

Polska Akademia Nauk
Instytut Podstawowych Problemów Techniki

**Lokalizacja robota mobilnego
względem automatycznie wybieranych obiektów**

mgr inż. Ireneusz Hallmann

Rozprawa doktorska

Promotor: Prof. dr hab. inż. Adam Borkowski

Warszawa, 2007

Spis treści

1	Wstęp	5
1.1	Teza i cel pracy	5
1.2	Charakterystyka proponowanej metody lokalizacji	5
1.2.1	Umiejscowienie pracy w robotyce	5
1.2.2	Motywacja pracy	7
1.3	Struktura pracy	8
2	Istniejące techniki lokalizacji	9
2.1	Wprowadzenie	9
2.2	Metody przyrostowe	12
2.2.1	Odometria	12
2.2.2	Nawigacja inercyjna	13
2.3	Triangulacja z wieloma znacznikami	14
2.3.1	Metoda Newtona-Raphsona	15
2.3.2	Geometryczne przecinanie się okręgów	15
2.3.3	Warunki dobrej dokładności metody triangulacyjnej	17
2.4	Filtry Bayesa	18
2.4.1	Filtr Kalmana	18
2.4.2	Filtr cząstkowy	19
2.5	Stereowizja	20
2.5.1	Elementy układu	20
2.5.2	Model boczno-kamerowy	22
2.5.3	Model osiowo-ruchowy	24
2.5.4	Przykład systemu opartego na jednej kamerze	25
2.6	Kamera dookólna	26
2.7	Kamera zainstalowana w otoczeniu	28
3	Reprezentacja i przetwarzanie obrazu	29
3.1	Reprezentacja obrazu	29
3.1.1	Reprezentacja obrazu szarego	29
3.1.2	Reprezentacja barw	29
3.1.3	Format RGB	29
3.1.4	Format HSV	30
3.1.5	Konwersje między modelami barw	30
3.2	Filtry liniowe	32
3.3	Detekcja krawędzi	34
3.3.1	Detektory Sobela i Prewitta	34

3.3.2	Operator Marra-Hildretha	34
3.4	Transformacje geometryczne	36
3.4.1	Transformacja afiniczna	36
3.4.2	Transformacja biliniowa	38
3.5	Transformata Hough'a	40
4	Kalibracja kamery	43
4.1	Parametry kamery	43
4.2	Klasyfikacja metod kalibracji kamery	45
4.2.1	Złożoność modelu	45
4.2.2	Pozostałe parametry	46
4.2.3	Przegląd publikacji dotyczących kalibracji kamery	47
4.3	Zastosowana metoda kalibracji	48
4.3.1	Przygotowanie do kalibracji	49
4.3.2	Znajdowanie położenia osi optycznej w obrazie	50
4.3.3	Zautomatyzowana kalibracja kamery	56
5	Detekcja i reprezentacja znaczników w literaturze	59
5.1	Momenty	59
5.1.1	Momenty geometryczne (kartezjańskie)	59
5.1.2	Niezmienniki	61
5.1.3	Niezmienniki przekształcenia afinicznego	61
5.2	Cechy niezmiennicze względem skali (SIFT)	62
5.3	Detekcja znaczników	63
5.3.1	Detekcja na podstawie barwy	64
5.3.2	Detekcja na podstawie kształtu	65
6	Charakterystyka proponowanej metody lokalizacji	66
6.1	Wprowadzenie	66
6.2	Zakres lokalizacji	69
6.3	Stosowane znaczniki	70
7	Zastosowana detekcja znaczników	72
7.1	Detekcja obiektów o określonej barwie	72
7.1.1	Wybór barw wzorcowych	72
7.1.2	Znajdowanie obszarów o charakterystycznej barwie	79
7.2	Detekcja charakterystycznych kształtów	83
7.3	Pomiar prostokąta opisanego na znaczniku	89
7.3.1	Detekcja linii w obrazie	90
7.3.2	Filtracja wykrytych krawędzi	90
7.3.3	Wyznaczenie równań linii	91
7.3.4	Określenie zewnętrznych krawędzi znacznika	95
7.3.5	Dokładność wykrywania krawędzi	98
8	Proponowana metoda lokalizacji	100
8.1	Wyznaczenie azymutu do obiektu	104
8.2	Wyznaczenie odległości do obiektu o znanej wysokości	107
8.2.1	Dokładność wyznaczenia odległości	108

8.3	Wyznaczenie kąta widzenia znacznika	116
8.3.1	Dokładność wyznaczenia kąta	118
8.4	Wyznaczanie położenia na podstawie jednej odległości	120
8.4.1	Dokładność wyznaczenia położenia	122
8.5	Wyznaczanie położenia na podstawie dwóch odległości	133
8.5.1	Dokładność wyznaczenia położenia	134
9	Pomiary i testy	141
9.1	Wyniki statyczne dla metody z pomiarem widocznej szerokości	141
9.1.1	Porównanie z wartościami oczekiwanymi	142
9.1.2	Własności statystyczne metody	148
9.2	Wyniki statyczne dla metody pomiaru dwóch odległości	151
9.2.1	Porównanie z wartościami oczekiwanymi	155
9.2.2	Własności statystyczne metody	155
9.3	Nawigacja oparta o kamerę jako sensor	159
10	Podsumowanie i wnioski	164
10.1	Dowód tezy	164
10.2	Porównanie z podobnymi metodami lokalizacji	166
	Literatura	167

Rozdział 1

Wstęp

1.1 Teza i cel pracy

Tezą niniejszej pracy jest sprawdzenie, czy jest możliwa lokalizacja robota mobilnego na podstawie jednego znacznika, w oparciu wyłącznie o dane pochodzące z jednej kamery zainstalowanej na robocie. Jedynym używanym sensorem jest pojedyncza kamera, którą można obracać niezależnie od robota. Niezależność obrotu kamery pozwala na lokalizację niezależnie od orientacji robota. Proponowana metoda pozwala traktować kamerę jak sensor odczytujący położenie względem widocznych znaczników.

W ramach dowodu tej tezy konieczne jest osiągnięcie następujących celów:

- wybór kształtu i typu znacznika umożliwiającego lokalizację tylko na podstawie jego widoczności – rozdział 6.3,
- wybór obiektów i cech odpowiednich do ich automatycznej identyfikacji – rozdział 7,
- identyfikacja obiektów na podstawie znalezionych cech – rozdział 7,
- opracowanie metody dokładnego wyznaczenia położenia kamery na podstawie jednego widocznego obiektu – rozdziały 8 i 9,
- eliminacja w jak największym stopniu błędów pomiarowych – rozdziały 4.3 oraz 7.3.1 – 7.3.5.

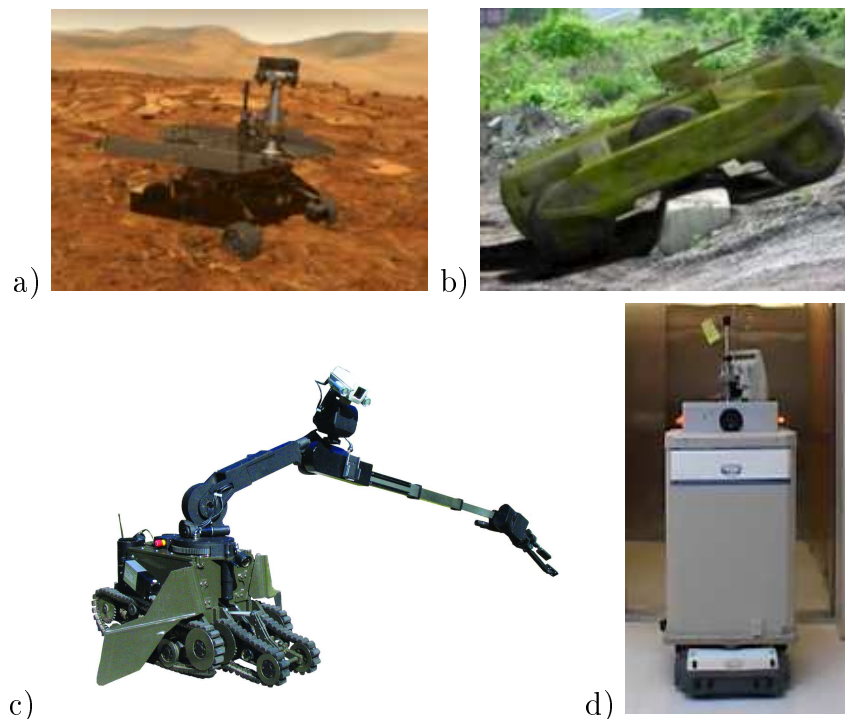
1.2 Charakterystyka proponowanej metody lokalizacji

1.2.1 Umiejscowienie pracy w robotyce

Pierwszymi stosowanymi praktycznie robotami były manipulatory, które do dziś stanowią większość przemysłowej produkcji robotów. Później pojawiły się roboty z ruchomą podstawą, które od razu można było podzielić na autonomiczne i zdalnie sterowane. Podział ten zachował się do dziś. Nie jest to podział ostry, ponieważ istnieją roboty zdalnie sterowane, które jednocześnie są w różnym stopniu autonomiczne. Roboty mobilne stosuje się dziś w:

- badaniach kosmosu jako samodzielne pojazdy (rys. 1.1.a);

- wojsku jako bezzałogowe pojazdy o różnych zastosowaniach, np. pojazdy zwiadowcze i inteligentne rakiety (rys. 1.1.b);
- wojsku i policji do przenoszenia i rozbrajania niebezpiecznych ładunków (rys. 1.1.c);
- zautomatyzowanych procesach produkcyjnych jako transport bezzałogowy;
- szpitalach do transportu (rys. 1.1.d);
- gospodarstwach domowych (pojawily się np. automatyczne kosiarki do trawy firmy Husquarna i odkurzacze firm Roomba i Electrolux).

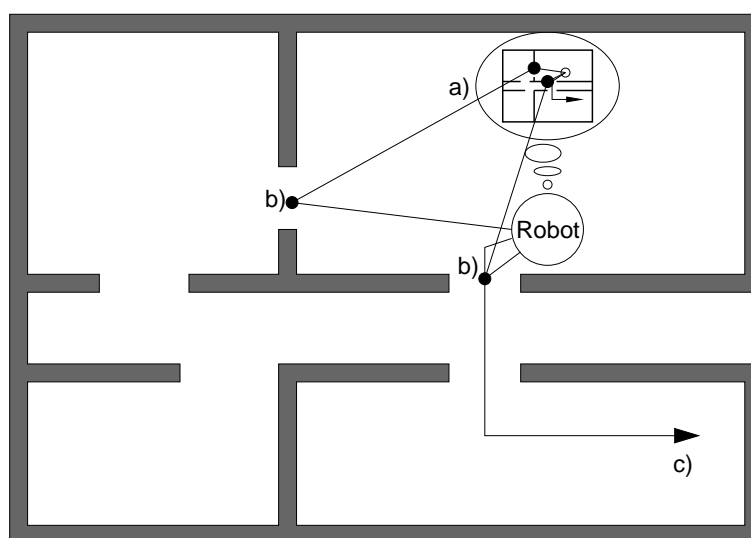


Rysunek 1.1: Przykładowe roboty mobilne: a) robot do badań kosmosu; b) bezzałogowy pojazd wojskowy; c) robot policyjny EXPERT ; d) szpitalny robot transportowy [51]

Wraz z upływem czasu roboty mobilne znajdowały zastosowanie w kolejnych dziedzinach.

W robotyce mobilnej wyznacza się dwie podklasy zadań dotyczących nawigacji: roboty działające w pomieszczeniach (*ang. indoor*) oraz poza pomieszczeniami (*ang. outdoor*). Niniejsza praca prezentuje metodę lokalizacji przeznaczoną dla robotów klasy *indoor*. W związku z mobilnością robotów pojawia się problem sterowania ich jazdą. Jeżeli robot jest zdalnie sterowany, problem ten może być rozwiązany przez człowieka sterującego ruchem robota. Roboty autonomiczne muszą wykonywać nawigację całkowicie samodzielnie. Wymaga to samodzielnego rozwiązywania przez robota zadań logicznych oraz sterowania ruchem. Nawigację można podzielić na pozyskiwanie mapy otoczenia i korzystanie z niej (*ang. mapping*), określanie aktualnej pozycji oraz planowanie i wykonanie trasy. Poglądowo przedstawia to rysunek 1.2.

Mapa otoczenia jest cyfrową reprezentacją otoczenia w pamięci komputera. Istnieje wiele sposobów reprezentacji otoczenia. Mapę otoczenia buduje się metodami statystycznymi i probabilistycznymi na podstawie wskazań sensorów. Zostały one krótko opisane w rozdziale 2.



Rysunek 1.2: Elementy nawigacji: a) pozyskanie i korzystanie z mapy; b) określenie własnego położenia; c) planowanie i przejazd trasą.

Lokalizacja robota mobilnego polega na wyznaczeniu położenia robota w otoczeniu na podstawie znanej mapy. Wyróżnia się lokalizację *lokalną* w układzie współrzędnych związanym z robotem lub *globalną* w układzie współrzędnych związanym z otoczeniem. W dalszej części tekstu pod pojęciem lokalizacji należy rozumieć lokalizację globalną. Pełne położenie robota opisują trzy wartości: współrzędne (*ang. position*) oznaczane najczęściej jako X i Y oraz orientacja Φ robota względem określonego kierunku w otoczeniu. Wyznaczone współrzędne robota wraz z orientacją określa się w literaturze nazwą *pose*. Większość sensorów i metod lokalizacji pozwala względnie łatwo wyznaczyć współrzędne robota, zaś dużo trudniej orientację. Wyjątkowym pod tym względem sensorem jest kamera dookólna (*ang. omniscam*), opisana w rozdziale 2.6, która pozwala na łatwe i dokładne wyznaczenie orientacji robota, przy czym wyznaczenie położenia robota jest obciążone dużym błędem.

