Chapter 20

A Mechatronic Platform for Robotic Educational Activities

Robotics courses can be efficiently taught only if self-involvement and development are encouraged early on. Robotics in education is mainly focused on the different disciplines that are put into practice in the respective field, including programming, hardware development, mechatronics, etc. This chapter discusses an end-to-end procedure to develop an autonomous robotic platform, as part of the educational procedure. First, the architectural design and the implementation of the platform are described in detail and then the educational activities that can be applied to this platform are detailed. The goal of this chapter is to highlight methods to stimulate the interest of students in robotics through their involvement in the development of a robotic platform and, later on, of the required software applications. In a nutshell, this chapter summarizes our experiments in mechatronics educational activities that actively involved students.

20.1. Introduction

During the past few decades, persistent research endeavors in the areas of robotics and artificial intelligence allowed new educational activities to take place in a real environment. In a course on intelligent tutoring systems (ITSs),

Chapter written by Ioannis KOSTAVELIS, Evangelos BOUKAS, Lazaros NALPANTIDIS and Antonios GASTERATOS.

first introduced in [CAR 70], and considering the advanced robotic technology, the simulations and the human–robot interaction have been extended to the environment of the robots. By exploiting various sensors and actuators, robots are capable of perceiving and interacting with the real world. Taking advantage of this intrinsic capacity, the robotic technology puts into effect a series of educational activities that have been developed to aid and foster learning of the relevant topics. Consequently, tutors have been able to increase the motivation and performance of their students by means of diverse robot building activities, such as robotic competitions and algorithm testing [MUR 01, NOU 04].

The majority of the research conducted in robotic educational activities can be categorized according to the interaction they provide with the teaching procedure. This also includes the approaches that aim to teach subjects closely related to the robotics. For example, the work presented in [PET 04] provides an effective understanding of robot programming and engineering principles, while in [KLA 03] the use of Lego Mindstorms robots for lab exercises and projects in classes is proposed. In a different approach [JAN 04], the authors presented the use of robotic technologies, for example application of a robot to escape from a maze to teach Personal Digital Assistant (PDAs) programming. The reader may refer to [WAN 01, WEI 01, AVA 02, JAN 04, AHL 02] for activities that include similar paradigms in robot programming, robot construction, artificial intelligence, algorithm development, embedded systems and mechatronics. Additionally, the growth of interest in mechatronics has identified a need for the provision of engineers whose education and training enables them to operate in an interdisciplinary manner [HAB 06]. In all the aforementioned works, the robot is used in the educational activity as just a passive tool, which can be considered as either the result of an assembling process of the students or a physical machine that executes a set of instructions and algorithms also provided by the students. Yet, alternative viewpoints exist, where autonomous robots could be used as active mediators of the educational activity, where they have an active role in the learning process [MIT 08].

In all such educational activities there is a plethora of robots that can be utilized, either as parts that should be assembled or as an out of the box solution ready for programming. The LEGO NXT robotic kit has been used for the construction of robots with up to three degrees of freedom [BAG 07]. Concerning the robots that are ready for use, the most representative example is the NAO humanoid robot, where the mechanics and electronics are already assembled and ready to run. It has been widely used in RoboCup projects by various universities, where the goal is to develop a team of three robots playing soccer against other teams [MÜL 11]. A humanoid robotic platform with less capabilities is the Robonova-1, which brings together both the

construction of the platform and its programming to accomplish simple robotic applications [LUP 10]. Apart from the humanoid robots there are also mobile platforms that are available for programming and development of sophisticated navigation algorithms. We could start with the Boe-Bot robot platform, which requires only basic skills in robotics and is suitable for accomplishing simple activities [BAL 08]. A more advanced and complicated robotic platform that enables the training of the users to enhanced robot navigation capabilities is the Pioneer platform. By using such a robot, the authors in [ANG 04] were able to detect and model doors in an indoor environment. Similarly, the authors in [SMI 09] describe a mobile robot platform equipped with vision sensors that is suitable for data acquisition under the limitations of the manual guidance. Additionally, there are also platforms that combine attributes from both humanoid and mobile platforms such as the PR2, which has been used to carry out advanced robotic applications such as grasping, simultaneous localization and mapping (SLAM) and path planning [MEE 10].

The main goal of this chapter is to show how educational activity can benefit by the development of a generic mobile robot platform. The contribution of this work can be divided into two objectives. The first objective involves the description of the activities that took place during the development and construction of a mobile robotic platform, suitable both for exploration of indoor and outdoor environments. This part includes all the prerequisite steps, namely requirement specification, the AutoCAD models, the construction of the platform and the mounting of the utilized sensors. The second objective concerns the description of the educational activities that will occur by utilizing this robotic platform. In particular, it will be used as a testbed in a series of undergraduate and graduate courses to assess students' capabilities in mechatronics and vision as well as to assist the students in acquiring knowledge in the respective topics. More specifically, reference is made in the different robotic applications that are supported by this platform, such as three-dimensional (3D) reconstruction, visual odometry (VO), visual SLAM (VSLAM), human–robot interaction tasks as well as robot control routines and hardware implementation. The latter provides a good insight for the potential of such a platform and the variety of the robotic educational activities that are supported.

20.2. System overview

The proposed work targets the development of a robotic agent equipped with high-resolution vision sensors ensuring great navigation and perception capacities of its environment. To effectively design and implement such a

robotic platform, several specification requirements should be posed in the first instance. The system should provide the means to the students to effectively use specific robotic activities. More specifically, the robot should be equipped with a wide stereoscopic camera mounted on the top of the robot, suitable both for the perception of the environment and the 3D reconstruction of the surroundings. Additionally, it should be equipped with a narrow stereoscopic camera mounted close to the base of the robot aiming at the ground exactly in front of it, which will be suitable for navigation activities such as VO. This second stereo camera is indeed useful due to the fact that localization algorithms tend to produce more accurate location estimations in static environments, that is when the sole moving object in the observed scene is the robot [NAL 11]. Consequently, the closer the stereo camera to the ground the less the observation of non-static objects in the scene. Moreover, the robot should be also equipped with an active vision sensor, that is, Kinect, that provides reliable in-depth information of the scene and a robust sensory input for indoor navigation and human–robot interaction procedures.
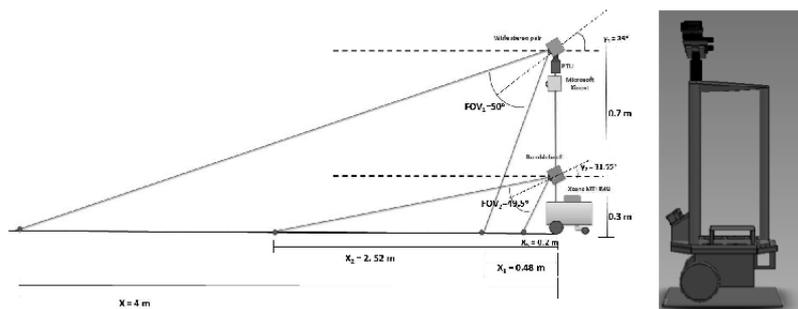
20.2.1. *Architectural design*

The first step towards the design of a robotic platform is the creation of a draft, which will comply with the entire system setup, as the one illustrated in Figure 20.1a. The proposed architecture comprises a robotic platform, a Bumblebee2 stereo camera (narrow stereo), a custom-made wide stereo camera which is mounted on a mast, a pan-tilt unit that holds the wide stereo camera and a Microsoft Kinect sensor. The assembly study behind the draft that is shown in Figure 20.1a is the following:

– The wide stereo pair, built from individual monocular cameras, placed 1 m above the ground with fixed tilt angle of 39°.

