

Barbara SIEMIĄTKOWSKA, Jacek SZKLARSKI, Michał GNATOWSKI, Arkadiusz ZYCHEWICZ  
 INSTYTUT PODSTAWOWYCH PROBLEMÓW TECHNIKI PAN

## Budowa hybrydowej semantyczno-rastrowej reprezentacji otoczenia robota mobilnego na podstawie wskazań dalmierza laserowego 3D

Dr Barbara SIEMIĄTKOWSKA

Barbara Siemiątkowska pracuje w Instytucie Automatyki i Robotyki Politechniki Warszawskiej i w Instytucie Podstawowych Problemów Techniki PAN. Autorka jest absolwentką Wydziału Matematyki Mechaniki i Informatyki Uniwersytetu Warszawskiego. Pracę doktorską z dziedziny zastosowań sztucznej inteligencji w robotyce obroniła w Instytucie Podstawowych Problemów Techniki PAN. Główną dziedziną zainteresowań naukowych autorki jest zastosowanie sieci komórkowych w nawigacji robotów mobilnych.

e-mail: bsiem@ippt.gov.pl



Dr inż. Jacek SZKLARSKI

Jacek SzklarSKI od 2008r. jest pracownikiem Instytutu Podstawowych Problemów Techniki PAN. Jest absolwentem Wydziału Informatyki Politechniki Szczecińskiej oraz Wydziału Matematyczno-Fizycznego Uniwersytetu Szczecińskiego. Doktorat z nauk fizycznych obronił na Uniwersytecie w Poczdamie, pracował też w ośrodku naukowym w Dreźnie.

e-mail: jszklar@ippt.gov.pl



Dr inż. Michał GNATOWSKI

Pracuje w Instytucie Podstawowych Problemów Techniki PAN. Jest absolwentem Wydziału Elektrycznego Politechniki Warszawskiej. Praca doktorska dotyczyła zastosowania metod wieloagentowych w współdziałaniu robotów mobilnych. Zainteresowania naukowe autora obejmują również budowanie map otoczenia robotów mobilnych.

e-mail: mignat@ippt.gov.pl



Mgr Arkadiusz ZYCHEWICZ

Jest doktorantem Instytutu Podstawowych Problemów Techniki PAN. Ukończył Wydziału Fizyki Uniwersytetu Warszawskiego. Praca doktorska dotyczy budowania map otoczenia robotów mobilnych na podstawie wskazań dalmierza laserowego 3D.

e-mail: azychew@ippt.gov.pl



### Streszczenie

W artykule przedstawiono metodę budowy hybrydowej rastrowo-obiektowej mapy otoczenia mobilnego na podstawie wskazań skanera laserowego 3D. Chmura punktów jest zapisywana w postaci zbioru wektorów normalnych. Składowe wektora są reprezentowane jako składowe RGB. Przeprowadzana jest segmentacja obrazu, a następnie dokonuje się klasyfikacji semantycznej. W procesie klasyfikacji wykorzystuje się cechy Haara oraz systemy regulowe. Każdy wykryty obiekt jest przypisywany do pewnej komórki mapy rastrowej. Mapa utworzona w ten sposób może być następnie wykorzystana w algorytmie nawigacyjnym – ułatwia współpracę robot-człowiek oraz planowanie trasy. Metoda jest kontynuacją algorytmu opisanego w artykule pt. "Segmentacja danych otrzymanych z dalmierza laserowego 3D".

**Słowa kluczowe:** nawigacja, reprezentacja środowiska.

### Construction of a dual metric-semantic map of the mobile robot environment based on data from a 3D laser scanner

#### Abstract

The major goal of our current research is to build a dual grid-based and semantic map of an unknown indoor environment based on data obtained from a 3D laser scanner. In this paper main steps concerning object classification are presented. A point cloud from the 3D scanner is transformed into a set of normal vectors, which are then represented as a RGB raster image where each color component corresponds to  $x$ ,  $y$ ,  $z$  coordinates of the vectors. In the next step we apply some standard methods from image analysis, like flood-filling and object detection using Haar-like features, in order to perform segmentation and find objects of our interest in the examined scene. Afterwards we use 3D geometrical information and relation between the detected objects in our classification process. If a semantic meaning can be assigned to any object it is used for building a dual metric-semantic map of the environment. Such a map should serve as a basic element for human-robot interaction.