W robotyce istnieje wiele technik lokalizacji, opartych na danych różnych typów z różnych źródeł (zob rozdz. 2). W zależności od okoliczności - dostępnych sensorów, czy lokalizacja jest wewnątrz budynku czy na zewnątrz, rozmiaru pomieszczeń, oświetlenia, zmienności otoczenia - różne metody okazują się lepsze lub gorsze. Wśród wielu metod lokalizacji nie ma idealnej lub uniwersalnej metody.

W ostatnich latach coraz popularniejsze staje się łączenie różnych technik lokalizacji oraz danych z różnych sensorów. Pozwala to na precyzyjniejsze określanie położenia robota. Wymaganie większej mocy obliczeniowej jest spełniane przez ciągły wzrost mocy obliczeniowej dostępnych komputerów. Z drugiej strony obecnie dostępna moc obliczeniowa pozwala na identyfikację obiektów widocznych z różnych miejsc, przy nie znanych wcześniej rozmiarach, położeniu i kształcie obiektów i budowanie map nieznanego otoczenia. Z tego powodu coraz popularniejsza staje się technika SLAM (*ang. Simultaneous Localization And Map building*). Jest to obecnie przedmiot bardzo intensywnych prac badawczych [76], [77], [28], [42], [41], [82], [87], [83], [37].

1.2.2 Motywacja pracy

W rozdziale 2 przedstawiono przegląd obecnie stosowanych metod lokalizacji. Motywacją do niniejszej pracy był fakt, że pojedyncza kamera często stanowi wyposażenie robota, zaś ist-

niejące techniki wymagały obecności co najmniej dwóch kamer (stereowizja - zob rozdz. 2.5) lub poszukiwania obiektów w rozległym obszarze (triangulacja - zob rozdz. 2.3). Wykorzystanie już zainstalowanej kamery do lokalizacji pozwala na obniżenie kosztów robota. W przypadku gdy kamera nie jest standardowym wyposażeniem robota, koszt jednej kamery jest niższy niż koszt układu dwu lub więcej kamer sprzężonych. W porównaniu z lokalizacją przy pomocy układu kamer (rozdz. 2.5) zastosowanie jednej kamery wymaga mniej skomplikowanego modelu, co upraszcza procedury kalibracji oraz wymaga mniejszej mocy obliczeniowej przeznaczonej dla potrzeb lokalizacji. Ponieważ układ kamer jest systemem bardziej skomplikowanym niż pojedyncza kamera, jest on bardziej narażony na awarie.

W niniejszej pracy zaproponowano metodę lokalizacji robota mobilnego na podstawie jednego znacznika. Opisano sposób przygotowania systemu do pracy, obróbki obrazu, rozpoznawania i ekstrakcji znaczników, lokalizację i oszacowanie dokładności lokalizacji. W rozdziale 9 przedstawiono wyniki testów dokładności oraz przebieg testu nawigacji robota wyłącznie na podstawie własnej metody lokalizacji. Uzyskane wyniki świadczą o możliwości zastosowania metody w rzeczywistych aplikacjach.

1.3 Struktura pracy

Układ pracy jest następujący:

- w rozdziale 2 przedstawiono przegląd metod lokalizacji, przy czym bardziej szczegółowo zostały omówione metody i techniki związane z metodą lokalizacji proponowaną w niniejszej pracy,
- w rozdziale 3 przedstawiono znane techniki obróbki obrazu, stanowiące podstawę techniczną proponowanej metody lokalizacji,
- w rozdziale 4 przedstawiono kilka technik kalibracji kamery, oraz zaproponowano bardzo efektywną technikę kalibracji bazującą na przekształceniu obrazu. Jedną z jej najważniejszych cech jest bardzo duża szybkość działania, co pozwala na lokalizację w czasie rzeczywistym. Ze względu na bardzo dużą ilość znanych technik kalibracji kamery trudno jednoznacznie powiedzieć, czy posiada ona nowe elementy,
- w rozdziale 8 przedstawiono krótką charakterystykę metody opracowanej w niniejszej pracy,
- w rozdziale 5 przedstawiono znane cechy reprezentujące obiekty, stosowane w proponowanej metodzie lokalizacji,
- w rozdziale 7 przedstawiono metodę detekcji znaczników przy zastosowaniu znanych technik obróbki obrazu (które opisano w rozdziale 3),
- w rozdziale 8 przedstawiono podstawy matematyczne oraz model dokładności proponowanej metody lokalizacji,
- w rozdziale 9 przedstawiono wyniki rzeczywistych pomiarów i prób lokalizacji przeprowadzonych z użyciem rzeczywistego robota. Wyniki obejmują lokalizację wykonywaną przy nieruchomej kamerze oraz zapis przebiegu jazdy robota przy lokalizacji wyłącznie proponowaną metodą.

Rozdział 2

Istniejące techniki lokalizacji

2.1 Wprowadzenie

Nie istnieje jedna uniwersalna metoda nawigacji i lokalizacji robota. Metody lokalizacji można podzielić na 2 grupy: przyrostowe (odometria i metody inercyjne) – opisane w rozdziale 2.2; oraz bezpośrednie, tzn. wyznaczające bezpośrednio określone parametry (lokalizacja z zastosowaniem kompasu, GPS, znaczników, nadajników).

Lokalizacja bezpośrednia może być wykonana z zastosowaniem różnych elementów. Elementami tymi są: znaczniki, sensory, nadajniki.

Sensor jest elementem pozwalającym wykryć i ustalić położenie obiektu. W robotyce istnieje wiele typów sensorów, m.in. zderzaki, czujniki podczerwieni, czujniki ultradźwiękowe (sonary), dalmierze laserowe, kamery. Wszystkie one są w stanie zarejestrować pewne sygnały i na tej podstawie pozwalają wnioskować o otoczeniu. Sensory można podzielić na aktywne i pasywne. W sensorach aktywnych stosuje się urządzenia wysyłające określone sygnały oraz detektory powracającego sygnału. Do tej grupy należą sonar i skaner laserowy. Czujniki pasywne nie wysyłają dodatkowych sygnałów - należą do nich kamera, licznik odometryczny i zderzak.

Sensory można podzielić w zależności od jednorazowo dostarczanych danych na **sensory wykonujące pojedynczy pomiar** (np. dalmierz, kamera) oraz **skanujące**, działające w trybie ciągłym. Sensory różnią się też **ilością dostarczanych danych** - np. dalmierz dostarcza pojedynczą wartość, podczas gdy kamera – jedną lub trzy macierze jasności o określonej wielkości, zależnej od rozdzielczości.

Znacznik (*ang. landmark*) jest obiektem pasywnym znajdującym się w otoczeniu robota. Jego podstawową cechą jest łatwość rozpoznania przy zastosowaniu określonego sensora. W przypadku czujnika laserowego mogą to być małe zwierciadła, w przypadku kamery - obiekty o określonych barwach (zob. rozdz. 7.1), kształcie (zob. rozdz. 7.2) lub innych łatwo rozpoznawalnych cechach. Znaczniki dzielą się na *naturalne* - stanowiące istniejący element otoczenia przed lokalizacją i *sztuczne* - wprowadzone tylko w celu lokalizacji.

Nadajnik (*ang. beacon*) jest obiektem aktywnym, emitującym sygnał dzięki któremu łatwo można go odnaleźć stosując sensor pasywny. Obiektem takim może być np. nadajnik ultradźwiękowy o określonej częstotliwości nadawania lub źródło światła o określonych cechach, np. barwie lub polaryzacji.

Rozmieszczenie elementów systemów lokalizacyjnego może być następujące:

- sensory są wyposażeniem robota, a obiekty charakterystyczne znajdują się w otoczeniu robota;

- sensory są zamontowane w otoczeniu robota, a elementy charakterystyczne stanowią wyposażenie robota.

W pierwszym przypadku przestrzeń w której może poruszać się robot, jest ograniczona do obszaru, w którym można rozpoznać obiekty, w drugim - do obszaru widzianego przez sensory. Obiekty charakterystyczne w otoczeniu robota mogą być różne, dzieli się je na znaczniki naturalne, znaczniki sztuczne i aktywne elementy. Jeżeli sensory znajdują się w otoczeniu robota – np kamery w pomieszczeniach – wówczas do identyfikacji robota wystarczy jeden znacznik. W tym przypadku robot może być wyposażony w sensory umożliwiające bezkolizyjną jazdę, zaś przemieszczenia nadzoruje system centralny.

Sensory umieszczone w pomieszczeniu z jednej strony pozwalają odczytać aktualną pozycję śledzonego robota bez efektów pamięciowych i przez to są o wiele łatwiejsze do modelowania (co powoduje że jest z tym związana mniejsza ilość zagadnień do rozwiązania). Z drugiej strony stacjonarny sensor ogranicza lokalizację do otoczenia, w którym się znajduje, co może być niewykonalne w niektórych warunkach. Do stacjonarnych sensorów należą: system GPS umożliwiający lokalizację w terenie i podwieszona po suficie kamera lub sensory ultradźwiękowe umożliwiające lokalizację w pomieszczeniach zamkniętych.

Sytuacja, w której sensory stanowią wyposażenie robota, jest bardziej skomplikowana. Zależnie od typu sensora, są one w stanie mierzyć przemieszczenie robota względem poprzedniej pozycji (licznik odometryczny) lub mierzyć wyznaczać i mierzyć cechy otoczenia, przez co można wnioskować o położeniu robota (skaner laserowy, kamera, czujnik podczerwieni i ultradźwiękowy). Czujniki te podają m.in. następujące dane:

- odległość do najbliższego obiektu w określonym kierunku (czujnik ultradźwiękowy i podczerwieni),
- odległość do najbliższego obiektu w wielu kierunkach (skaner laserowy),
- odległość do najbliższego obiektu w określonym kierunku z dużą rozdzielczością (układ kamer stereowizyjnych).

Sensory robota wyznaczają pewną przestrzeń, w której należy odnaleźć i zidentyfikować elementy charakterystyczne - znaczniki. W rozdziale 5 opisano znane cechy obiektów umożliwiające detekcję i identyfikację znaczników.

Nawigację wewnątrz pomieszczenia na podstawie kamery można podzielić na trzy klasy: nawigację opartą o zadaną mapę otoczenia (*ang. map-based*), nawigację opartą o stworzoną na bieżąco mapę otoczenia (*ang. map-building-based*) i nawigację bez mapy [26]. Pierwszy i drugi rodzaj nawigacji z mapą różnią się pochodzeniem mapy: zadaną mapę otoczenia tworzy człowiek i wprowadza ją do pamięci robota, zaś utworzoną mapę otoczenia tworzy robot podczas eksploracji budynku.

W przypadku nawigacji opartej na kamerze wyróżnia się 4 etapy przetwarzania danych [26]:

- Akwizycja danych z sensorów, czyli otrzymanie obrazu z kamery;
- Detekcja znaczników – zazwyczaj oznacza to ekstrakcję krawędzi, wygładzanie, filtrowanie i segmentację regionów na podstawie jasności, barwy, głębi lub ruchu;
- Dopasowanie obserwacji i oczekiwań – system próbuje dopasować znalezione znaczniki do informacji w bazie danych biorąc pod uwagę określone kryteria dopasowania;

- Wyznaczenie pozycji robota na podstawie widocznych znaczników oraz mapy otoczenia.

Można wyróżnić trzy metody lokalizacji na podstawie obrazu z kamery: lokalizację bezpośrednią, lokalizację przyrostową i lokalizację na podstawie śledzenia znaczników.

Lokalizacja bezpośrednia pozwala na wyznaczenie położenia robota generalnie na podstawie pojedynczego pomiaru. Problem lokalizacji można sprowadzić do:

- przypisania etykiet wykrytym znacznikom przez dopasowanie ich do posiadanej mapy oraz,
- wyznaczenia położenia robota na podstawie każdej kombinacji zidentyfikowanych znaczników. Zazwyczaj dokonuje się tego dla trójek znaczników metodą triangulacji (zob. rozdz. 2.3, [22] [65]) lub oddzielnie dla każdego znacznika dla układu sterowizyjnego (zob. rozdz. 2.6, [69]).

Jeżeli jednocześnie wyznaczonych zostanie wiele pozycji robota na podstawie wielu znalezionych znaczników, pozwala to na dodatkowe zwiększenie dokładności pomiaru. Obecnie znane są zaawansowane metody agregacji danych pomiarowych, pozwalające na redukcję błędów pomiarowych w czasie i ustalenie położenia globalnego robota w sytuacji, gdy nie można jednoznacznie stwierdzić które znaczniki są widoczne. Przykładem takiego narzędzia jest filtr cząstkowy (*ang. particle filter*), opisany w rozdziale 2.4.2.

Lokalizacja przyrostowa zakłada, że pozycja robota jest znana w przybliżeniu *a priori*. Konieczna jest tylko aktualizacja położenia na podstawie ostatnio znalezionych znaczników, do czego stosuje się wiele narzędzi probabilistycznych i statystycznych. Istotna jest znajomość statystycznej charakterystyki sensora. Aktualizację położenia przeprowadza się stosując filtr Kalmana, sieci neuronowe, logikę i sterowanie rozmyte. Stosowany aparat matematyczny jest o wiele bardziej skomplikowany niż w lokalizacji bezpośredniej [26].

