Application of semantic maps for mobile robot simulation

(Zastosowanie map semantycznych w symulacji robota mobilnego)

dr inż. JANUSZ BĘDKOWSKI, prof. dr hab. inż. ANDRZEJ MASŁOWSKI

Instytut Maszyn Matematycznych, Warszawa

In this paper a new concept of using semantic map for robot operator training purpose is described. The approach consists of 3D laser data acquisition, semantic elements extraction (using image processing techniques) and transformation to rigid body simulation engine, therefore the State Of the Art related to those research topics will be discussed. The combination of a 2D laser range finder with a mobile unit was described as the simulation of a 3D laser range finder in [1]. In this sense we can consider that several researches are using so called simulator of 3D laser range finder to obtain 3D cloud of points [2]. The common 3D laser simulator is built on the basis of a rotated 2D range finder. The rotation axis can be horizontal [3], vertical [4] or similarly to our approach (the rotational axis lies in the middle of the scanners field of view).

Semantic information extracted from 3D laser data is recent research topic of modern mobile robotics. In [5] a semantic map for a mobile robot was described as a map that contains, in addition to spatial information about the environment, assignments of mapped features to entities of known classes. In [6] a model of an indoor scene is implemented as a semantic net. This approach is used in [7] where robot extracts semantic information from 3D models built from a laser scanner. In [8] the location of features is extracted by using a probabilistic technique (RANSAC). Also the region growing approach [9] extended from [10] by efficiently integrating k-nearest neighbor (KNN) search is able to process unorganized clouds of points. The semantic map building is related to SLAM (Simultaneous Localization And Mapping) problem [11]. Most of recent SLAM (Simultaneous Localization And Mapping) techniques use camera [12], laser measurement system [13] or even registered 3D laser data [14]. Concerning the registration of 3D scans described in [15] we can find several techniques solving this important issue. The authors of [16] briefly describe ICP (Iterative Closest Points) algorithm and in [17] the probabilistic matching technique is proposed. In [18] the mapping system that acquires 3D object models of man-made indoor environments such as kitchens is shown. The system segments and geometrically reconstructs cabinets with doors, tables, drawers, and shelves, objects that are important for robots retrieving and manipulating objects in these environments.

A detailed description of computer based simulators for unmanned vehicles is shown in [19] [57]. Also in [20] the comparison of real-time physics simulation systems is given, where a qualitative evaluation of a number of free publicly available physics engines for simulation systems and game development is presented. Several frameworks are mentioned such as USARSim which is very popular in research society [21], Stage, Gazebo [22], Webots [23], MRDS (Microsoft Robotics Developer Studio) [24]. Some researchers found that there are many available simulators that offer attractive functionality, therefore they proposed a new simulator classification system specific to mobile robots and autonomous vehicles [25]. A classification system for robot simulators will allow researchers to identify existing simulators which may be useful in conducting a wide variety of robotics research from testing low level or autonomous control to human robot interaction.. To ensure the validity of robot models, NIST proposes standardized test methods that can be easily replicated in both computer simulation and physical form [26].

In this paper we propose a new idea of semantic map building, this map can be transformed into rigid body simulation. It can be used for several applications such as robot operator training. It is a new idea and can give an opportunity to develop training systems composed by real and virtual robots. We hope that it will improve multi robot system design and development. The paper is organized as follows: in section "Robot" robot and its data acquisition module is described. In section "Semantic map approach" we explained the semantic map concept applied to walls and stairs detection. Section "Semantic simulation engine" describes the core of proposed simulation system. The final discussion is given in the "Conclusion".

Robot

The robot used is an ActiveMedia PIONEER 3AT, equipped with SARA (Sensor Data Acquisition System for Mobile Robotic Applications). SARA is composed by 2 lasers LMS SICK 100 orthogonally mounted. Bottom laser can rotate, therefore it delivers 3D cloud of points in stop-scan fashion. Ethernet camera delivers image data available via web browser. *Figure 1* shows the hardware and data visualization.

