

dr Barbara Siemiątkowska
 Politechnika Warszawska, Wydział Mechatroniki
 Instytut Podstawowych Problemów Techniki PAN
 mgr Arkadiusz Zychewicz
 dr inż. Michał Gnatowski
 Instytut Podstawowych Problemów Techniki PAN,

ZASTOSOWANIE METOD PRZETWARZANIA OBRAZÓW W SEGMENTACJI DANYCH POCHODZĄCYCH Z DALMIERZA LASEROWEGO 3D

W poniższej pracy zaprezentowana zostanie metoda tworzenia map semantycznych otoczenia robota mobilnego. Robot jest wyposażony w dalmierz laserowy, który umożliwia zbieranie trójwymiarowej informacji o środowisku. Dane są zapamiętywane jako kolorowy obraz. Proces segmentacji danych i tworzenia trójwymiarowej mapy otoczenia składa się z następujących etapów: wyodrębnienie obszarów jednorodnych, opisanie krawędzi i powierzchni obszarów, klasyfikacja, tworzenie grafu opisującego scenę. Przyjęto założenie, że klasyfikacja odbywa się na podstawie wprowadzonych wcześniej do bazy wiedzy zbioru reguł i dodatkowej informacji przechowywanej w etykietowanych grafach. Przeprowadzone eksperymenty wykazały, że segmentacja powierzchni 3D przy pomocy klasycznych metod przetwarzania obrazów, które są stosowane w widzeniu maszynowym umożliwia przeprowadzanie obliczeń w sposób efektywny. Proces tworzenia mapy semantycznej jest nadal opracowywany, ale wstępne wyniki są zadawalające.

PICTURE SEGMENTATION METHODS IN 3D LASER DATA SEGMENTATION

In the article a method of building semantic map of robots' environment is presented. A robot is equipped with a laser sensor which enables to obtain 3 dimensional information of the scene. Data is stored in a colour image. The segmentation and map building processes consist of the following parts: selecting homogeneous areas in the image obtained from the data, the edges and the areas description, classification, creating a graph which describes the scene. It is assumed, that classification is performed based on previously stored rules database and additional information stored in labelled graphs. Experiments which have been done showed that 3D area segmentation, using classical pattern recognition methods which are widely used in machine vision enables to perform computations effectively. The semantic map creation process is still under development but the performed results are satisfied.

1. WSTĘP

Autonomiczny robot mobilny, aby wykonywać powierzone mu zadania powinien posiadać mapę otoczenia. Mapa taka może być zadana lub tworzona na podstawie wskazań sensorów. W literaturze wymienione są następujące sposoby przechowywania informacji: reprezentacja rastrowa [1], mapa cech [2], topologiczny opis sceny [3], reprezentacja hybrydowa rastrowo-topologiczna, reprezentacja hybrydowa metryczno-metryczna [4]. W większości przypadków tworzona jest mapa dwuwymiarowa. Mapa taka może być jednak niewystarczająca w sytuacji, gdy robot porusza się na zewnątrz budynków lub jeśli przeszkody umieszczone są na różnej wysokości. W takiej sytuacji konieczna jest reprezentacja 3D. W robotyce stosowane są układy stereowizyjne [5,6], ale wymagają one dużych mocy obliczeniowych oraz stabilnych warunków oświetleniowych. Często mapa środowiska tworzona jest na podstawie

