

mgr inż. Marcin Undak
Instytut Automatyki i Robotyki PW
dr Barbara Siemiątkowska
Instytut Automatyki i Robotyki PW
Instytut Podstawowych Problemów Techniki PAN

ZASTOSOWANIE SIECI KOMÓRKOWYCH DO TWORZENIA SEMANTYCZNYCH MAP OTOCZENIA ROBOTA MOBILNEGO

W pracy przedstawiono wyniki badania powierzchni, po której przemieszcza się robot mobilny. Dane uzyskiwane są ze skanera laserowego 3D, zmieniany jest sposób reprezentacji – z biegunowej na kartezjańską. W kolejnym kroku dane 3D są zamieniane na kolorowy obraz 2D, który jest następnie poddawany procesom segmentacji i analizy. Klasyfikacja danych pomiarowych 3D jest przeprowadzana na podstawie wyników analizy obrazów i informacji metrycznej. Wyodrębniono następujące rodzaje powierzchni w pomieszczeniach zamkniętych: podłoga, schody, podesty, progi oraz powierzchnie typu trawa i chodnik, które znajdują się na zewnątrz budynków. Testy przeprowadzono w rzeczywistym środowisku przy pomocy robota Elektron z zamontowanym dalmierzem laserowym 3D.

THE APPLICATION OF CELLULAR NEURAL NETWORK IN SEMANTIC MAP BUILDING

The paper presents the results of 3D data classification. The data obtained from 3D laser are transformed from the polar to the Cartesian coordinate system. The data is represented as a colored image. Then the image is analyzed. The classification of 3D data is performed based on the result of image segmentation and metric information. The experiments were performed in indoor and outdoor environments. The following objects were classified successfully: floor, stairs, grass, sidewalk.

1. WSTĘP

Jednym z podstawowych zagadnień robotyki mobilnej jest planowanie bezkolizyjnej trasy. Zwykle zadanie to jest wykonywane na podstawie mapy otoczenia, która jest zadana lub tworzona automatycznie przez pojazd na podstawie danych sensorycznych. Do niedawna powszechnie stosowano mapy dwuwymiarowe – metryczne [1, 2, 3, 4] (rastrowe, wektorowe) lub topologiczne [5, 6]. Zaletą tego typu reprezentacji jest to, że ich tworzenie może odbywać się w czasie rzeczywistym, ale jednocześnie elementy otoczenia, które znajdują się powyżej lub poniżej płaszczyzny skanowania nie są widoczne na mapie.

Od kilkunastu lat rozwijane są systemy skanowania 3D. Najbardziej naturalną metodą, bazującą na zasadzie działania ludzkiego wzroku jest stereowizja [7]. Do pomiaru wykorzystuje się dwie kamery, które są od siebie oddalone o znaną odległość. Wzajemna korelacja dwóch zarejestrowanych obrazów tego samego obiektu pozwala na wydobycie informacji metrycznej. Opisywana metoda jest tania, ale stosowanie stereowizji niesie ze sobą wiele problemów. Pierwszym z nich jest niewielki zasięg, w praktyce ograniczony do kilku metrów. Kolejnym jest długi czas analizy obrazów w przypadku zmiennego oświetlenia.

Kolejną metodą tworzenia reprezentacji 3D jest zastosowanie skanerów laserowych 3D. W tego typu urządzeniach obrotowe lustro, znajdujące się w głowicy skanującej odchyła wiązkę w zadanym kierunku. Taki mechanizm umożliwia osiągnięcie szybkości pomiaru do

100 000 punktów na sekundę. Podstawowym problemem są błędy pomiaru w sytuacji, gdy wiązka pada na krawędź obiektu. W takiej sytuacji część fali powraca, a część biegnie dalej, w wyniku czego odczyt odległości jest wartością znacznie większą niż wartość rzeczywista. Paradoksalnie zwiększenie rozdzielczości skanowania potęguje możliwość wystąpienia tego typu zakłóceń – im wyższa rozdzielczość tym większe prawdopodobieństwo błędu.