Rotated SICK

The rotational axis lies in the middle of the scanners field of view. To rotate the scanner we use a computer controlled PTU-D46-17.5 pan-tilt unit from Directed Perception. A horizontal scan point without rotation of pan tilt unit would have the coordinates

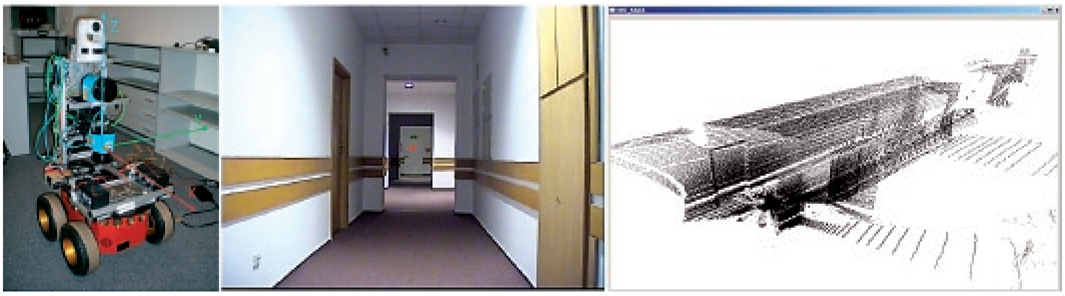

Fig. 1. Robot PIONEER 3AT equipped with SARA (Sensor Data Acquisition System for Mobile Robotic Applications). From left: robot, image from onboard camera, 3D cloud of points

Rys. 1. Robot PIONEER 3AT wyposażony w SARA (*Sensor Data Acquisition System for Mobile Robotic Applications*). Od lewej: robot, widok z kamery, chmura punktów 3D

$$x_i = \cos\alpha_i \cdot r_i ; y_i = \sin\alpha_i \cdot r_i ; z_i = 0; \quad (1)$$

with α_i as an angle between optical axis and the laser beam, r_i as corresponding distance. Taking into account the additional rotation around the optical axis X (Figure 1) leads to:

$$x_i = \cos\alpha_i \cdot r_i ; y_i = \sin\alpha_i \cdot \cos\beta \cdot r_i ; z_i = \sin\alpha_i \cdot \sin\beta \cdot r_i ; \quad (2)$$

with β as the rotation angle around X axis measured form Y counterclockwise.

Semantic map approach

The proposed approach is dedicated to structured INDOOR environment where floor is flat and walls are straight and orthogonal to floor and ceiling, stairs are parallel and ergonomic. It is obvious that if we try to use this approach in unstructured environment, algorithm will generate numerous not-labeled objects. To find semantic objects such as “wall”, “ceiling”, “floor”, “doors” (with joint), “stairs” we are using a new idea that is based on prerequisites generation from projected single 3D scan (onto OXY plane for “wall”, “door”, “stairs” and onto OXZ for “ceiling” and “floor”). Prerequisites are generated using image processing techniques such as Hough transform for line extraction. The prerequisites are checked in next step, if the constraint in satisfied, the semantic object is assigned. For instance if we assume 3D scan projected onto OXY plane, a single line is related to a “wall” prerequisite, 3 parallel lines are prerequisite of “stairs”, 2 connected lines are prerequisite of opened doors and the connection can be a joint. Single long line in 3D scan projected onto OXZ plane is a prerequisite of “ceiling” and “floor”. In next subsection we introduce the algorithm for line extraction from projected 3D scan.

Line extraction from projected 3D scan

In *Figure 2* the rectangle of 3D scan projection onto OXY plane is shown. We assume that robot is located on a surface (in other case additional pitch/roll sensors can be used), therefore projected wall, doors and stairs determine the line on the considered binary image. The same assumption is related to ceiling and floor detection in case of 3D scan projection onto OXZ plane. We consider the region 10 x 10 m because of the acceptable distance between closest measured 3D points. On the same figure the correspond image coordinates are shown for instance point Pxy=(-5, -5) corresponds to Puv=(0, 0). The image size is 512 pixels width, 512 pixels height, therefore one pixels occupies the rectangle region approximately 20 x 20 cm. The step of line extraction algorithm is to compute the sum of projected points for each pixels.