wskazań dalmierza laserowego 3D [7] lub kamer metrycznych, które oprócz informacji o barwie obiektów przekazują także informację o odległości. Tworząc mapę trójwymiarową podstawowy problem, jaki należy rozwiązać to bardzo duża liczba danych. Czas wykonywania operacji takich jak agregacja i filtracji danych powinien być krótki. W literaturze spotyka się wiele sposobów reprezentacji map trójwymiarowych [8-12], jedna z nich to tworzenie 2,5D lub trójmiarowych map rastrowych. Inny powszechnie stosowany w grafice komputerowej sposób polega na generowaniu siatki wieloboków opisujących powierzchnie obiektów znajdujących się w środowisku. W niedalekiej przyszłości roboty mobilne będą uczestniczyły w codziennym życiu człowieka, dlatego sposób opisu otoczenia powinien uwzględniać oprócz cech metrycznych także cechy semantyczne sceny. Dzięki temu można będzie wydawać robotowi polecenia typu: jedź do kuchni i przynieś talerz. Konieczne więc jest stworzenie algorytmów, które umożliwiają tworzenie map semantycznych. Opisywane w tej pracy metody są w fazie eksperymentów, ale otrzymane przez nas wyniki wydają się być na tyle interesujące, aby je zaprezentować na konferencji *Automation 2009*.

Proponowany w naszej pracy algorytm składa się z następujących etapów:

- zbieranie danych pomiarowych
- zapisanie danych w postaci obrazu kolorowego,
- segmentacja obrazu,
- trójwymiarowa reprezentacja obszarów jednorodnych,
- klasyfikacja obszarów, na podstawie bazy wiedzy,
- tworzenie semantycznej mapy otoczenia.

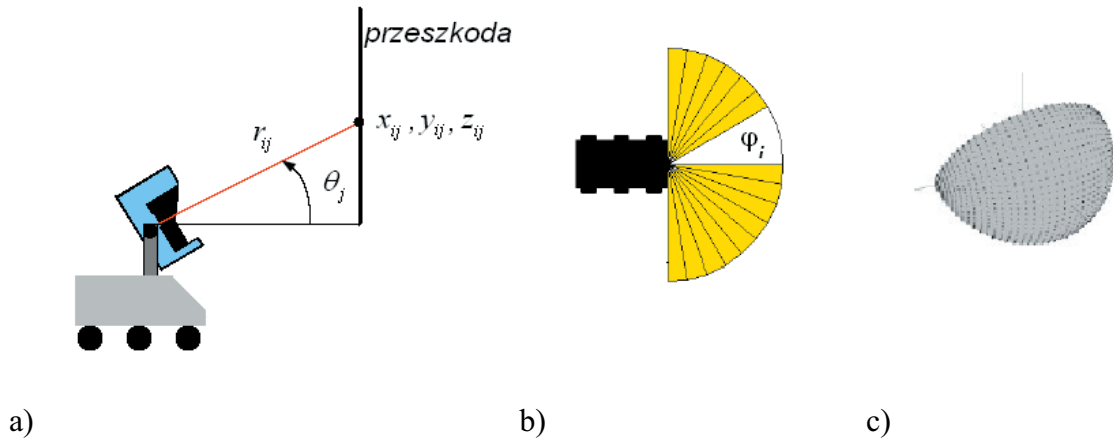
2. ZBIERANIE DANYCH POMIAROWYCH

Dane zbierane są przy pomocy dalmierza laserowego SICK LNS 200, zamontowanego na głowicy obrotowej umożliwiającej obrót lasera z pozycji poziomej do pionowej. Układ został wykonany i zaprojektowany na wydziałach: Mechatroniki oraz Elektroniki i Technik Informacyjnych Politechniki Warszawskiej. Dane otrzymywane z dalmierza możemy zapisywać w postaci trójki liczb: $\{ r_{ij}, \theta_j, \varphi_i \}$, gdzie r_{ij} jest odległością, θ_j - jest kątem obrotu serwomechanizmu, φ_i jest kątem skanowania dalmierza. Na rys. 1 przedstawiono przestrzeń skanowania oraz sposób zbierania danych.

Otrzymane w ten sposób dane mogą być przedstawione w postaci macierzy o rozmiarze $M \times N$. M określa liczbę różnych położeń głowicy skanującej. N jest liczbą odczytów otrzymanych z dalmierza laserowego w danym cyku skanowania. Ponieważ skaner zbiera dane z rozdzielczością 1° z przedziału $0^\circ - 180^\circ$ to $N = 181$. Parametr M zależy od przyjętej rozdzielczości obrotu serwomechanizmu.

