



WNIOSEK O PORTFOLIO: System analizy sceny trójwymiarowej dla robotów autonomicznych

Autorzy: Andrzej Bielecki, Tomasz Buratowski, Marzena Bielecka, Piotr Śmigielski

Centrum Inteligentnych Systemów Informatycznych Akademia Górniczo-Hutnicza im. Stanisława Staszica al. Mickiewicza 30, 30-059 Kraków
budynek C-2 pokój 426 tel: 12 617 44 53 www.isi.agh.edu.pl isi@agh.edu.pl



System analizy sceny trójwymiarowej dla robotów autonomicznych

Wprowadzenie

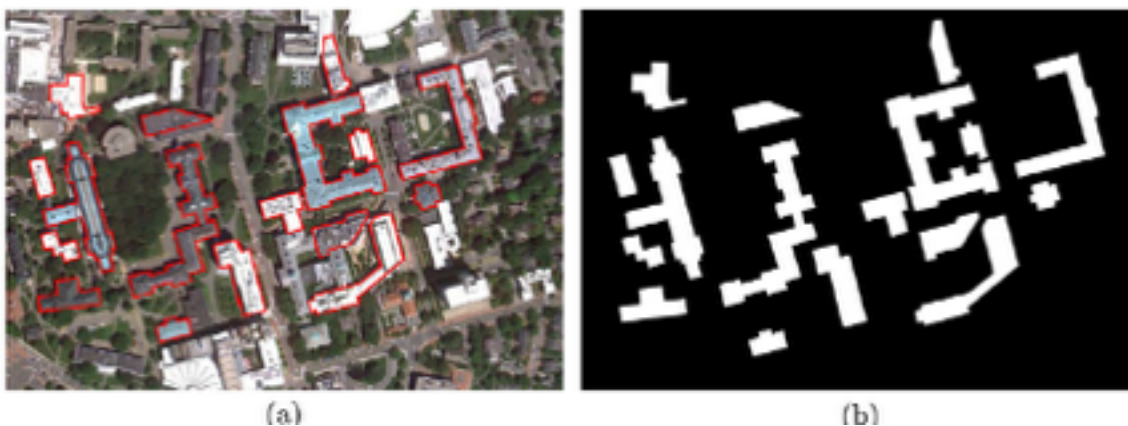
Systemy wizyjne robotów są obiektem badań od ponad dwudziestu lat. Systemy te mają szczególne znaczenie w przypadku robotów autonomicznych. Sensory, w które są wyposażone pozwalają na zbieranie informacji ze środowiska. Zebrane dane są wykorzystywane do zlokalizowania pozycji robota w przestrzeni, znalezienia bezkolizyjnej drogi do celu oraz zlokalizowania poszukiwanego obiektu. Do zlokalizowania robota może zostać użyty system GPS, jednakże są miejsca gdzie użycie go nie wchodzi w grę (pod wodą, w jaskiniach, na powierzchni innych planet) i tam właśnie systemy wizyjne powinny być alternatywą. Co więcej złożone zadania, jak przeprowadzanie inspekcji w trudno dostępnych miejscach czy badanie nieznanego środowiska wymaga nie tylko możliwości zlokalizowania robota w przestrzeni ale również możliwości rozpoznawania obiektów i analizowania obszaru, w którym robot się znalazł. Z tego powodu systemy rozpoznawania obrazów oraz analizy obszaru, bazujące na systemach wizyjnych, są intensywnie badane w kontekście analizy i rozumienia sceny i pozycjonowania robota. Aktualnie bezzałogowe roboty latające (Unmanned Aerial Vehicles – UAV-y) mogą być sterowane zdalnie, mogą latać automatycznie bazując na wgranym programie lotu bądź bazując na bardziej skomplikowanych automatycznych systemach. Misje wykonywane przez UAV-y są różnych rodzajów. Mogą one być wykorzystywane do robienia rekonesansu, ataku podczas konfliktu zbrojnego bądź w celach transportowych. Ich specyfikacje zależą głównie od rodzaju misji jaką wykonują.