**Keywords:** navigation, 3D mapping.

### 1. Wstęp

W robotyce mobilnej coraz więcej uwagi poświęca się tworzeniu robotów usługowych, które bezpośrednio komunikują się

z człowiekiem. Aby komunikacja była możliwa bardzo istotny jest sposób reprezentacji otoczenia. W literaturze najczęściej opisywane są algorytmy tworzenia metrycznych rastrowych lub wektorowych map 2D [1, 2]. Reprezentacja 2D jest jednak niewystarczająca w sytuacji, gdy robot przemieszcza się na zewnątrz budynków lub pomieszczeniach, w których przeszkody występują na różnej wysokości. Coraz częściej rozwijane są systemy, w których tworzone są mapy 3D. Najczęściej opisywane sposoby zapisu reprezentacji 3D to:

- chmura punktów [3] – w metodzie tej zapisywane są wszystkie punkty pomiarowe,
- reprezentacja 2.5D [4] – mapa rastrowa, w której dla każdej komórki zapisywana jest maksymalna obserwowana wysokość,
- siatka trójkątów [5] – reprezentacja powszechnie stosowana w grafice komputerowej, w której powierzchnie są opisywane za pomocą zbioru trójkątów,
- woksele [6] – trójwymiarowe komórki mapy rastrowej,
- mapy wielowarstwowe [7] są uogólnieniem map 2.5D,
- mapy semantyczne [8] – wykrytym przeszkodom przypisuje się etykiety określające znaczenie semantyczne.

Wadami reprezentacji za pomocą chmury punktów, siatki trójkątów i wokseli jest to, że wymagana jest duża pojemność pamięci. Reprezentacja 2.5D jest bardzo efektywnym sposobem zapisu postaci otoczenia, ale nie możemy reprezentować obiektów umieszczonych jeden nad drugim np. podłogi i wiszącej na dużej wysokości szafek. Rozwiązaniem tego problemu jest stworzenie reprezentacji wielowarstwowej, w której do każdej komórki mapy rastrowej 2D przypisywana jest lista „prostopadłościanów”, które opisują przeszkody znajdujące się na różnej wysokości. Zaletą tego typu reprezentacji jest mała zajętość pamięci, takie mapy są wygodne w planowaniu trasy, ale nie mogą być bezpośrednio wykorzystane w systemach interakcji człowiek-robot. Wiele prac z dziedziny robotyki jest poświęconych tworzeniu map semantycznych tzw. map obiektów. W większości zakłada się, że robot przemieszcza się we wnętrzu pomieszczenia, w którym większość obiektów możemy opisywać za pomocą fragmentów płaszczyzn [9]. W [10] przedstawiono system, w którym na podstawie danych ze stereowizji określane jest położenie ścian i podłogi w obserwowanym pomieszczeniu. W artykule [11] opisano metodę,

w której klasyfikowane są dane uzyskane ze skanera laserowego 3D i przypisywane są etykiety: ściana, podłoga, sufit, drzwi. Wadą map semantycznych jest jednak to, że mogą być wykorzystane w procesie planowania trasy dopiero po zapisaniu w postaci map np. 2.5D. Biorąc pod uwagę zalety i wady wszystkich wymienionych powyżej metod reprezentacji autorzy artykułu zdecydowali się na zapisywanie otoczenia robota w postaci mapy hybrydowej: rastrowo - obiektowej. Ten sposób zapisu możemy traktować jako uogólnienie map wielowarstwowych. Otoczenie robota jest dzielone na podobszary (klatki 2D), a następnie do każdego podobszaru przypisywana jest lista obiektów. Do każdego obiektu przypisywana jest etykieta semantyczna oraz informacja metryczna. W obecnej wersji systemu obiekty są opisywane przez podanie opisu krawędzi. Algorytm składa się z następujących etapów:

1. zebranie danych pomiarowych,
2. zapamiętanie danych w postaci kolorowego obrazu,
3. segmentacja obrazu,
4. klasyfikacja,
5. aktualizacja mapy.