Tak otrzymane dane są przeliczane z biegunowego do kartezjańskiego układu współrzędnych według wzorów:

$$\begin{aligned}
 x_{ij} &= r_{ij} \sin \theta_j \sin \varphi_i \\
 y_{ij} &= r_{ij} \sin \varphi_i \\
 z_{ij} &= r_{ij} \cos \theta_j \sin \varphi_i
 \end{aligned}
 \tag{1}$$



Rys. 1. Zbieranie danych pomiarowych: a) kąt obrotu serwomechanizmu, b) kąt skanowania, c) przestrzeń skanowania

Kolejnym etapem jest klasyfikacja punktów. Dla każdego punktu (z wyłączeniem punktów brzegowych) obliczany jest wektor \vec{p}_{ij} normalny do powierzchni [13], do której należy dany punkt [12]. Wektor ten jest wyznaczony jako iloczyn wektorowy p_{ij} określony następująco:

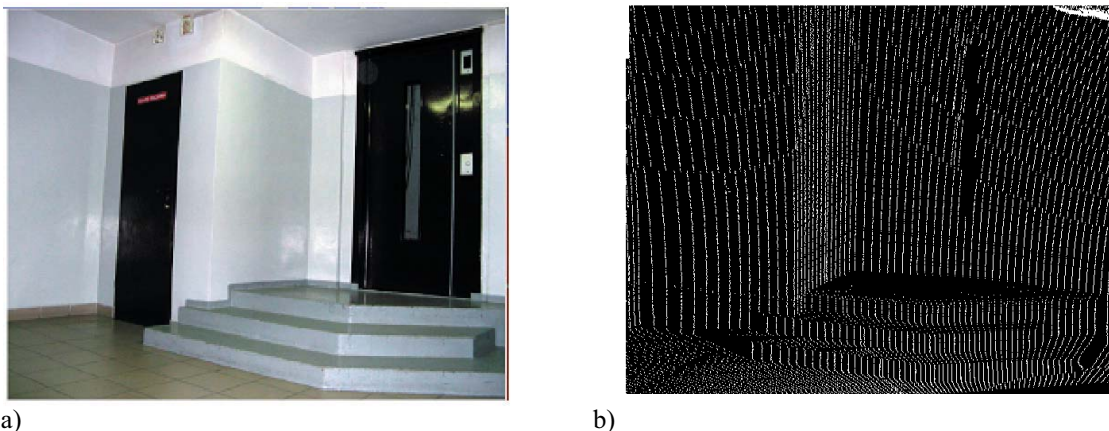
$$\begin{aligned}
 p^1_{ij} &= [x_{ij} - x_{i-1j}, y_{ij} - y_{i-1j}, z_{ij} - z_{i-1j}] \\
 p^2_{ij} &= [x_{ij} - x_{ij-1}, y_{ij} - y_{ij-1}, z_{ij} - z_{ij-1}] \\
 p_{ij} &= p^1_{ij} \otimes p^2_{ij}
 \end{aligned}
 \tag{2}$$

W przypadku, gdy odległości pomiędzy punktami są niewielkie, to obliczony iloczyn wektorowy, jest obarczony dużym błędem. W celu wyeliminowania tego błędu, zwiększany jest promień sąsiedztwa i do wyliczenia iloczynu wektorowego dla punktu (x_{ij}, y_{ij}, z_{ij}) brane są takie punkty (x_{kl}, y_{kl}, z_{kl}) , które spełniają warunek [14]:

$$\varepsilon_1 \geq \sqrt{(x_{kl} - x_{ij})^2 + (y_{kl} - y_{ij})^2 + (z_{kl} - z_{ij})^2} \geq \varepsilon_2,
 \tag{3}$$

gdzie ε_1 , ε_2 przyjęty próg dolny i górny.

