

CREATING SEMANTIC MAPS FROM LASER TERRESTRIAL DATA

TWORZENIE MAP SEMANTYCZNYCH NA PODSTAWIE DANYCH Z NAZIEMNEGO SKANINGU LASEROWEGO

Janusz Będkowski¹, Karol Majek¹, Paweł Musialik¹, Andrzej Masłowski¹, Artur Adamek²

¹ Instytut Maszyn Matematycznych

² Wydział Geodezji i Kartografii, Politechnika Warszawska

KEY WORDS: semantic maps, terrestrial scanning, mobile robot navigation, path planning, parallel computation

ABSTRACT: In this paper creating semantic maps based on laser terrestrial data is shown. Semantic map is based on transformed geometric data (3D laser range finder) into the data with assigned labels. This labels can help in several applications such as navigation of mobile robot by finding traversable and not traversable regions. Computation of large 3D data sets requires high computational power, therefore we proposed the GPU based (Graphic Processing Unit) implementation to decrease the computational time. As a result we demonstrate the computed semantic map for mobile robot navigation.

1. INTRODUCTION AND RELATED WORK

This work is related with 3D laser scanning, semantic mapping, data registration as an important part of 6DSLAM (6 Degrees of freedom Simultaneous Localization And Mapping), 3D data processing using parallel computation on modern GPUs and path planning, therefore the State of the Art will be described for the most important aspects. Using semantic information (Asada *et al.*, 1989) extracted from 3D laser data (Nüchter *et al.*, 2005) is a relatively recent research topic in mobile robotics. In (Nüchter *et al.*, 2008) 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 (Grau, 1997) a model of an indoor scene is implemented as a semantic net. This approach is used in (Nüchter *et al.*, 2003) where a robot extracts semantic information from 3D models built from a laser scanner. In (Cantzler *et al.*, 2002) the location of features is extracted by using a probabilistic technique (RANSAC: RANdom SAmple Consensus) (Fischler *et al.*, 1980). It was shown in (Eich *et al.*, 2010) that also the region growing approach, extended from (Vaskevicius *et al.*, 2007) using k-nearest neighbor (KNN) search, is able to process unorganized point clouds. The improvement of plane extraction from 3D Data by fusing laser data and vision is shown in (Andreasson *et al.*, 2005).

Dealing with 3D cloud of points is related with 3D data registration problem which was introduced by Besl and McKay in (Besl *et al.*, 1992), and from that moment on, many

researchers were trying to solve the problem of augmenting the accuracy and the performance of aligning two cloud of points. Based on the State of the Art, we can state that the key issues of GPU based 3D data registration and important contributions are very close to the optimum, but not yet satisfactory. We deal with the problem of so-called approximated GPU ICP methods which affects different accuracy requirements than CPU-based methods. The goal of the Iterative Closest Point algorithm is to find the transformation matrix that minimizes the sum of distances between the corresponding points in two different data sets, therefore two important aspects have to be solved: the nearest neighbor search (NNS) and choosing the proper optimization technique for the minimization of the mentioned function (estimation 3D rigid transformation).