– The localization stereo pair, which is based on the Bumblebee2 camera, placed 0.3 m above the ground with a fixed tilt angle of 31.55°.

– The Microsoft Kinect sensor mounted under the pan-tilt unit parallel to the robotic base.

– The entire vision system of the robot mounted on a specially designed metallic structure, mounted and stabilized over the base of the robot.

– The pan-tilt unit that allows the rotation of the wide stereo pair used to augment the vision capabilities of the wide stereo.

The next step comprises the drawing of the custom parts of the robot in a 3D Computer-Aided Design (CAD) environment. This task demands

mechanical skills and provides the opportunity for the students to design in AutoCAD and implement an integrated structure on a computer numerical control (CNC) machine tool. The virtual model of the designed robot is presented in Figure 20.1b.

Figure 20.1. *The robotic platform illustrated as a) schematic of an architectural design and b) conceptual draft with AutoCAD*

Characteristic	Value
Base platform size	40 cm (L) x 37 cm (W) x 18 cm (H)
Wheels	15 cm diameter (driven) 6.25 cm diameter (caster) Polymer core, soft non-marking rubber tread
Wheelbase	33 cm
Drive type	Differential, single rear caster
Maximum speed	2.0 m/sec, 720 deg/sec
Motors	DC reversible with gearhead 72 W continuous power
Encoder resolution	500 cycles per motor revolution
Controller	16 bit microcontroller Integrated controller / motor driver Analog, digital, and servo interfaces
Power	12V, 7AH lead-acid batteries (x3) 5A charger
Weight	4.5 Kg (base) 12 Kg (base + 3 batteries)
Payload	20 Kg

a)   b)

Figure 20.2. *a) The Videre Erratic mobile platform and b) the respective specifications*

20.2.2. *Subsystem equipment*

20.2.2.1. *Robotic base*

An off-the-shelf robotic platform was chosen as the base of the robot to be constructed. The Videre Erratic mobile platform was used, which is a full-featured, industrial mobile robot base. It is a compact and powerful platform, capable of carrying a full load of robotic equipment, including an integrated PC, laser rangefinder and stereo cameras. The model of the platform is illustrated in Figure 20.2a, while the respective specifications

are analytically presented in Figure 20.2b. Moreover, the platform bears an integrated mini-PC that is capable of performing even computationaly intensive tasks.

20.2.2.2. *Robot skeleton*

The skeleton of the robot is a custom-designed construction to bear all the sensors and equipment to be utilized. The first part of the skeleton is a metallic base that holds the Bumblebee2 camera in a specific position and orientation. More analytically, the base should hold the stereo camera 0.30 m above the ground ensuring a tilt angle of $31.55°$. Additionally, the camera should be aligned with the front part of the robot retaining the consistency between the environment perception and the directions to the servo-motors of the platform, therefore, the construction of the metallic parts should permit accuracy better than 1 mm. A mechanical interface has been designed consisting of three different parts, two of which are asymmetric, outlined to result in the desired tilt angle. The proposed topology is illustrated in Figure 20.3a. The rest of the skeleton consists of three identical metallic poles, which are fastened by a frame screwed on the platform and a triangular plate on top of the poles. The latter is also utilized as the base for the wide stereo rig. It should be mentioned that the design and the implementation of the metallic construction has taken into consideration the maximum payload of the Erratic platform, which should be less than 20 kg. Therefore, the materials that has been utilized for the implementation of the skeleton is aluminum and the overall weight was approximately 5 kg. The metallic construction with the respective dimensions is shown in Figure 20.3b.

a)   b)
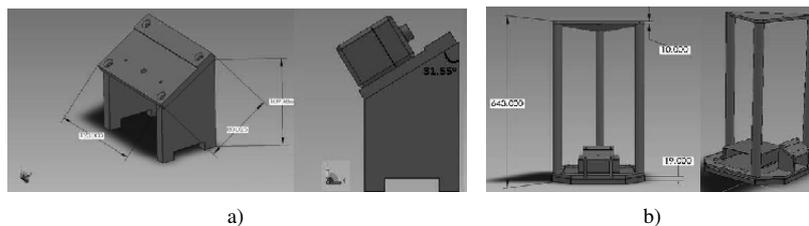
Figure 20.3. *a) The base for the Bumblebee2 stereo camera and b) the entire metallic construction as designed in AutoCAD*
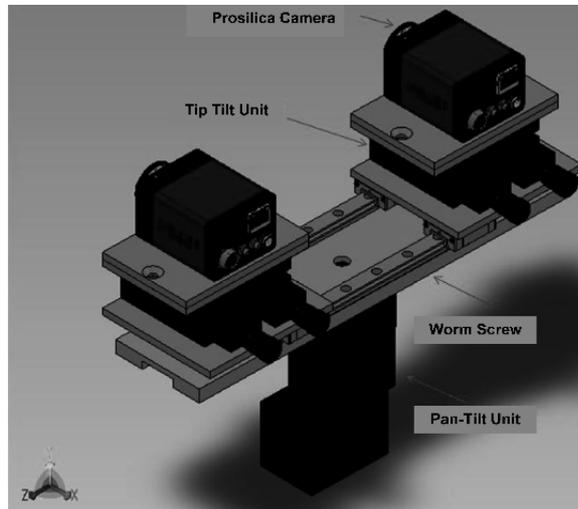
20.2.2.3. *Wide stereo camera*

The wide stereo camera consists of two identical Prosilica GE1660C cameras, an example of which is illustrated along with its specifications in Figure 20.4. The construction of the stereo rig is a demanding procedure

as it involves the assembly of several different parts and also demands the specification of many parameters. The most important one is the regulation of the baseline distance between the two cameras. This is an application-oriented parameter and influences the range within which the resolution is better than or equal to a desired value, that is by increasing the baseline the range is also increased, given that the rest of the intrinsic parameters are constant [KOS 11]. Therefore, a worm screw was used to ensure the regulation of the baseline distance within the range [15, 22] cm. An additional difficulty in the construction of the stereo rig is the precise alignment between the cameras, that is to parallelize the two respective principal axes. Therefore, a pair of additional regulating devices (tip-tilt units) have been placed beneath each of the cameras that adjust the rotation angle between the two axes (roll and pitch). The divergence of the two cameras on the yaw rotational angle is not of great importance as it can be easily minimized by the subsequent stereo calibration procedure. The tip-tilt unit bears a precision micrometer drive and is mechanically compatible to other interfaces. We considered the application scenario where the wide stereo vision system needed to cover a view range between 0.48 and 4 m and a horizontal field of view (FoV) of 120°. The wide stereo vision system was placed on the pan-tilt unit that is fixed on top of the metallic skeleton construction. The pan-tilt device used is the Directed Perception PTU-D46. This device has a pan range of 318° (−159°, +159°) and a tilt range of 78° (−47°, 31°). Given the fact that the vertical FoV of the used cameras in the wide stereo setup is 50°, the stereo rig should be tilted by an angle of 39° to accommodate the desired view range. Since the cameras horizontal FoV is smaller than the specified 120°, multiple pictures have to be captured by panning the wide stereo camera. More specifically, a solution for a camera with such a FoV, three pictures should be taken with panning angles of −35°, 0° and 35°, respectively. This results in an overlap of at least 25% between each successive couple, which is sufficient for a correct alignment and merging of the images into a wider image covering the total 120°. The wide stereo accompanied with all the aforementioned subordinate devices is illustrated in Figure 20.4.

