EU Horizon 2020 Research & Innovation Program
Advanced Robot Capabilities & Take-Up
ICT-25-2016-2017
Grant Agreement No. 779942

CROWDBOT

Safe Robot Navigation in Dense Crowds

http://www.crowdbot.org

Technical Report

D 5.1: System Architecture

Work Package 5 (WP 5)
System Architecture & Integration

Task Lead: INRIA, France
WP Lead: SoftBank Robotics Europe, France

Table of Contents

Executive Summary

The main objective of the CROWDBOT project is safe navigation of mobile robots in dense crowds. Operating a robot safely in a social setting requires overcoming current technical limitations and acquiring human-like intelligence as well as addressing ethical and safety concerns of machines co-existing among human crowds. The scope of this report is limited to the former: technical description of various sub-systems and intelligent modules, and when pieced together, forms a self-reliant, autonomous machine—i.e. the robotic system architecture.

Each CROWDBOT robotic system test requires 1) a mobile robot (i.e. "the bot"), 2) human participants (i.e. "the crowd") and 3) the physical venue (i.e. "the environment") in which the test takes place. Three different types of robots (a humanoid, a smart wheelchair and a futuristic transport machine) have been chosen due to their unique features and significant functional differences amongst them.

The Pepper robot [1] by SoftBank Robotics Europe is a humanoid robot that can be programmed for fully autonomous mode without human intervention. Due to its commercial success, Pepper is the only CROWDBOT robot that has interacted with humans in various real-world social environments. The smart wheelchair [2] prototype is a conventional motorized wheelchair with additional smart technology features. Compared to Pepper, the smart wheelchair has an open chassis for attachment and integration of additional sensory and processing modules and can move at higher speeds [5]. In terms of social context, operator-driven motorized wheelchairs are commonly found in public spaces and accepted by humans as part of the social fabric. The cuyBot [3], under development by Locomotec, is a self-contained differentially driven twin-wheel machine with integrated electronics. It is suitable for transporting both humans (similar to a wheelchair) and goods as a delivery service machine. Compared to the other two, the cuyBot is the most flexible option for augmenting the machine with the latest advances in electronics and computer technologies.

Key software-based technological innovations of CROWDBOT for safe navigation are the four functional blocks: 1) "tracker" (in charge of identifying and tracking of individuals and groups of people and obstacles around its sensory field of view, 2) "navigator" (responsible for mapping the physical space around the robot and localizing its position on the map and developing a navigation scheme for future movement, 3) "crowd predictor" (an intelligent machine that estimates crowd behavior (e.g. future motion aspects) based on past and current information (as supplied by the tracker and navigator) and by offline knowledge of crowd behavior, and 4) "decision-maker" (final authority in robotic motion control and interaction with a human controller). Further details of these functional blocks as well as their mutual interconnections are provided in the body using both text-based descriptions and graphical illustrations.

1. System Description

The main objective of the CROWDBOT project is safe navigation of mobile robots in dense crowds. Operating a robot safely in a social setting requires overcoming current technical limitations and acquiring human-like intelligence as well as addressing ethical and safety concerns of machines co-existing among human crowds. The scope of this report is limited to the former: technical description of various sub-systems and intelligent modules of a typical CROWDBOT robot and an integrated layout of such sub-systems and modules, when pieced together, forms a self-reliant, autonomous machine—i.e. the robotic system architecture.

In a typical CROWDBOT robotic test, a single robot navigates and traverses from its start position until it reaches the destination, and along its chosen paths, it interacts with humans it encounters. Here, the word "interact" implies 1) give-and-take coordination with humans to avoid collisions, 2) if touching is involved, the use of appropriate force or pressure exertion to not injure humans and 3) the ability not to block or impede the flow of human traffic. Three different types of robots (a humanoid, a smart wheelchair and a futuristic transport machine) have been chosen due to their unique features and significant functional differences amongst them. Details of CROWDBOT robots are elaborated in Section 2.1.

Each CROWDBOT robotic system test requires 1) a mobile robot (i.e. "the bot"), 2) human participants (i.e. "the crowd") and 3) the physical venue (i.e. "the environment") in which the test takes place. A suitable robot (one from the above list of three) is chosen for each test based on its unique features, technical capabilities and anticipated social role.

In this report, we provide a high-level system description of CROWDBOT robots and their integrated system architecture. We also provide high-level characterization of human crowds and social environments in Sections 2.2 and 2.3 to facilitate our discussion on different sub-system modules that constitute the overall system architecture.

1.1 Robots

Three different types of mobile robots have been selected for their complementary attributes and availability to the team.

1) Commercially available robot (a humanoid robot developed by Softbank Robotics)
2) Smart Wheelchair (a motorized wheelchair prototype with Artificial Intelligence features)
3) Smart-wheel robot (a highly customizable twin-wheel machine)

In principle, any robot that meets baseline requirements of a CROWDBOT robot prototype can be adapted and used as a qualifying candidate in any of CROWDBOT robot navigation tests.

Pictures of the Pepper humanoid robot and the smart wheelchair prototype (developed by University College London [3]) are shown in Figure 1.