Składowe wektora normalnego mogą być traktowane jako składowe *RGB* pikseli obrazu. W związku z tym, każdemu punktowi zostaje przypisany kolor, określający kierunek normalnej w danym punkcie. Przykład obrazu, który powstał na podstawie chmury punktów i wyodrębnionych normalnych jest zaprezentowany na rys. 2. Wstępną segmentację danych możemy więc przeprowadzić dla obrazu, a następnie otrzymane wyniki zapisać w postaci prymitywów geometrycznych w przestrzeni 3D.



Rys. 2. Dane z dalmierza laserowego 3D: a) zdjęcie otoczenia, b) chmura punktów



Rys. 3. Kolorowy obraz chmury punktów (rys. 2 b)

3. SEGMENTACJA OBRAZU

Celem segmentacji jest wydzielenie i klasyfikacja obszarów jednorodnych. W opisywanym przykładzie obszary jednorodne składają się ze zbioru punktów leżących na jednej płaszczyźnie. Punkty te na otrzymanym kolorowym obrazie są reprezentowane jako zbiory punktów o takiej samej barwie. Przed przystąpieniem do wyodrębniania tych obszarów należy wykonać klasyczne algorytmy przetwarzania obrazów.

W pierwszej kolejności obraz został poddany obróbce wstępnej jej celem jest usunięcie występujących szumów. Zastosowano filtry: medianowy, gaussowski i progujący.

Filtr medianowy – jest to nieliniowy filtr do przetwarzania obrazów, bardzo dobrze nadaje się do usuwania szumu drobnoziarnistego. Zaletą jego jest to, że zachowuje krawędzie. Filtr operuje na punktach sąsiednich szeregując je od najmniejszej do największej i wybierając wartość środkową (tj. medianę).

Filtr gaussowski – filtr bazujący na sąsiednich punktach, jest to splot odpowiedniego fragmentu obrazu z maską w której znajduje się funkcja Gaussa dla przykładu wykorzystywana maska dla okna 3x3. Filtr ten rozmywa krawędzie.

$$\frac{1}{12} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Filtr progujący – w klasycznej wersji wszystkie wartości poniżej zadanego progu są zamieniane na zero. Dzięki czemu łatwo jest usunąć tzw. biały szum.

Następnym etapem jest wykrycie krawędzi. W tym celu zostały zastosowane filtry Laplace'a i Canny'ego.

Filtr Laplace'a – szybki filtr do wykrywania krawędzi. Realizowany jest przez splot maski z obrazem. Przykład maski 3x3:

$$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

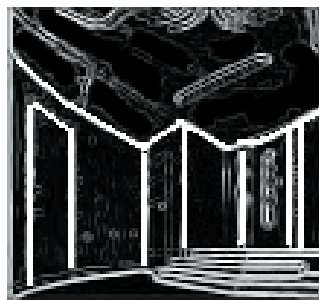
Filtr Sobel'a wykrywa krawędzie o określonym kierunku. Realizowany jest przez splot maski z obrazem. Przykład maski 3x3 dla kierunku 0°:

$$\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

Jest on używany w filtrze Canny'ego do wstępnego wykrywania krawędzi.



a)



b)



c)

Rys. 4. Etapy przetwarzania obrazu (rys. 2 b): a) wygładzanie, b) filtr Sobela, c) wykrywanie krawędzi