Characteristic	Value
Sensor	Kodak KAI 02050 (CCD) sensor 2/4"
Resolution	1120x1120 (rectified)
FOV (Horizontal x Vertical)	50° x 54°
Focal length	0.0060m
Baseline	0.7m
Pixel size	0.0000055m x 0.0000055m (5.5um x 5.5um)
Interface	Gigabit Ethernet (8pin RJ-45 jack)
Voltage	12V (provided by external source)
Power consumption	5W @ 12V

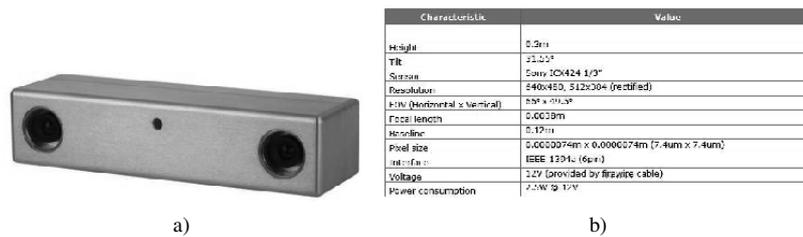a)   b)

Figure 20.4. *a) The Prosilica camera and b) the specifications of the Prosilica camera*

Figure 20.5. *The AutoCAD model of the wide stereo head with the pan-tilt unit*
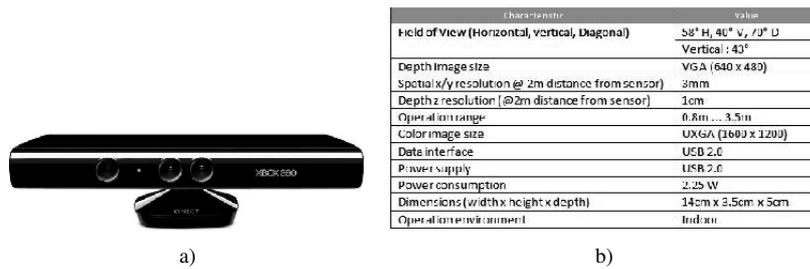
20.2.2.4. *Narrow stereo camera*

For the narrow stereo setup, the Bumblebee2 camera has been utilized. This camera has been exploited in various robotic applications both for indoor and outdoor exploration [NAL 09]. It has a very compact construction and retains precise alignment between the two vision sensors. This stereoscopic camera has a baseline distance among the two cameras that is 0.12 m and, therefore, we call it "narrow stereo" for the proposed robotic platform. The camera along with its specifications is illustrated in Figure 20.6.

Characteristic	Value
Height	0.3m
Tilt	31.55°
Sensor	Sony ICX424 1/3"
Resolution	640x480, 512x384 (rectified)
FOV (Horizontal x Vertical)	66° x 49.5°
Focal length	0.0038m
Baseline	0.12m
Pixel size	0.0000074m x 0.0000074m (7.4um x 7.4um)
Interface	IEEE 1394a (6pin)
Voltage	12V (provided by firewire cable)
Power consumption	2.1W @ 12V

a)   b)

Figure 20.6. *a) The Bumblebee2 camera and b) the specifications of the Bumblebee2 camera*

20.2.2.5. *Microsoft Kinect sensor*

Active vision sensors, such as time-of-flight (ToF) cameras, laser scanners and the Microsoft Kinect, are widely used in vision-based robotic applications due to the fact that they are able to directly obtain the depth map of the scene, without the need to spend additional computational resources that a stereo vision system would demand [STU 11]. The proposed robotic platform is thus additionally equipped with a Microsoft Kinect sensor, which is mainly suitable for indoor applications, due to the fact that its results are strongly affected by daylight. It can be used as a motion sensor that provides a natural user interface allowing users to interact intuitively and without any intermediary device, such as a controller. Among other attributes, the Kinect system can identify individual users through face recognition and voice recognition. In general, the Kinect sensor is widely used in education, enabling students to develop human–robot interfaces and various robotic applications [TÖL 10]. This sensor along with the respective specifications is shown in Figure 20.7.
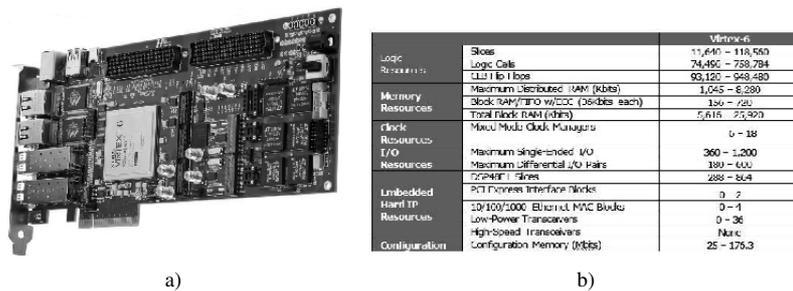

Characteristic	Value
Field of View (Horizontal, vertical, Diagonal)	58° H, 40° V, 70° D
	Vertical : 43°
Depth image size	VGA (640 x 480)
Spatial x/y resolution @ 2m distance from sensor)	3mm
Depth z resolution (@2m distance from sensor)	1cm
Operation range	0.8m ... 3.5m
Color image size	UXGA (1600 x 1200)
Data interface	USB 2.0
Power supply	USB 2.0
Power consumption	2.25 W
Dimensions (width x height x depth)	14cm x 3.5cm x 5cm
Operation environment	Indoor

a)   b)

Figure 20.7. *a) The Kinect sensor and b) the specifications of the Kinect sensor*

20.2.2.6. *FPGA platform*

The designed robotic platform also bears a field-programmable gate array (FPGA) device that is suitable for hardware programming and for parallelizing computational demanding tasks. An FPGA is a semiconductor device that can be programmed directly by students given the suitable software tools. As such FPGAs combine the best features of both hardware (i.e. speed and parallelization) and software (i.e. programmability and ease of update). Instead of being restricted to any predetermined hardware function, an FPGA allows the students to program their own features and functions, adapt to new standards and reconfigure the hardware for specific applications. FPGAs can be used to implement any logical function that an application-specific
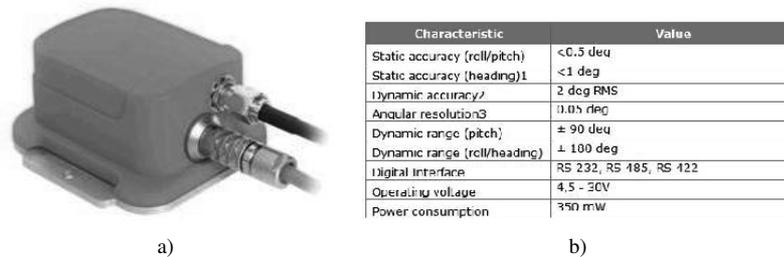
integrated circuit (ASIC) could perform, but the ability to update the functionality after shipping offers advantages for many robotic applications that need to be performed in real-time. Therefore, the selected FPGA device is the Virtex-6 LXT, which comprises the third-generation advanced silicon modular block (ASMBL) column-based architecture shown in Figure 20.8.