The Pepper robot [1] is a humanoid robot with a body similar to that of a human child of height ~ 120 cm and weight ~ 28 kg. It is holonomic and moves (both translationally and rotationally) via three multi-directional wheels at a maximum speed of 3 km/hr. It is equipped with a number of imagery, range, depth and touch sensors and processing engines for interacting with humans and navigating through obstacles without colliding or tipping over. Pepper can be programmed to operate in fully autonomous mode without human intervention. Due to its commercial success, Pepper is the only CROWDBOT robot that has interacted with humans in various real-world social environments. Being a humanoid, it is also capable of body and limb movements. Additionally, Pepper is equipped with full-spectrum Light-Emitting Diodes (LEDs), speakers and microphones for immersive interaction and voice communications with humans.

Source: SBR [1]

Source: UCL [3]

(a) Pepper interacting with humans (b) Smart wheelchair prototype

Figure 1: CROWDBOT robots: Pepper robot and Smart wheelchair

The smart wheelchair [2] prototype is a conventional motorized wheelchair with additional smart technology features embedded into the platform. As shown in Figure 1(b), the range sensors on its front side as well as at the tip of foot rests are visible. Additional sensors and computational modules are also strategically placed in the rear part of the wheelchair. The smart wheelchair operates in semi-autonomous or "shared control" mode where the human operator has the option to over-ride machine commands and manually control the wheelchair motion. A human operator maneuvers the wheelchair via a small joystick mounted on the armrest (see Fig 1(b)) or on the upper rear part of the frame. For operators who cannot manage a manual joystick, other available options are head switches, chin-operated joysticks and sip-and-puff controllers [6]. Recent advancements [7] may allow the operator to control the wheelchair via brain impulses. Compared to Pepper, the motorized wheelchair is non-holonomic, has an open chassis for attachment and integration of additional sensory and processing modules and can move at higher speeds of up to 6 km/hr on pavement and 12 km/hr on roads [5]. It is also much heavier: common motorized wheelchairs weigh more than 60 kg. Total gross weight (wheelchair plus human operator) is generally in excess of 120 kg. In terms of social context, operator-driven motorized wheelchairs are commonly found in public spaces and accepted by humans as part of the social fabric.

The cuyBot [3], currently under development by Locomotec, is a highly customizable robotic platform based on the concept of a smart wheel─a self-contained differentially driven twin-wheel machine with integrated electronics. It can support a payload weight of more than 125 kg and move at a maximum speed of 18 km/hr. Thus it is suitable for transporting both humans (similar to a wheelchair) and goods as a delivery service machine. Compared to Pepper and the wheelchair, the cuyBot is the most flexible option for augmenting existing robotic capabilities with the latest advances in electronics and computer technologies.

1.2 Human Crowds

A concise description is a human crowd generally requires elaboration and quantization of crowd size (head count), motion profile (speed, acceleration, direction), association (individual, pair, group) and mission or goal (e.g. marching in a demonstration, exiting a theatre). Here we describe a high-level characterization of crowd behavior as perceived by a robot:

- Motion: From a robot's perspective (as viewed via its camera "eyes"), it can only see, identify and track the position and movement of humans in its field of view (i.e. its vicinity). It perceives humans (both individuals and groups) as
  - Stationary (in absolute sense; zero translational and rotational motion)
  - Moving Along (i.e. human and robot motions in the same direction)
  - Moving Against (i.e. human and robot motion in opposition directions)
  - Moving Across (i.e. human and robot motion vectors at an angle)
  - Moving Randomly (i.e. Brownian motion type, direction-less movement)

  Crowd motion with respect to that of a robot is labeled as 1-dimensional (1D flow), 2-dimensional (2D flow) or cross-dimensional (Cross flow). All three cases are illustrated in Figure 2.

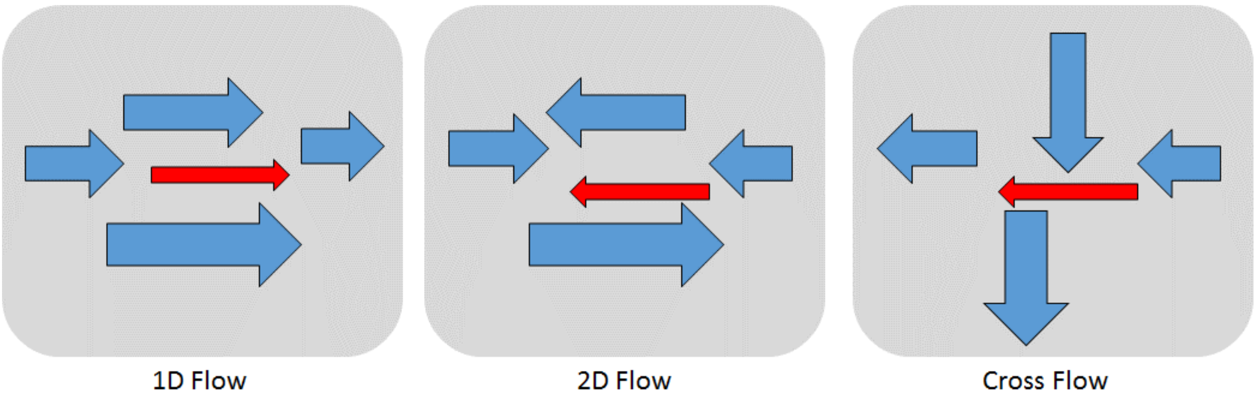

Figure 2: Various human motion models with respect to robot's movement (robot motion depicted using a red arrow and various crowd motions via blue arrows)