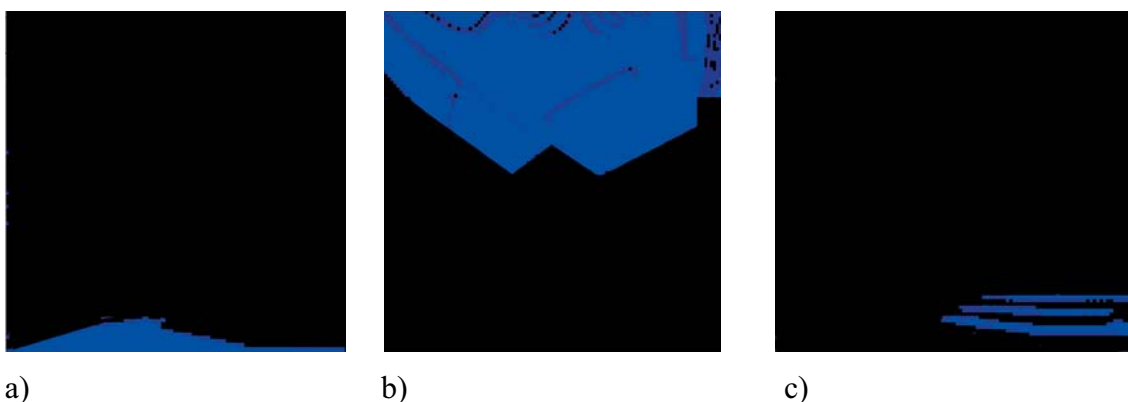
Filtr Canny'ego jest to wieloetapowy algorytm zoptymalizowany do wykrywania krawędzi. Zaletą filtru Canny'ego jest to, że w wyniku stosowania otrzymujemy krawędzie zamknięte o grubości jednego piksela. Na rys. 4 przedstawiono etapy przetwarzania obrazu.

4. KLASYFIKACJA I OPIS OBSZARÓW

Po wyodrębnieniu na podstawie analizy obrazów grup punktów jednorodnych, są one zapisywane w postaci reprezentacji trójwymiarowej, a następnie są klasyfikowane. Obiekt jest opisywany jako para atrybutów (*krawędź*, *powierzchnia*).

- *Krawędź* jest reprezentowana jako łamana w przestrzeni 3D, zapisywana jest jako ciąg wierzchołków $\{(x^w_i, y^w_i, z^w_i)\}_{i=0, \dots, K}$, gdzie K jest liczbą wierzchołków.
- *Powierzchnia* jest reprezentowana jako wektor $[A, B, C]$ normalny do płaszczyzny, która zawiera punkty należące do powierzchni.

Metoda opisu podstawowych prymitywów geometrycznych jest wystarczająca w przypadku pomieszczeń biurowych, ale może okazać się zbyt uboga w przypadku środowiska na zewnątrz budynków. W przyszłości planujemy rozszerzenie modelu o moduł analizy innych powierzchni. Zagadnienie to możemy sprowadzić do poszukiwania tekstur w obrazie.



Rys. 5. Wykryte obiekty: a) podłoga, b) sufit, c) elementy równoległe do podłogi

W kolejnym etapie, każdej powierzchni przypisujemy pewną etykietę. Część etykiet ma znaczenie semantyczne. Tego typu etykiety to:

- *Podłoga* – zbiór punktów leżących na jednej płaszczyźnie, wyznaczonej przez unormowany wektor normalny $[0,0,1]$, i wartości współrzędnej $z=0$.
- *Sufit* - zbiór punktów leżących na jednej płaszczyźnie, wyznaczonej przez unormowany wektor normalny $[0,0,1]$, i wartości współrzędnej $z=max$.
- *Ściana* - zbiór punktów leżących na jednej płaszczyźnie, wyznaczonej przez unormowany wektor normalny $[a,b,0]$ dla którego co najmniej jeden z wierzchołków ma wartość współrzędnej $z=0$, co najmniej jeden z wierzchołków ma wartość współrzędnej $z=max$ oraz powierzchnia wykrytego obiektu przekracza pewien przyjęty próg.

Pozostałym obiektom przypisujemy etykiety opisujące cechy geometryczne. Takie etykiety to:

Duża powierzchnia – powierzchnia, której pole jest większe niż zadany próg, ale nie styka się z sufitem lub z podłogą, i nie jest równoległa, ani do sufitu, ani podłogi.