There are several GPGPU based approaches for the NNS in ICP (Iterative Closest Point) algorithm. An approach from (Bedkowski *et al.*, 2012) is using the regular grid decomposition (Rozen *et al.*, 2008), the another approach is using kd-tree (Qiu *et al.*, 2009). The NNS procedure is dominant compared to the rest of ICP algorithm, therefore many of researchers are trying to optimize the time of its execution. Another promising approach - an octree based NNS is shown in (Elseberg *et al.*, 2011). Authors claim that the octree based NNS does not suffer considerably more from larger maximal distances than the k-d tree based NNS, unfortunately there is observed an increased variance for the computing time. Nearest Neighborhood Search problem is well known also in Computer Graphics, where point-model (often referred to as a point-cloud) usually contain millions of points. For example authors of (Sankaranarayanan *et al.*, 2006) claim that their NNS approach can be used for the surface normals computation, mollification and noise removal. Using NNS for normal vector estimation is also shown in (Park *et al.*, 2006), where authors use Elliptic Gabriel Graph (EGG) for finding neighbors. EGG provides balanced neighbors by considering both distance and directional spread. It is emphasized the fact that in recent years there has been a marked shift from using triangles to using points as object modeling primitives especially in computer graphics applications (Pauly *et al.*, 2003). It is related with improved scanning technologies (Zampa *et al.*, 2009), that have also meaningful impact into mobile robotic applications. An important contribution to NNS problem is so-called KD-Tree data decomposition. It is already shown that GPU is more efficient than CPU in NNS problem, for example in (Garcia *et al.*, 2008) authors claim that kNN GPU is up to 400 times faster than a brute force CPUbased implementation of NNS. The comparison of nearest neighbor search strategies and implementations for efficient shape registration is shown in (Elseberg *et al.*, 2012). Authors emphasized the fact that most spatial data structures are hierarchical in nature, such as k-d trees (Bentley, 1975) and octrees (Meagher, 1982). They mentioned that a special type of NNS employs the Morton order (a space-filling curve (SFC)) for arranging the point cloud (Connor *et al.*, 2009). They also discussed the R-tree based algorithm for nearest neighborhood search. The result of this comparison is that since most libraries implement only the k-d tree, it is hard to draw final conclusions as to what data structure is better suited for NNS. The octree implementation was amongst the best performing algorithms. Authors claim that they have contributed their own novel open-source implementation of NNS and have shown these to perform well on realistic as well as artificial shape registration problems. Choosing the proper optimization technique in ICP like algorithms has been a research topic during last decades. A comparison of four algorithms for estimating 3D rigid transformation is shown in (Lorusso *et al.*, 1995). First algorithm proposed in (Arun *et al.*, 1987) is using Singular

Value Decomposition(SVD) for derive matrix. Second approach based on orthonormal matrices and computation of eigensystem of derived matrix is proposed in (Horn *et al.*, 1988). The third algorithm is shown in (Horn, 1987), it finds the transformation for the ICP algorithm by using unit quaternions. The forth algorithm shown in (Walker *et al.*, 1991) is using so-called dual quaternions. Instead of these four closed-form solution methods a novel linear solutions to the scan registration problem is shown in (Nüchter *et al.*, 2010). The advantage of these new linear solutions is that they can be extended straightforward to n-scan registrations. It was stated that under the assumption that the transformation (R,t) that has to be calculated by the ICP algorithm is small, it can be approximated the solution by applying instantaneous kinematics. This solution was initially given in (Pottmann *et al.*, 2002; Hofer *et al.*, 2003). The reported experiments have shown that the helix transform performs qualitatively as good as the uncertainty-based algorithm using Euler angles.