Sposób zbierania danych pomiarowych oraz segmentacji chmury punktów został opisany w artykule [16] W tym opracowaniu jest opisana metoda klasyfikacji obszarów i tworzenia mapy.

Po przeprowadzaniu opisanej segmentacji danych trójwymiarowych obszary są klasyfikowane. Zastosowano trzy rodzaje klasyfikatorów:

- system regułowy, analizujący cechy geometryczne,
- system uczony pod nadzorem,
- system regułowy, analizujący teksturę obrazu i cechy geometryczne.

## 2. Klasyfikator regułowy

W otoczeniu typu wnętrze pomieszczenia, duża część chmury punktów należy do płaszczyzn dominujących – ścian, podłogi, sufitu. Obszary te są klasyfikowane przy pomocy systemu regułowego. Cechy które są uwzględniane w algorytmie to: szerokość obszaru, wysokość obszaru, położenie, relacja z innymi obszarami. W omawianym algorytmie zdefiniowano następujące reguły:

- *podłoga* – obszar poziomy (opisywany przez wartości: (128, 128, 255)), o wartościach  $z \approx 0$ .
- *sufit* – najwyżej położony obszar poziomy (wartości RGB=(128, 128, 0)).
- *ściana* – obszar pionowy opisywany przez parę wartości ( $h_s$ ,  $w_s$ ), gdzie  $h_s$  jest wysokością obszaru, a  $w_s$  – szerokością, dodatkowo spełnione są warunki:  
 $h_s > 0.9 \text{dist}(\text{podłoga}, \text{sufit})$  i  $w_s > w$ ,  
gdzie  $\text{dist}(\text{podłoga}, \text{sufit})$  jest odległością między podłogą, a sufitem,  $w$  jest parametrem, w prowadzonych eksperymentach przyjęto  $w = 2$  m.
- *drzwi* – obszar pionowy, sąsiadujący ze ścianą, opisywany przez parametry: ( $h_d$ ,  $w_d$ ),  $h_d$  – wysokość obszaru a  $w_d$  – szerokość. Przyjmujemy, że  $h_d \approx 2$  m,  $w_d \approx 0,8$  m

Ponieważ zwykle obraz dużych powierzchni składa się z wielu punktów pomiarowych, aby skrócić czas obliczeń określone są kontury fragmentów obrazu, a następnie obliczane są wartości wymienionych powyżej parametrów uwzględniając jedynie piksele należące do wykrytych krawędzi. Na rys. 1 przedstawiono kontury obszarów sceny przedstawionej na rys. 2 artykułu pt. "Segmentacja danych otrzymanych z dalmierza laserowego 3D".

Na rys. 2 przedstawiono wynik klasyfikacji obszarów. Widoczna jest ściana, drzwi, sufit, obszary nie zaklasyfikowane zaznaczono kolorem żółtym.

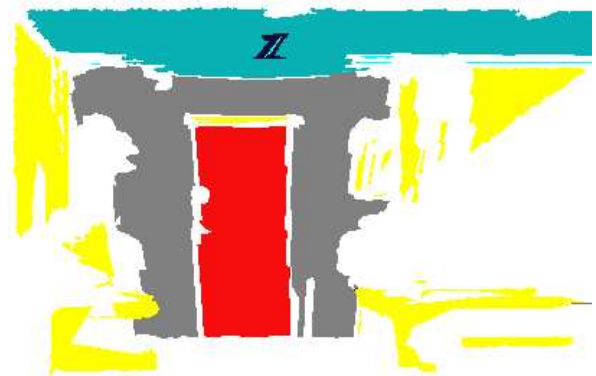
## 3. Cechy Haara

Przekształcenie pomiarów z lasera w obraz RGB daje możliwość wykorzystania standardowych metod przetwarzania i analizy obrazów w odniesieniu do danych reprezentujących odległość do przeszkód. W niniejszej pracy prezentujemy jak wykorzystać tzw. klasyfikator cech Haara do rozpoznawania i klasyfikacji określo-