This approach has a disadvantage caused by the 3D data acquisition technique. Because laser is rotated around its optical axis, the sum values are increased in front of the laser and decreased in other regions. To avoid this disadvantage we propose prerequisites boosting method. In *Figure 3* it can be seen that the obstacle located in front of laser (its optical axis) influences increased sum-values. On the contrary obstacle located near Y axis influences decreased sum-value that can result in omitting this prerequisite. The prerequisites boosting method is based on normalization technique that uses the model of laser beam intersection with rectangular prisms with base assigned by $pixel_{uv}$. The model of laser beam intersection is shown in *Figure 3* (centre). It is related to the computation of maximum available laser beam intersection with rectangular prism with the base in $pixel_{uv}$. Once the $model_{uv}$ is computed it can be used for normalization prerequisites image (sum_{uv}), where $sum_{new;\,uv} = sum_{uv}/model_{uv}$.

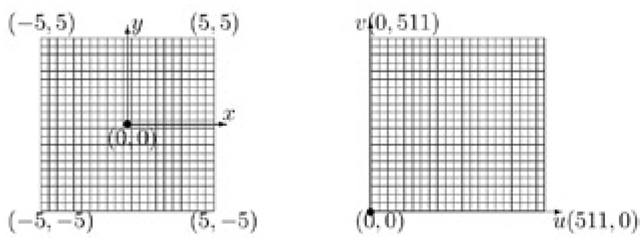

Fig. 2.The rectangle of 3D scan projection onto OXY plane – left, corresponding image coordinates – right

Rys. 2. Rzutowanie chmury punktów 3D na płaszczyznę OXY oraz współrzędne obrazu

Computed $sum_{new,uv}$ image (where values are real numbers from 0 to 1) is used for prerequisites generation based on image processing methods. The implementation is based on OpenCV image processing library [27]. *Figure 4* shows the procedure. Input image box represents the computed $sum_{new,uv}$ image transformed into binary image using simple threshold method. Filtering box reduces noise from image. The structuring element used for this operation is:

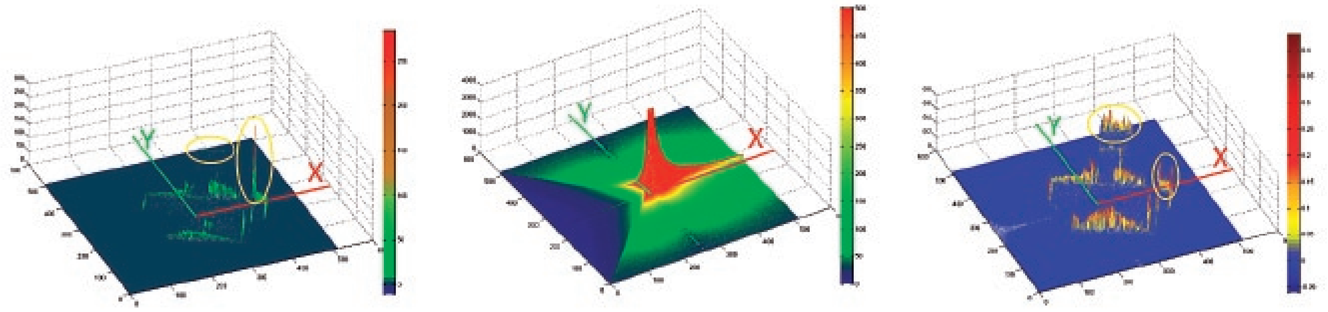

Fig. 3. Left: sum of projected points$_{xy}$ onto pixels$_{uv}$, middle: model of laser beam intersection with rectangular prisms with base assigned by pixel$_{uv}$, right the result of prerequisites boosting method