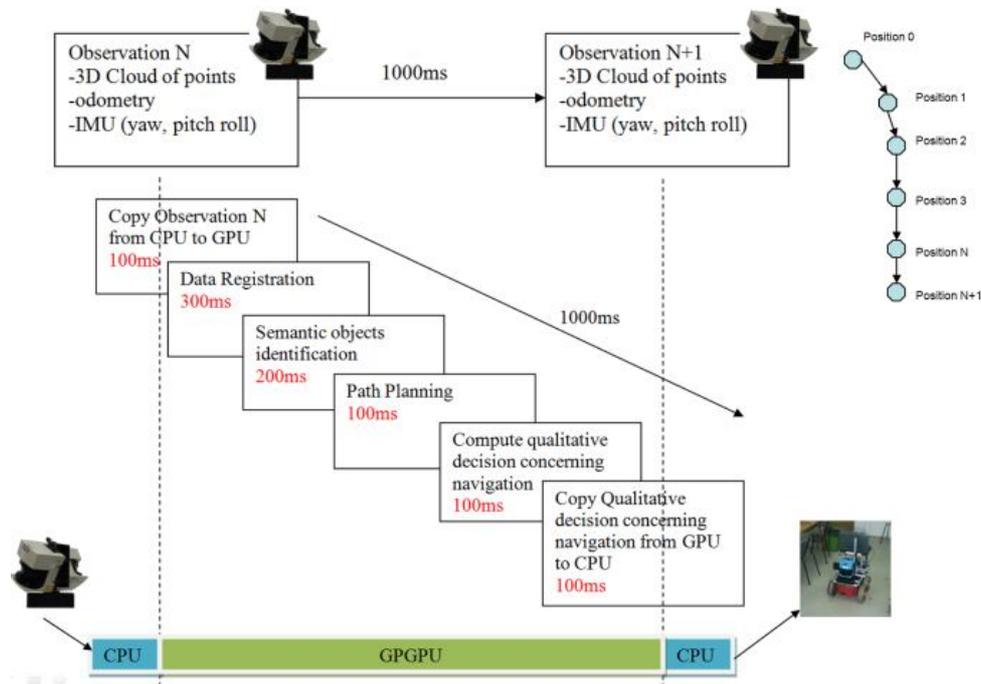


Fig. 1. GPGPU based robot control architecture.

It is important to mention some alternative methods for 3D data registration. Registration without ICP is shown in (Pottmann *et al.*, 2004). A competitive approach to ICP - Normal Distribution Transform is described in (Magnusson *et al.*, 2005). Multi-Scan alignment method based on the knowledge of partial correspondences is shown in (Teniente *et al.*, 2011). Instead of point to point scan matching there are another registration approaches such as point-to-plane (Teniente *et al.*, 2011). Furthermore, point-to-projection metrics are also possible (Blais *et al.*, 1995), it is done by matching points to ray indexes directly, inverting the ray casting process. The properties of these metrics are given in (Park S.Y., Su, 2003). It is important to emphasize the fact that point-to-plane method is faster

than point-to-point, but not always more accurate. The accuracy of point-to-plane like methods strongly depends on the surface approximation by local planes.

Complete 6DLSAM was analyzed in (Wulf *et al.*, 2008) and presented in (Sprickerhof *et al.*, 2009). Authors claim that to digitalize environments without occlusions, multiple 3D scans have to be registered. They discussed iterative ICP data registration (pairwise ICP) and so-called metascan technique. Both pairwise ICP and metascan ICP correct the robot pose estimates, but it is important to take into account that registration errors sum up. 6DSLAM algorithm uses the loop closing technique to bound these errors. Loop closing technique finds matchable scans, it appears when their Euclidean distance falls below a threshold (for example 5m), therefore these scans suppose to overlap. Local maps and robot poses are organized into a graph where edges correspond to each ICP registration. Once loop is detected, a 6-DoF graph optimization algorithm (Lu *et al.*, 1997) for global relaxation is employed. The result is globally consistent 3D map represented as a graph, where nodes are related with local 3D dense maps, edges are related with ICP registration modified if necessary by LUM. Similar approach can be found in (Pfaff *et al.*, 2007) where authors presented an approach to solve the SLAM problem with elevation maps generated also from three-dimensional range data acquired with a mobile robot. Their approach especially addresses the problem of acquiring such maps also with a ground-based vehicle. The approach classifies the individual cells of elevation maps into four classes representing parts of the terrain seen from above, vertical objects, overhanging objects such as branches of trees or bridges, and traversable areas. Authors claim that they extend of the ICP algorithm that takes this classification into account when computing the registration. Finally they claim that the consideration of the individual classes during the data association in the ICP algorithm provides more robust correspondences and more accurate alignments than classical approach.

In this paper we are focused on the semantic map obtained by the classification into several classes. This classes are used for robot navigation purpose.

2. ROBOT CONTROL ARCHITECTURE

GPU based robot control architecture is designed for 3D cloud processing during robot's motion. The architecture is shown on figure 1. There is constant interval of time between observations derived from 3D laser measurement unit. This time can be efficiently used for: transferring data from host (CPU) to device(GPU), data registration, semantic objects identification, path planning, computation of qualitative decision concerning navigation, transferring data(qualitative decision) from device to host.