Zdefiniowanie problemu

Przedmiotem badania jest bezzałogowy robot latający, który jest zaprojektowany do przeprowadzania inspekcji miejskiej zabudowy takiej jak kominy, mosty, wieże itp. Robot jest agentem autonomicznym. Musi zaplanować trajektorię z punktu startowego do obszaru misji, następnie przeanalizować obszar misji, zlokalizować rozpoznany obiekt i zaplanować przebieg inspekcji w kontekście struktury sprawdzanego obiektu i własności obszaru. W pamięci agenta znajduje się mapa, załadowana tam wcześniej, na przykład przetworzone zdjęcie satelitarne środowiska miejskiego. Robot jest wyposażony w kamerę, która może zostać skierowana w stronę ziemi w celu zrobienia zdjęć powierzchni nad którą się znajduje. Mapa, którą ma robot, prezentuje budynki wyekstrahowane z obrazu satelity (zob. Fig. 1(a)). Poprzez odnoszenie się do danych dwuwymiarowych map i obrazów z kamery, agent orientuje się w środowisku, równocześnie posługując się innymi informacjami, przykładowo GPS. W celu znalezienia lokalizacji na mapie robot robi zdjęcia obszaru nad którym się znajduje. Następnie porównuje je z kształtami budynków z obrazów i lokalizuje je na większej mapie. Przykład takiego obrazu znajduje się na Fig. 1(b). Metoda rozpoznawania

obiektów musi być niezależna od kąta i skali ponieważ zdjęcia ziemi będą robione z różnych wysokości i kierunków. Ten problem należy do grupy zadań, które składają się na rozpoznawanie i reprezentowanie wielobocznych obiektów przez system sensorów robota. Metoda jest rozszerzona o rozpoznawanie grup budynków na podstawie kształtów pojedynczych budynków oraz ich wzajemnego położenia. Następnie robot zniża lot aby zbadać środowisko bezpośrednio poprzez użycie kamer i sensorów. To pozwala agentowi stworzyć trójwymiarowy obraz środowiska. Następnie wymodelowana scena powinna zostać zrozumiana a żądany obiekt znaleziony. W ten sposób badania mogą zostać podzielone na następujące etapy:

1. Stworzenie pojedynczego, dwuwymiarowego obrazu budynku, bazującego na zdjęciu wykonanym na dużej wysokości.
2. Wymodelowanie dwuwymiarowego obszaru bazującego na zdjęciu z kamery wykonanym na dużej wysokości. Następnie otrzymany obszar powinien zostać rozpoznany jako fragment danej wcześniej mapy. Pojedynczy, żądany obiekt powinien zostać rozpoznany.
3. Stworzenie trójwymiarowego modelu obszaru, bazując na stworzonym wcześniej modelu dwuwymiarowym. Dwuwymiarowa reprezentacja pozwoli nam zlokalizować budynek. Kiedy jego lokalizacja jest znana, robot może zbadać konkretny obiekt w celu stworzenia trójwymiarowej reprezentacji.
4. Zrozumienie trójwymiarowej sceny.



Rys.1 (a) Zdjęcie satelitarne z zaznaczonymi budynkami (b) wyodrębnione budynki

Opracowane algorytmy i przeprowadzone testy

Algorytmem stosowanym do opracowania punktów 1 i 2 jest syntaktyczny algorytm wektoryzacji i rozpoznawania dwuwymiarowego obiektu. Składa się on z algorytmu wektoryzacji oraz algorytmu rozpoznawania obiektów.