Rys. 3. Lewy: suma punktów points$_{xy}$ rzutowanych na pixels$_{uv}$, środkowy: model laserowego systemu pomiarowego, prawy: rezultat zastosowanej metody wzmocnienia przesłanek

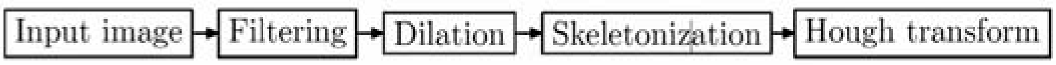

Fig. 4. Image processing methods used for prerequisites computation
Rys. 4. Algorytm wyznaczania przesłanek na bazie procedur przetwarzania obrazów

$$strel = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad (3)$$

For each pixel $p_{k;l}$ from binary image, where k = 1:510, l = 1:510, following equation is solved:

$$p_{res(k,l)} = \sum_{i=-1}^{1}\sum_{j=-1}^{1} strel_{i,j} \cdot p_{k+i,l+j} \cdot (|i|+|j|) \quad (4)$$

if $p_{res(k;l)} > 0$ and $p_{k;l} = 1$ then $p_{out(k;l)} = 1$, else $p_{out(k;l)} = 0$. Dilation box mathematical morphology operation increase the width of binary objects in the image. The operation dilation [27] dilates the source image using the specified structuring element that determines the shape of a pixel neighborhood over which the maximum is taken. Neighboring objects are going to be connected for better Hough transform result. Skeletonization based on classical Pavlidis [28] algorithm gives the output as thin lines that are used by Hough transform box to obtain line segments.

Walls and doors detection

The procedure of walls and doors prerequisites generation is using image processing methods. The result of this procedure is shown in *Figure 5*, where each line (marked by blue color) corresponds to wall prerequisite. The set of lines is used to obtain segmentation of 3D cloud of points, where different walls will have different labels. For each line segment the orthogonal plane$_{orth}$ to plane$_{OXY}$ is computed (*Fig. 5* – right). The intersection between this two planes is the same line segment. All 3D points which satisfy the condition of distance to plane$_{orth}$ have the same label. In the first step all prerequisites of walls are checked separately. To perform the scene interpretation semantic net is proposed. Nodes of a semantic net represent entities of the world and the relationships between them are defined. Possible labels of the nodes are L = {"wall", "wall above door", "floor", "ceiling", "door", "free space for door"}. The relationships between the entities are R = {"parallel", "orthogonal", "above", "under", "equal height", "available inside", "connected via joint"}. The semantic net can easily be extended to more entities and relationships which determine a more sophisticated feature detection algorithms. In our case the feature detection algorithm is composed by the method of cubes generation, where each cube should contain measured 3D point. In the second step of the algorithm wall candidates are chosen. From this set of candidates, based on relationships between them, proper labels are assigned and output model is generated.

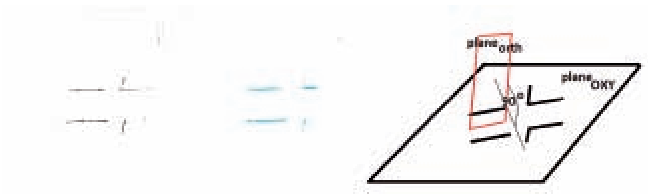

Fig. 5. Walls prerequisites detection. Left: Input image, centre: result as lines (prerequisites of walls), right: example orthogonal plane$_{orth}$ to plane$_{OXY}$.
Rys. 5. Przesłanki ścian, lewy: wejściowy obraz, środkowy: wynik jako linie będące przesłankami ścian, prawy: przykładowe prostopadłe płaszczyzny plane$_{orth}$ i plane$_{OXY}$

Stairs detection

The image processing methods are used for stairs prerequisites generation. The result of this procedure is shown in *Figure 6*, where red rectangle corresponds to stairs prerequisite. It is important to emphasize that the set of parallel lines in the same short distance between each other can be a projection of stairs. Possible label of the node is L = {"stair"}. The relationships between the entities are R = {"parallel", "above", "under"}.