3. SEMANTIC APPROACH FOR ROBOT PATH PLANNING

Semantic map is defined as a set of 3D points with assigned label. This map extends metric map with additional information (label), therefore we cope with four dimensional space $x,y,z,label$. This type of information can be used for traversability analysis and then for path planning. Motion planning is very important task for mobile robot working in unknown environment. Assuming that we have grid map describing robot environment, a trajectory of the robot can be computed using the modification of diffusion method described in (Siemiatkowska, 2008) improved by GPGPU computation. GPGPU implementation is using two dimensional grids of 512x512 cells. One grid is used for

initiation, where each cell can be free, occupied by the obstacle, occupied by a robot or occupied by a goal. Second grid is used for diffusion computation. The idea of usage GPGPU is to perform computation for each cell in parallel till diffusion reach robot position. To obtain robot trajectory we start from robot position and iteratively by finding local maximum in neighboring cells we are approaching the goal.

The idea of using semantics is to project 3D semantic map onto 2D grid. Algorithm 1 shows complete algorithm that results a path for given 3D cloud of points. First 3D cloud of points has to be registered. In the second step we compute normal vectors for each 3D data point using PCA/SVD (Principal Component Analysis/Singular Value Decomposition) method. We classify 3D points to assign semantic labels. For INDOOR we defined two classes flat region, not flat region, for OUTDOOR we defined four classes terrain, building, vegetation, unclassified. Traversability analysis is performed for each labeled 3D point that is projected onto grid, where white color corresponds to traversable region and black color corresponds to not traversable region. The trajectory is compute based on GPGPU implementation of the modified diffusion method.

Algorithm 1 Find path

registration of 3D data
 computation of normal vectors
 classification of 3D points
 traversability analysis
 path planning

1) Classification of 3D points in INDOOR environment: System classifies each point into one of two classes: flat region, not flat region. In first step vectors are computed, in second step we the condition of acceptable distance between points and approximation planes is checked. All points that violates this condition are marked as not flat region. This two classes enough to find traversable regions in indoor environment.

2) Classification of 3D points in OUTDOOR environment: System classifies each point into one of four classes: ground, building, vegetation, unclassified. The basis for the classification are vectors distributions. Next step finds the most populous horizontal plain, below the central point of the scan. We use RANSAC to find the exact parameters of the plain. After that each point which distance to the found plain is lower then a threshold is classified as ground. Classified points are then filtered out of the data 3D point cloud and are not used in the next steps. For the remaining points grouping into cells of 2x2 meters, in the horizontal plane is done. The groups are then preclassified based on these assumptions: buildings are mostly regular plains, trees and bushes are mostly irregular group of points. Therefore a cell in which most points are vegetation points should have a roughly uniform distribution of directions of normal vectors of the points, while cell that holds mostly building walls points should have most normal vectors pointing in one direction. Based on such reasoning the cells are classified in two groups: potential building or potential vegetation. The cell is considered potential building if normal vectors of at least half of the points in the cell are pointing in the same direction. For such cell RANSAC algorithm is used to find a plane from which at least half of the points in the cell lie in a certain threshold distance. If such a plain can be found the cell is classified as building. Each cell

that does not meet this condition is added to the group of potential vegetation. Vegetation hypothesis is checked by finding the dominant direction of normal vectors of each point in the cell. If the distribution of normal vectors directions for each point in the cell is roughly uniform the cell is considered vegetation. The cells that do not meet both conditions are classified as unclassified. After initial classification every cell is checked, by comparing it with its neighbors. The label is changed if it has a high chance of being an anomaly (a vegetation cell surrounded by building cells). Traversability is taken from the predefined properties of the classes. Buildings and vegetation are not traversable, while ground is traversable. For unclassified points the pessimistic version is chosen and they are considered not traversable.



Fig. 2. Indoor; Registered 3D cloud of points.



Fig. 3. Indoor: Reduced 3D cloud of points from figure 2 by 3D points above given threshold corresponding to the robot height.

4. EXPERIMENTS

Following set of figures demonstrate the result of proposed semantic approach for mobile robot path planning in indoor and outdoor environments. The goal was to register 3D clouds of points derived from different data source. In indoor data were gathered by mobile robot equipped with rotated SICK laser LMS100, in outdoor data were collected using Z+F IMAGER 5010 3D laser measurement system. The limitation is 64 millions of 3D points for data processing. Indoor environment is approximately 40m x 10m (2 millions of points), outdoor environment is approximately 200m x 200m (15 million of points). The computational time for each algorithm is considered as on-line, therefore it can be used by mobile robot executing task on the field with the assumption that it is equipped with NVIDIA GPGPU with compute capability 3.0.