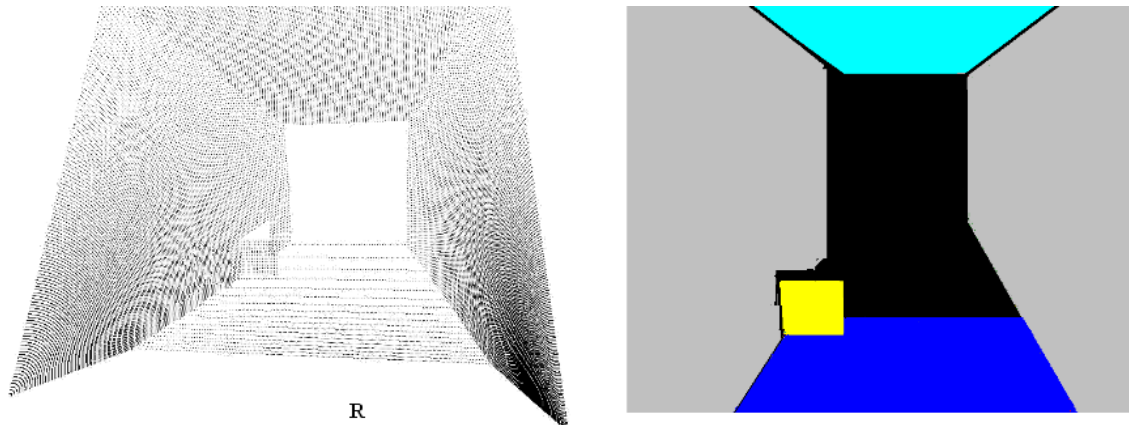
1.1 Metody reprezentacji 3D

W literaturze opisywanych jest wiele metod budowy reprezentacji otoczenia na podstawie danych pochodzących z sensorów 3D. Poniżej wymienione zostaną najważniejsze z nich:

- *Chmura punktów* [8] – w reprezentacji tej pomiary zapisywane są w postaci listy współrzędnych punktów, które opisują fragmenty przeszkód. Metoda ta wymaga ogromnej pojemności pamięci, nie jest wygodna w procesie planowania trasy.
- *Reprezentacja 2.5D* [9] – jest to mapa rastrowa 2D, której komórki przechowują informacje o wysokości wykrytych przeszkód. Ta forma zapisu ułatwia planowanie trasy, wymagana pojemność pamięci jest niewielka. W reprezentacji tej nie można zapisać otoczenia, w którym przeszkody umieszczone są na różnych wysokościach, jedna nad drugą.
- *Wielowarstwowa reprezentacja* [10] – jest to mapa rastrowa 2D, której komórki przechowują listę opisu przeszkód znajdujących się na różnych wysokościach. Metoda posiada zalety reprezentacji 2.5D, jednocześnie możliwe jest opisanie złożonego środowiska.
- *Woksele* [11] – jest mapą rastrową 3D, reprezentacja wymaga dużo pamięci, jest stosowana w przypadku trójwymiarowej wizualizacji obrazów medycznych. W robotyce może być przydatna w procesie planowania trasy po zamianie na inną reprezentację np. mapę przejezdności.
- *Siatka wieloboków* [12] – metoda stosowana w grafice komputerowej, mapa jest reprezentowana jako lista wieloboków, których wierzchołki wyznaczone są przez chmurę punktów.
- *Reprezentacje semantyczne* [13] – mapa jest zapisywana w formie listy obiektów znajdujących się w pomieszczeniu. Metoda umożliwia wydawanie robotowi poleceń zapisanych w języku zbliżonym do naturalnego, ale nie jest wygodna w procesie planowania trasy.

Na rys. 1 przedstawiono zdjęcie otoczenia (korytarza), zebraną chmurę punktów oraz reprezentację semantyczną mapy. Opis tworzenia tego typu reprezentacji znajduje się w [14].

W naszym algorytmie proponujemy zastosowanie reprezentacji, która jest połączeniem reprezentacji 2.5D i reprezentacji symbolicznej. Komórki mapy rastrowej oprócz informacji o wysokości komórek przechowują również dane dotyczące rodzaju obiektu, który w danej komórce występuje. Informacja taka ma bardzo duże znaczenie w procesie planowania bezkolizyjnej trasy np. schody stanowią przeszkodę dla pojazdów kołowych, ale są dostępne dla humanoidów. Trawa jest obszarem po którym powinna przemieszczać się automatyczna kosiarka, ale nie powinien po niej jechać robot kołowy. Powierzchnie znajdujące się na tej samej wysokości mogą się różnić kosztem przejazdu, wymagają także różnych parametrów kinematycznych.