Lokalizacja na podstawie śledzenia znacznika polega na wyznaczaniu położenia na podstawie wyznaczonych znaczników spośród wielu widocznych. Zazwyczaj ilość śledzonych znaczników jest tak dobrana, aby pogodzić dokładność lokalizacji z ograniczeniami czasowymi. Jeżeli któryś ze znaczników przestanie być widoczny, system wyznacza inny znacznik spośród aktualnie widocznych.

Przy lokalizacji za pomocą większej liczby sensorów możliwe jest stosowanie sensorów tego samego typu lub różnych typów. Sensory różnych typów mogą rozpoznawać różne obiekty, a w przypadku identyfikacji tych samych obiektów różnymi sensorami wzrasta dokładność lokalizacji.

Całkowicie odmienną techniką lokalizacji jest lokalizacja topologiczna, pozwalająca na określenie w przybliżeniu miejsca w którym znajduje się robot na podstawie obecności określonych obiektów w obrazie (*ang. appearance vision*). Nadaje się ona najlepiej np. do stwierdzenia w którym pomieszczeniu robot się znajduje. Pierwsze wersje tej metody pozwalały na określenie widoczności obiektów przez wyznaczenie korelacji wzajemnej obrazu z zapamiętanym wzorcem [2] [24] [25] [21]. Metody te wymagają przechowania dużej objętościowo bazy danych ze względu na przechowywane obrazy, choć widoczny jest stały postęp w tej dziedzinie, np w pracy [93]. Nowszą odmianą tej metody zaproponowano w pracach [68], [67]. Polega ona na zapamiętaniu zamiast serii obrazów referencyjnych części widma fourierowskiego obrazu pomieszczenia uzyskanego za pomocą kamery dookólnej.

2.2 Metody przyrostowe

2.2.1 Odometria

Odometria jest jedną z najstarszych i szeroko stosowanych metod nawigacji w robotyce mobilnej. Jej cechami są dobra dokładność w krótkim okresie czasu, niska cena i wysoka częstotliwość próbkowania. Podstawową ideą odometrii jest całkowanie przyrostowej informacji w czasie, co prowadzi do nieograniczonej akumulacji błędów wraz z upływem czasu i odległością przebytą przez robota. Błędy orientacji prowadzą do dużych błędów pozycji robota. Mimo tych ograniczeń odometria jest ważną częścią systemów nawigacyjnych. Istnieje wiele sposobów poprawy dokładności odometrii, np. przez połączenie jej z okresowym wyznaczaniem położenia metodą bezpośrednią.

Założmy, że w napędzie różnicowym po krótkim czasie I prawy i lewy licznik zliczyły odpowiednio N_R i N_L impulsów [19]. Założmy, że

$$c_m = \pi D_n / n C_e, \text{ gdzie :} \quad (2.1)$$

c_m – współczynnik przeliczenia impulsów licznika na przebytą odległość,

D_n – nominalna średnica kół,

C_e – rozdzielczość licznika w impulsach na obrót,

n – współczynnik biegu - przeliczenie obrotów silnika na obroty kół

Przebyty w tym czasie dystans $\Delta U_{R/L,I}$ wyraża się wzorem:

$$\Delta U_{R/L,I} = c_m N_{R/L,I} \quad (2.2)$$

Równanie (2.2) pozwala wyznaczyć przebytą odległość i rotację robota. Jeżeli obroty kół są bezbłędnie przeliczone na przesunięcie, wówczas odometria jest metodą dokładną. Jednak w praktyce występują m.in. poślizgi, powodujące błędy pomiaru. Błędy odometrii dzieli się na systematyczne i niesystematyczne [16]. Błędy systematyczne wynikają z niedokładności mechanicznych, np. niejednakowych rozmiarów kół. Błędy niesystematyczne wynikają ze zjawisk zachodzących podczas jazdy: poślizgów lub zderzenia z przeszkodą.

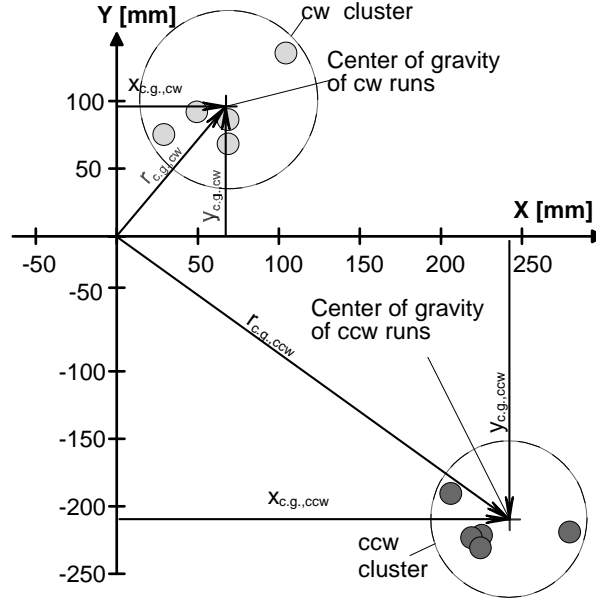
W typowym przypadku, gdy odometrię łączy się z lokalizacją znacznikową, gęstość rozmieszczenia znaczników należy wyznaczyć empirycznie na podstawie największych błędów systematycznych.

Wyznaczenie błędów odometrii

Metoda ilościowego pomiaru błędów systematycznych i związanej z tym kalibracji odometrii została opisana przez Borensteina i Fenga w pracach [19] i [20]. Ta metoda, zwana *UMBmark*, polega na pomiarze błędów odometrii podczas przejechania przez robota zaprogramowanej trasy. W teście robot jedzie po obwodzie kwadratu o boku długości 4 metrów - 5 razy zgodnie z kierunkiem ruchu wskazówek zegara, i 5 razy w przeciwną stronę.

Po przejechaniu zaprogramowanej trasy porównuje się rzeczywistą pozycję robota z wyznaczoną przez odometrię. Błędy grupują się w dwóch miejscach, podobnie jak na rysunku 2.1. Rozproszenie błędów w obu grupach wynika z błędów niesystematycznych. Z rysunku 2.1 wynika, że błędy systematyczne mogą być znacznie większe niż niesystematyczne. Asymetria błędów przy zmianie kierunku wynika z dwóch składowych błędów odometrii, nazywane błędami typu A i B. Błędy typu A mają te same wartości i znaki niezależnie od kierunku skręcania w teście, zaś błędy typu B – różne znaki i te same wartości, jeżeli robot skręca teście w raz w

prawo a drugi raz w lewo o ten sam kąt 90° . Błędy typu A są powodowane typowo niedokładną znajomością średnic kół, zaś typu B – różnicą średnic kół.



Rysunek 2.1: Typowy rozkład błędów w teście UMBmark [16].

Wyznacza się wartość średnią opisującą dokładność odometrii:

$$Err_{max,syst} = \max(r_{c.g.,cw}; r_{c.g.,ccw}), \quad (2.3)$$

gdzie $r_{c.g.,cw}$ oznacza błąd przy skręcaniu w prawo, zaś $r_{c.g.,ccw}$ - przy skręcaniu w lewo:

$$r_{c.g.,cw} = \sqrt{(x_{c.g.,cw})^2 + (y_{c.g.,cw})^2},$$

$$r_{c.g.,ccw} = \sqrt{(x_{c.g.,ccw})^2 + (y_{c.g.,ccw})^2}.$$

Na podstawie testu UMBmark Borenstein i Feng opracowali procedurę kalibracji w celu redukcji błędów systematycznych odometrii. Pozwala ona [16] na zmniejszenie błędów ok $10 \div 20$ krotnie.

Borenstein i Feng [20] opracowali też metodę pomiaru błędów niesystematycznych, nazwaną rozszerzony UMBmark (*extended UMBmark*). Pozwala ona na porównanie różnych robotów w podobnych warunkach, choć pomiar błędów niesystematycznych jest użyteczny ponieważ te błędy silnie zależą od podłoża. Stosując dobrze określone nierówności podłoża, można określić charakterystykę odometrii robota, co zostało opisane w pracy [18]. W pracy [17] opisano metodę korekcji niesystematycznych błędów odometrii, polegającą na wzajemnej korekcji położenia przez dwa roboty.

2.2.2 Nawigacja inercyjna

Główną cechą nawigacji inercyjnej jest całkowita autonomiczność – nie są wymagane żadne punkty odniesienia czy interakcja z otoczeniem. W nawigacji inercyjnej stosuje się akcelerometry do wyznaczenia przyspieszenia i rotacji. Pomiaru całości się w celu wyznaczenia położenia.

Wadą nawigacji inercyjnej jest, podobnie jak w odometrii, nieograniczony wzrost błędu położenia z upływem czasu. Mały stały błąd pomiaru jest w stanie spowodować bardzo duży błąd położenia w dłuższym czasie. Z tego powodu nawigacja inercyjna nie sprawdza się w robotyce.

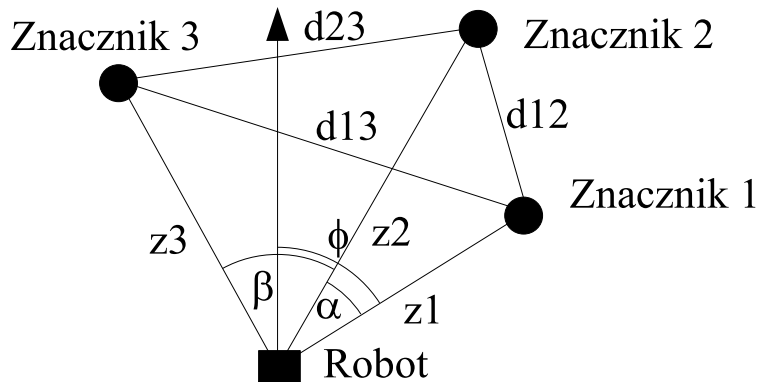
Wyniki stosowania akcelerometrów do nawigacji w robotyce są raczej słabe. W niektórych zastosowaniach nawigacja inercyjna sprawdza się znakomicie (np. dla łodzi podwodnych). W robotyce próbowano stosować tańsze akcelerometry niż te instalowane w łodziach podwodnych. Okazuje się, że dla małych wartości przyspieszeń (np. podczas powolnych obrotów) odczytywane wartości są obciążone dużym błędem względnym. Akcelerometry są też wrażliwe na nierówności podłoża. W pracy [10] Barshan i Durrant-Whyte opisali metodę eliminacji wpływu przyspieszenia ziemskiego na wskazania akcelerometru. Wyznaczane pozycje były obciążone błędem narastającym z prędkością $1 \div 8$ cm/s, co jest wartością zbyt dużą dla większości zastosowań w robotyce mobilnej.

Żyroskop jest szczególnie ważnym sensorem dla robotyki mobilnej, ponieważ można za jego pomocą skompensować największą słabość odometrii: błędy orientacji, powodujące narastające w czasie błędy pozycjonowania. Spadek cen żyroskopów powoduje, że są to urządzenia które można z powodzeniem stosować w robotyce mobilnej [16].

2.3 Triangulacja z wieloma znacznikami

Triangulacja jest przykładem lokalizacji bezpośredniej, gdzie na podstawie pojedynczego pomiaru wyznacza się położenie. Jest to bardzo stara metoda lokalizacji. Stosuje się ją od dawna np. w nawigacji morskiej. Określenie położenia wymaga znajomości mapy otoczenia oraz detekcji trzech znanych znaczników. Do każdego znacznika wyznacza się kąt widzenia względem określonego kierunku bazowego (azymut). Sytuację w rzucie z góry przedstawia rysunek 2.2. Można łatwo napisać układ równań (2.4) opisujący zależności między znanymi odległościami d_{12} , d_{13} , d_{23} , zmierzonymi kątami α i β oraz nieznanymi odległościami Z_1 , Z_2 i Z_3 :

$$\begin{cases} d_{12}^2 = z_1^2 + z_2^2 - 2 \cos(\alpha) \\ d_{23}^2 = z_2^2 + z_3^2 - 2 \cos(\beta) \\ d_{13}^2 = z_1^2 + z_3^2 - 2 \cos(\alpha + \beta) \end{cases} \quad (2.4)$$



Rysunek 2.2: Elementy lokalizacji metodą triangulacji

Zadanie wydaje się pozornie proste, jednak wyznaczenie położenia wymaga rozwiązania układu trzech równań nieliniowych. Istnieje wiele metod rozwiązywania układów równań nielin-

owych, jednak ich czas wykonania nie pozwala na zastosowanie w systemach czasu rzeczywistego.

W pracy [22] znajduje się analiza wyznaczania pozycji z zastosowaniem metod triangulacyjnych. Autorzy przedstawili 4 algorytmy, porównując ich stopień skomplikowania, złożoność obliczeniową oraz odporność na błędy. Proponowane metody iteracyjna i geometryczna okazują się mało praktyczne. Z praktycznych rozwiązań autorzy proponują metodę Newtona-Raphsona do rozwiązywania układu równań lub metodę przecinania się kół.