Powierzchnia równoległa – zbiór punktów leżących na powierzchni równoległej do podłogi.

Powierzchnia prostopadła – zbiór punktów należących do jednej płaszczyzny, które nie tworzą ściany, ale leżą na płaszczyźnie prostopadłej do podłogi.

Brak_klasyfikacji – zbiór punktów, których nie można opisać przy pomocy podanych reguł. Punkty należące do obiektów typu *podłoga*, *sufit* mogą zostać wykryte bezpośrednio na podstawie analizy obrazu. Na rys. 5 przedstawiono opisane powyżej grupy obiektów równoległych do podłogi, grupy te reprezentowane są przez punkty o barwie niebieskiej. Dokonując podziału wykorzystujemy oprócz informacji o barwie także informację metryczną (współrzędnych punktów).

Na rys. 6 przedstawiono wynik klasyfikacji pozostałych obszarów. Kolorem czerwonym zaznaczono obszary zakwalifikowane jako ściana, żółtym duże powierzchnie, zielonym małe powierzchnie.



Rys. 6. Obiekty prostopadłe do podłogi

W kolejnym kroku na podstawie wyodrębnionej grupy tworzony jest etykietowany graf. W grafie tym węzły reprezentują wyodrębnione fragmenty sceny. Dwa węzły są ze sobą połączone, jeśli obszary sąsiadują ze sobą. Etykieta gałęzi określa sposób połączenia obszarów. Wyróżniamy krawędzie: pionowe, poziome i ukośne.

Dalszy proces tworzenia semantycznej otoczenia jest bardzo podobny do procesu analizy zdań. Znaczenie danego słowa zależy nie tylko od ciągu liter, ale także od kontekstu, w jakim dany wyraz został użyty. Dlatego kolejny etap klasyfikacji obiektów możemy przeprowadzić, analizując nie tylko pojedynczy obszar, ale uwzględniając informację o obszarach sąsiednich.

5. BAZA WIEDZY

W większości baz wiedzy przyjmuje się założenie, obiekty reprezentowane są przy pomocy wektorów, których współrzędne opisują stopień potwierdzenia hipotezy o występowaniu pewnych cech lub przekazują pewną informację metryczną np. waga, wzrost itp. W wielu zadaniach praktycznych np. w zadaniach związanych z rozpoznawaniem języka naturalnego czy złożonych struktur biologicznych opis w postaci wektora cech nie jest wystarczający. Do tego typu zagadnień należy również semantyczny opis obiektów znajdujących się w otoczeniu robota. W pomieszczeniu typu wnętrze pomieszczenia powierzchnię większość obiektów możemy opisać przy pomocy ciągu wieloboków, ale ich powierzchnia i sposób połączenia okre-

śląją ten obiekt w sposób jednoznaczny. W proponowanym systemie wyróżniamy dwie klasy obiektów:

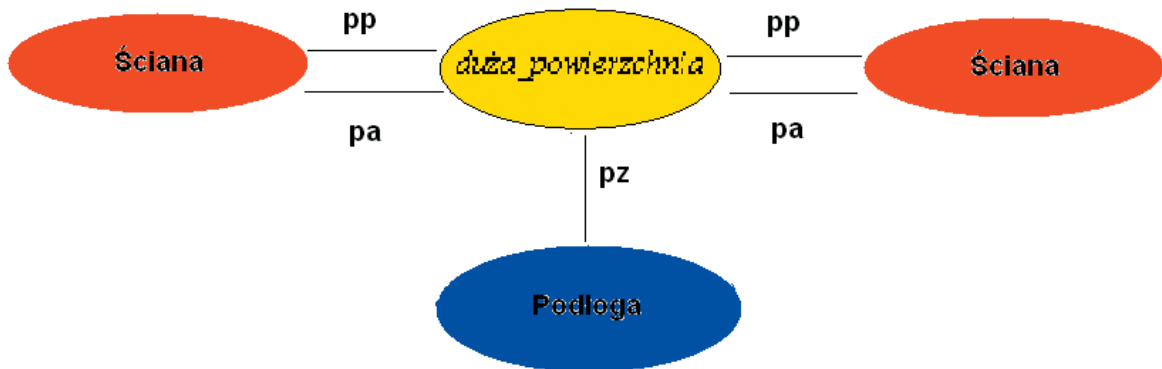
- obiekty proste - są to obiekty, które możemy w sposób jednoznaczny opisać przy pomocy zbioru reguł do nich możemy zaliczyć opisane w poprzednim rozdziale pojęcia typu: *podłoga*, *sufit*, *ściana*.
- obiekty złożone – są to obiekty, które składają się z kilku powierzchni (np. schody) lub wynik klasyfikacji zależy od etykiet obszarów sąsiednich (np. drzwi).

