

**Recenzja pracy doktorskiej mgr inżyniera Ireneusza Hallmanna
"Lokalizacja robota mobilnego względem automatycznie wybieranych
obiektów"**

Recenzowana praca dotyczy zagadnienia lokalizacji robota mobilnego względem znajdujących się w jego otoczeniu specjalnie przygotowanych znaczników. Praca ta jest napisana na 173 stronach i składa się z 9 rozdziałów. Na końcu pracy znajduje się bibliografia. Tekst pracy został mi dostarczony na piśmie oraz na CD ROMie, a również dołączono kopie artykułów napisanych przez Autora i związanych z doktoratem.

Zasadnicza teza pracy została przez Doktoranta sformułowana następująco:

"Tezą niniejszej pracy jest sprawdzenie, czy jest możliwa lokalizacja robota mobilnego na podstawie jednego znacznika, w oparciu o dane pochodzące z jednej kamery zainstalowanej na robocie."

W dalszym ciągu Autor pisze: "Jedynym używanym sensorem jest pojedyncza kamera, którą można obracać niezależnie od robota. Niezależność kamery pozwala na lokalizację niezależnie od orientacji robota. Proponowana metoda pozwala traktować kamerę jak sensor odczytujący położenie względem widocznych znaczników."

Powyższe sformułowanie tezy będzie przedyskutowane poniżej.

Szczegółowa treść rozprawy

Omawiana praca zawiera następujące rozdziały:

- 1) Wstęp – rozdział przedstawiający tezę oraz cele pracy.
- 2) Istniejące techniki lokalizacji – jest to przegląd literatury
- 3) Reprezentacja i przetwarzanie obrazu – jest to przegląd literatury.
- 4) Kalibracja kamery. Rozdział 4.2 zawiera przegląd metod kalibracji, natomiast rozdział 4.3 zawiera opis eksperymentów i obliczeń przeprowadzonych przez Doktoranta.
- 5) Detekcja i reprezentacja znaczników w literaturze.
- 6) Charakterystyka proponowanej metody kalibracji.
- 7) Zastosowana detekcja znaczników.
- 8) Proponowana metoda lokalizacji.
- 9) Pomiary i testy.
- 10) Podsumowanie i wnioski.

Ponadto praca zawiera bibliografię w objętości 94 pozycji.

Recenzja merytoryczna rozprawy

Podstawowe cele pracy zostały sformułowane następująco:

Cel 1. Wybór kształtu i typu znacznika umożliwiającego lokalizację robota na podstawie widoczności znacznika (rozdz. 6.3).

Cel 2. Wybór obiektów i cech odpowiednich do ich automatycznej identyfikacji (rozdz. 7).

Cel 3. Identyfikacja obiektów na podstawie znalezionych cech (rozdz. 7).

Cel 4. Opracowanie metody dokładnego wyznaczenia położenia kamery na podstawie jednego widocznego obiektu (rozdz. 8 i 9).

Cel 5. Eliminacja błędów pomiarowych (rozdz. 4.3 i 7.3).

Autor nie podał, jakie są jego główne osiągnięcia. Pewną wskazówką mogą być jednak uwagi zamieszczone przez Doktoranta w rozdz. 10 – Podsumowanie i wnioski. Na tej podstawie można powiedzieć, że głównymi osiągnięciami są:

1. Potwierdzenie teoretyczne i doświadczalne, że możliwe jest estymowanie położenia i orientacji robota mobilnego na podstawie pojedynczego zdjęcia zrobionego kamerą umieszczoną na robocie, przy czym na zdjęciu tym powinien być widoczny znacznik mający formę rysunku naniesionego na płaską powierzchnię. W tym celu kamera powinna być uprzednio skalibrowana przy użyciu kratki pomiarowej przedstawianej kamerze.

2. Opracowanie metody estymacji położenia i orientacji robota mobilnego na podstawie obrazu pomocniczego znacznika umieszczonego na scenie, po której porusza się robot. Estymację tę przeprowadza się na podstawie współrzędnych czterech narożników prostokąta wyznaczonego przez obraz znacznika. W tym zakresie Doktorant opracował dwie metody:

- metodę, w której jako zmiennej pomocniczej używa się odległości od kamery do jednej krawędzi znacznika (rozdz. 8.4)

oraz

- metodę, w której jako zmiennych pomocniczych używa się odległości od kamery do dwóch krawędzi znacznika (rozdz. 8.5).