- Size: Concise definitions of crowd size and density measure will be provided in a future deliverable D1.1 (due end of June 2018); here, we use a simple parametric formula: From a robot's perspective (via imagery and range sensors), the human crowd in its vicinity (e.g. 360° field-of-view and 1-meter radius) is dense if the number of human body count detected surpasses a threshold (say, 4).
- Speed: In all CROWDBOT tests, human crowd speed is limited to that of pedestrians (~ 5 km/hr [4]) or less. This requirement is mainly due to self-imposed safety and precautionary measures and maximum speed limits of the robots. Hence, all test participants (both robot and humans) can vary their speeds between zero and the maximum speed limit. Each participant's variable speed profile at a particular time and space is highly dependent on its pre-defined mission objectives (see next bullet) and the collective behavior of the remaining crowd.

- Mission/Goal: Just as a robot is programmed to carry out a specific task (e.g. escort a hospital patient, deliver a package), all human participants in a CROWDBOT test are also assigned specific tasks—also known as missions or goals. A thorough discussion of this subject is outside the scope of this report but we provide a brief outline here: Due to spatial and temporal aspects of motion for both human crowd and robot, a mission is defined in terms of a sequence of goals where each is marked by spatial (e.g. start-stop markers, see next section for "entry-exit points") and temporal (e.g. elapsed time from start of mission or clock time-stamp) parameters. During the course of the test, each agent (human or robot) follows its assigned mission plan (a sequence of goals), which is defined from the starting location and time until the finish point and time.

1.3 Test Environment

Real-world social environments such as pubic parks, train stations and sporting venues are highly dynamic spaces and are thus difficult to articulate using mathematical models due to both spatial and temporal variations arising from both human and non-human object motion. Fortunately, based on practical considerations, the CROWDBOT team has introduced a number of requirements to simplify robotic navigation tests. They are:

- Indoor Controlled Space: Due to ethical and safety concerns, all CROWDBOT robotic tests will be held at a controlled venue such as a gymnasium, storage warehouse or university laboratory/office where the test area can be cordoned off and entrance/exit points can be manned. The indoor space requirement is due to multiple concerns: 1) privacy of human participants and team members during a test, 2) robot motion constraints (e.g. Pepper may trip and fall on a rough outdoor surface) and 3) inclement weather conditions.

- Well-defined and bounded physical space: The test environment roughly conforms to one of the following geometrical shapes:
  - rectangular area (area dimension and border extents defined via physical walls or delimited via traffic cones),
  - tubular/tunnel area (where human and robot motions are confined to the space between two physical walls or traffic cones),
  - a bend (similar to a tunnel but with an abrupt change in direction)
  - an intersection (two tunnels merging at a right angle)

  Illustrations of envisioned test spaces are shown in Figure 3.

- Strategically placed objects: Objects (building structures such as columns and doors as well as furniture items such as tables and chairs) are placed in a well-defined physical space such that the resulting test environment mimics the spatial properties of a real-world social environment. Examples are illustrated in Figure 4.

- Well-defined entry and exit points: The entry ("start") and exit ("finish") point pair is defined for each agent (the robot or human person/crowd) as a critical part of its overall mission/goal profile. In the example shown in Figure 5 [9], the robotic wheelchair moves from its start position in a room, out the door to another room through a corridor and then back to its original position via the same corridor, and finally entering from the adjacent room. A single entry-exit point set can be assigned for the entire test or multiple, segmented entry-exit point sets can be set up to test out different areas of the environment. For example, in Figure 5, we can run several segmented tests: one from the initial start to the first door, another from the corridor to an adjacent room and then on to the final position.