2.3.1 Metoda Newtona-Raphsona

Ponieważ równania triangulacji są nieliniowe, znalezienie rozwiązania metodą Newtona-Raphsona nie jest proste, ani nawet gwarantowane ze względu na możliwość zbieżności do minimum lokalnego. Jednak wiadomo, że rozwiązanie istnieje, ponieważ robot istnieje w świecie rzeczywistym i ma określone położenie i orientację. Autorzy zakładają, że jest znane przybliżone położenie robota. Dla układu trzech równań ogólną formę równań nieliniowych przedstawia równanie:

$$f_i(x_1, x_2, x_3) = \tan(Lo_i + Ro) + \frac{Ly_i - Ry}{Lx_i - Rx} = 0; \quad i = 1, 2, 3 \quad (2.5)$$

gdzie Lx_i i Ly_i oznaczają współrzędne znacznika i , Rx i Ry – współrzędne robota, Lo_i – azymut znacznika i względem przodu robota, zaś Ro – orientację robota. Gdy $X = [x_1, x_2, x_3]^T$, przyjęte rozwinięcie funkcji f_i w sąsiedztwie punktu X przedstawia równanie:

$$f_i(X + \partial X) = f_i(X) + \frac{\partial f_i}{\partial x_1} * \partial x_1 + \frac{\partial f_i}{\partial x_2} * \partial x_2 + \frac{\partial f_i}{\partial x_3} * \partial x_3. \quad (2.6)$$

Pochodne wyższych rzędów są pominięte, i w ten sposób otrzymuje się układ równań liniowych. Podczas rozwiązywania otrzymuje się poprawki ∂X zbliżające wszystkie funkcje jednocześnie do rozwiązania. Wyznacza się je podstawiając do równania (2.5) 0 za $f_i(X + \partial X)$:

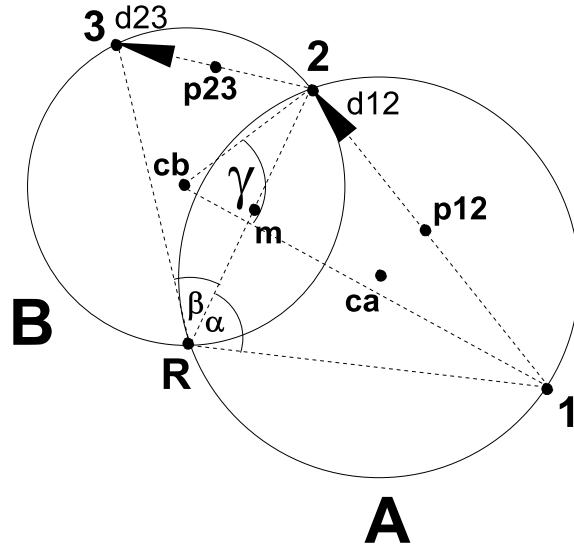
$$\frac{\partial f_i}{\partial x_1} * \partial x_1 + \frac{\partial f_i}{\partial x_2} * \partial x_2 + \frac{\partial f_i}{\partial x_3} * \partial x_3 = -f_i(X). \quad (2.7)$$

Autorzy piszą o stosowaniu dekompozycji LU do rozwiązania równania (2.7). Z przeprowadzonych badań wynika, że metoda nie zawsze jest zbieżna nawet jeżeli punkt X leży blisko rozwiązania. Z kolei jeżeli otrzyma się rozwiązanie, łatwo można sprawdzić czy jest ono prawidłowe. Jeżeli układ równań (2.5) jest zbieżny rozwiązanie uzyskuje się bardzo szybko. Jeżeli układ jest rozbieżny, zazwyczaj daje się to wykryć w pierwszych 10 iteracjach. Metoda Newtona-Raphsona jest dobrym testem przed zastosowaniem innych, bardziej przewidywalnych metod triangulacji.

2.3.2 Geometryczne przecinanie się okręgów

Według tej metody, robot R oraz znaczniki 1 i 2 jednoznacznie określają okrąg A (rys. 2.3). Nawet jeżeli położenie robota nie jest znane, możliwe są tylko dwa okręgi, ponieważ kąt między znacznikami jest znany. Jeżeli uwzględni się kolejność znaczników, może istnieć tylko jeden okrąg. Analogicznie robot oraz znaczniki 2 i 3 tworzą okrąg B. Okręgi przecinają się w dwóch punktach : w położeniu robota R oraz w znaczniku 2. Rysunek 2.3 przedstawia układ trzech znaczników i robota oraz podstawowe oznaczenia geometryczne zastosowane w algorytmie. Oto lista symboli użytych w algorytmie:

- ca cb – środki okręgów A i B,
- d12 – odległość między znacznikami 1 i 2,
- d23 – odległość między znacznikami 2 i 3,
- z2 – odległość między robotem i znacznikiem 2,
- ra – promień okręgu A,
- rb – promień okręgu B,
- la – odległość między ca a odcinkiem łączącym znaczniki 1 i 2,
- lb – odległość między cb a odcinkiem łączącym znaczniki 2 i 3,
- m – środek odcinka z2 łączącego robota ze znacznikiem 2,
- c2m – odległość między cb a m.



Rysunek 2.3: Trzy znaczniki pozwalają na jednoznaczne wyznaczenie położenia robota – metoda geometrycznego przecięcia okręgów [22]

Poniżej zamieszczono schemat wyznaczenia położenia i orientacji robota metodą geometrycznego przecięcia okręgów. Równania są wymienione w odpowiednich miejscach schematu.

$$\alpha = Lo2 - Lo1 \quad (2.8)$$

$$\beta = Lo3 - Lo2 \quad (2.9)$$

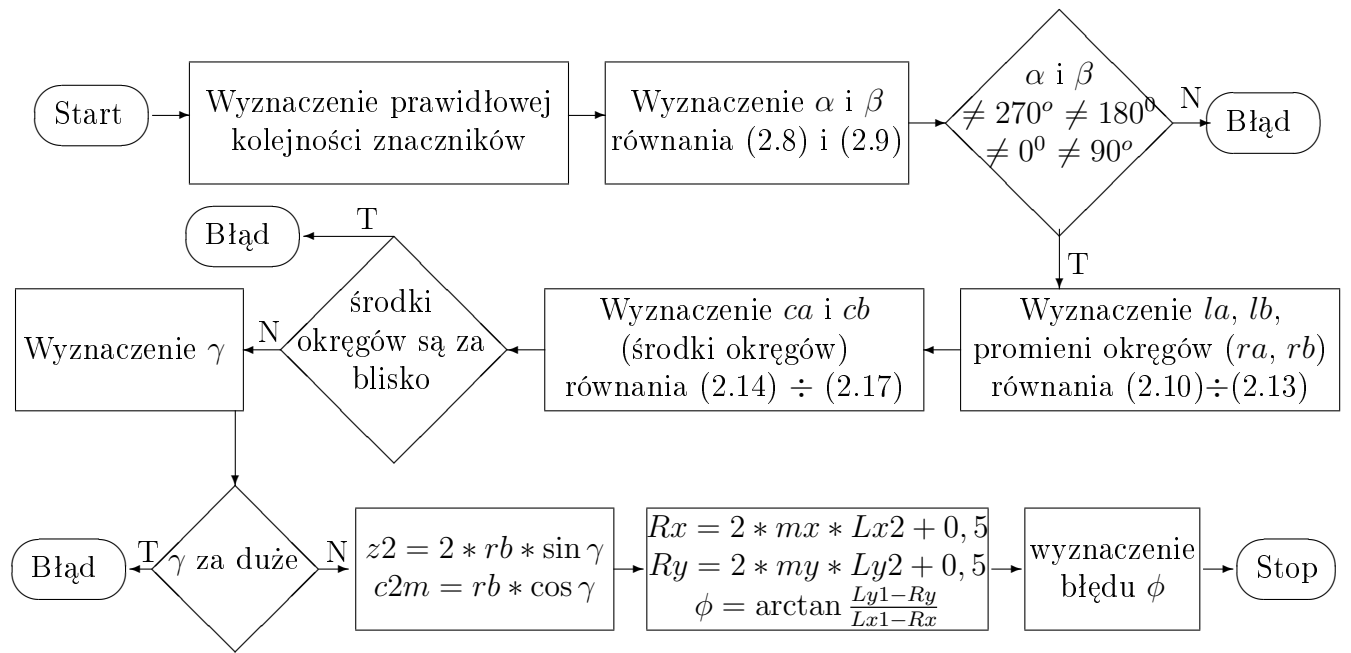
$$ra = \frac{d12}{2 \sin \alpha} \quad (2.10)$$

$$rb = \frac{d23}{2 \sin \beta} \quad (2.11)$$

$$la = \frac{d12}{2 \tan \alpha} \quad (2.12)$$

$$lb = \frac{d23}{2 \tan \beta} \quad (2.13)$$

$$cax = p12x - la * v12y \quad (2.14)$$



Rysunek 2.4: Schemat algorytmu wyznaczania położenia metodą geometrycznego przecięcia okręgów

$$cay = p12y - la * v12x \quad (2.15)$$

$$cbx = p23x - lb * v23y \quad (2.16)$$

$$cby = p23y - lb * v23x \quad (2.17)$$

2.3.3 Warunki dobrej dokładności metody triangulacyjnej

Metoda triangulacyjna opiera się na pomiarach kąta widzenia trzech znaczników. Wymaga ona, aby wszystkie znaczniki znajdowały się na tym samym obrazie lub aby robot był nieruchomy między wykonaniem kolejnych obrazów jeżeli znajdują się na różnych obrazach. Zakłada przy tym, że istnieje możliwość wykonania obrotu kamery o 360° wokół osi pionowej niezależnie od napędu robota. Na uzyskanym obrazie może znajdować się wiele znaczników. Konieczna jest znajomość mapy otoczenia, rozmieszczenia znaczników oraz możliwość estymacji aktualnego położenia innymi metodami lub identyfikacji znaczników.

W pracy [65] znajduje się szczegółowa analiza dokładności lokalizacji metodą triangulacji. Znaczniki w tych pracach są małe i traktuje się je jak obiekty punktowe, czyli o nieistotnych (zerowych) rozmiarach. Autorzy skupili się na oszacowaniu dokładności dla znaczników widocznych w jednym obrazie, co daje kąt widzenia 53° . Teoretyczna analiza powstających błędów prowadzi do rezultatu, że nie wszystkie trójki znaczników jednakowo nadają się do lokalizacji.

Pierwszy sposób uzyskania dobrej dokładności w pracy [65] polega na możliwości swobodnej rotacji kamery wokół własnej osi. Następujące czynniki wpływają na dokładność pomiarów:

- jeżeli środkowy znacznik jest bliżej robota niż linia łącząca dwa skrajne, wówczas dokładność pomiaru jest największa;
- współliniowe znaczniki dają najgorszą dokładność;
- dokładność lokalizacji maleje ze wzrostem odległości robota od znaczników;
- znaczniki powinny być rozproszone w otoczeniu;

- zwiększenie długości ogniskowej kamery zwiększa dokładność.

W pomieszczeniu 8x10 metrów powierzchnia elipsy niepewności dla najgorszej konfiguracji znaczników wyniosła w konkretnych przypadkach średnio $2m^2$ i maksymalnie $12m^2$. W tej samej pozycji robota, najlepsze konfiguracje znaczników pozwalały uzyskać elipsy niepewności o powierzchni $0,03m^2$ i $0,06m^2$. Metoda ma wadę w postaci konieczności wyszukiwania wielu znaczników. W pracy [60] proponuje się przeszukanie otoczenia tworząc stereowizyjną parę obrazów panoramicznych z mapą znaczników.

W rzeczywistej lokalizacji dane pomiarowe są zawsze obarczone błędami. Zależnie od tych błędów, położenia robota względem znaczników oraz znaczników względem siebie, otrzymuje się różną wrażliwość wyników na błędy pomiarów. W pracach [13] i [14] proponuje się metodę pozwalającą na dokładniejsze wyznaczenie położenia robota przy większej ilości znaczników. Metoda ma złożoność liniową względem ilości znaczników i opiera się na rozwiązaniu układu równań nieliniowych z ilością równań przekraczającą ilość niewiadomych. Dla lepszego przybliżenia rozwiązania stosuje się metodę najmniejszych kwadratów.

2.4 Filtry Bayesa

Filtr Bayesa stosuje się do ważonej agregacji danych pomiarowych, zmiennych w czasie, w celu uzyskania większej dokładności pomiarów. Do najbardziej znanych implementacji filtru należą różne odmiany filtru Kalmana – oznaczane skrótami KF (ang. Kalman Filter), EKF (ang. Extended Kalman Filter), UKF (ang. Unscented Kalman Filter), testowanie hipotez oraz filtr cząstkowy (ang. Particle Filter).

2.4.1 Filtr Kalmana

Filtr Kalmana w ogólnej postaci nadaje się do przewidywania położenia na podstawie wektorów obserwowanych danych oraz znajomości funkcji transformujących wektor danych na wartość śledzoną, którą w danym przypadku jest położenie robota. Wektor stanu (o rozmiarze n) w chwili t oznacza się symbolem x_t , wektor sterowania (o rozmiarze m) oznacza się jako u_t , zaś wektor obserwacji (o rozmiarze k) oznacza się symbolem z_t . Dla filtru Kalmana w wersji podstawowej równanie stanu ma postać:

$$\mathbf{x}_t = \mathbf{A}_t \mathbf{x}_{t-1} + \mathbf{B}_t \mathbf{u}_t + \varepsilon_t, \quad (2.18)$$

gdzie mierzy się wartości z_t :

$$\mathbf{z}_t = \mathbf{C}_t \mathbf{x}_t + \delta_t. \quad (2.19)$$

Znaczenie symboli jest następujące:

- \mathbf{A}_t jest macierzą ($n \times n$) opisującą zmianę stanu od chwili $t - 1$ do chwili t przy braku zakłóceń.
- \mathbf{B}_t jest macierzą ($n \times l$) opisującą jak zmienia się sterowanie ze stanu $t - 1$ do t .
- \mathbf{C}_t jest macierzą ($k \times n$) opisującą przejście z wektora obserwacji z_t na wektor stanu x_t .
- ε_t i δ_t są zmiennymi losowymi reprezentującymi odpowiednio zakłócenia przetwarzania i pomiaru, przy założeniu że są niezależne i mają rozkłady gaussowskie z kowariancją odpowiednio R_t i Q_t .