Uwagi szczegółowe

W związku z tytułem pracy nasuwa się następująca uwaga. Tytuł "Lokalizacja robota mobilnego względem automatycznie wybieranych obiektów" sugeruje więcej niż zostało w pracy wykonane. Automatycznie wybierane obiekty mogłyby być dowolne, np. każda osoba wchodząca do pomieszczenia. Natomiast w pracy chodzi nie o dowolne obiekty, a ściśle określone i odpowiednio skonstruowane znaczniki, co jest też w rozprawie powiedziane już na str. 5.

Przytoczona powyżej teza pracy doktorskiej jest dla mnie zbyt mało asertatywna. Tezą mogłoby być np. stwierdzenie, że możliwe estymowanie położenia robota na podstawie pojedynczego zdjęcia w jakichś warunkach. Natomiast sformułowanie "Tezą jest sprawdzenie, czy jest możliwa lokalizacja ..." niczego nie stwierdza. Fakt użycia pojedynczego znacznika nie ma chyba zasadniczego znaczenia. Można np. użyć pojedynczy rozległy znacznik lub dwa mniejsze położone w pobliżu siebie i zarówno metoda jak i jej efekt będzie identyczny.

Str. 4. Brak jest określenia, co stanowi nowość w pracy w porównaniu do istniejącej literatury.

W tytule pracy jest mowa o detekcji automatycznie wybieranych obiektów, natomiast na str. 4 jest mowa o znacznikach.

Str. 16. Brak objaśnienia symboli, np. co to jest $v_{12x}, v_{12y}, p_{12x}, p_{12y}$?

Str. 30. Co to jest r i g we wzorze nad wzorem (3.1)?

Tekst na str. 29-32 dotyczący transformacji między przestrzeniami kolorów jest zapewne zaczerpnięty z literatury. Jednak Autor nie podaje źródeł, a jest to ważne ponieważ występują pewne niedokładności.

Co to jest r i g we wzorze powyżej wzoru (2.1)? Jeżeli przeprowadzamy konwersję z przestrzeni RGB do HSV, to należy obliczyć składowe H, S, V. Podane obliczenia są niejasne i brakuje obliczenia składowej H. Podobna uwaga odnosi się do transformacji z przestrzeni HSV do RGB. Najbardziej dostępnym źródłem dotyczącym transformacji między RGB i HSV jest Wikipedia. Przepuszczalnie źródło to byłoby dla Autora wystarczające.

Wzór (3.5) na rozkład Gaussa na str. 33 jest błędny. Komentarz nie wyjaśnia, w jaki sposób znajduje się współczynniki na rys. 3.4 ze wzoru (3.5). Co to są liczby 8, 9, 16 na rys. 3.4?

Rozdz. 3.3 dotyczący detekcji krawędzi ma elementarny, podręcznikowy charakter i mógłby być skrócony.

Str. 36. Równanie (3.11) wymaga stwierdzenia, że macierz A nie jest dowolna, a musi być odwracalna. W przeciwnym przypadku transformacja nie jest afiniczna.

Wzór (3.17) nie zgadza się z (3.14), (3.15).

Rys. 3.15 i 3.16 – co jest na której osi?

Str. 45. Tekst między równaniem (4.2) i (4.3) jest niejasny. Jak należy rozumieć, że wektor parametrów składa się ze wszystkich punktów przestrzeni i punktów obrazu (oprócz parametrów kamery). Jaki jest związek μ ze zmienną losową U ? Wartość średnia, a raczej wektor μ w rozkładzie Gaussa u dołu przed wzorem (4.3) nie ma nic wspólnego z μ zdefiniowanym wyżej. Występuje wyraźna kolizja symboli.

Model fizyczny kamery opisany wzorami (4.4) i (4.5) powinienn być uzgodniony z równaniem (4.1), ponieważ w zasadzie przedstawiają to samo.

Wzór (5.5) na moment $M_{j,k}$ jest niejasny. Otrzymujemy z niego $M_{0,0} = (X + 1)(Y + 1)$ niezależnie od rozpatrywanej figury.

Wzór (5.7) nie zgadza się ze wzorem spotykanym w literaturze, np. w [33] i kilku innych artykułach. Niejasne jest też występowanie czynnika XY oraz odmienna definicja wykładnika potęgi w mianowniku we wzorze (5.7).