W klasycznej teorii graf składa się z węzłów, które są ze sobą połączone jeśli odpowiadające węzłom obiekty są ze sobą w relacji, w przypadku grafów etykietowanych etykiety połączeń między węzłami określają typ połączenia.

W naszym systemie obiekty złożone są reprezentowane przy pomocy grafów etykietowanych, informacja ta jest uzupełniona przez informację metryczną.

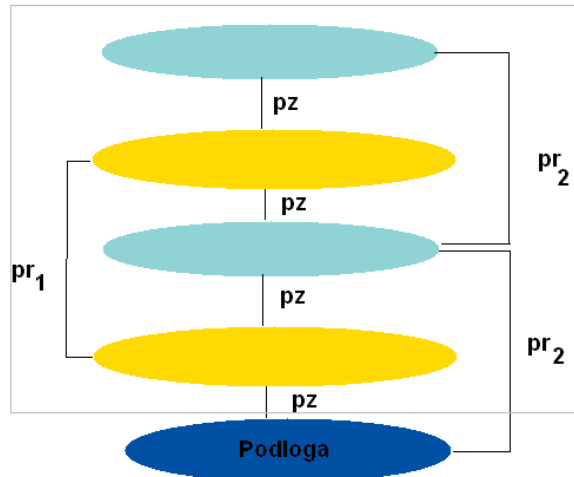
Obiekt *drzwi* jest opisywany przy pomocy reguł które uwzględniają pewną informację metryczną:

Obiekt d należy do klasy *drzwi_zamknięte*, jeśli należy do klasy *duża_powierzchnia* oraz $s_1 < szerokość < s_2$, $wysokość > h_1$, gdzie s_1, s_2, h_1 są parametrami. Dodatkowo powinny zachodzić relacje zapisywane przy pomocy grafu przedstawionym na rys. 7, *pp* – jest etykietą określającą krawędź pionową, a *pz* – krawędź poziomą, *pa* – określa relację równoległości płaszczyzn.



Rys. 7. Graf etykietowany opisujący obiekt typu drzwi zamknięte

Innym sposobem zapisu obiektów jest zapis relacji między prymitywami geometrycznymi obiektów. Na rys. 8 przedstawiono graf opisujący obiekt typu schody. Kolorem jasno niebieskim zaznaczono obszary równoległe do podłogi, a kolorem żółtym obszar prostopadły do podłogi. Etykieta *pz* określa że między obiektami występuje krawędź pozioma, etykiety pr_1, pr_2 określają że powierzchnie są równoległe i znajdują się w określonej odległości.