Jak widać z równań (2.18) i (2.19), zakłada się liniowość wszystkich transformacji. Gęstość prawdopodobieństwa dla wielowymiarowego wektora stanu jest następująca:

$$p(x) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} \exp -\frac{1}{2}(x - \mu)^T \Sigma^{-1}(x - \mu) \quad (2.20)$$

Transformacje liniowe z początkowymi dystrybucjami gaussowskimi dają w wyniku również dystrybucje Gaussa, co symbolicznie przedstawiają równania (2.21) i (2.22):

$$\left. \begin{array}{l} X \sim N(\mu, \sigma) \\ Y = aX + b \end{array} \right\} \Rightarrow Y \sim N(\mathbf{A}\mu + \mathbf{B}, \mathbf{A}\Sigma\mathbf{A}^T), \quad (2.21)$$

$$\left. \begin{array}{l} X_1 \sim N(\mu_1, \Sigma_1^2) \\ X_2 \sim N(\mu_2, \Sigma_2^2) \end{array} \right\} \Rightarrow p(X_1) * p(X_2) \sim N\left(\frac{\Sigma_2}{\Sigma_1 + \Sigma_2}\mu_1 + \frac{\Sigma_1}{\Sigma_1 + \Sigma_2}\mu_2, \frac{1}{\Sigma_1^{-1} + \Sigma_2^{-1}}\right). \quad (2.22)$$

W kolejnym kroku t wyznacza się dwie wartości: przewidywaną w kolejnym kroku $t + 1$ wartość stanu (równanie (2.23)) i poprawkę (równanie (2.24)):

$$\bar{bel}(x_t) = \int p(x_t|u_t, x_{t-1})bel(x_{t-1})dx_{t-1}, \quad (2.23)$$

$$bel(x_t) = \eta p(z_t|x_t) \bar{bel}(x_t). \quad (2.24)$$

Dla wielowymiarowych zmiennych równanie (2.24) przekształca się do postaci:

$$bel(x_t) = \left\{ \begin{array}{l} \mu_t = \bar{\mu}_t + \mathbf{K}_t(\mathbf{z}_t - \mathbf{C}_t\bar{\mu}_t) \\ \Sigma_t = (\mathbf{I} - \mathbf{K}_t\mathbf{C}_t)\bar{\Sigma}_t \end{array} \right. ; \text{gdzie } \mathbf{K}_t = \bar{\Sigma}_t\mathbf{C}_t^T(\mathbf{C}_t\bar{\Sigma}_t\mathbf{C}_t^T + \mathbf{Q}_t)^{-1}, \quad (2.25)$$

zaś równanie (2.23) przekształca się do postaci:

$$\bar{bel}(\mathbf{x}_t) = \left\{ \begin{array}{l} \bar{\mu}_t = \mathbf{A}_t\mu_{t-1} + \mathbf{B}_t\mathbf{u}_t \\ \bar{\Sigma}_t = \mathbf{A}_t\Sigma_{t-1}\mathbf{A}_t^T + \mathbf{R}_t \end{array} \right. . \quad (2.26)$$

Zazwyczaj funkcje wiążące obserwację z wektorem stanu nie są liniowe. Wówczas zastępuje się je pierwszym wyrazem szeregu Taylora w punkcie μ_t (EKF) lub dwoma pierwszymi wyrazami (UKF). Wyniki są zadowalające, jednak nadal wymagane jest założenie dotyczące gaussowskiego rozkładu prawdopodobieństwa.

2.4.2 Filtr cząstkowy

Filtr cząstkowy pozwala na obliczenia przy dowolnej dystrybucji pozycji robota na podstawie danych z sensora. Jego najważniejszą zaletą jest dowolność charakterystyki sensora oraz możliwość stosowania nieliniowych funkcji transformujących obserwacje na położenie robota. W pierwszej iteracji algorytm tworzy dużą liczbę pozycji robota, według rozkładu równomiernego. W kolejnych iteracjach tworzy się dystrybucję prawdopodobieństwa, że robot znajduje się w punkcie x przy danych obserwacjach, na podstawie ostatniej obserwacji. Funkcja ta jest stosowana jako funkcja wagowa dla rozkładu otrzymanego w poprzedniej iteracji. W celu uniknięcia rozbieżności wartości, po każdym wyznaczeniu funkcji ważności na podstawie dopasowania pomiarów do mapy otoczenia, wyznacza się nowy rozkład położenia, według gęstości otrzymanej jako znormalizowany iloraz wagi danego położenia (czyli prawdopodobieństwa, że

robot znajduje się w danym miejscu na podstawie obserwacji) i obszaru należącego do danego położenia (czyli odwrotności gęstości położeń w danym miejscu). Po pewnym czasie, tzn. pewnej ilości iteracji, zależnym m.in. od charakterystyki sensora, otrzymuje się mały, zwarty obszar w którym znajdują się wszystkie prawdopodobne położenia robota. Dystrybucja docelowa opisana jest wzorem:

$$f : p(x|z_1, z_2, \dots, z_n) = \frac{\prod_k p(z_k|x)p(x)}{p(z_1, z_2, \dots, z_n)}, \quad (2.27)$$

gdzie $p(x)$ oznacza prawdopodobieństwo danego położenia, zaś $p(z_i)$ prawdopodobieństwo otrzymanej obserwacji znacznika i . Dystrybucja próbek (położeń) wyraża się wzorem:

$$g : p(x|z_l) = \frac{p(z_l|x)p(x)}{p(z_l)}, \quad (2.28)$$

zaś funkcja ważności - wzorem:

$$w : \frac{f}{g} = \frac{p(x|z_1, z_2, \dots, z_n)}{p(x|z_l)} = \frac{p(z_l) \prod_{k \neq l} p(z_k|x)}{p(z_1, z_2, \dots, z_n)} \quad (2.29)$$

2.5 Stereowizja

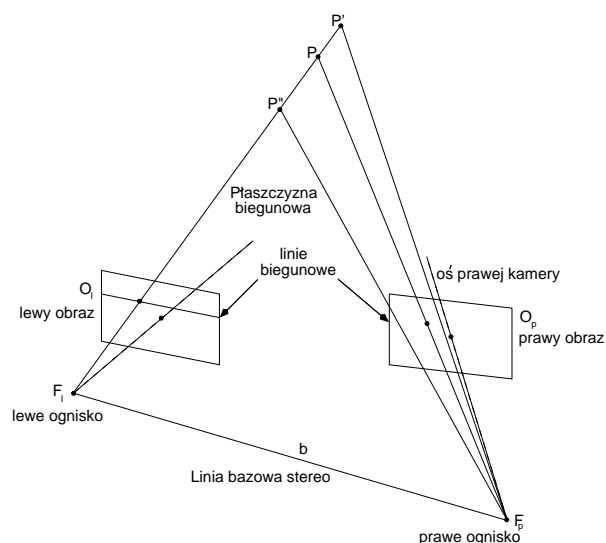
Układ dwóch lub więcej sprzężonych kamer pozwala na pomiar odległości. Jest to system przeznaczony do lokalizacji bezpośredniej. Poniżej omówiono układ dwóch sprzężonych kamer. Zamieszczony opis systemów stereowizyjnych został zaczerpnięty z [69].

Układ dwóch kamer przypomina widzenie dwuoczne, spotykane u większości zwierząt zamieszkujących Ziemię. Podobnie jak dwoje oczu, układ dwóch kamer pozwala na określenie odległości do widocznych obiektów. Podstawowym warunkiem pomiarów odległości obiektu jest jego widoczność przez obie kamery układu. Poza tym warunkiem wzajemne usytuowanie i orientacja kamer mogą być dowolne.

2.5.1 Elementy układu

Najważniejsze elementy układu kamer pokazano na rys 2.5. Są to: ognisko prawej i lewej kamery, linia bazowa, osie kamer oraz linie biegunowe i płaszczyzny biegunowe. **Ognisko kamery** jest miejscem, w którym zbiegają się promienie świetlne skupiane przez soczewki kamery. Choć w praktyce ogniska nie są idealne, tzn. nie wszystkie promienie zbiegają się w jednym punkcie, jednak do obliczeń należy przyjąć taki model. **Linia bazowa stereo** jest linią łączącą ogniska dwóch kamer układu. Środek tej linii wygodnie jest przyjąć za początek układu współrzędnych świata rzeczywistego.

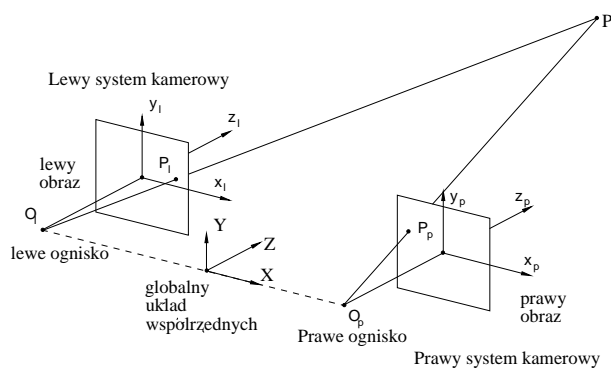
Linia biegunowa jest linią, na której mogą znajdować się punkty widoczne w określonym miejscu drugiego obrazu. Dokładniej, linie te – po jednej z każdego obrazu – tworzą pary. Punkty odpowiadające sobie w obrazach prawym i lewym zawsze leżą na odpowiadających sobie liniach biegunowych. W ten sposób do znalezienia punktów odpowiadających sobie, lub stwierdzenia że punkt jest tylko w jednym z obrazów, wystarczy przeszukać linie biegunowe. Powoduje to istotne zmniejszenie ilości obliczeń potrzebnych do dopasowania obrazów stereoskopowych. Na parze linii biegunowych znajdują się rzuty punktów leżących na wyznaczonej przez nie **płaszczyźnie biegunowej**.



Rysunek 2.5: Geometria w stereoskopowym systemie kamerowym [69]

Część wspólna otoczenia, widziana przez obie kamery, powinna być jak największa. Najczęściej stosowanym układem są dwie kamery o osiach skierowanych równolegle, i współpłaszczyznowych obrazach. Gdy kamery stereo są tak usytuowane i zorientowane, że różnią się tylko współrzędne poziome, wówczas na obrazach również obiekty mogą być przesunięte tylko poziomo. O takim układzie kamer stereo mówi się że **są w odpowiedniości**. Jeżeli obrazy stereo są w odpowiedniości, wówczas linie biegunowe pokrywają się z liniami poziomymi w obrazie, co pozwala na dodatkową optymalizację obliczeń.

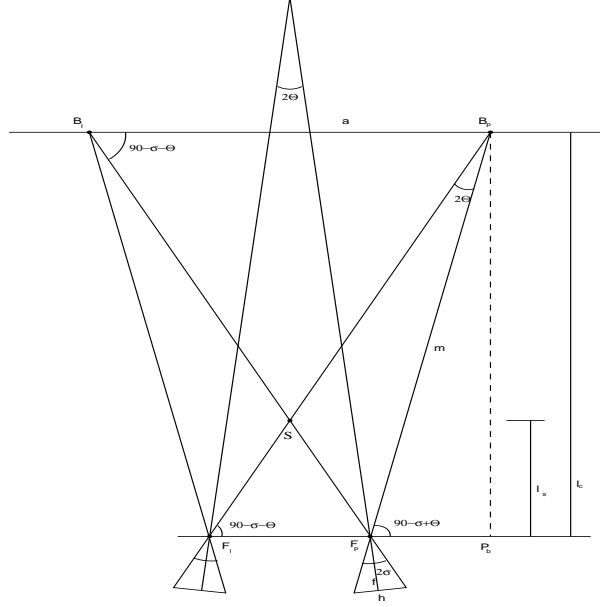
Jako początek nadrzędnego (lokalnego, tzn. związanego z układem kamer) układu współrzędnych przyjmuje się środek linii bazowej. Osie współrzędnych kamer i globalnego układu współrzędnych XYZ mają zgodne kierunki i zwroty. Współrzędne Y i Z układu i obu kamer są jednakowe, zaś współrzędne X kamer i globalnego układu współrzędnych różnią się o połowę długości linii bazowej. Ilustruje to rysunek 2.6.



Rysunek 2.6: Układy współrzędnych: nadrzędny oraz stereoskopowego systemu kamerowego [69]

System kamer zbieżnych

Układ ten pozwala na zwiększenie wspólnego pola widzenia w określonym zakresie odległości od kamer. Położenie kamer przedstawiono na rysunku 2.8.