		Virtex-6
Logic Resources	Slices	11,640 – 118,560
	Logic Cells	74,490 – 758,784
	CLB Flip Flops	93,120 – 948,480
Memory Resources	Maximum Distributed RAM (Kbits)	1,045 – 8,280
	Block RAM/FIFO w/ECC (36Kbits each)	156 – 720
	Total Block RAM (Kbits)	5,616 – 25,920
Clock Resources	Mixed Mode Clock Managers	6 – 18
I/O Resources	Maximum Single-Ended I/O	360 – 1,200
	Maximum Differential I/O Pairs	180 – 600
Embedded Hard IP Resources	DSP48E1 Slices	288 – 864
	PCI Express Interface Blocks	0 – 2
	10/100/1000 Ethernet MAC Blocks	0 – 4
	Low-Power Transceivers	0 – 36
	High-Speed Transceivers	None
Configuration	Configuration Memory (Mbits)	25 – 176.3

a)	b)

Figure 20.8. *a) The Virtex-6 FPGA and b) the specifications of the Virtex-6 FPGA*

20.2.2.7. *Inertial measurement unit*

On top of the metallic plate, behind the wide stereo head, a compact inertial sensor platform has been placed, which is able to measure translational and rotational movements about all three axes. The device used here is the MTi-G, a small size and low-weight Global Positioning System (GPS) aided MEMS-based inertial measurement unit (IMU) and static pressure sensor Figure 20.9. The platform possesses an internal signal processor to provide accurate drift-free signals with highly dynamic response and long-term stability. Moreover, the embedded GPS sensors renders this device an excellent navigation and control aid for outdoors environments. MTi-G smart design allows its utilization in a wide range of output modes and sophisticated settings for specific usage situations, optimizing the sensor fusion (s.f.) algorithm routine for different applications. Indeed, the use of the IMU on our test robotic platform is twofold: (1) to provide a heterogeneous to the visual source of localization measurement, thus to enable the application and testing by the student of robust s.f. algorithms and (2) to allow the implementation and testing of image stabilization algorithms by fusing the rate gyro signals from the roll and pitch angle of the inertial sensors platform and the optic flow vectors from the image planes of the two cameras.

Characteristic	Value
Static accuracy (roll/pitch)	<0.5 deg
Static accuracy (heading)1	<1 deg
Dynamic accuracy2	2 deg RMS
Angular resolution3	0.05 deg
Dynamic range (pitch)	± 90 deg
Dynamic range (roll/heading)	± 180 deg
Digital interface	RS 232, RS 485, RS 422
Operating voltage	4.5 - 30V
Power consumption	350 mW

a)   b)

Figure 20.9. *a) The MTi-G and b) the specifications of the MTi-G*

20.2.3. *Robot montage*

This section describes the integration of all the aforementioned sensors and hardware devices into a single architecture, as shown in Figure 20.1a. This is a time-consuming task as it involves the efficient cooperation between software and hardware. As far as the hardware part is concerned, the entire construction should make use of mechanical interfaces designed to match exactly the sensors mountings. The entire construction has been assembled bearing an accuracy as high as possible. Specifically, the metallic parts have been designed and implemented with accuracy better than 1 mm. The wide stereo rig is a custom-made device and its accurate alignment is mandatory to produce accurate 3D reconstructions. Therefore, it has been implemented with an accuracy better than 0.1 mm. Moreover, the selected mechanical units (worm screw and tip-tilt) also retain strict specifications from the constructor. The exact knowledge of the precise positions of the sensors onboard the robot allows us to perform the accurate transformations that are demanded during the development of software applications, thus avoiding systematic errors that might occur due to misalignments. Concerning the software part, which is also important for the integration of the robot, several procedures have to be calibrated. The most demanding procedure is the calibration of the wide stereo camera setup, as it requires simultaneously the regulation of the roll and pitch rotation angles of the cameras. The epipolar error after the stereo calibration procedure that has been performed is $0.35°$ for the wide stereo camera and $0.15°$ for the narrow stereo camera (Bumbebee2), which in both cases is within the acceptable ranges, that is less than $0.50°$ [SZE 10]. The Kinect sensor has also been calibrated; this procedure is slightly different as it comprises the registration and mapping of the pixels between the color and infrared camera [SMI 11]. The integrated robotic platform is shown in Figure 20.10.
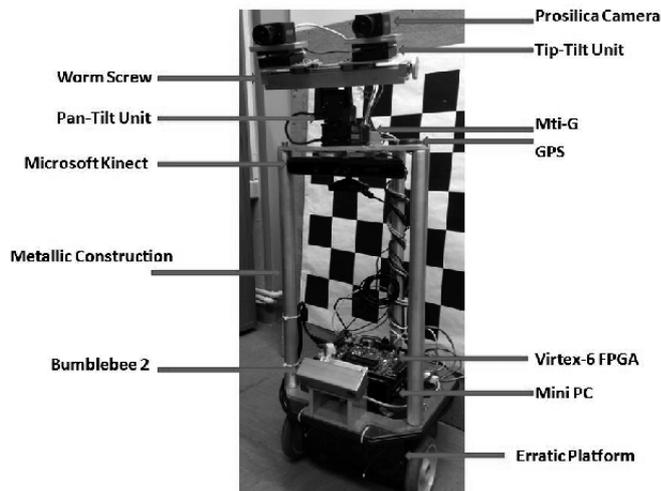
Figure 20.10. *The integrated robot platform with its subordinate parts*

20.3. Educational activities

The proposed robotic platform is suitable for a great variety of robot navigation activities, as it bears various sensors suitable either for indoor or outdoor exploration. More precisely, the platform is utilized for educational purposes on different courses, such as *Measurements* and *Robotics*. In accordance with these lessons, the students obtain hands on experience on data acquisition activities and robotic applications. This section describes the potential of the robotic platform to be used in various, but not limiting, application paradigms that are implemented and tested by the students.

20.3.1. *3D reconstruction*

The 3D reconstruction can be performed either with the wide stereo camera if it concerns outdoor exploration or with Microsoft Kinect for indoor applications. In the outdoor exploration scenario, the stereo camera is accompanied by a stereo correspondence algorithm capable to produce dense disparity images, while in the indoor applications, where the Kinect sensor is utilized, the depth map of the scene expressed in meters is directly obtained by the sensor.