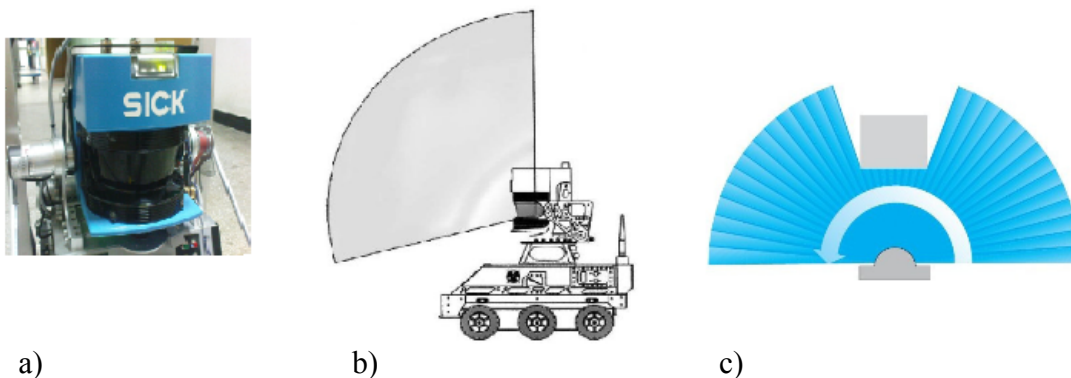


b) c)

Rys. 1. Reprezentacja semantyczna: a) zdjęcie otoczenia, b) chmura punktów, c) mapa semantyczna, kolorem ciemno-niebieskim zaznaczono obszar zaklasyfikowany jako podłoga, jasno-niebieskim sufit, kolorem szarym zaznaczono ściany, a żółtym fragmenty nie przypisane do żadnej z klas.

2. SKANER LASEROWY SICK LMS 200

W prowadzonych eksperymentach stosowano skaner laserowy LMS 200 firmy Sick. Urządzenie zostało umieszczone na głowicy, która umożliwia odchylenie skanera w pionie (rys. 2b) w zakresie od -10° do 90° . Zakres skanowania w poziomie wynosi 180° , a rozdzielczość może być ustawiona na $0,25^{\circ}$, $0,5^{\circ}$ lub $1,0^{\circ}$. Skaner laserowy 3D został zaprojektowany i wykonany na Wydziałach Elektroniki i Technik Informatycznych oraz Mechatroniki Politechniki Warszawskiej. Zdjęcie skanera, sposób skanowania przedstawiono na rys. 2.



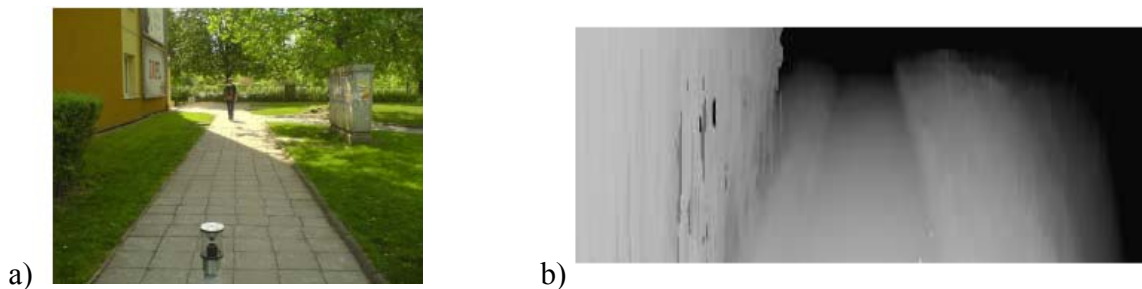
a) b) c)

Rys. 2. Skaner laserowy: a) zdjęcie skanera LMS 200, b) skanowanie w pionie, c) skanowanie w poziomie.