Rysunek 2.8: Zbieżnoosiowy system stereowizyjny [69]

Gdy osie kamer w odpowiedniości zostaną obrócone o ten sam kąt wokół osi y przechodzących przez ogniska kamer, wówczas współrzędne punktu XYZ w układach współrzędnych kamer lewej i prawej wyrażają się wzorami:

$$X_l = \left(X + \frac{b}{2}\right) \cos \Theta - Z \sin \Theta, Y_l = Y, Z_l = \left(X + \frac{b}{2}\right) \sin \Theta + Z \cos \Theta, \quad (2.33)$$

$$X_p = \left(X - \frac{b}{2}\right) \cos \Theta + Z \sin \Theta, Y_p = Y, Z_l = -\left(X - \frac{b}{2}\right) \sin \Theta + Z \cos \Theta. \quad (2.34)$$

Wyznaczając wzory na współrzędne obrazowe otrzymuje się:

$$x_l = f \frac{\left(X + \frac{b}{2}\right) \cos \Theta - Z \sin \Theta}{\left(X + \frac{b}{2}\right) \sin \Theta + Z \cos \Theta}, \quad (2.35)$$

$$x_p = f \frac{\left(X - \frac{b}{2}\right) \cos \Theta + Z \sin \Theta}{-\left(X - \frac{b}{2}\right) \sin \Theta + Z \cos \Theta}. \quad (2.36)$$

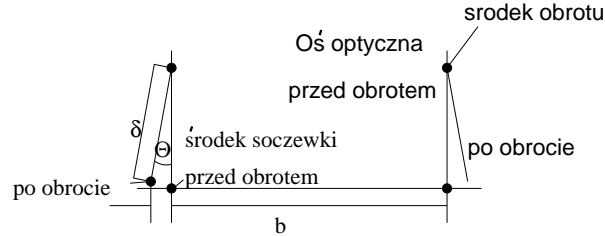
Z powyższych wzorów otrzymuje się następujące wzory na współrzędne XYZ na podstawie współrzędnych ekranowych:

$$X = \frac{bf(x_l + x_p)}{2(f^2 + x_l x_p) \sin 2\Theta}, \quad (2.37)$$

$$Y = \frac{by_l}{d}, \quad (2.38)$$

$$Z = \frac{b(f \cos \Theta - x_l \sin \Theta)(f \cos \Theta + x_p \sin \Theta)}{fd \cos 2\Theta + (f^2 + x_l x_p) \sin 2\Theta}. \quad (2.39)$$

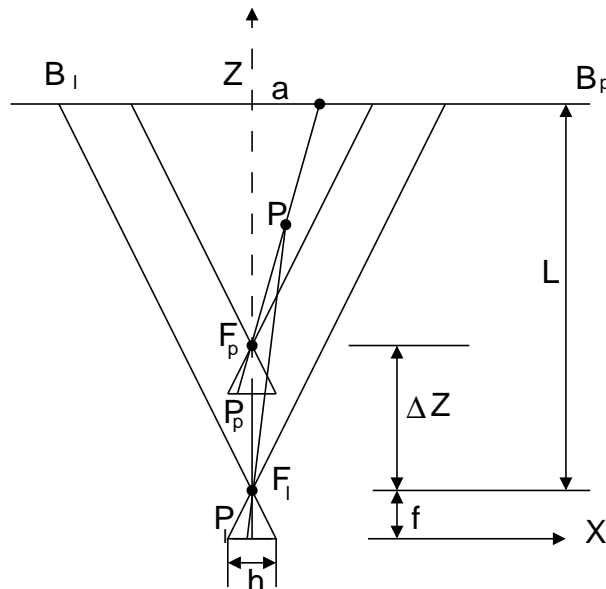
Zmiana kąta Θ może powodować zmianę odległości bazowej b , jak pokazano na rysunku 2.9. Jeżeli kamery obracają się wokół innej osi pionowej niż przechodząca przez ognisko soczewki, odległej od ogniska o Δl , wówczas odległość bazowa zmienia się o $\pm 2\Delta l \sin \Theta$.



Rysunek 2.9: Zmiana długości linii bazowej b spowodowana zmianą zbieżności kamer [69]

2.5.3 Model osiowo-ruchowy

Wadą modelu boczno-kamerowego jest niemożliwość określenia, czy dopasowanie punktów jest poprawne. Przesunięcie kamer względem siebie powoduje, że nie wszystkie punkty, nawet znajdujące się we wspólnym polu widzenia, są widoczne w obu kamerach. Nie ma też zadowalająco dobrych zabezpieczeń przed tego typu błędnym dopasowaniem. W modelu osiowo-ruchowym przyjmuje się, że para obrazów stereoskopowych pochodzi z jednej kamery, przesuniętej jedynie wzdłuż osi optycznej. Powoduje to, że obiekty widoczne w obrazie "bliższym" prawie na pewno są obecne w obrazie dalszym, wyjątek mogą stanowić obszary częściowo zasłonięte sąsiadujące w obrazie z obiektami położonymi obok bliższych obiektów. Linie biegunowe rozchodzą się promieniście od rzutu osi optycznej na obraz i są do siebie równoległe w obu obrazach. Model osiowo-ruchowy pokazano na rysunku 2.10.



Rysunek 2.10: Geometria systemu osiowo-ruchowego [69]

Z podobieństwa trójkątów otrzymuje się zależności:

$$\frac{X}{X_p} = \frac{Y}{Y_p} = \frac{f + \Delta Z - Z}{f}, \quad (2.40)$$

$$\frac{X}{X_t} = \frac{Y}{Y_t} = \frac{f - Z}{f}. \quad (2.41)$$

Z wzorów (2.40) i (2.41) wynika wzór:

$$\frac{Y_t}{X_t} = \frac{Y_p}{X_p}. \quad (2.42)$$

Z wzorów tych uzyskuje się wzory na współrzędne XYZ:

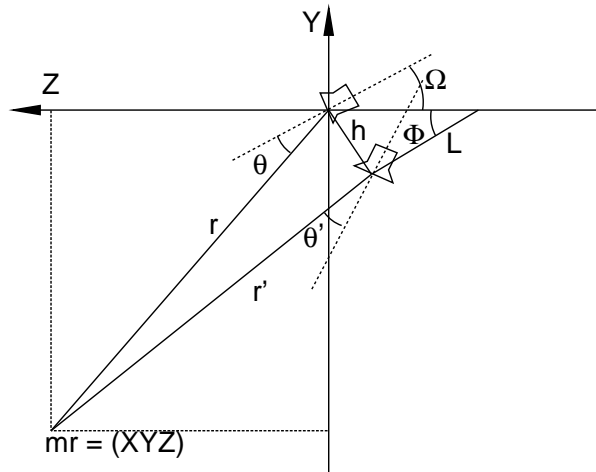
$$X_t = \frac{Z - f - \Delta Z}{Z - f} X_p, \quad (2.43)$$

$$Y_t = \frac{Z - f - \Delta Z}{Z - f} Y_p, \quad (2.44)$$

$$Z = f + \frac{X \Delta Z}{d}. \quad (2.45)$$

2.5.4 Przykład systemu opartego na jednej kamerze

Interesujący przykład systemu stereowizji opartego na jednej kamerze opisuje praca [39]. Kamera jest ruchoma, można ją obracać o 360° wokół osi pionowej oraz przechylać zmieniając kąt patrzenia na podłoże. Dzięki temu można uzyskać obraz stereowizyjny, jak z systemu o zbieżnym układzie osi, jednak nie jest to tak prosty system jak opisane modele. Układ sterowania kamerą pozwala na precyzyjne ustawienie kamery. Położenia i kierunki osi kamery podczas wykonywania zdjęć są stałe, a kąt między osiami wynosi $7,7^\circ$. Schemat działania układu przedstawiono na rysunku 2.11. Kamera pracuje w niskiej rozdzielczości się 160×120 , uznawanej za optymalną ze względu na jakość danych i prędkość obliczeń. System jest w stanie śledzić 200-500 obiektów jednocześnie.



Rysunek 2.11: Przykład pojedynczej kamery pracującej w systemie stereowizyjnym [39]

W opisywanym systemie stałe wartości wynosiły odpowiednio: $\Phi = 7,7^\circ$, $L = 21\text{cm}$, z czego wynikała długość linii bazowej $h \approx 3\text{cm}$. Taki rozmiar linii bazowej powoduje przeciętne przesunięcie obiektu w parze obrazów o około 10 pikseli. W celu wyznaczenia zgodności wykonuje się skanowanie wzdłuż linii epipolarnych o grubości 3 pikseli w oknie o rozmiarach 5

na 5 piksli. Obiekty uznaje się za dopasowane, jeżeli różnica między najlepszym dopasowaniem i oryginalnym obiektem jest mniejsza niż określony próg. Odległość r do obiektu wyznacza się z równania:

$$\vec{m}r - \vec{h} = R\vec{m}'r', \quad (2.46)$$

gdzie $\vec{m} = [i, j, F]$ i $\vec{m}' = [i', j', F]$ określają wektory jednostkowe wyznaczone przez położenie obiektów podczas wykonywania pierwszego i drugiego obrazu. F jest ogniskową kamery, R – macierzą obrotu, zaś h – wektorem przesunięcia kamery. Wartości te wyznacza się jednokrotnie.

Ponieważ kamera porusza się tylko w płaszczyźnie pionowej, równanie (2.46) można zmodyfikować używając metody następujących równań:

$$r' \sin(\Phi + \Omega - \theta') = Z \tan(\Omega - \theta) - h \cos \frac{\Phi}{2}, \quad (2.47)$$

$$r' \cos(\Phi + \Omega - \theta') = Z + h \sin \frac{\theta}{2} \quad (2.48)$$

Ostateczne równania współrzędnych mają postać przedstawioną na równaniach (2.49), (2.50), (2.51):

$$X = Z \tan \Theta, \quad (2.49)$$

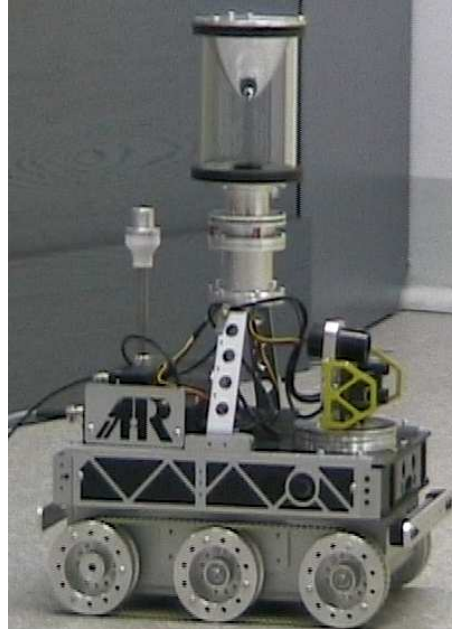
$$Y = Z \tan \frac{i}{F}, \quad (2.50)$$

$$Z = \frac{h \cos \frac{\Phi}{2} + h \sin \frac{\Phi}{2} \tan(\Phi + \Omega - \Theta')}{\tan(\Omega - \Theta) + \tan(\Phi + \Omega - \Theta')} \quad (2.51)$$

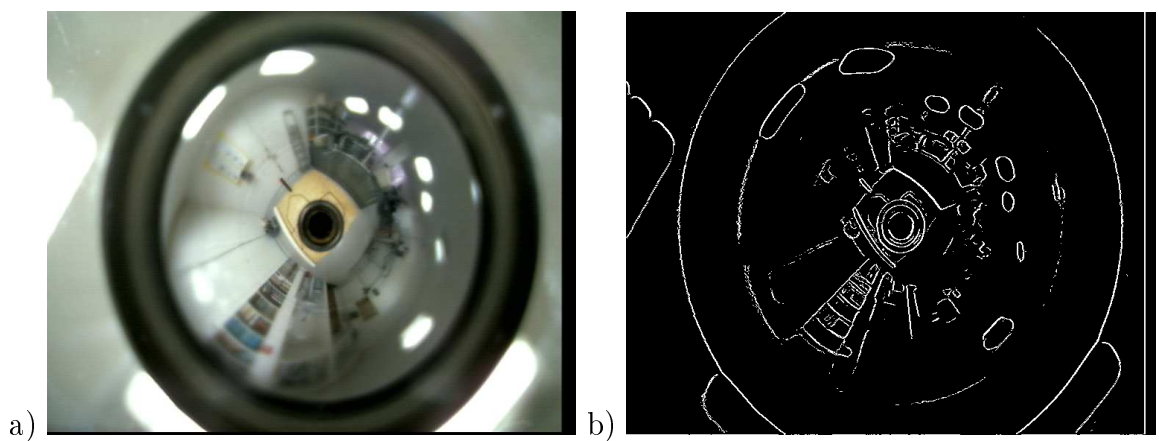
2.6 Kamera dookólna

Kamera dookólna (*ang. omniscam*) jest sensorem względnie nowym. Jego najważniejszą cechą jest uzyskiwanie obrazu panoramicznego otoczenia. Sensor składa się z kamery oraz lustra stożkowego, parabolicznego lub sferycznego umożliwiającego uzyskanie obrazu dookólnego. Przykładowego robota z zainstalowaną kamerą dookólną przedstawia rysunek 2.12.

Rozdzielczość liniowa sensora zmienia się w zależności od odległości, natomiast rozdzielczość kątowna jest stała. Sensor ten nadaje się w szczególności do nawigacji kierunkowej, tzn. określania orientacji robota w układzie biegunowym. Typowy obraz dookólny przedstawia rysunek 2.13.a, zaś rysunek 2.13.b – ten sam obraz po ekstrakcji krawędzi. Linie pionowe są widoczne jako linie radialne, co umożliwia ich łatwą detekcję, identyfikację oraz śledzenie [5]. Linie poziome są widoczne jako fragmenty elips, co powoduje trudności z ich identyfikacją. Sensor charakteryzuje się zmienną rozdzielczością liniową zależną od współrzędnych [81]. Uzyskanie informacji geometrycznej innej niż linie pionowe (radialne) jest dość czasochłonne, co jest związane z rzutowaniem obrazu na powierzchnię walca. Można rozpoznawać obszary wg barwy, ale z uwzględnieniem zmiennej rozdzielczości liniowej. Kamera dookólna nadaje się do śledzenia obiektów i ludzi w pomieszczeniach [75] [50], utrzymania robotów w szyku i wspólnej nawigacji w bliskim otoczeniu [23] [1] [31] oraz do łączenia z innymi metodami lokalizacji. Nawigacja wyłącznie w oparciu o kamerę dookólną jest możliwa w pomieszczeniu o niewielkich rozmiarach ze względu na zmienną rozdzielczość liniową sensora.