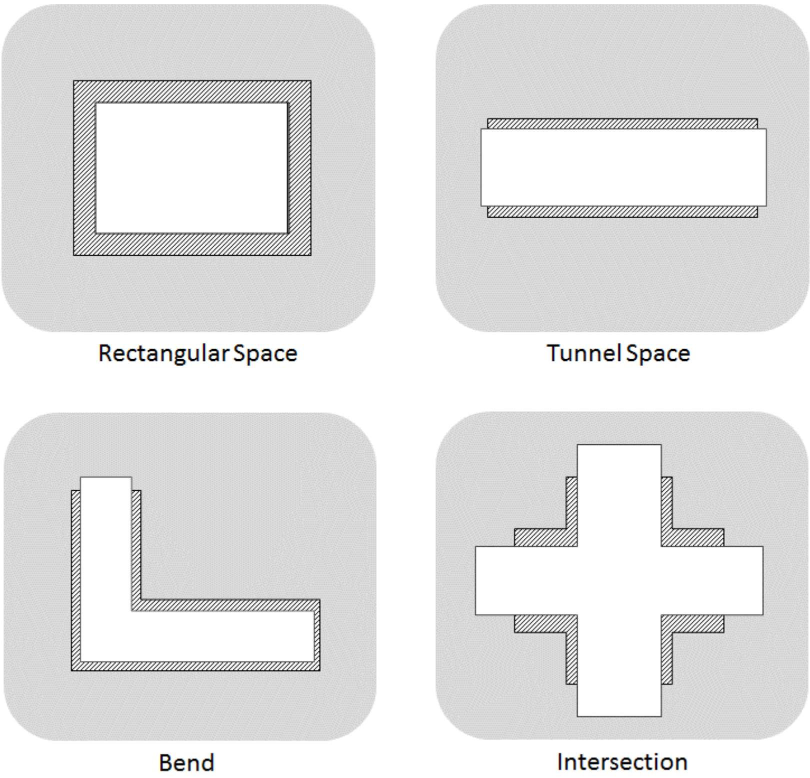

Figure 3: Examples of well-defined and bounded physical spaces (top view)

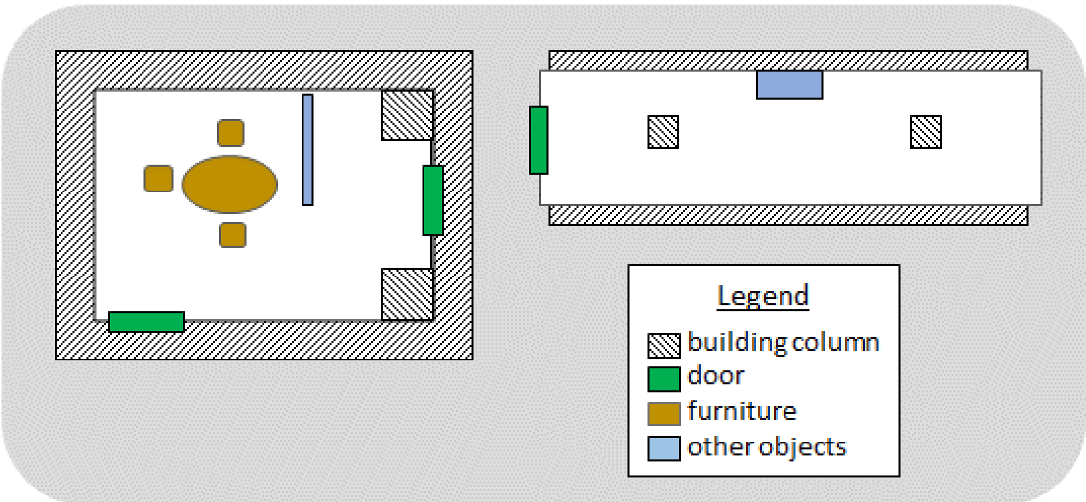

Figure 4: Examples of strategically placed objects in physical spaces (top view)

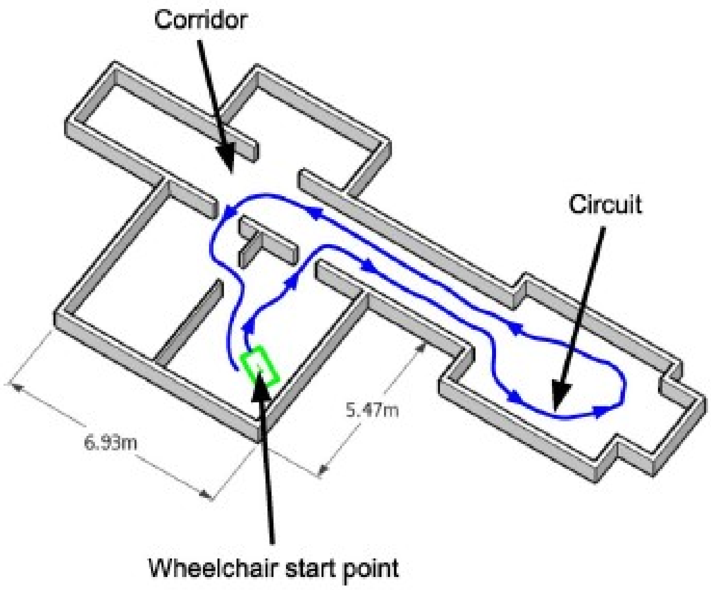

Figure 5: Example of well-defined entry and exit points (isometric view). The green rectangular block indicates both the start and finish points of the test vehicle—in this case, the robotic wheelchair. Source: [9]

In a majority of cases, a social environment can be modeled using a combination of test environmental features. This is the case in Figure 6 for the robotic wheelchair office environment of Figure 5 [9]. A rectangular space A with blue object partition is used to represent two adjacent rooms. The corridor is denoted via a tunnel space C and the larger room is modeled as rectangular space B. Doors are placed at strategic locations using green door objects. The rectangular building structure objects in B complete the corner layout of the larger room. A red triangular icon is used to denote the entry-exit point.