5. CONCLUSIONS

Results can be summarized as follows: parallel computing drastically improves 3D cloud of points processing, semantic map can be used efficiently for mobile robot's path planning in indoor and outdoor environments, semantic mapping approach can be efficiently integrated with state of the art path planner such as based on wave propagation, we can process up to 64 million of 3D points in parallel in single step using modern GPGPU, proposed approach can be integrated with on-board computer of mobile robot. Proposed approach can be improved by extending set of classes. Semantic map is crucial for traversable analysis what extends the State of the Art path planning. Presented implementation can efficiently solve the problem of path planning for indoor and outdoor environments.

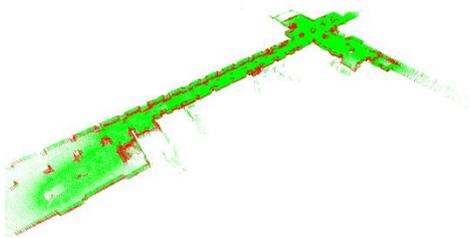


Fig. 4. Indoor: Semantic map as the result of classification of 3D points into two classes: green - flat, red - not flat

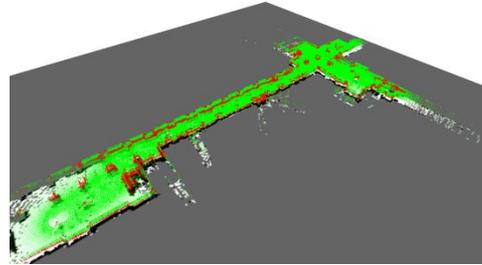


Fig. 5. Indoor: The result of traversability analysis. Green points are projected as traversable (white pixels), red as not traversable (black pixels).

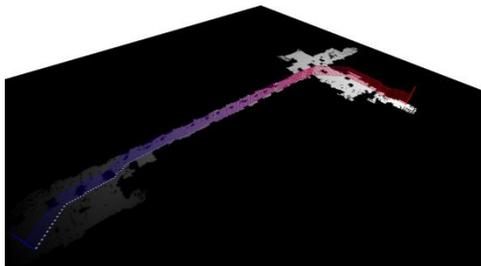


Fig. 6. Indoor: The result of path planning using wave propagation method.

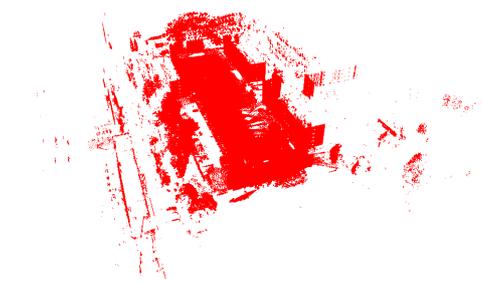


Fig. 7. Outdoor: Registered 3D cloud of points.

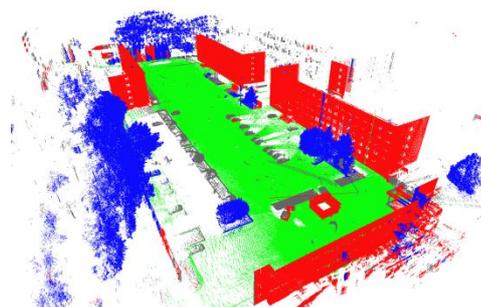


Fig. 8. Outdoor: Semantic map as a result of classification of 3D points into four classes: ground-green, building-red, vegetation-blue, unclassified-gray.

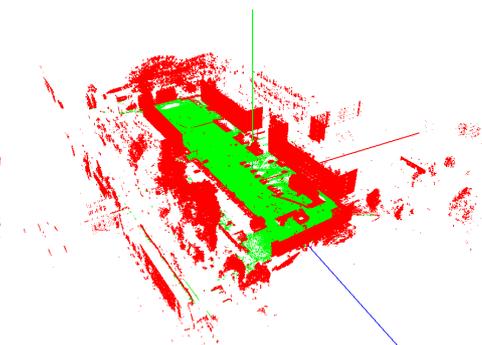


Fig. 9. Outdoor: The result of traversability analysis based on semantic map from figure 8. Green-traversable regions, red - not traversable regions.