Algorytm 1: Generacja reprezentacji pojedynczych obiektów

Algorytm wektoryzacji pracuje na przygotowanych danych, bazujących na grafice rastrowej. Są to płaskie, dwuwymiarowe dane gdzie 0 reprezentuje piksele tła, a 1 piksele budynku. Algorytm ten przeszukuje macierz dopóki nie znajdzie pierwszego nieoznaczonego budynku. Przetwarza następnie kształt tego budynku zmieniając go na wektorową reprezentację i oznacza jego powierzchnię poprzez zmianę wartości "1" na inny numer aby uniknąć badania tego samego budynku dwa razy. Należy również wspomnieć, że podczas przeszukiwania macierzy każdy nieoznaczony punkt, który został znaleziony jest rogiem budynku lub znajduje się bardzo blisko tego rogu. Pierwszym krokiem wektoryzacji jest generacja konturów pojedynczych budynków. Zaczynając od pierwszego punktu konturu (rogu budynku) kolejne punkty są znajdowane przy użyciu *okna* – kwadratu używanego do zdefiniowania podmacierzy. Jeżeli początkowy punkt *okna* leży poza konturem budynku to poszukiwania granicy są przeprowadzane zgodnie z ruchem wskazówek zegara, w przeciwnym wypadku w stronę przeciwną. W efekcie, badania wokół obiektu są prowadzone zawsze zgodnie z ruchem wskazówek zegara.

Wynikiem badania budynku jest sekwencja punktów, które mogą zostać zinterpretowane jako sekwencja wektorów wyznaczających kontur budynku.

Danymi, na których pracuje algorytm rozpoznający są dwie zwektoryzowane mapy. Jedna mapa pokazuje duży obszar wypełniony budynkami, natomiast druga pokazuje tylko jeden budynek, podobny do jednego z tych pokazanych na mapie, ale obrócony i w innej skali. Algorytm ma za zadanie porównać obie mapy i znaleźć na dużej mapie obiekt przedstawiony na małej. Porównanie przebiega zgodnie z następującymi krokami:

1. Porównanie liczby wierzchołków obu budynków. Jeśli liczby są różne, algorytm zatrzymuje się, ponieważ są to różne budynki; jeżeli liczby wierzchołków się zgadzają, algorytm przechodzi do następnego kroku.
2. Reprezentacja budynku będąca liczbą wierzchołków zostaje zmieniona na reprezentację używającą kątów i długości ścian. Tak więc każda ściana budynku jest opisana parą (F_i, N_i) , gdzie F to kąt a N to długość ściany.
3. Dla obu powyższych reprezentacji ścian budynku, tworzymy nową listę składającą się z par (α_i, N_i) . Pierwsza wartość wyraża kąt pomiędzy sąsiednimi ścianami (tzn. $\alpha_i = (F_{i+1} - F_i)$), druga to długość ściany.
4. Następnie, w celu odnalezienia poszukiwanego budynku odpowiednie listy są porównywane. Jeśli wszystkie analizowane współczynniki są mniej więcej równe to budynki są traktowane jako podobne i algorytm kończy swoją pracę. Jeśli współczynniki się wyraźnie różnią to poszukiwania są kontynuowane.

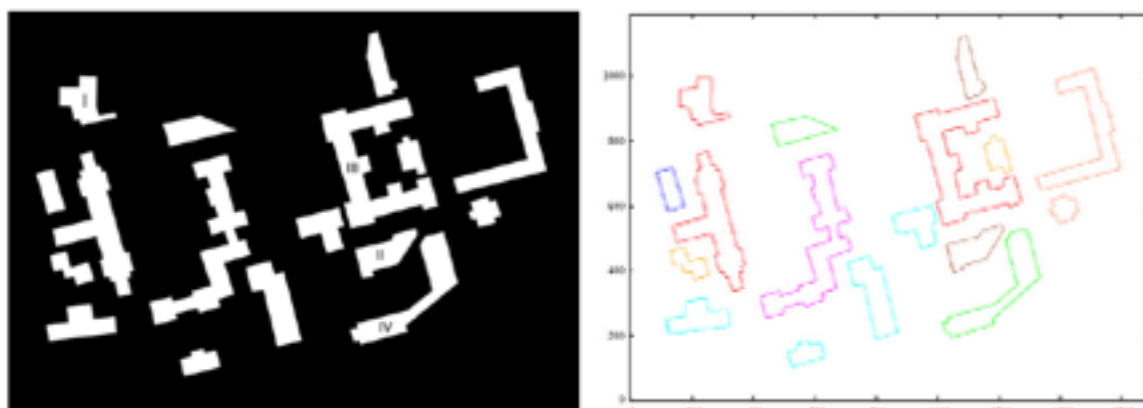
Wyniki działania algorytmu zostały przedstawione dla danych wygenerowanych oraz danych rzeczywistych.