Dane przekazywane są przez dalmierz w układzie pseudo-biegunowym: podawany jest kąt odchylenia serwomechanizmu, kąt skanowania w poziomie oraz odległość. Dane są zapisywane na trzy sposoby:

- **Mapa głębi**

Kąty skanowania w pionie i poziomie wyznaczają indeksy tablicy, w której zapamiętywane są dane, natomiast komórki przechowują wartości odległości, które zostały otrzymane ze skanera 3D. Na rys. 3 przedstawiono zdjęcie otoczenia oraz mapę głębi.



Rys. 3. Dane uzyskane ze skanera 3D: a) zdjęcie otoczenia, b) mapa głębi

- **Chmura punktów**

Dane mogą być przedstawione w postaci chmury punktów. Wcześniej musimy jednak zamienić reprezentację biegunową na kartezjańską. Sposób zapisu został opisany w [14].

- **Kolorowy obraz 3D**

Dane przechowywane w postaci chmury punktów są w następnym kroku przetwarzane na reprezentację w postaci kolorowego obrazu. Dla każdego punktu (p_x, p_y, p_z) poszukujemy pary punktów p^1, p^2 , takich, że punkty (p, p^1, p^2) nie są współliniowe i leżą na jednej powierzchni, tzn. odległość m.in. jest mniejsza niż przyjęty próg. W kolejnym kroku obliczane są współrzędne dwóch wektorów n^1, n^2 :

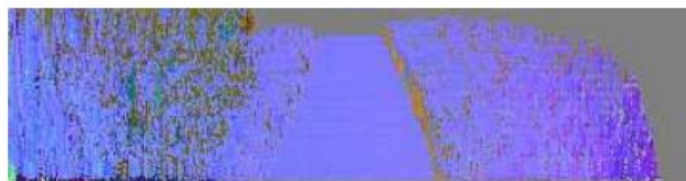
$$\begin{aligned} n^1 &= [p_x^1 - p_x, p_y^1 - p_y, p_z^1 - p_z] \\ n^2 &= [p_x^2 - p_x, p_y^2 - p_y, p_z^2 - p_z] \end{aligned} \quad (1)$$

Następnie obliczany jest wektor n określony następująco:

$$n = n^1 \times n^2 \quad (2)$$

W kolejnym kroku składowe wektora $n=[n_x, n_y, n_z]$ są reprezentowane jako składowe RGB.

W ten sposób na podstawie chmury punktów otrzymujemy kolorowy obraz. Wynik stosowania opisywanego algorytmu dla danych przedstawionych na rys. 3 przedstawia rys. 4.

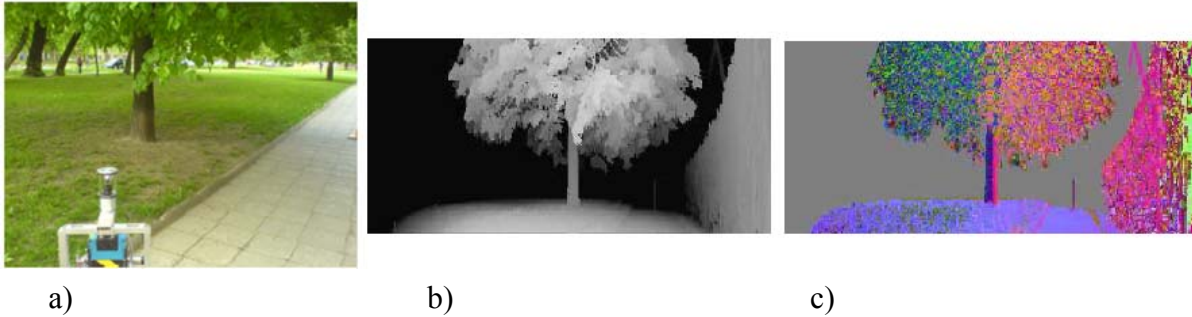


Rys. 4. Kolorowy obraz utworzony dla danych przedstawionych na rys. 3

Na rys. 5 i rys. 6 przedstawiono wyniki dla innego typu środowiska.