Niezmienniki H_u na str. 61 są podane z błędami, np. H_5 .

Uwagi na str. 62 w pobliżu wzoru (5.25) chyba są nie na temat, ponieważ wzór ten nie jest dalej używany.

Histogram na rys. 7.1(b) powinien mieć dokładniejszy opis. Jakie są jednostki na osi poziomej? Powinno to być uzgodnione z rozdz. 1.4. Co oznaczają kolory na rys. 7.1.(b)?

Rys. 7.17 na str. 88 ma zawartość niezgodną z podpisem. Cały schemat efektywnie sprawdza, czy dany piksel ma 4-sąsiadów należących do tego samego obiektu. Nie widać jednak żadnego obliczenia momentów.

Rozdz. 7.1 o detekcji znaczników o określonej barwie jest niejasny. Najprawdopodobniej Doktorantowi chodzi o wykrycie znaczników mających określony kolor, np. niebieski. Do czego jednak jest potrzebne uczenie barw charakterystycznych. Jeżeli mamy próbkę barwy (a tak właśnie jest), to

należy zebrać z niej odpowiednie statystyki, a następnie znaleźć wszystkie piksele w obrazie, których kolor będzie dostatecznie bliski kolorowi zadanemu. Odpowiednio duża plama tego koloru będzie kandydatem na znacznik. Odległość między kolorami można liczyć w przestrzeni RGB, ale odległość w przestrzeni RGB nie zgadza się z intuicyjnym pojęciem odległości kolorów i lepiej jest zastosować obliczenie odległości w przestrzeni HSI lub HSV. Zamiast tego, Autor proponuje algorytm uczenia barw. Algorytm ten w rzeczywistości jest po prostu poszukiwanie wydatniejszych kolorów w obrazie sceny wokół robota. Z tych kolorów trzeba następnie ręcznie wybrać coś nam odpowiadającego. Mówiąc innymi słowami należałoby obraz z przestrzeni RGB przekształcić do przestrzeni HSV i pokazać 3 składowe na 3 obrazach. Następnie należałoby wybrać operacje (m.in. progowania), które pozwolą wyodrębnić znacznik.

Rys. 7.3 jest niejasny. Obraz 7.3(b) jest kolorowy. Natomiast wynikający z niego obraz 7.3(c) jest z gradacją szarości. W jaki sposób znajdowanie maksimum zamienia wektor w każdym pikselu obrazu 7.3(b) na liczbę w obrazie 7.3(c)?

Na rys. 7.4 na jednej osi jest odcień, a na drugiej nasycenie. Poziom szarości w każdym pikselu prawdopodobnie jest proporcjonalny do liczby pikseli o zadanym odcieniu i nasyceniu. Przy podwyższeniu progu obraz zbinaryzowany ma coraz mniej białego, co potwierdzają rys. 7.4(b)-(f). Jednak podane progi 20, 30 itd. nie mają związku z obrazem. Pokazanie wyników dla różnych progów jest zbędne, a czytelnik nadal nie wie, jak dobrać progi. Tekst o wyborze barw wzorcowych bez uczenia dotyczy właściwie bardzo prymitywnej segmentacji obrazu i jest mało przydatny dla dalszych rozważań.

Na str. 74 "wykonanie histogramu" powinno być zastąpione "obliczeniem histogramu".

W pracy brak jest kompletnego algorytmu estymacji położenia i orientacji robota mobilnego (tj. algorytmu lokalizacji robota). Np. informacje o obliczeniu momentów służących do identyfikacji znacznika są rozrzucone w rozdz. 5 i 7.2, co w efekcie utrudnia określenie, co weszło w skład algorytmu, a co było tylko dyskusją możliwości opisanych w literaturze. Podobnie transformata Hougha jest omawiana w rozdz. 3.5 i ponownie w rozdz. 7.3.3.