Fig. 6. Stairs prerequisites detection
Rys. 6. Detekcja przesłanek wystąpienia schodów

Semantic simulation engine

The concept of semantic simulation engine is a new idea, and its strength lies on the semantic map integration with mobile robot simulator. The engine basic elements are: semantic map nodes(entities) Lsm={"wall", "wall above door", "floor", "ceiling", "door", "free space for door", "stairs"}, robot simulator nodes(entities) Lrs={"robot", "rigid body object", "soft body object"}, semantic map relationships between the entities Rsm = {"parallel", "orthogonal", "above", "under", "equal height", "available inside", "connected via joint"}, robot simulator relationships between the entities Rrs = {"connected via joint", "position"}, semantic map events Esm = robot simulator events Ers = {"movement", "collision between two entities started", "collision between two entities stopped", "collision between two entities continued", "broken joint"}. Robot simulator is implemented in NVIDIA PhysX. The entities from semantic map correspond to actors in PhysX. Lsm is transformed into Lrs based on spatial model i.e. "walls", "doors" and "stairs" correspond to actors with "BOX" shapes (details concerning NVIDIA PhysX shapes of actors are available at [58]). Rsm are transformed into Rrs with remark that doors are connected to walls via revolute joints. All entities/relations Rsm has the same initial location in Rrs, obviously the location of each actor/entity may change during simulation. The transformation from Esm to Ers effects that events related to entities from semantic map correspond to the events related to actors representing proper entities. Following events can be noticed during simulation: robot can touch each entity, open/close the door, climb the stairs, enter empty space of the door, damage itself (broken joint between actors in robot arm), brake joint that connects door to the wall. It is noteworthy to mention that all robot simulator semantic events are useful for operator training, where computer has to monitor simulation events, judge them, and report the result to the instructor.

Conclusion

In the paper the semantic simulation engine that is used for merging real semantic data with NVIDIA PhysX mobile robot simulation is shown. The approach can be used for further development of sophisticated training tools i.e. AR (Augmented Reality), where real robots will be used for environment modeling. New approach of image processing techniques in the process of semantic entities identification is shown. Available State of The Art INDOOR objects recognition techniques can improve proposed implementation. The approach can be extended to build large environment model