In case of the wide or narrow stereo configuration, the 3D reconstruction module is responsible for extracting the depth, or equivalently the

disparity, information from the stereo cameras arrangements and provide 3D representations of the scene. The image acquisition should be accompanied by a stereoscopic algorithm capable of computing consistent disparity maps [NAL 08]. Every stereo correspondence algorithm makes use of a matching cost function in order to establish correspondence between two pixels. The results of the matching cost computation comprises of the disparity space image (DSI). DSI is a 3D matrix containing the computed matching costs for every pixel and for all its potential disparity values. Usually, the matching costs are aggregated over support regions. These regions could be 2D or even 3D regions within the DSI cube. Due to the aggregation step it is not single pixels that will be matched, but image regions. Aggregation of the matching cost values is a common and essential technique in order to suppress the effect of noise that usually leads to false matching. The selection of the optimum disparity value for each pixel is performed afterwards. It can be a simple winner-takes-all process or a more sophisticated process. In many cases, it is an iterative process as discussed in section 20.3.1. An additional disparity refinement step is frequently adopted. It is usually intended to interpolate the calculated disparity values, giving sub-pixel accuracy or assign values to not calculated pixels. The general structure of the majority of stereo correspondence algorithms is shown in Figure 20.11. Moreover, a bidirectional consistency check is applied. The selected disparity values are approved only if they are consistent, irrespectively to which image is the reference and which one is the target. Thus, even more false matches are disregarded. If multiple complementary stereo pairs are available simultaneously, the disparity maps from each of them can be merged into a single larger depth map.

In the case of Microsoft Kinect, the disparity map is obtained directly from the sensor. More precisely, the Kinect sensor consists of an infrared laser emitter, an infrared camera and an color camera. The laser source emits a single beam that is split into multiple beams by a diffraction grating to create a constant pattern of speckles projected onto the scene. This pattern is captured by the infrared camera and is correlated against a reference pattern. The reference pattern is obtained by capturing a plane at a known distance from the sensor, and is stored in the memory of the sensor. When a speckle is projected on an object whose distance to the sensor is smaller or larger than that of the reference plane, the position of the speckle in the infrared image will be shifted in the direction of the baseline between the laser projector and the perspective center of the infrared camera. These shifts are measured for all speckles by a simple image correlation procedure, which yields a disparity image. For each pixel, the distance to the sensor is then retrieved from the corresponding disparity.
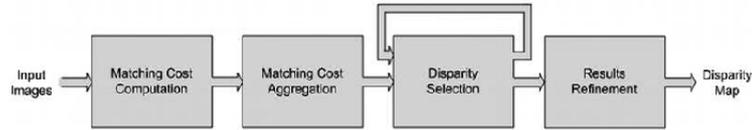
Figure 20.11. *General structure of stereo correspondence algorithms*

Once the depth map is computed in the case of the stereo rig or obtained immediately by the sensor in the case of Microsoft Kinect, then the reconstruction of the scene includes the transformation of the depth map to world coordinates by exploiting the triangulation. Specifically, by utilizing the depth information calculated in the disparity procedure, a dense point of the scene could be obtained and expressed in 3D world coordinates. More specifically, pixels expressed in camera coordinates $(x_c, y_c, disp(x_c, y_c))$, with respect to the stereo geometry, are transformed in 3D points (x, y, z). The relation between the world coordinates of a point $P(x, y, z)$ and the coordinates on the image plane $(x_c, y_c, disp(x, y))$ is expressed in case we have a stereoscopic camera by the pin-hole model and the stereo setup as follows:

$$x = \frac{x_c \cdot z}{f} \tag{20.1}$$

$$y = \frac{y_c \cdot z}{f} \tag{20.2}$$

$$z = \frac{b \cdot f}{disp(x_c, y_c)} \tag{20.3}$$

where z is the depth value of a feature depicted in (x_c, y_c), b is the stereo camera's baseline, f the focal length of the lenses expressed in pixels and $disp(x_c, y_c)$ the corresponding pixel's disparity value. In equations [20.1] and [20.2], x and y denote the abscissa and the ordinate in 3D world coordinates and corresponds to the (x_c, y_c) pixel in the image plane, respectively. The 3D reconstruction of a scene captured by a Microsoft Kinect is performed in a slightly different manner due to differentiations among the stereo and the RGB-D geometry. In the RGB-D scenario, a scale parameter is also considered in the triangulation procedure.

20.3.2. *Visual odometry*

The VO component estimates the robot displacement relative to a starting location. It is obtained by examining and correlating consecutive color images

captured by the narrow stereo or the Kinect sensor. The first step in a VO method comprises the detection of the salient landmarks between successive images. The success of the solely vision-based localization algorithms is, to a large extent, owed to the development of robust feature detection and description methods. The most common feature detector is the Harris corner detector while more recently developed feature detection methods such as the scale-invariant feature transform (SIFT), speed up robust features (SURF), and even the Center Surround Extremas (CenSurE) comprise both feature detection and description practices. A depth estimation of these features is then obtained by means of an enhanced stereo correspondence algorithm in case of the stereo rig or by exploiting the depth map obtained directly by the sensor in case of the Microsoft Kinect sensor. The next step comprises an outlier detection methodology responsible to discard both the mismatches between the features and the inserted errors due to the 3D reconstruction procedure. Assuming that the robot observes a specific point ^{t}P in 3D space, such as $^{t}P = [x_t, y_t, z_t]^T$. In time $t + 1$, the robot undergoes a specific motion with rotation $^{t}_{t+1}R$ and translation $^{t}_{t+1}T = [T_x, T_y, T_z]^T$, and the corresponding point ^{t}P is next observed as $^{t+1}P = [x_{t+1}, y_{t+1}, z_{t+1}]^T$. The transformation from point ^{t}P to ^{t+1}P is given as follows:

$$^{t}P = {}^{t}_{t+1}R \cdot {}^{t+1}P + {}^{t}_{t+1}T \qquad [20.4]$$

The calculation of the aforementioned transformation comprises the motion estimation procedure and maps the matched 3D reconstructed features between two successive time instances. In this module, only the correspondences that have passed the outlier detection procedure are utilized. Let us consider the resulting two 3D point clouds that correspond to times t and $t + 1$. The local coordinates feature position vectors ^{t+1}P on the reference image of the stereo pair (or the color image of the Kinect sensor) at time $t + 1$ are related to the position vectors ^{t}P in the reference image of the corresponding stereo pair (or the color image of the Kinect sensor) at time t according to the equation [20.4]. Ideally, six perfectly matched features should be sufficient to compute $^{t}_{t+1}T$ and $^{t}_{t+1}R$. However, in reality, error-suffering situations in several independent 3D points are needed for an efficacious calculation of $^{t}_{t+1}T$ and $^{t}_{t+1}R$, which should minimize a specific cost function, for example the least square one.

Students can obtain a deeper knowledge in the computer vision field with their involvement in the development of such a framework, as the VO can be divided into several subordinate modules, that is landmark detection, landmark 3D reconstruction, landmark matching and motion estimation. All the aforementioned routines constitute an active research topic in the field of robotics.