nej klasy obiektów. Użyto algorytmów zaimplementowanych w bibliotece OpenCV, zaproponowanych w pracach [11] dla podstawowych cech Haara i [12] dla zestawu obróconych cech. Po wytrenowaniu klasyfikatora do rozpoznawania określonych obiektów, możliwa jest bardzo szybka klasyfikacja dla zebranych pomiarów. Dzięki temu omawiana metoda może mieć praktyczne zastosowania w systemach klasyfikacji działających w czasie rzeczywistym. Ponadto dzięki natychmiastowo dostępnej informacji o geometrycznych rozmiarach rozpoznanych obiektów możliwa jest dodatkowa weryfikacja takich obiektów i odrzucenie fałszywie rozpoznanych obszarów (ang. *false positives*). Przykładowo, jeśli klasyfikator rozpoznaje obiekt typu „umywalka”, wiadomo, że rozmiary rzeczywiste powinny leżeć w określonym zakresie.



Rys. 1. Wykryte krawędzie obszarów przedstawionych na rys. 6 pracy [16]  
 Fig. 1. Contours of polygons formed by edge-detection method applied for each area shown in Fig. 6 presented in [16]



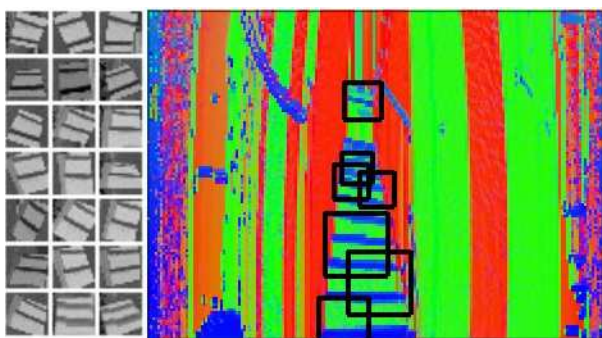
Rys. 2. Wynik klasyfikacji. Widoczne są drzwi (czerwone), ściana (szara), sufit (zielony)  
 Fig. 2. 3D scene with areas labeled by the rule base classification system

Obrazy wygenerowane z pomiarów z lasera mają oczywiście nieco inne właściwości niż zwykłe obrazy z kamery wizyjnej. Przykładowo iluminacja i zmiany oświetlenia nie mają tu żadnego wpływu na obraz końcowy – zarówno w jasnym świetle z dużą ilością cieni, jak i w ciemnościach będzie on taki sam. Z drugiej strony, na skutek przemieszczania się robota mobilnego niektóre składowe obrazy mogą ulegać zmianie – zależne jest to od zastosowanego modelu barw i metody mapowania normalnych na kolory. W przypadku analizy cech Haara nie ma to zasadniczego znaczenia, gdyż rozpatrywane obrazy i tak najpierw przekształcają się na odcienie szarości.

Przeprowadzone badania wykazały, że pewne metody konwersji normalnych na kolory dają lepsze wyniki w rozpoznawaniu obiektów za pomocą cech Haara niż inne: np. użycie wartości bezwzględnych składowych normalnych  $|nx|$ ,  $|ny|$ ,  $|nz|$  zamiast  $nx$ ,  $ny$ ,  $nz$ , i zamiana ich na RGB. W takim przypadku wszystkie powierzchnie równoległe do osi patrzenia robota, w szczególności

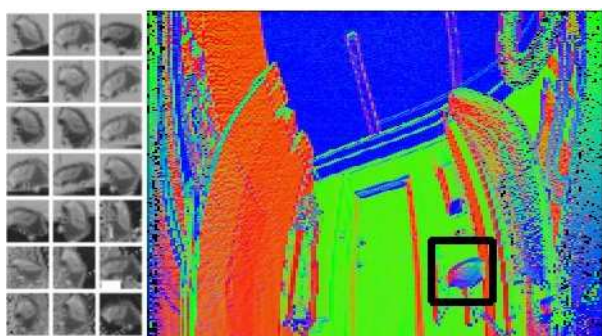
sufit i podłoga, będą niebieskie. Z drugiej strony wszystkie powierzchnie prostopadłe do robota, takie jak ściany, będą miały niebieską składową równą zeru. Ponadto ściany biegnące wzdłuż osi patrzenia robota będą zawsze czerwone.