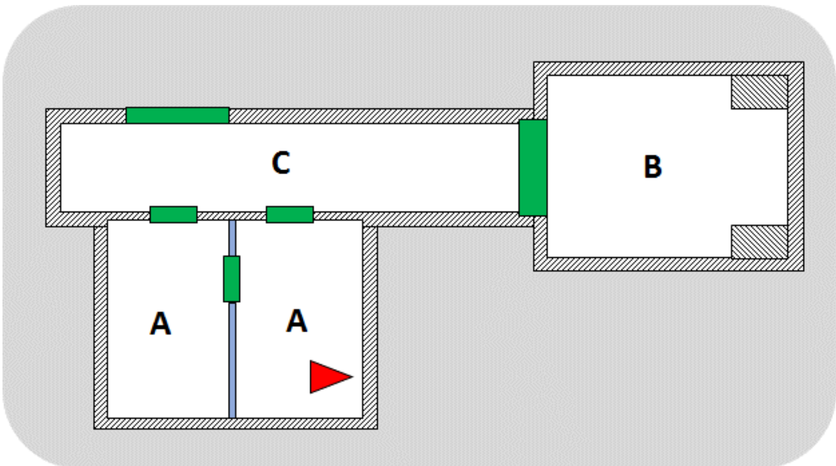

Figure 6: Representation of a social environment using a variety of test environment constructs: The office environment of Figure 5 is represented here using rectangular and tunnel spaces and door objects (top view)

2. System Architecture

For modeling, test and evaluation purposes, a mobile robot can be classified in terms of its

1. modular composition (hardware components and software modules)
2. main functional blocks (sensory information capture, data processing & interpretation, decision-making and motion control)

We first start with the modular composition as shown in Figure 7. For both hardware and software, six different high-level sub-systems are identified. For example, the power electronics sub-system may consist of an energy source (a battery or electrical cable to wall outlet), power conditioning module (AC-to-DC conversion and voltage regulation to supply power to motors and digital electronics) and electronic drivers and actuators for motor control. Similarly, a human-machine interaction sub-system may contain various software modules: communication protocols for remote control, text/voice recognition and interpretation, text-to-speech translation, voice synthesizer, etc.

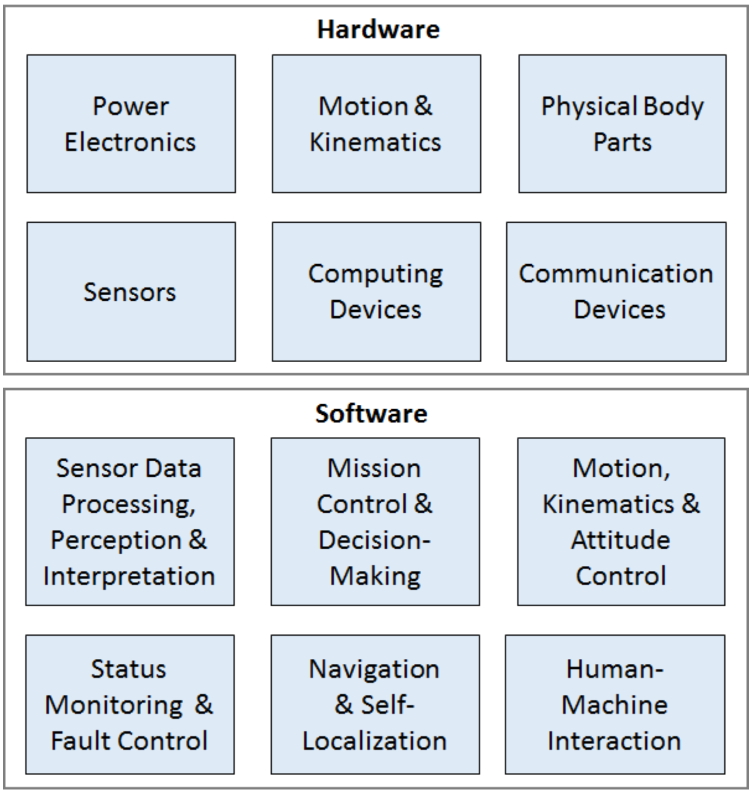

Figure 7: High-level association of hardware and software sub-systems of a mobile robot

The sub-system list in Figure 7 is not exhaustive; some robot platforms may have additional features not listed. On the other hand, some robot models may be equipped with fewer sub-systems than those listed in Figure 7. Furthermore, component details of a sub-system may vary greatly from one mobile robot platform to another. For example, in the sensor hardware sub-system, a plethora of options exist for color imagery, depth and range finding, contact and pressure sensing, inertial measurements, etc. Hence, the hardware-software component description of a mobile robot is one convenient option for comparing features and capabilities of two different CROWDBOT robots or between a CROWDBOT robot and an external model. Such an approach becomes useful when CROWDBOT robots evolve or when CROWDBOT technologies are transferred to another external platform. In Figure 8, we use the Pepper robot as an example to highlight component-level details of its two hardware sub-systems. Similarly, in Figure 9, we present details of a smart wheelchair prototype as tested by researchers at University College, London [2]. The wheelchair is equipped with a number of electronic sub-systems: 1) a variety of sensors ─sonar, scanning laser─ to detect nearby objects (and thus avoid collisions), 2) an optical camera above the headrest for self-localization, 3) an onboard computer underneath the seat and 4) a number of communication devices (joystick, joypad, tablet and eye-tracking system) for human-machine interfacing and interaction.

