15	0.0
K_d	0.45	2.0
v_{max}	2.5 m/s	
v_{min}	-2.5 m/s	
Dead zone	±0,15 m	

In addition, the architecture incorporates a Recovery state to handle temporary losses of visual information. When skeletal data are unavailable, the robot maintains the last valid control state for a limited interval while the perception module attempts to recover the target. This mechanism prevents abrupt reactions and contributes to safe operation in clinical settings. In case the recovery is not complete, the robot stops, and the experiment is canceled.

5.2. 6-Meter Walk Test (6MWT) Validation

One of the most common functional evaluation tests in clinical environments is the 6-Meter Walk Test (6MWT). This test was used to analyze the behavior of the presented approach using two observation geometries: Frontal 0° and lateral tracking 90°, and diagonal views 45° and 135°.

Figure 4 illustrates a single representative trial for lateral (Robot 1, 90°) and frontal (Robot 2, 0°) follow-up. During the test, both platforms synchronized their velocity smoothly with the participant's pace. For the frontal case, the controller triggered a smooth acceleration maneuver (reaching 1.3 m/s) to adjust the distance from 3.0 m to 2.3 m without abrupt transitions. The success of the tracking is assessed by the absence of oscillations or tracking losses, validating that the decentralized infrastructure supports stable data flow for clinical monitoring.

The convergence of these results into a single stable response demonstrates the effectiveness of the PID-based controller to ensure a safe human–robot interaction.

5.3. *Timed Up and Go (TUG) and Diagonal Performance*

The presented architecture was also tested by emulating the well-known Time Up and Go (TUG) test, which is widely used in clinical settings to assess mobility, balance, and gait performance (Podsiadlo and Richardson, 1991).

In the standard TUG procedure, a patient begins in a seated position, stands up, walks forward, performs a 180° turn, returns to the starting location, and sits down again, at the highest possible speed. In the present study, the participant started from a standing position; therefore, no chair was incorporated into the experimental setup.

Furthermore, the participant was a healthy adult with no known biomechanical impairments. Consequently, the observed gait pattern represents an ideal reference scenario, rather than the movement characteristics typically exhibited by clinical populations.

As shown in Figure 5, a single representative trial verifies the diagonal follow-up (45° and 315°). The robots successfully adapted to the 180° turn and the subsequent backward motion phase. The PID controllers maintained the 3-meter reference distance throughout the maneuver without introducing significant oscillations or tracking errors. This demonstrates the functional stability of the RoboMesh architecture during dynamic clinical maneuvers where the patient acts as the primary reference.

6. Conclusions

This work presents RoboMesh, a decentralized multi-robot architecture for cooperative perception and clinical gait analysis. The proposed framework establishes a solid technological foundation to address limitations of RGB-D based gait analysis, such as temporary occlusions. Employing multiple mobile platforms equipped with RGB-D sensors, the system expands spatial coverage and provides complementary perspectives of the patient, increasing the potential robustness of observation in clinical environments.

The architecture follows a decentralized approach in which each robot independently estimates the relative position of the patient and generates its own control actions, using the patient as the primary dynamic reference of the system, without relying on a central coordinator. This design improves the scalability, fault tolerance, and future integration of additional robotic platforms within the cooperative network.

Experiments based on representative clinical assessment trajectories verify the functional integration of perception and control modules. Furthermore, the frontal, lateral and diagonal observation configurations demonstrated the ability of the proposed model to maintain a consistent representation of the relative position of the patient in different geometries.

Overall, the results support the feasibility of the proposed architecture as a foundation for cooperative gait-analysis systems. The platform establishes a solid technological framework for future developments that involve automatic orientation estimation, cooperative robot triangulation, and shared localization strategies aimed at improving the accuracy and robustness of clinical monitoring applications.

Future evaluations will focus on dynamic scenarios where the patient performs unconstrained movements, requiring the system to automatically estimate and adapt to varying θ angles in real-time.