Rys. 5. Kolorowy obraz: a) zdjęcie otoczenia, b) mapa głębi, c) obraz normalnych

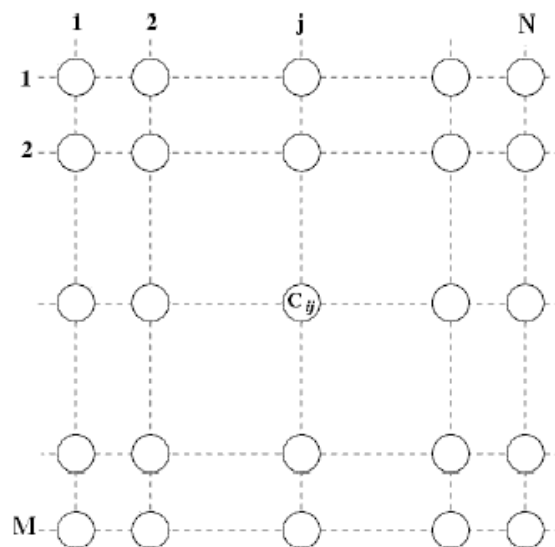


Rys. 6. Kolorowy obraz: a) zdjęcie otoczenia, b) mapa głębi, c) obraz normalnych

W kolejnym kroku powierzchnia, po której przemieszcza się robot, jest klasyfikowana na podstawie tekstury oraz informacji metrycznej. Analiza danych jest przeprowadzana w dużym stopniu przy pomocy klasycznych metod przetwarzania obrazu. Klasyczne algorytmy to usuwanie szumów, rozrost ziarna. W pracy [14] przedstawiono zastosowanie cech Haara, które umożliwiają rozpoznawanie obiektów typu schody. Metodami rozrostu ziarna możemy wyodrębnić obszary jednobarwne. Stosując systemy regułowe możemy wyodrębnić pochylnie, drzwi, podłogę, sufit [14]. Jeśli jednak robot przemieszcza się na zewnątrz budynków, klasyfikację należy przeprowadzać na podstawie tekstury. Obszary jednobarwne są najczęściej fragmentami obiektów stworzonych przez człowieka: są to fragmenty domów, chodnik czy jezdnia. Obiekty występujące w naturze np. liście, trawa są na kolorowych obrazach utworzonych na podstawie wartości wektorów normalnych reprezentowane przez zbiory pikseli o barwach zmiennych. Na tej podstawie możemy odróżnić powierzchnie typu trawa od powierzchni typu chodnik, a analizując informację metryczną możemy wykryć obiekty, które są fragmentami przeszkód np. budynki, drzewa.

3. SIECI KOMÓRKOWE

Narzędziem informatycznym, które jest od wielu lat stosowane w przetwarzaniu obrazu są sieci komórkowe[15]. Strukturę sieci przedstawia rys. 7. Neurony ułożone są w formie prostokątnej siatki i połączone są w obrębie pewnego sąsiedztwa.

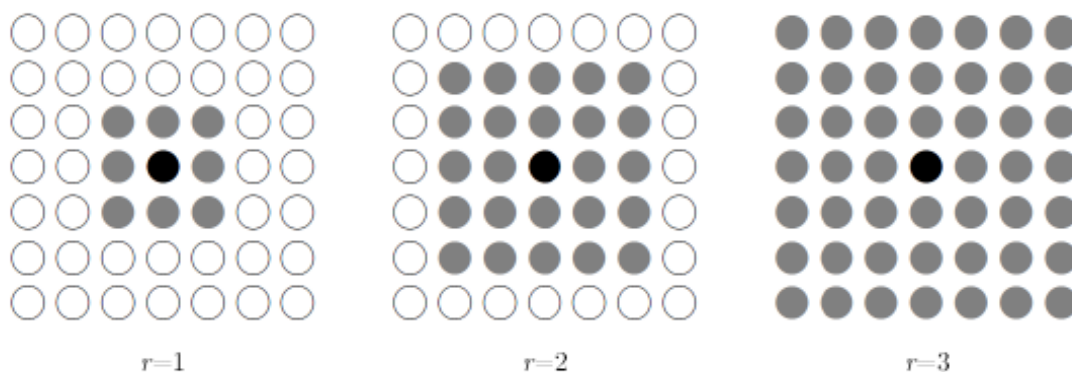


Rys. 7. Struktura sieci

Wszystkie komórki przetwarzają sygnał w identyczny sposób. Sygnałami sterującymi są:

- sygnały wejściowe u_{ij} komórek należących do r -sąsiedztwa
- sygnały wyjściowe y_{ij} komórek należących do r -sąsiedztwa
- sygnał I zwany polaryzacją.