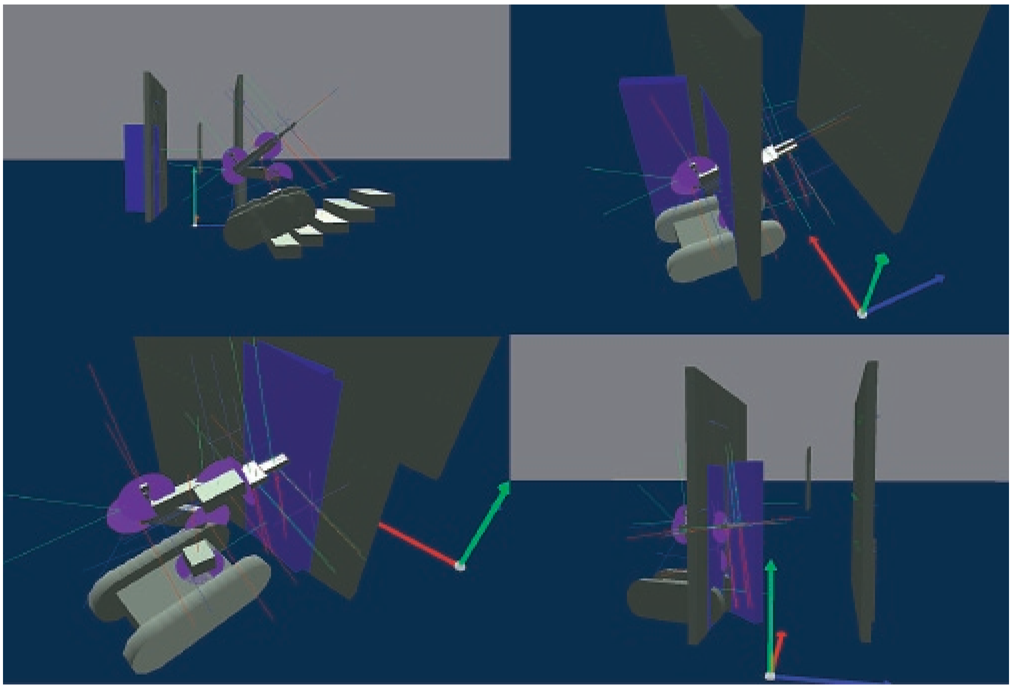

Fig. 7. Robot simulator semantic events. Rys. 7. Semantyczne zdarzenia w symulatorze robota

by integrating 3D data registration such as ICP. We have shown new application for semantic mapping - the mobile robot operator training semantic simulation engine, where identified and later modeled semantic objects interact with predefined simulation entities. In our opinion the approach deliver powerful tool for INDOOR environment inspection and intervention in which operator can use semantic information to interact with entities. Future work will be related to the integration of data registration techniques and augmented reality approach.

References

[1] Klimentjew D., M. Arli, J. Zhang: 3D scene reconstruction based on a moving 2D laser range finder for service-robots, in: Proceedings of the 2009 international conference on Robotics and biomimetics, 2009.

[2] Magnusson M., T. Duckett, A. J. Lilienthal: 3d scan registration for autonomous mining vehicles, Journal of Field Robotics 24 (10) (2007) 803-827.

[3] Nuchter A., H. Surmann, J. Hertzberg: Automatic model refinement for 3D reconstruction with mobile robots, in: Fourth International Conference on 3-D Digital Imaging and Modeling 3DIM 03, 2003, p. 394.

[4] Montemerlo M., S. Thrun: A multi-resolution pyramid for outdoor robot terrain perception, in: AAAI'04: Proceedings of the 19th national conference on Artificial intelligence, AAAI Press, 2004, pp. 464–469.

[5] Nuchter A., J. Hertzberg: Towards semantic maps for mobile robots, Robot. Auton. Syst. 56 (11) (2008) 915–926.

[6] Grau O.: A scene analysis system for the generation of 3-d models, in: NRC '97: Proceedings of the International Conference on Recent Advances in 3-D Digital Imaging and Modeling, IEEE Computer Society, Washington, DC, USA, 1997, p. 221.

[7] Nuchter A., H. Surmann, K. Lingemann, J. Hertzberg: Semantic scene analysis of scanned 3d indoor environments, in: in: Proceedings of the Eighth International Fall Workshop on Vision, Modeling, and Visualization (VMV 03), 2003.

[8] Cantzler H., R. B. Fisher, M. Devy: Quality enhancement of reconstructed 3d models using co planarity and constraints, in: Proceedings of the 24th DAGM Symposium on Pattern Recognition, Springer-Verlag, London, UK, 2002, pp. 34–41.

[9] Eich M., M. Dabrowska, F. Kirchner: Semantic labeling: Classification of 3d entities based on spatial feature descriptors, in: IEEE International Conference on Robotics and Automation (ICRA2010) in Anchorage, Alaska, May 3, 2010.

[10] Vaskevicius N., A. Birk, K. Pathak, J. Poppinga: Fast detection of polygons in 3d point clouds from noise-prone range sensors, in: IEEE International Workshop on Safety, Security and Rescue Robotics, SSRR, IEEE, Rome, 2007, pp. 1–6.

[11] Oberlander J., K. Uhl, J. M. Zollner, R. Dillmann: A region-based slam algorithm capturing metric, topological, and semantic properties, in: ICRA'08, 2008, pp. 1886–1891.

[12] Castle R. O., G. Klein, D. W. Murray: Combining monoslam with object recognition for scene augmentation using a wearable camera 28 (11) (2010) 1548–1556.