Doktorant rozważa użycie trzech rodzajów cech do określenia, czy dany obiekt jest poszukiwanym znacznikiem. Cechy te to 8 znormalizowanych momentów centralnych, 2 niezmienniki zaproponowane przez Hu oraz jeden niezmiennik AMI (Affine Moment Invariant) zaproponowany przez Flussera. Znormalizowane momenty centralne i momenty niezmiennicze Hu charakteryzują się niezależnością od translacji obiektu, obrotu obiektu oraz zmiany skali obiektu w płaszczyźnie obrazu. Momenty te mają zastosowanie głównie do obiektów płaskich. Momenty niezmiennicze Flussera dopuszczają transformacje afiniczne w obrazie i gwarantują, że wynik będzie niezależny od wykonania tych transformacji. W oryginalnej publikacji Flussera [33] łatwo to prześledzić na zmianach kształtu liter, które mają być zidentyfikowane. Doktorant przebadał zmienność wszystkich znormalizowanych momentów centralnych, momentów Hu oraz momentów Flussera dla znaczników w kształcie liter U, H oraz O przy dokonywaniu zdjęć z różnych odległości i różnych kierunków. Na tej podstawie wytypował do zastosowania znacznik w kształcie litery O i wspomniany wyżej zestaw momentów $8+2+1=11$. Dany obiekt jest poszukiwanym znacznikiem, jeśli jego momenty mieszczą się w przedziałach znalezionych doświadczalnie. Metoda taka jest typowo heurystyczna. Nie jest ona zła i daje wyniki podane w pracy. Nasuwa się jednak wątpliwość, czy nie można było spróbować postąpić w sposób lepiej odpowiadający występującej tu sytuacji. Podstawowym źródłem zniekształcenia obrazu znacznika jest występowanie transformacji perspektywicznej. Suk i Flusser w artykule "Projective Moment Invariants", Trans. PAMI, vol 26, No 10, pp. 1364-67, 2004 dają przykład obliczenia momentów niezmienniczych w przypadku takiej transformacji. Chociaż opisana przez nich metoda ma charakter dość przyczynkowy, to biorąc pod uwagę, że znacznik jest obiektem sztucznie wygenerowanym, a więc w pewnym sensie dowolnym, można by spróbować zbadać możli-

wość udoskonalenia algorytmu opisanego w rozprawie. Powyższa uwaga ma charakter dyskusyjny, ale Doktorant powinien być w stanie w miarę dokładnie wyjaśnić tę sprawę.

Na rys. 7.23 i 7.24 powinny być opisane osie na wykresie transformaty Hougha. Dokładnie biorąc w wyniku tej transformaty powinniśmy znaleźć stosowną liczbę prostych krawędzi odpowiadających danemu znacznikowi. Jednak transformata może dać za mało lub za dużo krawędzi, albo w niewłaściwej kolejności. Wynika stąd dość skomplikowany algorytm obliczeniowy. Ponieważ algorytmu ani jego wyników nie podano, więc można domniemywać, że odpowiednie obliczenia są wykonywane ręcznie. Sytuacja taka znacznie spowalnia wykonanie obliczeń, które skądinąd Doktorant starał się przyspieszyć.

Wyniki w tabeli 7.6 są niedokładnie przedstawione. Co to są X_0 , Y_0 , itd? W jakich jednostkach są mierzone. Jakie są prawdziwe rozmiary znacznika? Dlaczego akurat Y_2 ma najgorszy rozkład. Rozkład ten w niczym nie przypomina normalnego - skąd on się może brać?

Wzór (8.6) wygląda na niepoprawny. Występuje w nim niezdefiniowany symbol X_{img} .

Na rys. 8.8 na osi pionowej jest prawdopodobnie błąd wielkości β . Natomiast na osi poziomej jest x , mierzone w cm . Zmienna x nie jest zdefiniowana. Do jakich wielkości wejściowych odnosi się podany błąd 1 piksla?

W tekście u dołu strony 107 występuje $X - DLX_p$, tzn. uwzględniona jest poprawka pozioma odnosząca się do odchylenia poziomego osi optycznej kamery. Jednak poprawki pionowej nie uwzględniono i występuje tylko V_p . Brak tu jest jakiegoś wyjaśnienia.

We wzorze (8.12) występuje $H_p - DLX_{cm}$, tzn. od długości w pikslach odejmujemy długość w cm . Wygląda to na błąd.

Na rys. 8.13 na osi pionowej jest prawdopodobnie odległość D liczona wg wzoru (8.12).

Na rys. 8.12 i 8.13 są różne krzywe, ale nie podano, co jest parametrem i jaka jest jego wartość.