20.3.3. *Visual SLAM*

The visual SLAM problem is that of progressively building a map of the environment and at the same time localizing the robot on this map using vision. This module utilizes the output of the motion estimation and the calculated map to produce a global consistent map of the explored area. The robot could operate in indoor or outdoor environments, making different sensors more suitable for each scenario and posing different issues. The preferred sensor to obtain consistent and accurate depth maps could be, depending on the scenario, either one of the stereo camera setups or the Kinect sensor. The created map could be either a 2D grid in the simplist case or a 3D map in the most complex case. Additional sensors can be mounted on the robot by the students. This gives them opportunity to try out the use of different kinds of sensors to solve the SLAM problem. Other RGB-D sensors or ToF cameras or laser scanners can be used if the scenario favors them. In general, the different kinds of sensor provides heterogeneous outputs. Moreover, there are basic types of SLAM algorithms: the probabilistic and non-probabilistic SLAM algorithms. In our case, the students of undergraduate level, first use the non-probabilistic aspect as described in [NAL 11] obtaining a first experience of such a complex framework. Later on, on a more sophisticated course, which complies with the graduate level, the students use the probabilistic SLAM algorithms. As a result, students need to greatly understand the underlying principles of the system they work with before they can obtain meaningful representations of the world.

20.3.4. *Human tracking using RGB-D sensor*

This application comprises the detection and tracking of a human target from the robotic platform utilizing the mounted Microsoft Kinect sensor. The capabilities of this system concern, first, the recognition of the human body exploiting the skeleton detection and tracking application [BOU 11] and straight after the fixation on the trunk of the human body. Then, the mobile platform follows the human target by retaining a constant distance. The system decides about its movements by analyzing the data obtained directly from the Kinect sensor. The human is free to move in any direction and the robotic platform is responsible for retaining the human body at the center of the scene at a fixed minimum distance. The main limitation of the system is that in direct exposure to sunlight it influences the quality of the information grabbed from the infrared camera and the algorithm fails to execute correctly. In any case, the information derived from the RGB-D senor is further processed on the embedded PC, which decides about the steering commands that are passed

to the wheels of the platform. Frequent controls for the distance between the platform and the human target ensure the direct response of the system in any movement of the human body. An example of this application is shown in Figure 20.12, where Figure 20.12a shows an instance from the color camera of the RGB-D sensor and Figure 20.12b presents the skeleton tracking processes, which is executed on the mini-PC of the mobile robot platform.

a)   b)

Figure 20.12. *a) The color reference image of the Microsoft Kinect and b) the skeleton detection and tracking processes as illustrated on the mini-PC*

20.3.5. *Hardware acceleration*

Autonomous robots behavior greatly depends on the accuracy of their decision-making algorithms. Vision-based solutions are becoming more and more attractive due to their decreasing cost, as well as their inherent coherence with human-imposed mechanisms. However, in the case of stereo vision-based navigation, the accuracy and the refresh rate of the computed disparity maps are the cornerstone of success. Robotic applications place strict requirements on the demanded speed and accuracy of vision depth computing algorithms. The proposed robotic platform is equipped with a Virtex-6 FPGA suitable for hardware implementation of the computer vision algorithms. In many applications [KOS 11], the hardware implementation of the vision algorithms is mandatory. This robotic platform provides the students with the opportunity to create even faster robotic applications by mapping the existing algorithms into VHSIC hardware description language (VHDL) and execute on an FPGA, which is easily integrated with the robot.

20.3.6. *Sensor fusion algorithms*

The data fusion exploits different data sources with regard to their synchronized use for the extraction of useful information. In the case, that the data sources derive from different sensors, then sensors fusion is concerned. This is a very useful tool in robotics enabling the utilization of diverging sensor data to extract accurate and robust measurements for navigation and control. Several expressions of the Kalman filter, fuzzy logic and Bayes law are among the s.f methods as shown in equation [20.5] that can be utilized and tested on a students project to assess their application in navigation and control. In equation [20.5], the $P(x, y, z)$ stands for the estimated robot position after the fusion of different sensors and estimation routines. Moreover, different sensor setups can also be assessed depending on the scenario. In particular, the navigation can be evaluated either with the narrow stereo camera, the GPS and the IMU for outdoor exploration or with Microsoft Kinect and the IMU for indoor applications.

$$P(x, y, z) = s.f.\{GPS(z, y, z), IMU(x, y, z), VO(x, y, z)\} \quad [20.5]$$

20.3.7. *User interfaces*

The development of the aforementioned algorithms by the students is performed in different interfaces. The algorithms are first created in high-level programming languages and right after are passed into robotic and application-oriented tools. We may start from the rapid prototype platforms such as the Matlab® to confirm the accuracy of the designed algorithms. Latter on more complex frameworks are designed by utilizing the LabView toolbox as well as generic compilers for hardware implementation. Once the developed algorithms have been accurately designed, they are integrated in the robot by utilizing the robot operating system (ROS), which is an interface that provides the capacity to the students to interact immediately with the robotic platform and the embodied sensors. The ROS provides libraries and tools to help in developing software applications for robots. It provides hardware abstraction, device drivers, libraries, visualizers, message passing and package management. ROS comprises a straight forward selection for robot integration applications as it retains an open source license. The fundamental concepts of the ROS implementation are nodes, messages, topics and services. A node is a process that performs computations and is meant to operate at a fine-grained scale; a robot control system will usually comprise many nodes. For example, one node can control the wide stereo camera, the Bumblebee2, the wheel motors, while some nodes are responsible for more

advanced robotic frameworks such as localization and path planning. Nodes communicate with each other by passing messages. A message is a strictly typed data structure. Standard primitive types (integer, floating point, boolean, etc.) are supported, as are arrays of primitive types and constants. Messages can be composed of other messages, and arrays of other messages, nested arbitrarily deep. A node sends a message by publishing it on a given topic, which is simply a string such as *odometry* or *map*. Topics have anonymous publish/subscribe semantics, which decouples the production of information from its consumption. Although the topic-based publish-subscribe model is a flexible communications paradigm, its *broadcast* routing scheme is not appropriate for synchronous transactions, which can simplify the design of some nodes. In ROS, this is called a service, defined by a string name and a pair of strictly typed messages: one for the request and one for the response. This is analogous to Web services, which are defined by uniform resource identifiers (URIs) and have request and response documents of well-defined types. Additionally, the ROS also provides a graphical view of the system, which can be utilized as a simulation tool in corresponding learning activities. The proposed robotic platform comprises computational nodes for each subordinate sensor, endowing the students with the capacity to create a complete application for robot navigation by combining the rent sensory inputs.
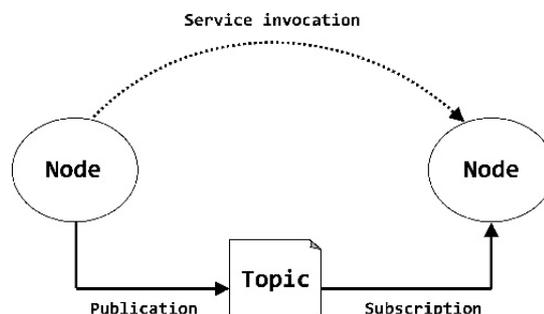

Figure 20.13. *The robotic operating system*

20.4. Experiences from educational activities

The developed robotic platform has been compared with similar constructions such as the one described in [SMI 09]. The specifications of the two compared platforms are summarized in Table 20.1, where the different sensor types are described. Specifically, the proposed system bears more sensors than the compared one, such as an extra stereoscopic camera, which

enables navigation with a different camera geometry setup, a Kinect sensor that enables the use of the platform in indoor environments and a hardware acceleration platform that provides the students with the capacity to involved with hardware development procedures. The presented platform lacks the existence of a panoramic camera that is a very useful tool to implement advanced image processing techniques such as panorama-stitching, however this issue could be bypassed with a top-down reprojection of the existing stereo pairs of the wide stereo rig.