[13] Pedraza L., G. Dissanayake, J. V. Miro, D. Rodriguez-Losada, F. Matia: Bs-slam: Shaping the world, in: Proceedings of Robotics: Science and Systems, Atlanta, GA, USA, 2007.

[14] Magnusson M., H. Andreasson, A. Nuchter, A. J. Lilienthal: Automatic appearance-based loop detection from 3D laser data using the normal distributions transform, Journal of Field Robotics 26 11-12 (2009) 892–914.

[15] Andreasson H., A. J. Lilienthal: Vision aided 3d laser based registration, in: Proceedings of the European Conference on Mobile Robots (ECMR), 2007, pp. 192–197.

[16] Besl P. J., H. D. Mckay: A method for registration of 3-d shapes, Pattern Analysis and Machine Intelligence, IEEE Transactions on 14 (2) (1992) 239–256

[17] Hahnel D., W. Burgard: Probabilistic matching for 3D scan registration, in: In.: Proc. of the VDI - Conference Robotik 2002 (Robotik), 2002.

[18] Rusu R. B., Z. C. Marton, N. Blodow, M. Dolha, M. Beetz: Towards 3d point cloud based object maps for household environments, Robot. Auton. Syst. 56 (11) (2008) 927–941.

[19] Craighead J., R. Murphy, J. Burke, B. Goldiez: A survey of commercial and open source unmanned vehicle simulators, in: Proceedings of ICRA, 2007.

[20] Boeing A., T. Bröaunl: Evaluation of real-time physics simulation systems, in: GRAPHITE '07: Proceedings of the 5th international conference on Computer graphics and interactive techniques in Australia and Southeast Asia, ACM, New York, NY, USA, 2007, pp. 281–288.

[21] Wang J., M. Lewis, J. Gennari: Usar: A game-based simulation for teleoperation, in: Proceedings of the 47th Annual Meeting of the Human Factors and Ergonomics Society, Denver, CO, Oct. 13–17, 2003.

[22] Rusu R. B., A. Maldonado, M. Beetz, I. A. Systems, T. U. Munchen: Extending player/stage/gazebo towards cognitive robots acting in ubiquitous sensor-equipped environments, in: in Accepted for the IEEE International Conference on Robotics and Automation (ICRA) Workshop for Network Robot System, 2007, April 14, 2007.

[23] Hohl L., R. Tellez, O. Michel, A. J. Ijspeert: Aibo and Webots: Simulation, Wireless Remote Control and Controller Transfer, Robotics and Autonomous Systems 54 (6) (2006) 472–-485.

[24] Buckhaults C.: Increasing computer science participation in the first robotics competition with robot simulation, in: ACMSE 47: Proceedings of the 47th Annual Southeast Regional Conference, ACM, New York, NY, USA, 2009, pp. 1–4.

[25] Craighead J., R. Murphy, J. Burke, B. Goldiez: A robot simulator classification system for HRI, in: Proceedings of the 2007 International Symposium on Collaborative Technologies and Systems (CTS 2007), 2007, pp. 93–98.

[26] Pepper C., S. Balakirsky, C. Scrapper: Robot simulation physics validation, in: PerMIS '07: Proceedings of the 2007 Workshop on Performance Metrics for Intelligent Systems, ACM, New York, NY, USA, 2007, pp. 97–-104.

[27] http://opencv.willowgarage.com/wiki/.

[28] Lee S.W., Y. J. Kim: Direct extraction of topographic features for gray scale character recognition, IEEE Trans. Pattern Anal. Mach. Intell. 17 (7) (1995) 724–729.

[29] Bedkowski J., M. Kacprzak, A. Kaczmarczyk, P. Kowalski, P. Musialik, A. Maslowski, T. Pichlak: Rise mobile robot operator training design, in: 15th International Conference on Methods and Models in Automation and Robotics (CD-ROM), Miedzyzdroje, Poland, 2010.

[30] NVIDIA PhysX SDK 2.8 - Introduction