Pierwszym etapem przy budowie klasyfikatora cech Haara jest wytrenowanie go w celu detekcji określonych obiektów. W niniejszej pracy pokazujemy przykłady jak wykrywać obiekty typu „umywalka” i „schody” widziane z perspektywy robota mobilnego. Aby wytrenować klasyfikator niezbędny jest duży zbiór obrazów-próbek zawierających rozpoznawany obiekt („pozytywne”) i obrazów nie zawierających takiego obiektu („negatywne”). Wszystkie takie obrazy powinny mieć jednakowy rozmiar, tutaj używamy 20x20. Jako zbiór „negatywnych” używamy wycinki z dużej losowej próby obrazów, przy czym żaden wycinek nie reprezentuje rozpoznawanego obiektu (czyli „umywalkę” lub „schodów” w tym przypadku). W celu otrzymania zbioru „pozytywnych” próbek reprezentujących np. „umywalkę” branych jest kilka obrazów z różnymi umywalkami, następnie poddawane są one całej serii różnych przekształceń takich jak obroty wokół trzech osi czy losowej zmianie jasności i nanoszone na losowe tło. Na skutek tych przekształceń otrzymywany jest duży zbiór tysięcy próbek pozytywnych wykorzystywanych dalej do treningu klasyfikatora.



Rys. 3. Wynik klasyfikacji – rozpoznawanie schodów. Z lewej strony znajdują się przykłady uczące, z prawej wynik klasyfikacji

Fig. 3. Classification results – stairs recognition. Patterns on the left, results of classification on the right



Rys. 4. Wynik klasyfikacji – rozpoznawanie umywalki. Z lewej strony znajdują się przykłady uczące, z prawej wynik klasyfikacji

Fig. 4. Classification results – washbasin recognition. Patterns on the left, results of classification on the right

Po zakończeniu procesu trenowania, klasyfikator można zastosować do dowolnie wybranego fragmentu obrazu wygenerowanego na podstawie nowych pomiarów z lasera. Statystyczne właściwości fragmentu wykorzystywane są w kaskadowym klasyfikatorze i jeśli wszystkie fazy klasyfikacji zakończą się pozytywnie, oznaczać to będzie, że wskazany fragment zawiera obraz rozpoznawanego obiektu. Analiza fragmentu obrazu jest bardzo szybka, więc może być zastosowana do wielu różnych regionów o różnych rozmiarach w celu przeszukania całego obrazu. Rys. 3 i 4 pokazują testowe wyniki takiej analizy dla klasyfikatorów wytrenowa-

nych na rozpoznawanie obiektu „umywalka” i „schody”. Do testowania oczywiście nie są wykorzystywane obrazy zawierające obiekty wykorzystane już wcześniej do treningu.

Obszary oznaczane jako zawierające wyszukiwany obiekt przez klasyfikator cech Haara są w dalszych etapach użyte do klasyfikacji semantycznej i tworzenia mapy semantycznej. Oczywiście w takim podejściu każdy obiekt wymagający rozpoznania musi mieć swój własny, wytrenowany klasyfikator cech Haara. W bardziej złożonym podejściu planuje się dodatkowo wykorzystanie w klasyfikacji klasycznych obrazów wizyjnych, tak aby obie metody uzupełniały się.