- Eich M., Dabrowska M., Kirchner F., 2010. 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.
- Elseberg J., Borrmann D., Nüchter A., 2011. Efficient processing of large 3D point clouds, in: Information, Communication and Automation Technologies (ICAT), 2011 XXIII International Symposium on, s. 1–7.
- Elseberg J., Magnenat S., Siegwart R., Nüchter A., 2012. Comparison of nearest-neighbor-search strategies and implementations for efficient shape registration, *Journal of Software Engineering for Robotics (JOSER)* 3 (1), s. 2–12.
- Fischler M. A., Bolles R., 1980. Random sample consensus. a paradigm for model fitting with applications to image analysis and automated cartography, in: Proc. 1980 Image Understanding Workshop, s. 71–88.
- Garcia V., Debreuve E., Barlaud M., 2008. Fast k nearest neighbor search using GPU, 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (2), s. 1–6.
- Grau O., 1997. 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, p. 221.
- Hofer M., Pottmann H., 2003. Orientierung von Laserscanner-punktwolken, *Vermessung und Geoinformation* 91, s. 297–306.
- Horn B. K. P., 1987. Closed-form solution of absolute orientation using unit quaternions, *Journal of the Optical Society of America* 4 (4), s. 629–642.
- Horn B. K. P., Hilden H., Negahdaripour S., 1988. Closed-form solution of absolute orientation using orthonormal matrices, *Journal of the Optical Society America A* 5 (7), s. 1127–1135.
- Lorusso E. D., Fisher A., Fisher R., 1995. A comparison of four algorithms for estimating 3-D rigid transformations, in: Proceedings of the 1995 British Conference on Machine Vision (BMVC95), Birmingham, Press, Guildford, Vol. 1.
- Lu F., Milios E., 1997. Globally Consistent Range Scan Alignment for Environment Mapping, *Autonomous Robots* 4, s. 333–349.
- Magnusson M., Duckett T., 2005. A Comparison of 3D Registration Algorithms for Autonomous Underground Mining Vehicles, in: Proceedings of European Conference on Mobile Robots ECMR, s. 86–91.
- Meagher D., 1982. Geometric modeling using octree encoding, *Computer Graphics and Image Processing* 19 (2), s. 129–147.
- Nüchter A., Surmann H., Lingemann K., Hertzberg J., 2003. Semantic scene analysis of scanned 3d indoor environments, in: Proceedings of the Eighth International Fall Workshop on Vision, Modeling, and Visualization (VMV 03).
- Nüchter A., Wulf O., Lingemann K., Hertzberg J., Wagner B., Surmann H., 2005. 3D mapping with semantic knowledge, in: IN ROBOCUP INTERNATIONAL SYMPOSIUM, s. 335–346.
- Nüchter A., Hertzberg J., 2008. Towards semantic maps for mobile robots, *Robot. Auton. Syst.* 56 (11), s. 915–926.