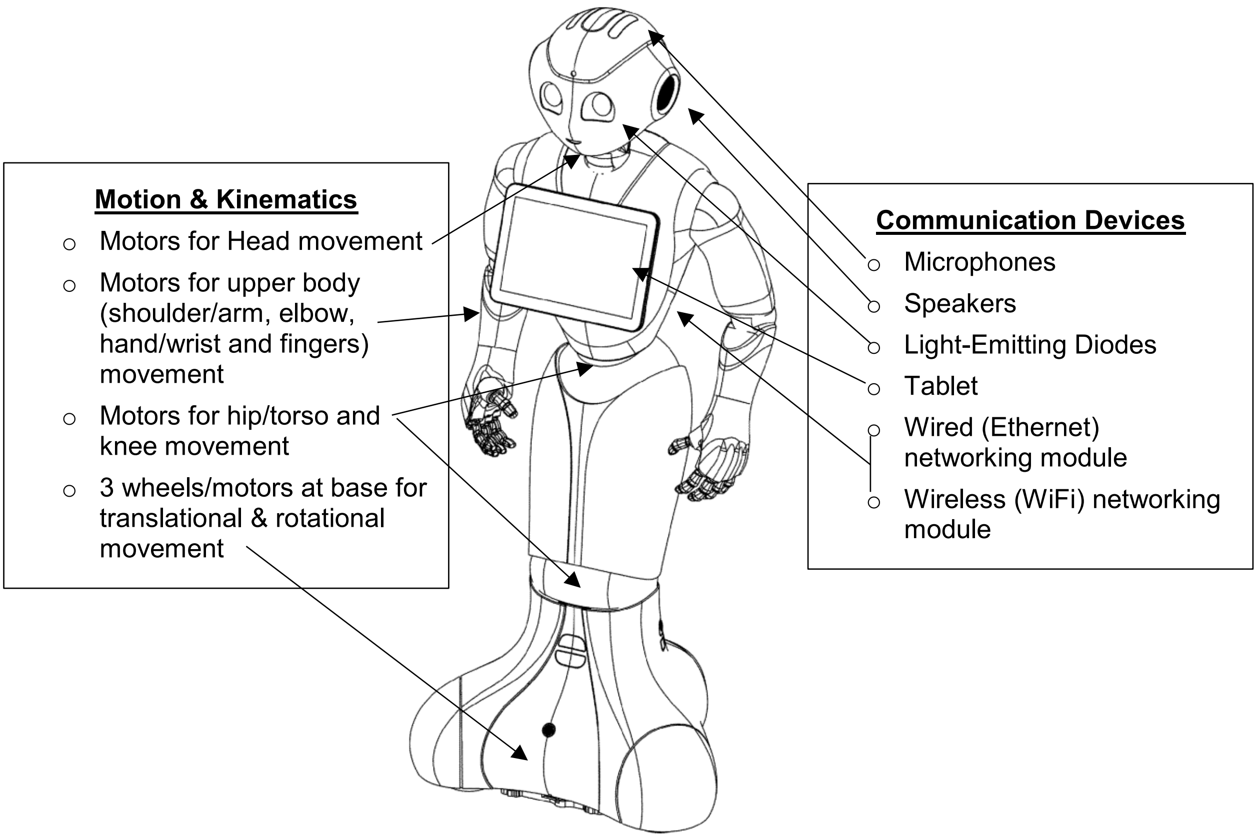

Figure 8: Details of motion/kinematics and communication devices hardware sub-systems of the Pepper robot. Source: [8]

An alternative description of a mobile robotic system is in terms of its key functional blocks. This paradigm is better suited for delineating new machine-intelligent technologies being developed by the CROWDBOT team.

As shown in Figure 10, the CROWDBOT robot gathers information about its immediate environment (crowds, objects, physical space) from its embedded and attached sensors. Various types of commonly used sensors are listed in the "Sensors" functional block. Next, these raw sensor data undergo further processing; for example, the optical vision sensor output is used to construct video images at a certain frame rate and the resulting file is stored in memory. Another functional block "Integration & Perception" is used to convert and combine sensor data into a format suitable for machine interpretation.

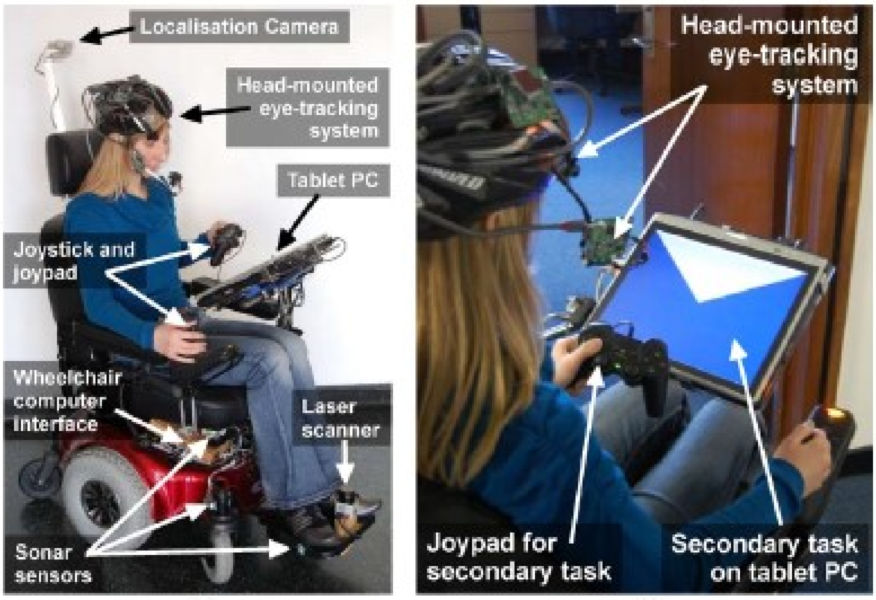

Figure 9: Details of sensors (sonar, laser and camera) and communication devices (joystick, joypad and tablet) hardware sub-systems of a robotic wheelchair. Source: [9]