#### 4. Analiza tekstury

Oprócz testów przeprowadzanych wewnątrz budynku testowano również system przypisywania znaczeń semantycznych obszarom znajdującym się na zewnątrz. Zbudowano system, który umożliwi określenie na podstawie danych sensorycznych rodzaj podłoża, po którym porusza się robot. Wstępne testy umożliwiły przeprowadzenie klasyfikacji obszarów na trzy grupy: chodnik (preferowany obszar przejezdny), trawa (obszar przejezdny, ale nie zalecany) oraz teren niedostępny dla robota. Na rys. 5a przedstawiono zdjęcie otoczenia, na rys. 5b kolorowy obraz utworzony na podstawie wartości wektorów normalnych.

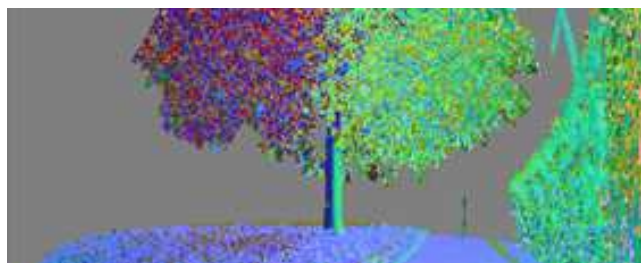
Piksele będące obrazem chodnika mają wartość  $z$  (wysokość) oraz barwę bardziej wyrównaną, niż będące obrazem liści. Piksele będące obrazem liści są bardziej zróżnicowane kolorystycznie niż należące do obszarów „ściana budynków”, czy „pień”.

Informacja o typie powierzchni, po której przemieszcza się robot jest bardzo wygodna w procesie planowania bezkolizyjnej trasy – robot może unikać przemieszczania się po pewnych obszarach np. trawie, ale w sytuacji gdy jest to konieczne może po niej przejechać. Tworząc klasyfikator powierzchni zastosowano analizę tekstury obrazu przy pomocy sieci komórkowych [14].

a)



b)



Rys. 5. Wektory normalne dla otoczenia znajdującego się na zewnątrz budynków

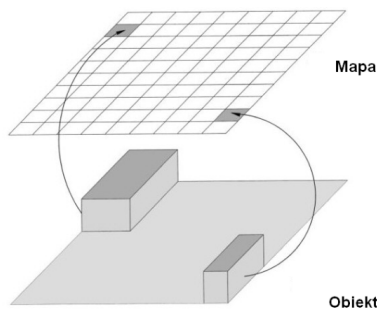
Fig. 5. 3D Normal vector for the outdoor environment



## 5. Budowa reprezentacji otoczenia

Otoczenie jest reprezentowane w postaci hybrydowej, rastrowo-semantycznej. Po wykryciu i przyporządkowaniu etykiet obiektów zostają one przypisane do odpowiednich komórek mapy rastrowej.

Idea metody jest przedstawiona na rys. 6.



Rys. 6. Rastrowo-obiektowa reprezentacja otoczenia  
Fig. 6. Dual grid and object representation of the environment

Z każdym obiektem oprócz znaczenia semantycznego wiążemy pewną wagę z przedziału  $[0; 1,0]$ , która określa przejezdność danego obszaru. Dla obiektów typu podłoga, chodnik waga wynosi 1,0, dla obszarów zajętych przez przeszkody waga wynosi 0,0. Taki sposób zapisu sceny z jednej strony umożliwia wydawanie poleceń typu *jeźdź do drzwi*, czy *jeźdź wzdłuż ściany*. Z drugiej strony wartości wag mogą być wykorzystywane w procesie planowania trasy. Metody planowania trasy umożliwiające uwzględnienie różnego typu podłoża zostały opisane w [15].

## 6. Podsumowanie

W artykule opisano metodę tworzenia hybrydowej semantyczno-rastrowej mapy otoczenia na podstawie danych pochodzących ze skanera laserowego 3D. Taki sposób zapisu umożliwia stworzenie systemu nawigacyjnego, w którym polecenia będą wydawane w języku naturalnym np. *jeźdź do drzwi*, a robot planując trasę uwzględni nie tylko, to czy dany obszar jest zajęty przez przeszkody czy wolny, ale również rodzaj mijanych przeszkód. Pojazd może więc, np. zmniejszać prędkość zbliżając się do drzwi, unikać pewnych obszarów (np. jazdy po trawie). Przedstawione badania będą nadal kontynuowane. Planujemy z jednej strony rozszerzać system klasyfikacji obiektów, wprowadzając większą bazę etykiet. Z drugiej strony chcemy rozwijać system planowania trasy tak, aby robot mógł sprawnie przemieszczać się w środowisku zmieniającym się dynamicznie.