Acknowledgements

This work was funded by the 'Project PID2023-150967OB-I00, financed by MICIU/AEI/10.13039/501100011033 and by FEDER, EU.', and by the Madrid Government (Comunidad de Madrid-Spain) under the Multiannual Agreement 2023-2026 with Universidad Politécnica de Madrid in the Line A, Emerging PhD researchers.

Referencias

Álvarez, D., Guffanti, D., Brunete, A., Hernando, M., Gambao, E., 2023. RoboGait: un sistema robótico no invasivo para el análisis de la marcha humana. Revista Iberoamericana de Automática e Informática Industrial 21 (2), 137–147.
DOI: 10.4995/riai.2023.20066

Bonanno, M., De Pasquale, P., Fonti, B., Gjonaj, E., De Salvo, S., Quartarone, A., Calabrò, R. S., 2025. Neural control meets biomechanics in the motor assessment of neurological disorders: A narrative review. Frontiers in Neural Circuits 19.
DOI: 10.3389/fncir.2025.1608328

Bravo, M., Álvarez, D., 2016. Comparison of two Motion Capture Systems by means of Joint Trajectories of Human Gait. Revista Mexicana de Ingeniería Biomédica 37 (2), 149–160.
DOI: 10.17488/RMIB.37.2.2

Chen, W., Chi, W., Ji, S., Ye, H., Liu, J., Jia, Y., Yu, J., Cheng, J., 2025. A survey of autonomous robots and multi-robot navigation: Perception, planning and collaboration. Biomimetic Intelligence and Robotics 5 (2), 100203.
DOI: 10.1016/j.birob.2024.100203

Chovet, L. P., Garcia, G. M., Bera, A., Richard, A., Yoshida, K., Olivares-Mendez, M. A., 2025. Performance comparison of ros2 middlewares for multi-robot mesh networks in planetary exploration. Journal of Intelligent & Robotic Systems 111, 18.
DOI: 10.1007/s10846-024-02211-2

Dahiya, A., Aroyo, A. M., Dautenhahn, K., Smith, S. L., 2023. A survey of multi-agent human–robot interaction systems. Robotics and Autonomous Systems 161.
DOI: 10.1016/j.robot.2022.104335

Fareh, R., Baziyad, M., Khadraoui, S., Brahmi, B., Bettayeb, M., 2023. Logarithmic potential field: A new leader–follower robotic control mechanism to enhance the execution speed and safety attributes. IEEE Access 11, 85451–85466.
DOI: 10.1109/ACCESS.2023.3303873

Hu, J., Bhowmick, P., Jang, I., Arvin, F., Lanzon, A., 2021. A Decentralized Cluster Formation Containment Framework for Multirobot Systems. IEEE Transactions on Robotics 37 (6), 1936–1955.
DOI: 10.1109/TRO.2021.3071615

Olfati-Saber, R., Murray, R. M., 2004. Consensus Problems in Networks of Agents with Switching Topology and Time-Delays. IEEE Transactions on Automatic Control 49 (9), 1520–1533.
DOI: 10.1109/TAC.2004.834113