The next four functional blocks represent key technological innovations of CROWDBOT for safe navigation. The "tracking" block is in charge of identifying and tracking of individuals and groups of people close to the robot as well as other obstacles within its sensory field of view (i.e. camera sensors). Since the robot itself is moving, static objects also undergo displacement from one video frame to another. The "mapping, localization and navigation" block is responsible for several functions: 1) it has to map the static physical space around the robot (dimension of the space, location of objects, walls, doors, etc.), 2) It has to localize the position of the robot within its map, and 3) it must develop a navigation scheme for future movement using inputs from the "tracker" and the "predictor". The "crowd predictor" is an intelligent machine that estimates crowd behavior (e.g. future motion aspects) based on past and current information (as supplied by the tracker and navigator) and by offline knowledge of crowd behavior. The last functional block is the "decision-maker" which is the final authority in robotic control. Its sub-modules are 1) motion control, which initiates commands to steer and control the robot's motors and other devices (such as LEDs, voice synthesizer, alarm sounding, etc.) and 2) human interfacing (via a joystick/joypad, touchpad/tablet or a remote control console). In a semi-autonomous wheelchair, the human operator has communication devices such as a joystick and a kill pushbutton to manually override robotic decision-making. Likewise, manual override capabilities of a robot are available under remote control by a human.

Figure 11 provides a more granular view of the intelligent functional blocks detailed in Figure 10. Each functional block is also labeled with a Work Package (WP) number that identifies the association of this particular work product within the CROWDBOT project. Each arrow indicator (with a corresponding letter label) is used to denote information flow from a functional block to another. Hence, each functional block is a "black box" with well-defined inputs and outputs. This modular paradigm allows each software developer/engineer to develop his or her own black box separately for a given input-output set and later execute component/block level tests independently. After several blocks are developed, the team can run a more integrated sub-system test, and then followed by a complete integrated system test.

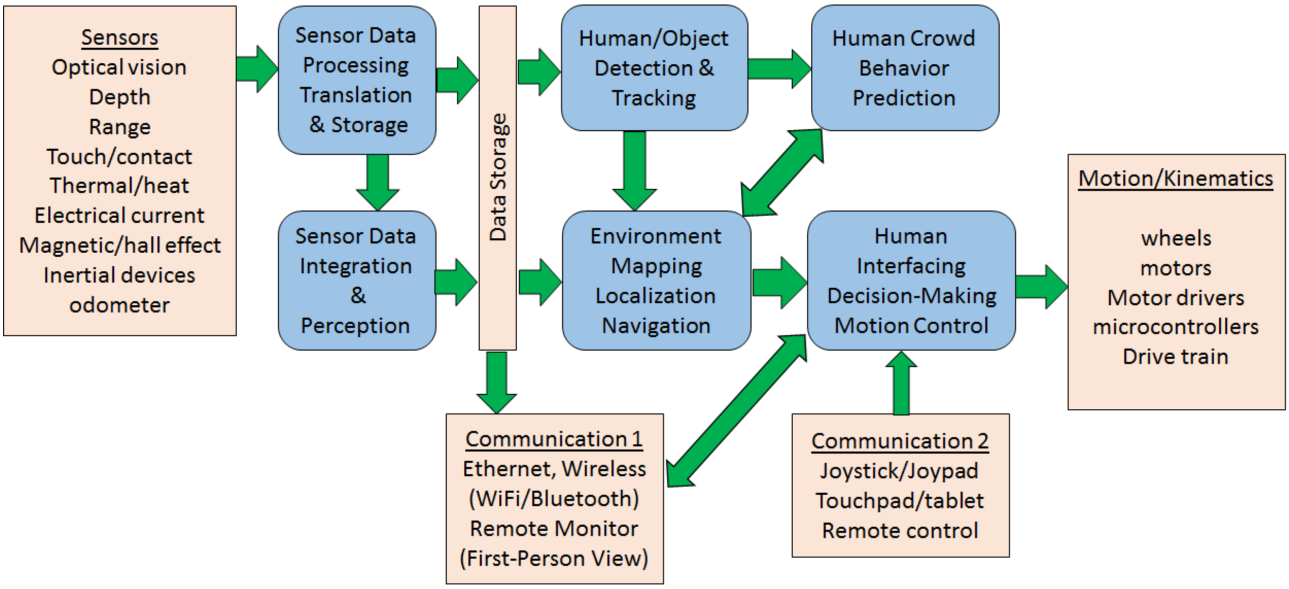

Figure 10: Functional blocks of a mobile robot

References

[1] Pepper robot: Softbank Robotics, https://www.ald.softbankrobotics.com/en/robots/pepper

[2] University College London, http://www.ucl.ac.uk/surgery/research/materials-tissue/aspire-create

[3] Locomotec GmbH, http://www.locomotec.com/de/

[4] Browning, R. C., Baker, E. A., Herron, J. A. and Kram, R. (2006). *"Effects of obesity and sex on the energetic cost and preferred speed of walking*". Journal of Applied Physiology. 100 (2), pp. 390–398.