## 7. Literatura

- Prowadzone badania są finansowane przez Ministerstwo Nauki, Szkolnictwa Wyższego i Informatyzacji w ramach grantu nr 4311/B/T02/2007/33.*
- [1] Elfes A.: Sonar-based real-world mapping and navigation. IEEE Trans. on Robotics and Automation, pp. 249–265, 1987.
  - [2] Thrun S., Burgard W., and Fox D.: Probabilistic Robotics. MIT-Press, Cambridge, Massachusetts, USA, 2005.
  - [3] Rusu R. B., Marton Z., C., Blodow N., Dolha M., and Beetz M.: Towards 3d point cloud based object maps for household environment. Journal of Robotics and Autonomous Systems, 56:927–941, 2008.
  - [4] Gu J., Cao Q., and Huang Y.: Rapid traversability assesment in 2.5d grid based map on rough terrain. Int. Journal of Advanced Robotic Systems, 5(4):389–394, 2008.
  - [5] Schroeder W., Zarge J., and Lorensen W.: Decimation of triangle meshes. Computer Graphics, pp. 65–70, 1992.
  - [6] Sakas G. and Hartig J.: Interactive visualization of large scalar voxel fields. In Vizualization, pp. 29–36, 1992.
  - [7] Triebel R., Pfaff P., and Burgard W.: Multi-level surface maps for outdoor terrain mapping and loop closing. Proc. of IROS, pp. 1 - 2, 2006.
  - [8] Mozos O. M., Triebel R., Jensfelt P., Rottman A., and Burgard W.: Supervised semantic labeling of places using information extracted from sensor data. Robotics and Autonomous Systems, 5(55): 392 – 402, 2007.
  - [9] Weingarten J., and Siegwart R.: EKF-based 3D SLAM for structured environment reconstruction. In Proc. of IROS 2005, 2005.
  - [10] Iocci L., Konolige K., Bajracharya M.: Visually realistic map ping of a planar environment with stereo. Experimental robotics 7, pp. 521-532, 2001.
  - [11] Intel Corp. Opencv library, 1999-present.
  - [12] Lienhart R., and Maydt J.: An extended set of Haar-like features for rapid object detection. IEEE ICIP 2002 1:900–903, 2002.
  - [13] Gonzales R. C., Woods R. E.: Digital Image Processing. Prentice Hall, 2002.
  - [14] Siemiątkowska B., Gnatowski M., and Chojecki R.: Cellular neural networks in 3D laser data segmentation. In 9th WSEAS Int. Conf. on NEURAL NETWORKS, pp. 84–88, 2008.
  - [15] Siemiątkowska B.: Uniwersalna metoda planowania działań robota mobilnego wykorzystująca architekturę uogólnionych sieci komórkowych. Oficyna Wydawnicza PW, Warszawa, 2009.
  - [16] Siemiątkowska B., Szklarski J., Gnatowski M., Zychiewicz A.: Segmentacja danych otrzymanych z lasera 3D. PAK, 03, 2010.

otrzymano / received: 27.09.2009  
przyjęto do druku / accepted: 03.02.2010

artykuł recenzowany

## INFORMACJE

# Zapraszamy do publikacji artykułów naukowych w czasopiśmie PAK

WYDAWNICTWO POMIARY AUTOMATYKA KONTROLA  
ul. Świętokrzyska 14A, pok. 530, 00-050 Warszawa,  
tel./fax: 22 827 25 40

Redakcja czasopisma POMIARY AUTOMATYKA KONTROLA  
44-100 Gliwice, ul. Akademicka 10, pok. 30b,  
tel./fax: 32 237 19 45, e-mail: wydawnictwo@pak.info.pl