- Nüchter A., Elseberg J., Schneider P., Paulus D., 2010. Study of parameterizations for the rigid body transformations of the scan registration problem, *Comput. Vis. Image Underst.* 114 (8), s. 963–980.
- Park S.Y., Subbarao M., 2003. A fast point-to-tangent plane technique for multi-view registration, in: *Proceedings of Fourth International Conference on 3-D Digital Imaging and Modeling, 3DIM*, s. 276–283.
- Park J., Shin H., Choi B., 2006. Elliptic Gabriel graph for finding neighbors in a point set and its application to normal vector estimation, *Computer-Aided Design* 38 (6), s. 619–626.
- Pauly M., Keiser R., Kobbelt L. P., Gross M., 2003. Shape modeling with point-sampled geometry, *ACM Trans. Graph.* 22 (3), s. 641–650.
- Pfaff P., Triebel R., Burgard W., 2007. An efficient extension to elevation maps for outdoor terrain mapping and loop closing, *Int. J. Rob. Res.* 26 (2), s. 217–230.
- Pottmann H., Leopoldseder S., Hofer M., 2002. Simultaneous registration of multiple views of a 3D object, *ISPRS Archives* 34 (3A), s. 265–270.
- Pottmann H., Leopoldseder S., Hofer M., 2004. Registration without ICP, *Computer Vision and Image Understanding* 95 (1), s. 54–71.
- Qiu D., May S., Nüchter A., 2009. GPU-Accelerated Nearest Neighbor Search for 3D Registration, in: *Proceedings of the 7th International Conference on Computer Vision Systems: Computer Vision Systems, ICVS09*.
- Rozen T., Boryczko K., Alda W., 2008. GPU bucket sort algorithm with applications to nearest-neighbor search, *WSCG* 16 (1-3), s. 161–167.
- Sankaranarayanan J., Samet H., Varshney A., 2006. A Fast k-Neighborhood Algorithm for Large Point-Clouds, in: *Eurographics Symposium on Point-Based Graphics 2006*.
- Siemiatkowska B., 2008. Coordinating the motion of mobile robots using cellular neural network, *Journal of Automation, Mobile Robotics and Intelligent Systems* 2 (2), 65–69.
- Sprickerhof J., Nüchter A., Lingemann K., Hertzberg J., 2009. An Explicit Loop Closing Technique for 6D SLAM, in: I. Petrovic, A. J. Lilienthal (Eds.), *Proceedings of the 4th European Conference on Mobile Robots, ECMR 09, September 23-25, 2009, Mlini/Dubrovnik, Croatia, KoREMA*, s. 229–234.
- Teniente Avils E. H., Andrade Cetto J., 2011. FaMSA: Fast multi-scan alignment with partially known correspondences, in: *Proceedings of European Conference on Mobile Robots ECMR*, s. 139–144.
- Vaskevicius N., Birk A., Pathak K., Poppinga J., 2007. 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*, s. 1–6.
- Walker M. W., Shao L., Volz R. A., 1991. Estimating 3-d location parameters using dual number quaternions, *CVGIP: Image Underst.* 54 (3), s. 358–367.
- Wulf O., Nüchter A., Hertzberg J., Wagner B., 2008. Benchmarking urban six-degree-of-freedom simultaneous localization and mapping, *J. Field Robot.* 25 (3), s. 148–163.
- Zampa F., Conforti D., 2009. Mapping with Mobile Lidar, *GIM International* 23 (4), s. 35–37.

TWORZENIE MAP SEMANTYCZNYCH NA PODSTAWIE DANYCH Z NAZIEMNEGO SKANINGU LASEROWEGO

SŁOWA KLUCZOWE: mapy semantyczne, skanowanie naziemne, nawigacja robota mobilnego, planowanie trasy, obliczenia równoległe

Streszczenie

W niniejszej pracy zostało przedstawione tworzenie map semantycznych na podstawie danych z naziemnego skanowania laserowego. Mapa semantyczna bazuje na danych pomiarowych z przypisanymi etykietami. Te etykiety mogą zostać wykorzystane w wielu aplikacjach, jak nawigacja robota mobilnego z wykorzystaniem podziału na regiony przejezdne i nieprzejezdne. Obliczenia dużych trójwymiarowych zbiorów danych wymaga zastosowania dużej mocy obliczeniowej, dlatego zaproponowaliśmy implementację wykorzystującą GPU (*Graphic Processing Unit*), by zmniejszyć czas obliczeń. W rezultacie prezentujemy mapę semantyczną do nawigacji robota mobilnego.

Dane autorów:

dr inż. Janusz Będkowski
e-mail: januszbedkowski@gmail.com

inż. Karol Majek
e-mail: karolmajek@gmail.com

mgr inż. Paweł Musialik
e-mail: PJMusialik@gmail.com

prof. dr hab. inż. Andrzej Masłowski
e-mail: a.maslowski@imm.org.pl

mgr inż. Artur Adamek
e-mail: artex_pl@yahoo.com