Zestaw wygenerowanych danych składał się z dwóch obrazów budynków wyekstrahowanych z dużego obrazu. Ich kształty zostały przeskalowane i obrócone – rys.3.



Rys.3 (a) Wygenerowana mapa z zaznaczonymi dwoma budynkami; (b) i (c) Budynki I i II przeskalowane i obrócone.

W teście bazującym na realnych danych, 4 budynki zostały wybrane z mapy – rys. 1. Na rysunku 4 jest przedstawiona zwektoryzowana mapa tych budynków. Wybrane budynki, również przeskalowano i poddano rotacji.



Rys.4 Zwektoryzowana mapa z rys.1, z zaznaczonymi 4 budynkami, wybranymi do testu.

Algorytm 2: Rozpoznawanie mapy

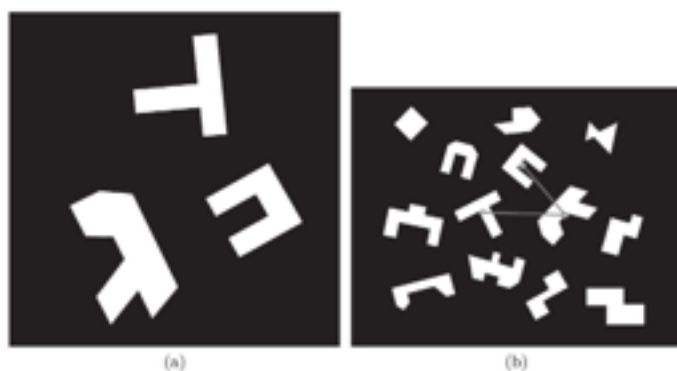
Algorytm rozpoznawania pojedynczego obiektu jest punktem wyjścia algorytmu rozpoznawania mapy. Celem tego algorytmu jest zlokalizowanie podmapy na większym obrazie. Podmapa składa się z grupy budynków, które są częścią dużej mapy. Metoda ta pozwala na to, aby obraz był skalowany i obracany. W pierwszym etapie została stworzona odpowiednia struktura danych. W wyniku analizy obu map wszystkie budynki z mapy B zostały sprawdzone pod kątem podobieństwa do budynków z mapy b. W efekcie każdemu obiektowi z b przyporządkowano listę obiektów z B. Dodatkowo relacja między obiektami z b i B została opisana przez dwie wartości: skalę i rotację. Działanie algorytmu polega na przeszukiwaniu utworzonej struktury danych w celu znalezienia pasujących obiektów.

Połączenie pomiędzy podmapą i dużą mapą zostaje znalezione kiedy na każdej liście pasujących elementów jest co najmniej jeden odpowiedni budynek. Poszukiwania są zatrzymane jeśli:

1. Którakolwiek z list pasujących obiektów jest pusta
2. Algorytm napotkał pierwszą listę na której nie znaleziono połączenia.