Wagi połączeń określające oddziaływanie sygnałów wyjściowych komórek sąsiedztwa a_{ij}^{kl} jest indeksem komórki, która jest sterowana, a kl jest indeksem źródła sterowania. Wagi określające stopień sterowania sygnałami wejściowymi oznaczamy symbolami b_{ij}^{kl} . W swoich dalszych pracach Chua rozszerzył swoją teorię i wprowadził uogólnione sieci komórkowe, w których nie wymagany jest powtarzalny szablon oraz sygnały wejściowe i wyjściowe mogą być wielowymiarowe. Dzięki temu rozwiązaniu znacznie rozszerzył się zakres możliwych zastosowań CNN.



Rys. 8. Rodzaje sąsiedztwa

W kolejnych pracach definicja sieci komórkowej została zmodyfikowana. Neuronową siecią komórkową nazywamy wielowymiarową strukturę elementów połączonych lokalnie.

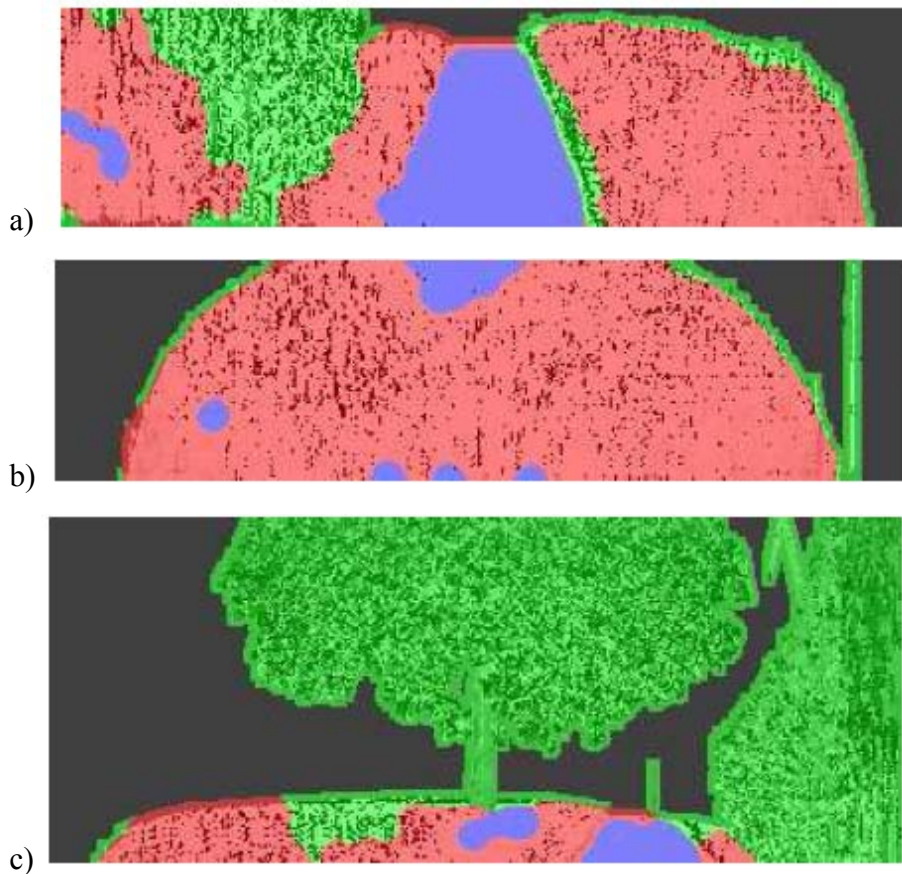
4. ZASTOSOWANIE SIECI KOMÓRKOWYCH DO ANALIZY POWIERZCHNI

Sygnałem wejściowym do sieci jest wektor, w którym przechowywane są wartości $[z, R, G, B]$, gdzie z jest informacją metryczną i określa wysokość na której obserwowany jest dany punkt pomiarowy, a $[R, G, B]$ są składowymi wektora normalnego. Jeśli powierzchnia wokół badanego punktu jest fragmentem płaszczyzny (chodnika), to wektory normalne do powierzchni w danym punkcie pomiarowym i punktach sąsiednich są do siebie równoległe. Na obrazie są one widoczne jako obszary jednobarwne. Sytuację tą obserwujemy na rys. 11. Na rys. 11a przedstawiono