The Perception and Manipulation Group (P&M) (http://www.iri.upc.edu/research/perception) is a research group from the Institut de Robòtica i Informàtica Industrial (IRI), CSIC-UPC, a joint University Research Institute participated by the Spanish National Research Council (CSIC) and the Technical University of Catalonia (UPC). The P&M Group is one IRI's four research groups and its members belong to the Consolidated Research Group RobIRI (2021 SGR 00514) recognized by the Government of Catalonia.

In September 2024, the group is formed by the following members:
Permanent Research Staff and Faculty:
Carme Torras Genís (Group Leader, Research Professor CSIC), Antonio Agudo Martínez (Tenured Scientist CSIC), Guillem Alenyà Ribas (IRI Director, Research Scientist CSIC), Maria Alberich Carramiñana (Full Professor UPC), Júlia Borràs Sol (IRI Deputy Director, Tenured Scientist CSIC), Adrià Colomé Figueras (Tenured Scientist CSIC), Maria Dimiccoli (Tenured Scientist CSIC), Sergi Foix (Tenured Scientist CSIC), Pablo Jiménez Schlegl (Research Scientist CSIC), Jordi Sánchez Riera (Research Scientist CSIC).

RESEARCH ACTIVITIES: The research of the P&M group focuses on enhancing the perception, learning, and planning capabilities of robots to achieve higher degrees of autonomy and user-friendliness during everyday manipulation tasks for assistive, service, and collaborative robotics. Domestic, service and industrial robotics require the easy acquisition of 3D object models and user-friendly programming of manipulation tasks. Therefore, efforts are oriented toward 3D active perception combining several sensory modalities, such as colour vision, depth from time-of-flight, structured light, stereo, and haptics; teaching manipulation skills to a robot under a learning-by-demonstration approach, as well as teaching the robot to accomplish tasks requiring planning through rule learning; and computer vision and deep learning techniques for the detection of human body configurations and recognition of user actions and intentions.

The main areas of interest are:
Learning by demonstration: We devise methods to learn object-action relations to accomplish tasks at different levels of abstraction, where object models are generated from visual and depth information, and actions, involving manipulation skills, are learned from demonstrations provided by a human using multimodal algorithms that combine vision and haptics.
Planning for perception and manipulation: We are interested in planning algorithms for object modelling, with a special focus on deformable objects. High-level task formulations are integrated with low-level geometry-based methods and simplified physical models, as well as with an online sensory-based treatment of uncertainty so as to come up with specific sequences of motion commands.
Camera pose estimation and reconstruction of rigid and non-rigid objects: We develop solutions to recover the camera pose and to reconstruct the shape of deformable objects from monocular images and video sequences. For the pose estimation problem, we explore algebraic methods that can be efficiently solved and can handle corrupted observations. For deformable shape estimation, we explore both solutions based on statistical methods and techniques that incorporate physical priors. In the latter case, we also develop approaches to estimate material properties from video.
High-level scene interpretation: The goal is to exploit recent advances in Deep Learning to perform scene interpretation considering several input domains: only RGB images/video, RGB+depth Images, and multimodal data including image and text.

RESEARCH INFRASTRUCTURE: The Perception & Manipulation group has a well-equipped robotic lab including a life-scale mock-up of a small apartment. Two PAL TIAGo mobile robots dwell within the apartment, that occasionally are taken out for experimentation elsewhere. This area is also equipped with the robots ARI, Stretch, Temi and Obi. There is also a manipulation area with two WAM robot arms, two UR5e arms and two KINOVA manipulators. It is further equipped with commercial and self-developed grippers, sensing devices, motion capture and augmented reality systems. Lab service offers quick experimental setup, several standardised software tools, GPU servers, and expertise in robot control and perception algorithms.

SCIENTIFIC PUBLICATIONS: During the period 2019-2023, the group has published more than 60 journal articles, 100 publications in conferences and 15 books and book chapters. Below, we list a selection of the publications with the highest impact in the most important journals and conferences in our fields.

General Objective of the Project:

The overarching aim of this project is to advance the field of 3D human digitization, interaction modeling, and related technologies. Through cutting-edge research and innovation, we seek to contribute to the development of highly realistic and personalized avatars with applications spanning virtual reality, entertainment, e-commerce, healthcare, and fitness monitoring.

Specific Objectives:

1. Enhance 3D Human Digitization: Develop novel algorithms that enable high-fidelity 3D human digitization from RGB images, particularly in non-controlled settings. Address challenges related to noisy observations, non-calibrated cameras, and diverse body poses to improve robustness in real- world scenarios.
2. Bridging the Gap in 3D-Aware Editing Tools: Integrate diffusion models and advance 3D-aware editing tools to empower content creators in the entertainment industry. Enable seamless application of these tools to human avatars for more realistic and captivating storytelling experiences.
3. Learning Precise Human Dynamics: Pioneer algorithms for learning individual-specific traits beyond dynamics, focusing on facial features, body structures, and other distinctive attributes. Ensure that 3D avatars accurately represent the unique nuances of each individual.
4. Modeling Interaction with the Environment: Investigate and model a wide range of avatar interactions with the environment, including changes influenced by environmental lighting, geometric interactions with tools and clothing, and interactions between different avatars in coordinated tasks or group activities.
5. Addressing Uncertainty in Modeling: Develop neural models specifically designed to incorporate and manage uncertainty associated with predictions, particularly when modeling avatars in diverse and less controlled real-world settings.
6. Resource-Efficient Algorithms for Accessibility: Devise algorithms that demand minimal data and computational resources, addressing challenges posed by data scarcity and limited computational capabilities. Aim to democratize advanced technologies by making them more accessible, particularly in academic and resource-constrained contexts.