[5] Powered wheelchair speed regulations: https://www.gov.uk/guidance/the-highway-code/rules-for-users-of-powered-wheelchairs-and-mobility-scooters-36-to-46

[6] Operator maneuver options for a wheelchair: https://en.wikipedia.org/wiki/Wheelchair

[7] Wheelchair control via brain impulses: https://cnbi.epfl.ch/software/index.html

[8] United States Design Patent No. USD 719620 S., Dec. 16, 2014.

[9] T. Carlson and Y. Demiris, "*Collaborative Control for a Robotic Wheelchair: Evaluation of Performance, Attention and Workload*," IEEE Trans. Sys, Man & Cybernetics, vol.42, no.3, June 2012, pp. 876-888.

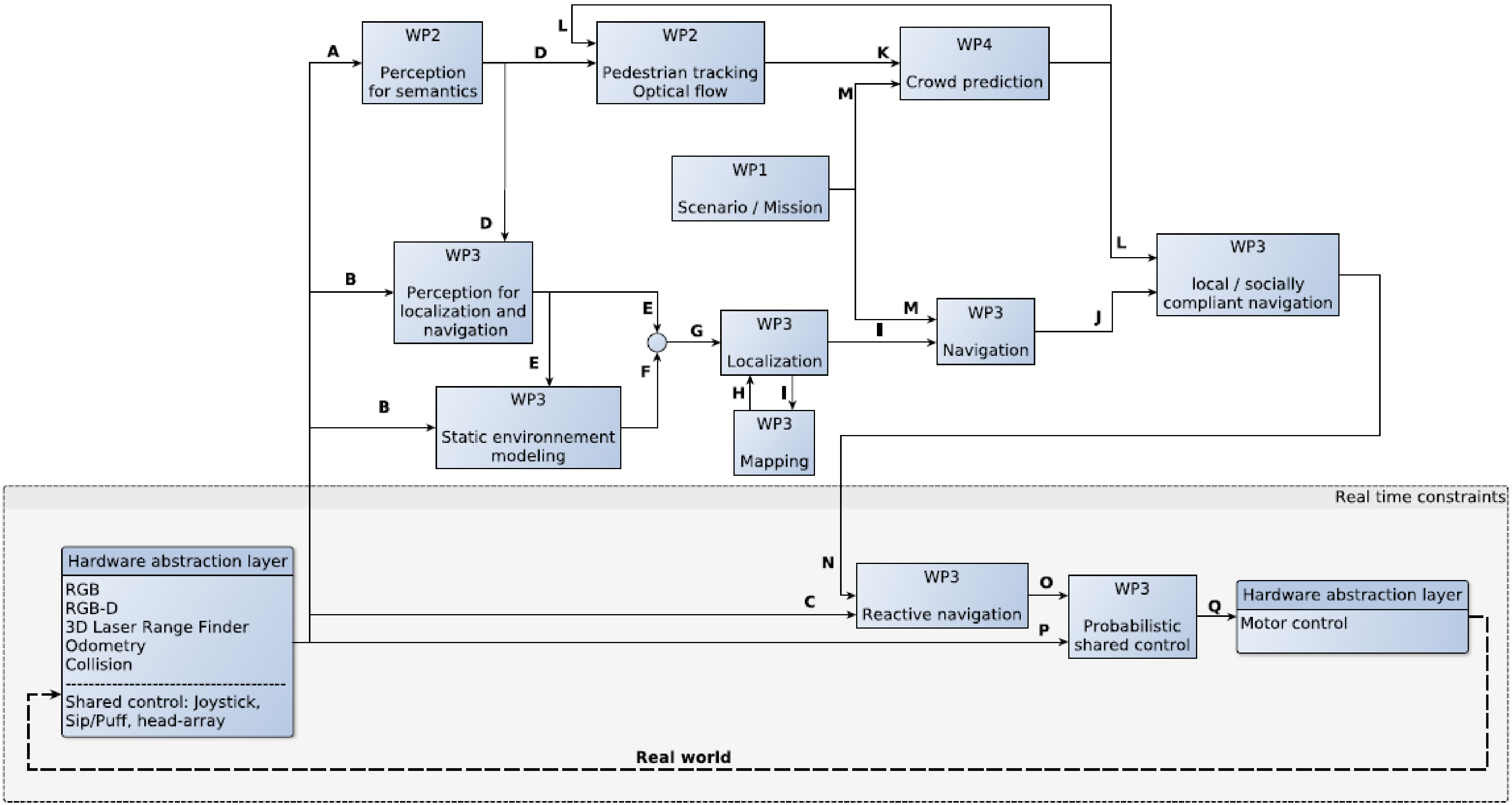

Figure 11: System architecture layout of a mobile robot using functional blocks and their interconnection