Na str. 107, wiersz 5 od dołu Autor pisze "odległość od osi optycznej oznaczono $H_p - DLX_p$ ". Na str. 109, wiersz 6 od dołu jest napisane "odległości obrazowej od osi optycznej H_p ". Jak widać występuje niezgodność terminologii. Duża szkoda, że Autor nie ujednotwił terminologii i nie zrobił zestawienia symboli na początku pracy.

Na str. 113 zamiana D i β ma D_n i β_n jest dla mnie niezrozumiała. Czy są to te same wielkości?

W związku z opisem obliczenia błędów pomiaru na str. 113 nasuwają się następujące uwagi. Wyznaczenie D_n i β_n przypuszczalnie następuje wg wzorów (8.11) i (8.12). Co jednak zrobić z α , które we wzorach nie występuje.

Na rysunkach na str. 114 i 115 występuje wymienione wyżej α_n , które nie jest zdefiniowane.

Na str. 118 wzory (8.29) i (8.30) mają niezgodnione symbole i trudno sprawdzić poprawność podstawienia do wzorów. Niewątpliwie utrudnieniem jest też nie uzgodnienie jednostek. W nienumerowanych wzorach na dole strony 117 B_1 i B_2 są liczone w cm , a na rysunkach na str. 119 i 120 w pikslach.

Ocena ogólna rozprawy

Najważniejsze fakty przemawiające za pozytywną oceną rozprawy mgra inż. Ireneusza Hallmanna doktorskiej są:

1. Przeprowadzenie obszernych badań teoretycznych i doświadczalnych

2. Opublikowanie 9 artykułów, w tym 5 artykułów przedstawionych na konferencjach serii AUTOMATION, organizowanych przez PIAP w latach 2001-2005, artykułu na konferencję międzynarodową SIRS'2001, artykułu na konferencję międzynarodową CORES'2005, artykułu na Krajową Konferencję Robotyki(2001) oraz artykułu w piśmie Pomiar, Automatyka, Robotyka (2002).

Po stronie negatywnej jest chaotyczna prezentacja tezy, celów, osiągnięć oraz przebiegu obliczeń.

Ocena dorobku Autora w zakresie nie przedstawionym w pracy doktorskiej

W omawianym zakresie należy wymienić pracę Doktoranta na stanowisku programisty zarówno przed doktoratem jak i po. Prócz tego w okresie 2001-2006, a więc równoległe z doktoratem mgr I. Hallmann prowadził zajęcia w Wyższej Szkole Informatyki Stosowanej i Zarządzania w Warszawie.

Wniosek końcowy

Przedstawiona rozprawa doktorska i dorobek naukowy mgr inż. Ireneusza Hallmanna reprezentują poziom naukowy potwierdzony kilkoma publikacjami w literaturze krajowej i międzynarodowej.

W obecnej wersji rozprawa wymaga intensywnych poprawek, gdyż ze względu na niedopracowanie jest zbyt trudna do przesłania poprawności.

W szczególności:

- Sformułowanie tezy jest niefortunne.
- W pracy nie sprecyzowano, jakie są osiągnięcia Doktoranta i brak jest określenia, co Autor sam uważa za nowość w stosunku do istniejącej literatury.
- Brak jest zestawienia użytych symboli. Jeżeli objaśnienie symbolu jest zbyt długie i niemożliwe bez rysunku, to w zestawieniu powinna być podana strona i dodatkowy wzór definiujący dany symbol. Wszystkie oznaczenia powinny być ujednolicone, a zmienne oznaczone raczej pojedynczymi literami, ewentualnie z odpowiednimi indeksami.
- Brak jest całościowego algorytmu (grafu przepływu sygnałów lub schematu blokowego) ilustrującego we właściwej kolejności czynności, które należy wykonać w celu lokalizacji, tj. estymacji położenia i orientacji robota mobilnego na scenie.
- Występujące liczne usterki skutecznie utrudniają dokładne sprawdzenie poprawności wzorów.
- Praca zawiera wiele detali znanych z podręczników analizy obrazów. Detale te można było pominąć, albo co najwyżej wstawić do załącznika.

Zarówno recenzowana rozprawa jak i ogólny dorobek naukowy Doktoranta mogłyby spełniać wymagania Ustawy o stopniach i tytułach naukowych pod warunkiem jednak usunięcia wymienionych wyżej niedociągnięć.

Mariusz Aieniewski