Rysunek 2.12: Robot NAVIGATOR - prototyp z zainstalowaną kamerą dookólną

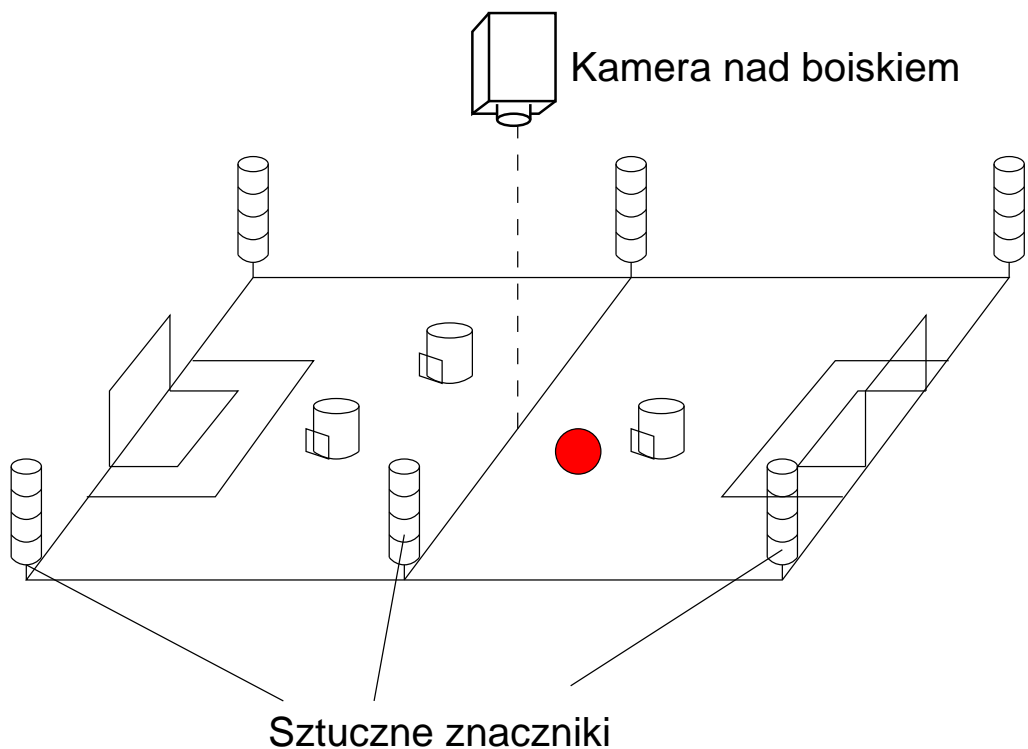


Rysunek 2.13: Przykładowy obraz pomieszczenia uzyskany z kamery dookólnej: a) obraz; b) ten sam obraz po ekstrakcji krawędzi

2.7 Kamera zainstalowana w otoczeniu

Lokalizacja za pomocą kamery zamocowanej na stałe jest coraz częściej brana pod uwagę metodą, wraz z malejącą ceną kamer. Od kilku lat rośnie liczba kamer instalowanych głównie w celach monitorowania. W literaturze spotyka się metody lokalizacji o różnym stopniu skomplikowania. W większości przypadków robot jest oznaczony sztucznym znacznikiem, który znajduje się na górnej powierzchni robota. System taki jest przedmiotem badań w kilku miejscach m.in. na Wydziale Mechatroniki Politechniki Warszawskiej i na Wydziale Automatyki i Robotyki Politechniki Poznańskiej. W publikacjach [74] i [12] opisano wieloagentowy system lokalizacji robota mobilnego, którego elementem lokalizującym jest podwieszona kamera. Rolę znaczników pełnią diody sygnalizacyjne umieszczone na pokrywie robota, umieszczone w narożnikach czworokąta. Dokładność lokalizacji zależy przede wszystkim od rozmiarów znacznika.

Kamera umocowana w otoczeniu stanowi bardzo często jeden podstawowych elementów nawigacyjnych w systemach rozgrywania piłki nożnej robotów, zwanych RoboCup. Przykłady takich systemów opisano m.in. w [8], [89]. Często też stosuje się sztuczne znaczniki przy liniach i w narożnikach boiska. Schemat typowego systemu RoboCup przedstawia rysunek 2.14. W systemach tych stosuje się specjalne barwy dla łatwego rozpoznania obiektów: teren boiska jest zielony z białymi liniami (jak w prawdziwej piłce nożnej), piłka jest czerwona, zaś roboty należące do przeciwnych drużyn mają różne barwy, np. żółte i niebieskie. Oprócz zainstalowanej nad boiskiem kamery roboty są wyposażone w różne sensory, w tym coraz częściej w kamerę zwykłą lub dookólną.



Rysunek 2.14: Schemat systemu lokalizacji w systemie RoboCup na podstawie kamery i sztucznych znaczników w określonych miejscach

Rozdział 3

Reprezentacja i przetwarzanie obrazu

3.1 Reprezentacja obrazu

3.1.1 Reprezentacja obrazu szarego

Obraz szary przechowuje jedynie informację o jasności piksli. W zależności od rozdzielczości jasności wyróżnia się obrazy czarno-białe, oraz obrazy o N poziomach szarości. Najczęściej spotykane ilości odcieni to 16, 64 i 256. Większej ilości odcieni nie stosuje się, ponieważ oko ludzkie jest w stanie wyróżnić ok. 64 poziomy. Pamięć obrazu może być różnie zorganizowana. Jeżeli ilość bitów niezbędna do przedstawienia piksła jest mniejsza niż 8 (ilość odcieni szarości nie przekracza 128), można spotkać się ze stosowaniem kompresji danych obrazu polegającej na umieszczaniu bitów różnych piksli w jednym bajcie. Kompresję tego typu stosuje się powszechnie przy zapisie obrazów 2-barwnych, np. czarno-białych, oraz 4- i 16-barwnych.

W obrazie w 256 odcieniach szarości jeden bajt przechowuje jasność jednego piksła.

3.1.2 Reprezentacja barw

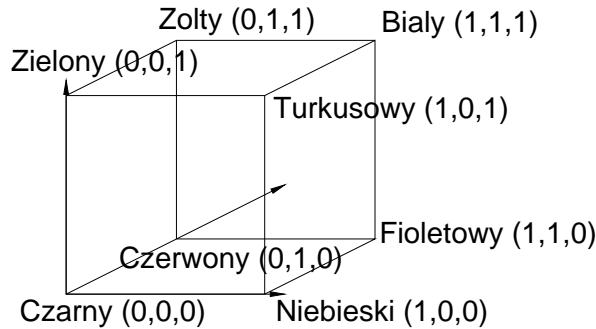
Obraz barwny przechowuje informację o jasności oraz barwie piksli. W większości formatów dane te są ze sobą skorelowane, tzn. informacja o jasności i barwie jest ze sobą powiązana. Zakodowanie barwy wymaga co najmniej dwóch danych, jako że barwa jest zmienną dwuwymiarową. Przechowanie pełnej informacji o jasności i barwie wymaga trzech liczb, i dlatego wszystkie formaty kodowania barw są przynajmniej trójwymiarowe.

Formaty różnią się od siebie ilością barw możliwych do przedstawienia. Najczęściej spotykane formaty to: RGB, CMYK, HSV, CIE, YUV (YCbCr). W modelach RGB i CMYK jasność jest kodowana przynajmniej częściowo ze składowymi barwy. W pozostałych jasność jest teoretycznie nie powiązana z barwą, choć w modelu HSV zakres możliwych barw zależy od jasności.

3.1.3 Format RGB

W formacie RGB barwę reprezentuje suma trzech składowych barw: czerwonej (**R**ed), zielonej (**G**reen) i niebieskiej (**B**lue). Wszystkie trzy składowe mają identyczne zakresy wartości, w standardzie komputerów PC przyjmuje się że minimalną wartością jest 0, a maksymalną – największa liczba jaką można zapisać na ilości bitów przeznaczonych do tego. Jasność jest średnią wartości składowych, barwa czarna składa się z trzech składowych o wartości 0, zaś

biała z trzech składowych o wartości maksymalnej. Przestrzeń barw opisuje sześcian, którego krawędziami są barwy dwuskładnikowe. Schemat barw RGB przedstawia rysunek 3.1. Wewnątrz sześciany znajdują się barwy pośrednie. Linia łącząca wierzchołki biały i czarny reprezentuje odcienie szarości.



Rysunek 3.1: Schemat modelu barw RGB

3.1.4 Format HSV

W grupie formatów HSV barwę opisuje się przez odcień (**H**ue), nasycenie (**S**aturation) oraz jasność (**V**alue). Pomimo pozornego rozdzielenia zależności składowych, w rzeczywistości zakresy wartości nasycenia są zależne od odcienia i jasności.

Skala odcienia jest cykliczna, najczęściej przedstawia się ją jako zakres kątów $0^\circ - 360^\circ$. Wartości 0° odpowiada "czysty" kolor czerwony, 120° – "czysty" zielony, zaś 240° – "czysty" niebieski. W pełni nasycone barwy są osiągalne tylko dla jasności pośredniej pomiędzy białym i czarnym. Wartości nasycenia zmieniają się w zakresie 0–100, zaś jasności 0–255. Możliwa przestrzeń barw tworzy podwójny stożek, lub podwójny ostrosłup o podstawie sześciokąta. Zakres barw przestrzeni HSV/HSL przedstawia rysunek 3.2. Ilość możliwych reprezentowanych barw jest zależna od jasności, maksimum jest dla jasności pośredniej. Wraz z odległością od osi symetrii bryły rośnie nasycenie barw.

3.1.5 Konwersje między modelami barw

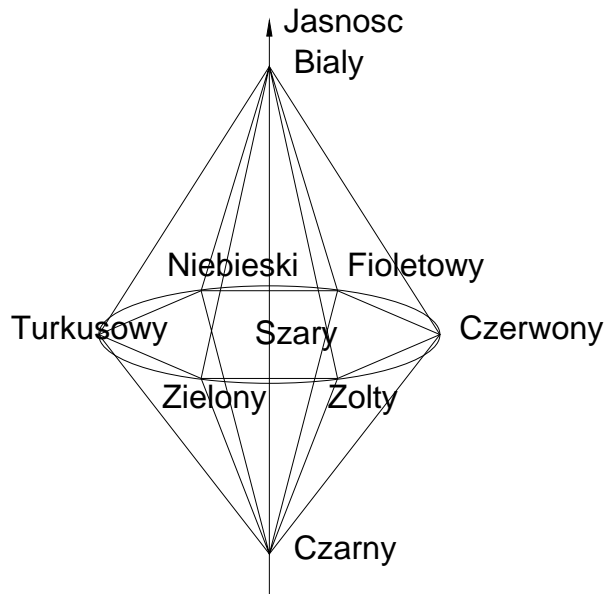
Najczęściej spotykanym formatem jest RGB lub jego dopełnienie CMYK. W praktyce informacja zapisana w tym formacie jest trudna do interpretacji. Jedną z metod uzyskania lepszej reprezentacji barw jest wyodrębnienie składowej luminancji, wyznaczonej jako średnia trzech barw oraz dwóch wartości opisujących barwę. Układ ten nazywa się układem r'g'. Wartości opisujące barwę wyznacza się na podstawie składowych RGB wg następujących wzorów:

$$L = \frac{R + G + B}{3}; \quad r' = \frac{r}{R + G + B}; \quad g' = \frac{g}{R + G + B}.$$

Poszczególne modele obejmują różne zakresy barw, choć dla każdego z nich istnieje zakres barw wspólnych, tzn. możliwych do przedstawienia w obu modelach.

Przekształcenia pomiędzy modelami RGB i HSV nie da się opisać prostymi równaniami. Można łatwo opisać jasność i nasycenie:

$$V = \max(R, G, B), \quad S = \frac{\max(R, G, B) - \min(R, G, B)}{\max(R, G, B)} * 100. \quad (3.1)$$



Rysunek 3.2: Schemat modelu barw HSV/HSL

O odcieniu decydują wartości wszystkich trzech składowych barw. Wartość maksymalna wyznacza **bazę**. Następnie wyznacza się przesunięcie odcienia :

$$OH = \frac{\text{med}(R, G, B) - \min(R, G, B) * 60^\circ}{\max(R, G, B) - \min(R, G, B)}. \quad (3.2)$$

Ma ono wartości z zakresu $-60 - 60$. Odcień jest sumą wartości bazy i przesunięcia, przy czym zbiór wartości jest cykliczny modulo 360. Wartości bazy i znaku przesunięcia dla różnych relacji między składowymi barwami przedstawia tabela 3.1.5.