zdjęcie otoczenia, a na rys. 11b obraz normalnych. Obraz obiektów stworzonych przez człowieka (skrzynka sterownicza, słupek) składa się z obszarów jednobarwnych, natomiast obraz naturalnych elementów otoczenia cechuje duża zmienność tekstury (trawa, liście). Klasyfikacja składa się z dwóch etapów:

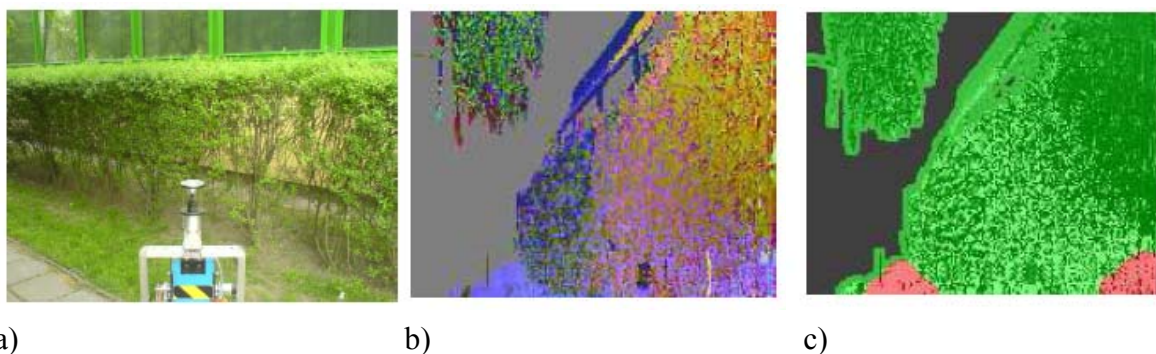
- podział pikseli obrazu na należące do powierzchni należącej do podłoża po którym przemieszcza się robot i pozostałych elementów, klasyfikacja ta jest przeprowadzana na podstawie wartości z .
- W kolejnym kroku na podstawie wartości $[R, G, B]$ klasyfikowane są wszystkie punkty należące do powierzchni, po której przemieszcza się robot i dzielone są na dwie grupy: chodnik, trawa i pozostałe.

5. EKSPERYMENTY

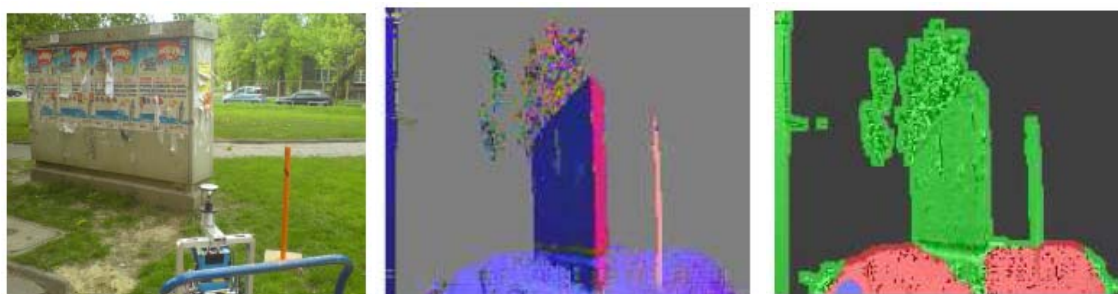
Eksperymenty były prowadzone z wykorzystaniem rzeczywistego skanera laserowego, we wnętrzu i na zewnątrz budynku. Wyniki klasyfikacji wewnątrz pomieszczeń zostały opisane w [14]. W eksperymentach prowadzonych na zewnątrz punkty były dzielone na cztery klasy: chodnik (niebieski), trawa (czerwony), przeszkody (zielony), pozostałe (czarne). Kolorem czarnym zaznaczono punkty dla których brak danych (odległość przekracza zadany próg). Wynik klasyfikacji był zależny od przyjętego sąsiedztwa w sieci komórkowej. Na rys. 9 – 11 przedstawiono wynik klasyfikacji dla promienia sąsiedztwa = 3.