The realized robotic platform has been utilized in various educational activities, which among other, includes the collection of data sets exploiting the vision sensors, which can be used for further assessment in other academic activities. So far three different data sets have been collected from the wide stereo camera, the Bumblebee2 and the Kinect sensor, respectively. Additionally, the robot setup has been evaluated by the students on already existing stereoscopic algorithms, such as the algorithm proposed in [NAL 09].

Sensor	Platform developed in *Oxford New College*	Proposed Platform
Robotic Platform	RMP200 from Segway	Videre Erratic mobile platform
Narrow Stereo	Point Grey BumbleBee	Point Grey BumbleBee
Wide Stereo	Non	Custom Made Stereo-Rig with 0.1mm accuracy
Panoramic Camera	LadyBug 2	Non
RGB-D Senor	Non	Microsoft Kinect Sensor
Laser Scanner		Non
Hardware Platform (FPGA)	Non	Virtex-6 LXT
IMU	Non	Mti-G
Pan Tilt Unit	Non	PTU-D46
GPS	CSI Series 5 Hz	Magellan ProMark-500 Rover
Operation	Outdoor	Indoor and Outdoor

Table 20.1. *The integrated robot platform with its subordinate parts*

Wide stereo camera. Several stereo pairs have been captured both in indoor and outdoor environments. The images acquired immediately from the cameras possess a resolution of 1,200 pixels, while after the calibration and rectification procedure the resolution was reduced to 1,120 pixels. Figure 20.15 depicts an indoor and outdoor scene captured by the wide stereo rig and the respective disparity images. Since the utilized stereo correspondence algorithm is a line-scan one, and given the fact that the quality of the resulting disparity images is very good, it can be deduced that the hardware regulation and the calibration procedure resulted in an accurately aligned stereo rig.

Figure 20.14. *The first row depicts a stereo pair and the respective disparity of an indoor scene, while the second row depicts a stereo pair and the respective disparity map of an outdoor scene*

a)   b)   c)
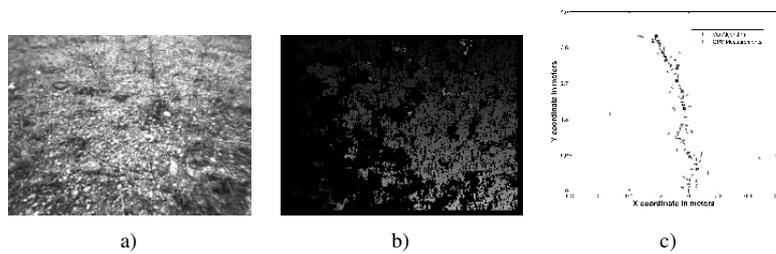
Figure 20.15. *a) A left reference image of the Bumblebee2, b) the respective disparity map and c) the output of a VO algorithm along with the GPS measurements*

Narrow stereo camera. The narrow stereo camera, that is Bumblebee2, has been utilized for the acquisition of a data set that corresponds to a short outdoor route. The main attribute of this data set is that it contains successive stereo pairs with close spatial proximity and, therefore, it is suitable for the evaluation of localization algorithms. It consists of 100 consequent stereo pairs and is accompanied by GPS location measurements. In the course of a robotic activity, this set of images has been employed by students for the evaluation of a 2D VO algorithm. Figure 20.15a and 20.15b depict a left reference image and a disparity map, respectively, while Figure 20.15c depicts the output of a VO algorithm along with the GPS measurements.

Kinect sensor. The Kinect sensor has been widely utilized as part of educational activities in indoor scenarios. A noteworthy task for students is the acquisition of a large data set known as the cognitive navigation data set, intended for use in indoor localization and semantic mapping. This data set has been captured with the Kinect RGB-D sensor mounted on the constructed robot, 0.8 m above the ground and tilted by $2.54°$. It consists of three different parts:

– Figure 20.16a involves the representation of a corridor, an office, a laboratory, a bathroom and an elevator area, captured in natural illumination conditions. This means that the first part has been acquired in day light. This part contains 1,286 images including the respective depth maps.

– Figure 20.16b involves exactly the same places as Figure 20.16a, but it has been captured during night and, therefore, the illumination conditions are artificial. This part contains 1,292 images including the respective depth maps.

– Figure 20.16c has also been acquired under artificial illumination conditions and comprises a continuous exploration of all the aforementioned places. The movements of the robot are very smooth with common FoV between the successive frames. This data set contains a continuous sequence of 557 images and corresponds to a 70 m route.
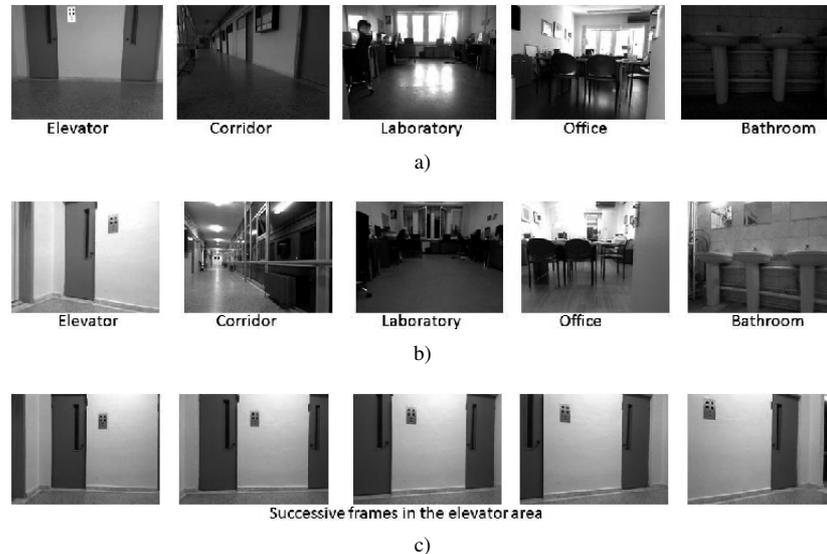
Figure 20.16. *Image samples a) comprising natural illumination conditions, b) comprising artificial illumination conditions and c) comprising a small root of the robot in artificial illumination conditions*

The first two parts of the data set are suitable for the validation of cognitive algorithms. The third part of the data set is suitable for indoor localization and mapping, due to the fact that the consecutive frames have close spatial and time proximity. Some examples of the captured images are shown in Figure 20.16. The cognitive navigation data set is freely available to the research community and can be retrieved, along with further technical information in [KOS 12].