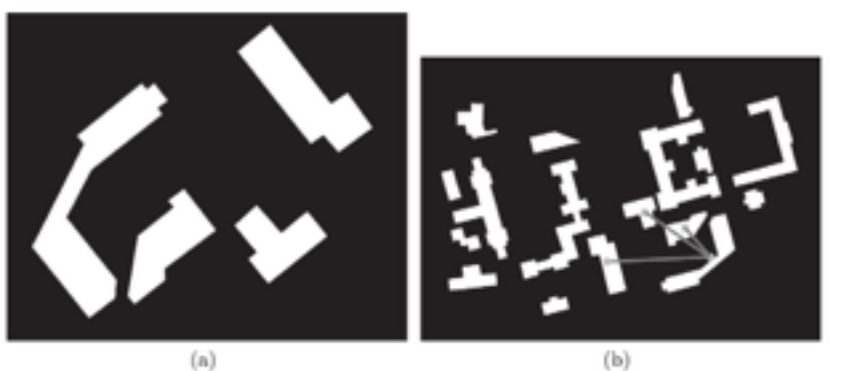
W celu przetestowania algorytmu wykorzystano dwa zestawy danych: dane wygenerowane oraz dane rzeczywiste tj. zdjęcie satelitarne z serwisu Google Maps (zdjęcie pokazuje obszar Harvardu w Cambridge, Massachusetts).

Dla danych wygenerowanych wyniki działania algorytmu są przedstawione na rys.5. Grupa budynków tworząca podmapę dla mapy z rys. 3(a), została odnaleziona na mapie większej - rys.5 (b), na której szarymi kropkami oznaczono znalezione budynki natomiast szare linie wyznaczają kąty, które wykorzystano w obliczeniach.



Rys.5 (a) Podmapa mapy większej z rys.3(a), (b) większa mapa z zaznaczonymi poszukiwanymi budynkami .

Dla danych rzeczywistych wyniki działania algorytmu zostały przedstawione na rys.6.



Rys.6 (a) Podmapa mapy z rys.4, (b) większa mapa z zaznaczoną poszukiwaną grupą budynków z punktu (a).

Na rys.6(b) zostały zaznaczone szarymi kropkami odnalezione budynki, natomiast szare linie wyznaczają kąty wykorzystane do obliczeń.

Algorytm 3: Tworzenie trójwymiarowej reprezentacji obiektu

Centrum Inteligentnych Systemów Informatycznych Akademia Górniczo-Hutnicza im. Stanisława Staszica al. Mickiewicza 30, 30-059 Kraków
budynki C-2 pokój 426 tel: 12 617 44 53 www.isi.agh.edu.pl isi@agh.edu.pl

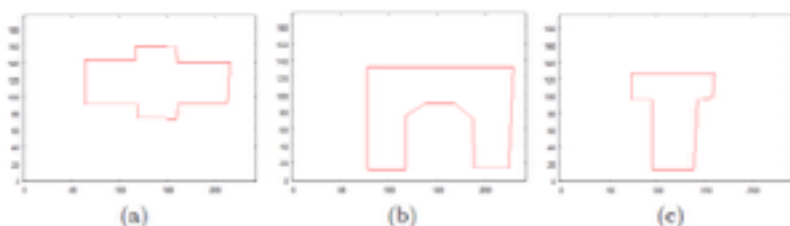
W algorytmie pozwalającym na uzyskanie trójwymiarowej reprezentacji badanego obiektu zostały wykorzystywane trzy zwektoryzowane obrazy – strona górna, frontowa i prawy bok budynku. Algorytm został podzielony na trzy części. W każdej z nich jedna projekcja budynku jest przyjmowana jako wzór, a pozostałe dwie projekcje są traktowane jako modele, które są poddawane transformacji. Na każdym etapie są wykonywane następujące kroki:

1. Pobieramy dwa kolejne punkty z projekcji, która aktualnie jest wzorcową: $A=(x_i, y_i)$, $B=(x_{i+1}, y_{i+1})$; każda taka para tworzy odcinek wzorcowy.
2. Odcinek wzorcowy zostaje wykorzystany do wycięcia odpowiednich części z dwóch pozostałych projekcji (modeli).
3. Uzyskane, w wyniku powyższej operacji ściany leżą na tym samym planie, prostopadle do ścian wzorca, których nachylenie jest zdefiniowane przez odcinek wzorcowy.