Rys.8. Graf etykietowany opisujący obiekt schody

6. PODSUMOWANIE

W powyższym artykule przedstawiono koncepcję i wstępne eksperymenty systemu automatycznego tworzenia semantycznych map otoczenia robota mobilnego. Do wstępnego przetwarzania danych pochodzących ze skanera 3D wykorzystano klasyczne metody analizy obrazów płaskich. W opracowanym systemie wykorzystano procedury biblioteki OpenCV, dzięki temu czas przetwarzania obrazów nie przekracza 200 ms. Czas niezbędny do klasyfikacji obiektów jest zależny od liczby obiektów, które powinny być zapamiętanie w bazie wiedzy. W obecnym stanie systemu obiektów jest kilkanaście, w przyszłości planujemy proces przeszukiwania skrócić przypisując do każdego obiektu funkcję haszującą. Dzięki temu przeszukiwane są jedynie elementy bazy o zadanej wartości funkcji. Nawet jeśli w pewnym pomieszczeniu nie wszystkie elementy zostały sklasyfikowane możliwe jest określenie typu pomieszczenia: np. korytarz, pokój. Możliwy jest też topologiczny opis sceny. Opisywany system jest niezbędny w przypadku współpracy robota z człowiekiem. Umożliwia wydawanie poleceń typu: jedź do drzwi. Mapy semantyczne mogą być wykorzystywane w procesie lokalizacji robota. W przypadku planowania trasy oprócz informacji topologicznej potrzebna jest także informacja metryczna, dlatego oprócz informacji topologicznej z każdym elementem wiązana jest informacja metryczna. Opisywany system jest nadal rozwijany i testowany konieczne jest wprowadzenie klasyfikatorów w których jest określana niepewność wyniku klasyfikacji. Planujemy także stworzenie opisu hierarchicznego w którym oprócz obiektów będą również klasyfikowane całe pomieszczenia oraz budynki.

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

LITERATURA

1. Moravec H. , Elfes A.: High resolution maps from wide angle sonar, Proc. IEEE International Conference on Robotics and Automation, s. 116 - 121, 1985.
2. Andersen C.S. Jones S. and Crowley J.L. Appearance Based Processes for Visual Navigation. Proceedings of Symposium on Intelligent Robotics Systems. s. 227 -236, 1999.

3. Triebel R. Frank B. Meyer J. Burgard W.: First Steps Towards a Robotic System for Flexible Volumetric Mapping of Indoor Environments, CD-ROM. IAV04, 2004.
4. B. Siemiątkowska. Hybrydowa reprezentacja otoczenia robota mobilnego. PAR vol. CD ROM, 2006.
5. Floryczyk R., Robot Vision, Wiley-Vch, 2005.
6. Gonzalez R. C., Digital Image Processing, Adison-Wesley, Reading MA, 1992.
7. Sakas G., Hartig J., Interactive Visualization of Large Scalar Voxel Fields. Visualization, Boston, USA, s. 29--36, 1992.
8. Sawwa R., Siemiątkowska B., Racz J., 2.5D Map Building Based on LRF Readouts. III-d Int. Symp. on Methods and Models in Automation and Robotics. s. 13--15. 1997.
9. Sawwa R., Siemiątkowska B., Racz J.: A Laser Range Finder for Mobile Robot Navigation. 28-th International Symposium on Robotics. Detroit, MI USA. s. 13--15, 1997.
10. Siemiątkowska B., Gnatowski M., Zychewicz A. Fast Method of 3D Map Building Based on Laser Range Data. Jamris. s. 65-70, 2007.
11. Schroeder W. , Zarge J. , Lorensen W., Decimation of Triangle Meshes. Computer Graphics. s. 65-70, 1992.
12. J. Weingarten, Siegart R.: EKF-based 3D SLAM for Structured Environment Reconstruction. IROS 2005.
13. D.L. Page, Y. Sun, A.F. Koschan, J.Paik, M. A. Abidi, Normal Vector Voting: Crease Detection and Curave Estimation on Large, Noisy Mesh. Graphical Models, vol. 64, 2002, s. 199-229.
14. Siemiątkowska B., Gnatowski M., Chojecki R.: Zastosowanie sieci komórkowych w procesie segmentacji danych pochodzących z dalmierza laserowego 3D Krajowa Konferencja Robotyki, Oficyna Wydawnicza Politechniki Warszawskiej, 2008, s. 121-130.