20.5. Conclusions

In this chapter, the development of a robotic platform suitable for educational activities in indoor and outdoor scenarios has been presented. The work emphasized two directions. First, it goes through the description of the demanded steps that have taken place during the construction and assembly of the platform. This procedure has allowed students to be actively involved in the hardware development phases and includes the steps such as the requirement specification, the design of 3D CAD models, the construction of the platform and the physical placement and connection of the used sensors. Additionally, each utilized sensor has been described analytically by providing its specifications. The second focus direction concerns the description of the robotic educational activities that can be performed by exploiting the robotic platform. The integrated system will be used to assess students' capabilities in several topics of the mechatronics discipline. More specifically, reference is made in the variant robotic applications that can be supported from such a platform including 3D reconstruction, navigation, human–robot interaction tasks or even the low-level robot control routines. Moreover, some experiences of the educational activities have been shown providing the opportunity for the students to get involved and come one step closer to the state-of-the-art robotic technologies. Such student projects include data sets acquisition both in indoor and outdoor environments, while on the other hand, the students have exploited these acquired data to evaluate already existing algorithms, both for the perception of its environment and for robot navigation tasks.

20.6. Acknowledgments

20.7. Bibliography

[AHL 02] AHLGREN D., VERNER I., "An international view of robotics as an educational medium", *International Conference on Engineering Education (ICEE'2002)*, Citeseer, 2002.

[ANG 04] ANGUELOV D., KOLLER D., PARKER E., *et al.*, "Detecting and modeling doors with mobile robots", *IEEE International Conference on Robotics and Automation*, IEEE, vol. 4, pp. 3777–3784, 2004.

[AVA 02] AVANZATO R., ABINGTON P., "Mobile robot navigation contest for undergraduate design and K-12 outreach", *ASEE Annual Conference Proceedings*, June 2012.

[BAG 07] BAGNALL B., *Maximum Lego NXT: Building Robots with Java Brains*, Variant Press, 2007.

[BAL 08] BALOGH R., "Basic activities with the Boe-Bot mobile robot", *In Proceedings of Conference DidInfo*, 2008 FPV UMD Banska Bystrica, Slovakia 2008.

[BOU 11] BOULOS M., BLANCHARD B., WALKER C., *et al.*, "Web GIS in practice X: a Microsoft Kinect natural user interface for Google Earth navigation", *International Journal of Health Geographics*, vol. 10, no. 1, pp. 45, 2011.

[CAR 70] CARBONELL J., "AI in CAI: an artificial-intelligence approach to computer-assisted instruction", *IEEE Transactions on Man-Machine Systems*, vol. 11, no. 4, pp. 190–202, 1970.

[CRA 04] CRAIG J., *Introduction to Robotics: Mechanics and Control*, Prentice Hall, 2004.

[HAB 06] HABIB M., "Mechatronics engineering the evolution, the needs and the challenges", *32nd Annual Conference on Industrial Electronics, IECON 2006*, IEEE, pp. 4510–4515, 2006.

[JAN 04] JANSEN M., OELINGER M., HOEKSEMA K., *et al.*, "An interactive maze scenario with physical robots and other smart devices", *2nd IEEE International Workshop on Wireless and Mobile Technologies in Education*, IEEE, pp. 83–90, 2004.

[KLA 03] KLASSNER F., ANDERSON S., "Lego MindStorms: not just for K-12 anymore", *IEEE Robotics & Automation Magazine*, vol. 10, no. 2, pp. 12–18, 2003.

[KOS 11] KOSTAVELIS I., BOUKAS E., NALPANTIDIS L., *et al.*, "SPARTAN system: towards a low-cost and high-performance vision architecture for space exploratory rovers", *IEEE International Conference on Computer Vision Workshops (ICCV Workshops)* 2011, IEEE, pp. 1994–2001, 2011.

[KOS 12] KOSTAVELIS I., GASTERATOS A., "Cognitive navigation dataset, group of robotics and cgnitive systems", Available at http://robotics.pme.duth.gr/kostavelis/Dataset.html, 2012.

[LUP 10] LUPIÁN L., ROMAY A., MONROY P., *et al.*, "Cyberlords robocup 2010 humanoid kidsize team description paper", *Workshop Robocup Singapore*, June 2010.

[MEE 10] MEEUSSEN W., WISE M., GLASER S., *et al.*, "Autonomous door opening and plugging in with a personal robot", *IEEE International Conference on Robotics and Automation*, IEEE, pp. 729–736, 2010.

[MIT 08] MITNIK R., NUSSBAUM M., SOTO A., "An autonomous educational mobile robot mediator", *Autonomous Robots*, vol. 25, no. 4, pp. 367–382, 2008.

[MÜL 11] MÜLLER J., LAUE T., RÖFER T., "Kicking a ball-modeling complex dynamic motions for humanoid robots", *RoboCup 2010: Robot Soccer World Cup XIV*, Springer, pp. 109–120, 2011.

[MUR 01] MURPHY R., "Competing for a robotics education", *IEEE Robotics & Automation Magazine*, vol. 8, no. 2, pp. 44–55, 2001.

[NAL 08] NALPANTIDIS L., SIRAKOULIS G., GASTERATOS A., "Review of stereo vision algorithms: from software to hardware", *International Journal of Optomechatronics*, vol. 2, no. 4, pp. 435–462, 2008.

[NAL 09] NALPANTIDIS L., KOSTAVELIS I., GASTERATOS A., "Stereovision-based algorithm for obstacle avoidance", *Intelligent Robotics and Applications*, Springer, pp. 195–204, 2009.

[NAL 11] NALPANTIDIS L., SIRAKOULIS G., GASTERATOS A., "Non-probabilistic cellular automata-enhanced stereo vision simultaneous localization and mapping", *Measurement Science and Technology*, vol. 22, no. 11, p. 114027, 2011.

[NOU 04] NOURBAKHSH I., HAMNER E., CROWLEY K., *et al.*, "Formal measures of learning in a secondary school mobile robotics course", *Proceedings of the IEEE International Conference on Robotics and Automation 2004 (ICRA'04)*, IEEE, vol. 2, pp. 1831–1836, 2004.

[PET 04] PETRE M., PRICE B., "Using robotics to motivate 'back door' learning", *Education and Information Technologies*, vol. 9, no. 2, pp. 147–158, 2004.

[SMI 11] SMISEK J., JANCOSEK M., PAJDLA T., "3D with Kinect", *IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*, IEEE, pp. 1154–1160, 2011.

[SMI 09] SMITH M., BALDWIN I., CHURCHILL W., *et al.*, "The new college vision and laser data set", *The International Journal of Robotics Research*, vol. 28, no. 5, pp. 595–599, 2009.

[STU 11] STURM J., MAGNENAT S., ENGELHARD N., *et al.*, "Towards a benchmark for RGB-D SLAM evaluation", *Proceedings of the RGB-D Workshop on Advanced Reasoning with Depth Cameras at Robotics: Science and Systems*, California, USA, July 2011.

[SZE 10] SZELISKI R., *Computer Vision: Algorithms and Applications*, Springer, 2010.

[TÖL 10] TÖLGYESSY M., HUBINSKỲ P., "The Kinect sensor in robotics education", *Proceedings of 2nd International Conference on Robotics in Education*, Austria, 2010.

[WAN 01] JERRY B., WEINBERG GEORGE L., ENGEL, KEQIN GU, *et al*, *A Multidisciplinary Model for Using Robotics in Engineering Education*, American Society for Engineering Education, 2001.

[WEI 01] WEINBERG J., ENGEL G., GU K., *et al.*, "A multidisciplinary model for using robotics in engineering education", *ASEE Annual Conference and Exposition*, Albuquerque, New Mexico, June 2001.