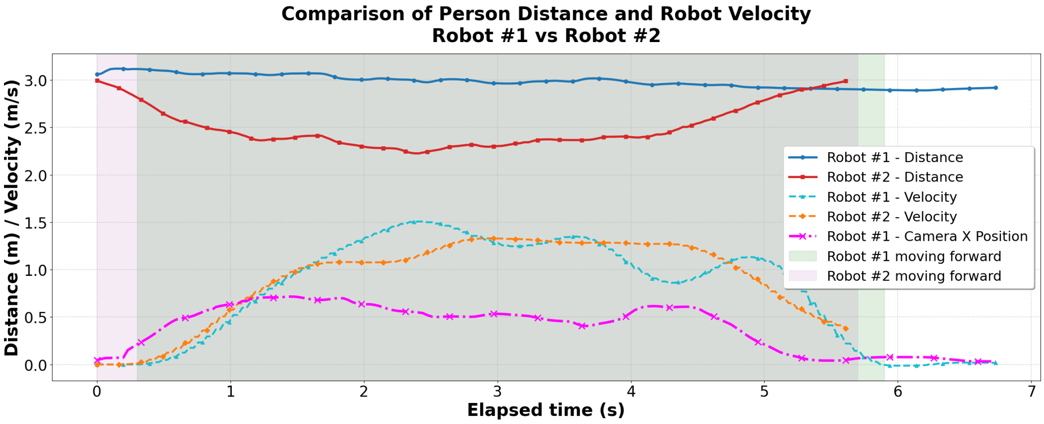

Figura 4: 6MWT tracking performance: robot 1 at 90° (lateral) and robot 2 at 0° (frontal).

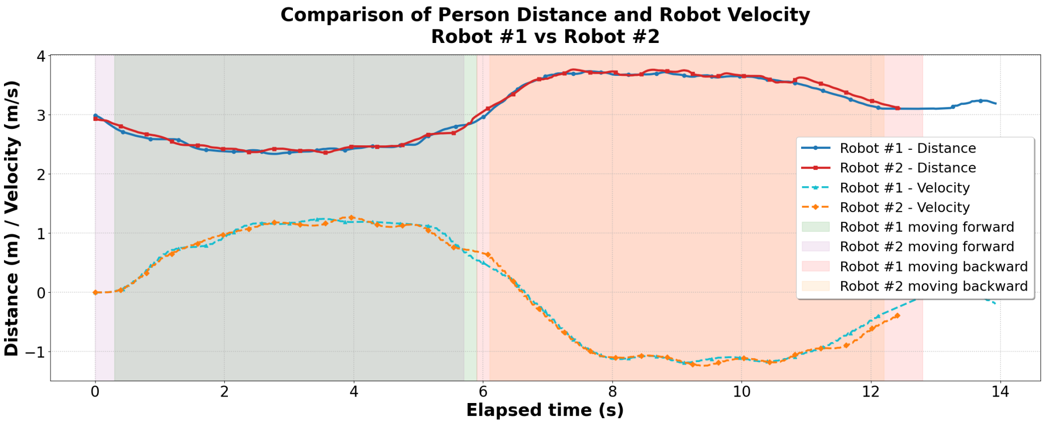

Figura 5: Diagonal Frontal tracking performance and approach maneuver (45° and 315º configuration).

Pfister, A., West, A. M., Bronner, S., Noah, J. A., 2014. Comparative abilities of Microsoft Kinect and Vicon 3D motion capture for gait analysis. Journal of Medical Engineering & Technology 38 (5), 274–280.
DOI: 10.3109/03091902.2014.909540

Podsiadlo, D., Richardson, S., 1991. The Timed "Up & Go": A test of basic functional mobility for frail elderly persons. Journal of the American Geriatrics Society 39 (2), 142–148.
DOI: 10.1111/j.1532-5415.1991.tb01616.x

Scataglini, S., Abts, E., Van Bocxlaer, C., Van Den Bussche, M., Meletani, S., Truijen, S., 2024. Accuracy, validity, and reliability of markerless camera-based 3D motion capture systems versus marker-based 3D motion capture systems in gait analysis: A systematic review and meta-analysis. Sensors 24 (11), 3686.
DOI: 10.3390/s24113686

Spletzer, J., Das, A. K., Fierro, R., Taylor, C. J., Kumar, V., Ostrowski, J. P., 2001. Cooperative localization and control for multi-robot manipulation. In: IEEE/RSJ International Conference on Intelligent Robots and Systems. pp. 631–636.
DOI: 10.1109/IROS.2001.976240

Talebpour, Z., Martinoli, A., 2018. Multi-robot coordination in dynamic environments shared with humans. Proceedings of the IEEE International Conference on Robotics and Automation, 4593–4600.
DOI: 10.1109/ICRA.2018.8460978

Wang, S., Wang, Y., Li, D., Zhao, Q., 2023. Distributed Relative Localization Algorithms for Multi-Robot Networks: A survey. Sensors 23 (5), 2399.
DOI: 10.3390/s23052399