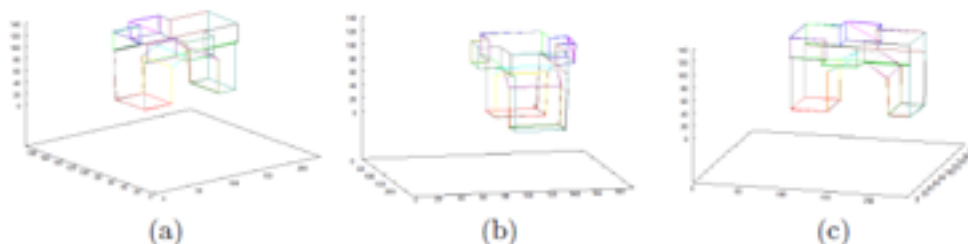
Poniżej są zaprezentowane 3 przykłady, wykorzystujące zaproponowany algorytm.

Przykład 1 – Łuk tryumfalny

W pierwszym przypadku kształt budynku to łuk tryumfalny z dwoma parapetami wystającymi z podstawy. Na rys.7 zostały przedstawione odpowiednie strony budynku po wektoryzacji.



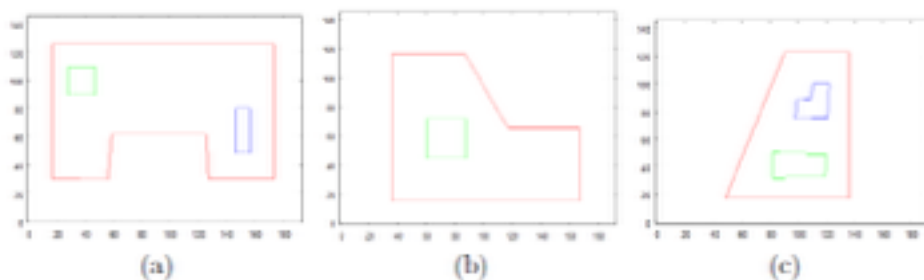
Rys.7 Łuk tryumfalny po wektoryzacji. (a) widok z góry (b) widok od frontu (c) widok z boku



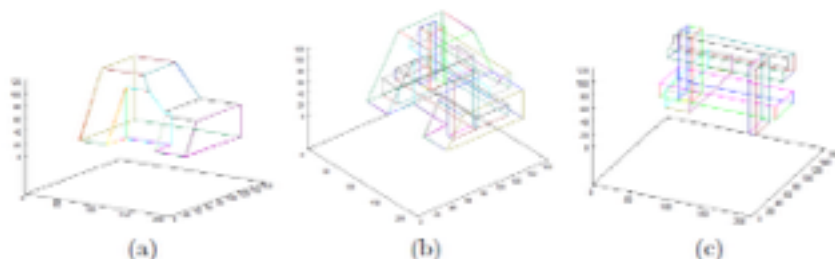
Rys.8 Rekonstrukcja trójwymiarowa łuku, zrealizowana przy pomocy zaproponowanego algorytmu, widziana pod różnymi kątami.

Przykład 2 – Struktura z otworami w każdej ścianie

Analizowana struktura posiada co najmniej jeden otwór w każdej ścianie. Widok z góry, od frontu i z prawej strony jest pokazany na rys.9. W wyniku zastosowania algorytmu otrzymujemy strukturę trójwymiarową przedstawioną na rys.10.