Rys. 9. Wynik klasyfikacji



Rys. 10. Klasyfikacja powierzchni: a) zdjęcie otoczenia, b) obraz normalnych, c) wynik klasyfikacji



a)

b)

c)

Rys. 11. Klasyfikacja powierzchni: a) zdjęcie otoczenia, b) obraz normalnych, c) wynik klasyfikacji

6. WNIOSKI

W pracy przedstawiono koncepcję klasyfikacji podłoża na podstawie danych pochodzących ze skanera laserowego. Dane 3D pochodzące ze skanera laserowego zostały przedstawione w formie kolorowego obrazu. Proces klasyfikacji danych jest przeprowadzany na podstawie analizy tekstury obrazu. Przeprowadzone badania mają charakter wstępny i zastosowano wiele uproszczeń – powierzchnia po której przemieszcza się robot jest płaska (brak wzniesień), nie klasyfikujemy powierzchni typu kamienie, wgłębienia itd. Niemniej uzyskane wstępne wyniki prac są zachęcające. Przewidujemy przeprowadzać klasyfikację większej liczby obiektów. Informacja ta będzie przydatna w procesie nawigacji semantycznej.

Prace zostały wykonane w ramach grantu MEiN. 4311/B/T02/2007/33.

7. LITERATURA

- 1 Elfes A.: Sonar-based real-world mapping and navigation. *IEEE Transactions on Robotics and Automation*, pp. 249–265, 1987.
- 2 Thrun S., Burgard W., Fox D.: Probabilistic Robotics. *MIT-Press*, Cambridge, Massachusetts, USA, 2005.
- 3 Siemiątkowska B.: Uniwersalna metoda planowania działań robota mobilnego wykorzystująca architekturę uogólnionych sieci komórkowych, *Oficyna Wydawnicza PW*, Warszawa, 2009.
- 4 Simmons R., Koenig S.: Probabilistic navigation in partially observable environments. *Proceedings of IJCAI'95*. Morgan Kaufman Publishing, 1995.
- 5 Shatkay H., Kaelbling L. P.: Learning topological maps with weak local odometric information. In *Proceedings of the 15th International Joint Conference on Artificial Intelligence*, 1997.
- 6 Dedeoglu G., Mataric M., Sukhatme G. S.: Incremental, online topological map building with a mobile robot. in: *Proceedings of Mobile Robots XIV – SPIE*, 1999.
- 7 Iocci L., Konolige K., Bajracharya M.: Visually realistic mapping of a planar environment with stereo, *Experimental robotics 7*, pp. 521-532, 2001.

- 8 Rusu R. B., Marton Z. C., Blodow N., Dolha M., Beetz M.: Towards 3D point cloud based object maps for household environment. *Journal of Robotics and Autonomous Systems*, 56:927–941, 2008.
- 9 Siemiątkowska B., Gnatowski M., Zychewicz A.: Fast Method of 3D Map Building Based on Laser Range Data, *Journal of Automation Mobile Robotics & Intelligent Systems*, (1), nr 2, pp. 35 – 39, 2007.
- 10 Triebel R., Pfaff P., Burgard W.: Multi-level surface maps for outdoor terrain mapping and loop closing. in: *Proc. of IROS*, pp. 1 - 2, 2006.
- 11 Sakas G. and Hartig J.: Interactive visualization of large scalar voxel fields. in: *Vizualization*, pp. 29–36, 1992.
- 12 Schroeder W., Zarge J., Lorensen W.: Decimation of triangle meshes. *Computer Graphics*, pp. 65–70, 1992.
- 13 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): pp. 392 – 402, 2007.
- 14 Siemiątkowska B., Szklarski J., Gnatowski M., Borkowski A.: Towards Semantic Navigation System. *Proc. of Intelligent Information System*, pp. 711–720, 2009.