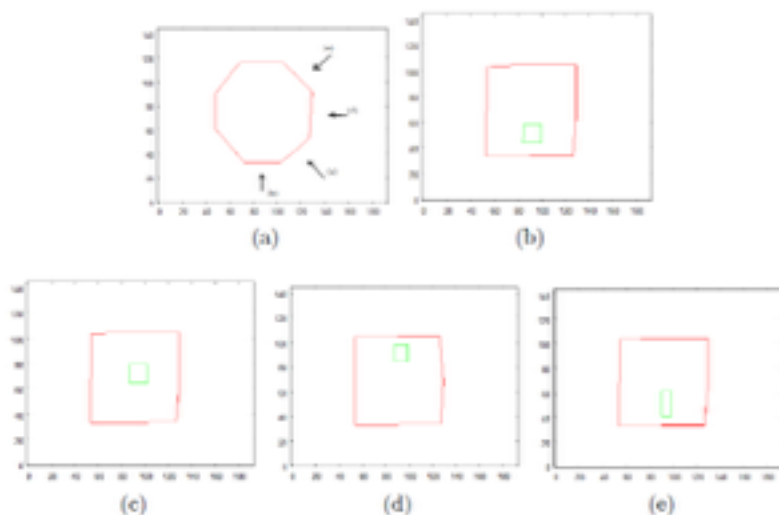
Rys.9 Obiekt z otworami – widok z trzech stron.



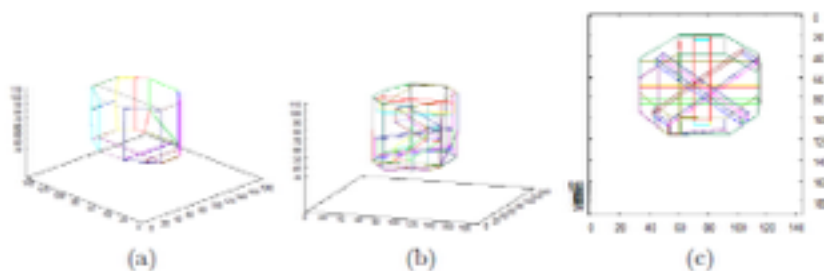
Rys.10 Rekonstrukcja trójwymiarowa obiektu z drugiego przykładu, (a) widok bez otworów, (b) widok z otworami (c) reprezentacja samych otworów

Przykład 3 – Obiekt wielościenny z otworami

Analizowany obiekt jest ośmiościenny z otworami przechodzącymi na wylot. Widok obiektu jest przedstawiony na rys.11. Na rys.12 pokazano trójwymiarową reprezentację budynku.



Rys.11 (a) widok od góry (b) widok od frontu (c) widok pod kątem 45 stopni (d) widok prawego boku (e) widok pod kątem 135 stopni



Rys.12 Rekonstrukcja trójwymiarowa budynku z przykładu trzeciego, (a) obiekt bez otworów, (b) obiekt z otworami, (c) reprezentacja samych otworów.

Plan przyszłych badań

W niniejszym raporcie zostały opisane algorytmy pozwalające na stworzenie trójwymiarowej reprezentacji pojedynczego obiektu. Otrzymane trójwymiarowe reprezentacje powinny zostać przypisane do odpowiednich obiektów na dwuwymiarowej scenie odpowiadającej zdjęciom satelitarnym.

Należy wspomnieć, że uzyskane rezultaty, włączając w to wykonane eksperymenty, mają tylko aspekt wirtualny – software nie został osadzony na robocie. Jednakże wypracowana trójwymiarowa reprezentacja sceny pozwala nam na przetestowanie UAV-a wyposażonego w moduł analizy wizyjnej w następnym etapie badań.

W przyszłości całkowicie autonomiczne roboty będą wykorzystywane do badania nieznanego środowiska. Inspekcja przemysłowych i miejskich obiektów jest więc stosunkowo łatwym zadaniem, w kontekście analizy sceny. Jeśli robot jest wykorzystywany do badania nieznanego naturalnego środowiska jak Księżyc lub planety, możliwości zrozumienia sceny przez agenta muszą być znacznie większe niż syntaktyczne metody analizy sceny. W takim wypadku agent powinien zostać wyposażony w algorytmy o charakterze kognitywnym.