Składowa maksymalna	Czerwony		Zielony		Niebieski	
Wartość bazy	0		120		240	
Składowa minimalna	Zielony	Niebieski	Czerwony	Niebieski	Czerwony	Zielony
Znak przesunięcia	-	+	+	-	-	+

Tablica 3.1: Wartości bazy i przesunięcia odcienia dla różnych maksymalnych składowych barw

Konwersja HSV->RGB jest równie złożona. Jeżeli nasycenie ma wartość 0, barwa jest szara i należy wyznaczyć tylko jasność : $R = G = B = V$. Jeżeli nasycenie nie jest zerowe, wówczas należy wykonać następujące kroki:

$$\min = V - \frac{S * V}{255}. \quad (3.3)$$

W zależności od parametru Hb wyznaczanego wg wzoru $Hb = H/60$ określa się dwie z trzech składowych modelu RGB oraz parametr $hoff$ dla obliczenia pośredniej pod względem jasności składowej. Zależności pokazuje tabela 3.1.5. Następnie trzecią składową wyznacza się wg wzoru (3.4):

$$\text{med} = \min + \frac{\text{hoff} - (V - \min)}{60}, \quad (3.4)$$

Hb	0	1	2	3	4	5
V->	Czerwony	Zielony	Zielony	Niebieski	Niebieski	Czerwony
min->	Niebieski	Niebieski	Czerwony	Czerwony	Zielony	Zielony
hoff=	H	120 - H	H - 120	240 - H	H - 240	360 - H

Tablica 3.2: Zależności składowych RGB od odcienia H w modelu HSV

gdzie *min* oznacza składową RGB o najmniejszej jasności.

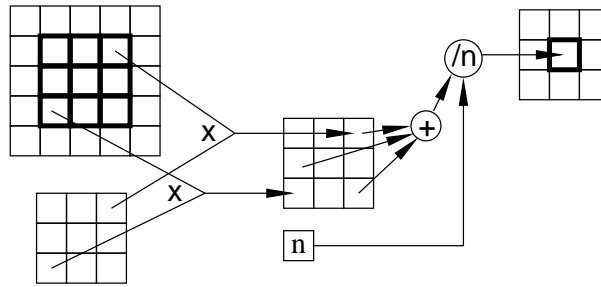
Podczas konwersji z modelu RGB na HSV powstają fałszywe barwy nasycone dla prawie czarnych piksli. Przykładowo, barwa praktycznie czarna, o zerowych składowych R i G oraz niebieskiej składowej równej 1 (w skali 0-255) zostanie uznana za barwę niebieską w pełni nasyconą. Dla uniknięcia tego problemu w pracy [7] zaproponowano minimalne wartości progowe jasności, poniżej których nasycenie uznaje się za nieokreślone. Wartości te są zastosowane w implementacji wyszukiwania obszarów o określonej barwie (zob 7.1). Wartości tych progów przedstawia tabela 3.3.

Zakres jasności (0-255)	Nasycenie (0%-100%)
0 - 10	Brak
11 - 15	$\geq 55\%$
16 - 20	$\geq 40\%$
21 - 30	$\geq 30\%$
31 - 40	$\geq 22\%$
41 - 50	$\geq 17\%$
51 - 100	$\geq 12\%$
101 - 180	$\geq 7\%$
181 - 255	$\geq 3\%$

Tablica 3.3: Minimalne wartości jasności dla różnych przedziałów nasycenia barw

3.2 Filtry liniowe

Filtry liniowe są często stosowanym narzędziem obróbki wstępnej. Pozwalają uzyskać różne efekty, np. rozmycie, wyostrenie, efekt płaskorzeźby lub nawet detekcję krawędzi. Elementem charakteryzującym filtr jest **maska filtru**, z którą splota się obraz źródłowy. Ogólnie działanie filtru liniowego polega na uzyskaniu splotu obrazu wejściowego z maską filtru, przy zachowaniu średniej jasności obrazu. Na rysunku 3.2 przedstawiono schemat działania filtru liniowego. Maska składa się z współczynników, przez które mnoży się wartości sąsiednich piksli, z zachowaniem względnego położenia piksli obrazu i maski. Następnie uzyskane wartości dodaje się. Ponieważ suma współczynników nie zawsze wynosi 1, należy podzielić uzyskaną wartość przez sumę współczynników (o ile jest różna od 0). W ten sposób zachowuje się granice jasności piksli. Wielkość maski może być dowolna, przedstawiona na schemacie ma rozmiar 3 piksli w poziomie i 3 w pionie. Otrzymany wynik jest jasnością piksela wynikowego. W celu uzyskania całego obrazu wynikowego należy wykonać opisaną operację dla wszystkich piksli obrazu źródłowego.



Rysunek 3.3: Schemat filtru liniowego o rozmiarach maski 3x3

W niniejszej pracy stosowano filtr liniowy Gaussa dla w obróbce wstępnej przed detekcją krawędzi (zob rozdz. 7.3.1). Należy on do filtrów dolnoprzepustowych. W przypadku obrazu wysoka częstotliwość oznacza drobne elementy, ostre krawędzie oraz szum drobnoziarnisty. Filtracja dolnoprzepustowa powoduje częściowe lub całkowite usunięcie tych elementów i efekt "rozmycia" obrazu.

Filtracja dolnoprzepustowa polega na uśrednieniu wartości sąsiednich piksli. Wszystkie współczynniki maski filtru są nie ujemne. Z tego powodu dzielnik maski jest zawsze większy od 1. Przykładowe maski przedstawia rysunek 3.4. Filtr Gaussa uzyskuje się z rozkładu Gaussa:

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left(\frac{-x^2y^2}{2\sigma_x^2\sigma_y^2}\right), \quad (3.5)$$

skalując wartości tak, aby najniższe współczynniki wynosiły 1. Przykład zastosowania filtrów dolnoprzepustowych przedstawia rysunek 3.4. Stopień rozmycia jest tym większy, im większy jest wpływ wag sąsiadów piksela na jego wartość ostateczną, czyli większe wartości wzmocnienia centralnego piksela powodują słabsze rozmycie.



0	1	0	1	1	1	1	2	1
1	4	1	1	1	1	2	4	2
0	1	0	1	1	1	1	2	1

8

9

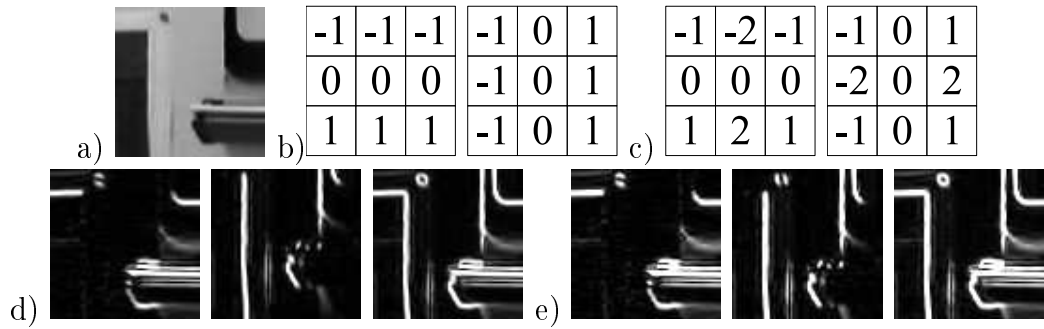
16

a)

b)

c)

Rysunek 3.4: Przykłady liniowych filtrów dolnoprzepustowych: a) słaba filtracja b) bardzo silna filtracja c) filtr Gaussa - umiarkowana filtracja



Rysunek 3.5: Detektory krawędzi: **a)** Obraz bazowy; Detektor Prewitta: **b)** maski pozioma i pionowa, **d)** efekt działania filtrów składowych i końcowy; Detektor Sobela: **c)** maski pozioma i pionowa, **e)** efekt działania filtrów składowych i końcowy

3.3 Detekcja krawędzi

Detekcja krawędzi jest bardziej zaawansowaną niż filtry techniką obróbki obrazu. Wszystkie znane techniki detekcji wymagają zastosowania filtrów liniowych, opisanych w rozdziale 3.2. Detektory krawędzi można podzielić na: jednoprzebiegowe (np. Laplace'a), dwufazowe (detektory Sobela i Prewitta) oraz bardziej skomplikowane (np. detektory Marra-Hildretha i Cumaniego), które umożliwiają dokładniejszą detekcję. Poniżej przedstawiono detektory dwufazowe i detektor Marra-Hildretha.

3.3.1 Detektory Sobela i Prewitta

Detektory Sobela i Prewitta są dwuprzebiegowe. W pierwszym przebiegu wyznacza się oddzielnie linie przebiegające w dwóch prostopadłych kierunkach, np. poziomym i pionowym. Maski detektorów mają są kwadratami o nieparzystym wymiarze. W drugim przebiegu wyznacza się sumaryczne krawędzie na podstawie krawędzi kierunkowych. Ponieważ jasność wykrytych krawędzi jest zależna od ich kierunku, ostateczną sumaryczną jasność otrzymuje się według równania: $I_{sum} = \sqrt{I_x^2 + I_y^2}$. Otrzymaną wielkość można znormalizować, np. dzieląc przez wielkość maski.

Maski detektora Prewitta spotyka się tylko w rozmiarze 3x3, przedstawiono je na rysunku 3.3.1.b. Filtr Prewitta promuje krawędzie o kierunkach wykrywanych przez detektory, np. przedstawione maski wykrywają przede wszystkim krawędzie pionowe i poziome, zaś ukośne w mniejszym stopniu.

W detektorze Sobela w pierwszym przebiegu można stosować maski o różnej wielkości. Maski o rozmiarze 3x3 przedstawiono na rysunku 3.3.1.c. Detektor Sobela waży wykrywane krawędzie w zależności od ich kierunku. Choć różnica w porównaniu z detektorem Prewitta wydaje się niewielka, krawędzie ukośne oraz słabsze są wykrywane lepiej przez detektor Sobela. Przykłady działania obu detektorów przedstawia rysunek 3.3.1.c i 3.3.1.e.

3.3.2 Operator Marra-Hildretha

Opisaną poniżej metodę stosuję do wykrywania krawędzi znacznika w celu ustalenia ich przebiegu oraz położenia narożników. Metoda jest znana jako operator Marra-Hildretha (*ang. Marr-Hildreth operator*) i szczegółowy opis można znaleźć w [66]. Danymi wejściowymi są:

obraz wyjściowy z detektora Sobela oraz oryginalny obraz. Z obrazu wejściowego sporządza się tablice:

- modułu i kierunku cyfrowej pierwszej pochodnej, wyznaczanych na podstawie cyfrowych pierwszych pochodnych cząstkowych w kierunkach poziomym i pionowym;
- wszystkich cyfrowych drugich pochodnych cząstkowych: w kierunkach poziomym $\frac{\partial^2 f}{\partial x^2}$ i pionowym $\frac{\partial^2 f}{\partial y^2}$ oraz mieszanej $\frac{\partial^2 f}{\partial x \partial y}$.

Pierwszą i drugą pochodną cyfrową w punkcie x, y aproksymuje się jako:

$$\frac{\partial f(x, y)}{\partial x} \approx 0,5f(x + 1, y) - 0,5f(x - 1, y), \quad (3.6)$$

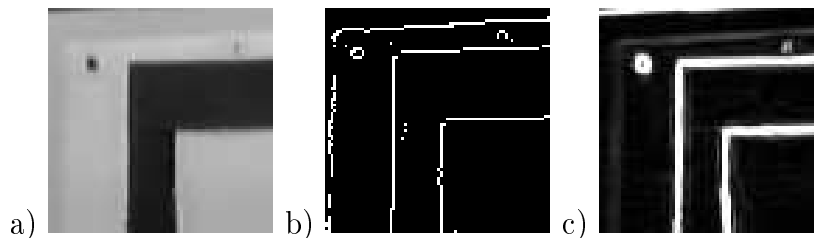
$$\frac{\partial f(x, y)}{\partial y} \approx 0,5f(x, y + 1) - 0,5f(x, y - 1), \quad (3.7)$$

$$\frac{\partial^2 f(x, y)}{\partial x^2} \approx 0,25f(x + 2, y) - 0,5f(x, y) + 0,25f(x, y - 2), \quad (3.8)$$

$$\frac{\partial^2 f(x, y)}{\partial y^2} \approx 0,25f(x, y + 2) - 0,5f(x, y) + 0,25f(x - 2, y), \quad (3.9)$$

$$\frac{\partial^2 f(x, y)}{\partial x \partial y} \approx 0,25f(x + 1, y + 1) - 0,25f(x + 1, y - 1) + 0,25f(x - 1, y - 1) - 0,25f(x - 1, y + 1). \quad (3.10)$$

Szczegółową analizę przeprowadza się dla tych punktów krawędzi, których jasność przekracza określony próg. W pierwszym kroku wyznacza się wartości pierwszych pochodnych, oraz na ich podstawie moduł i kierunek pierwszej pochodnej. Kierunek ogranicza się do wyznaczenia kierunku w siatce 3×3 , czyli jednej z 8 części kąta pełnego. Kierunek pierwszej pochodnej jest prostopadły do maksymalnego gradientu jasności w danym punkcie. Dany piksel należy do wykrytej linii, jeżeli jego wartość sumy drugich pochodnych dla tego piksela w porównaniu z sąsiadami leżącymi w kierunku prostopadłym do kierunku pierwszej pochodnej ma wartość najbliższą zeru, a ponadto zachodzi zmiana znaku sumy drugich pochodnych, tzn. suma drugich pochodnych sąsiada z jednej strony piksela ma inny znak niż dla drugiego sąsiada. Porównanie wyników detekcji krawędzi algorytmem Sobela i Marra-Hildretha przedstawia rysunek 3.6.



Rysunek 3.6: Detektor krawędzi Marra-Hildretha: a) obraz wyjściowy; b) efekt działania detektora Marra-Hildretha; c) efekt działania detektora Sobela

3.4 Transformacje geometryczne

3.4.1 Transformacja afiniczna

W wielu systemach wizyjnych, pozyskiwany obraz zawiera zniekształcenia geometryczne wprowadzone przez nieregularności perspektywy, gdzie pozycja kamery względem oglądanych obiektów ma wpływ na ich rozmiary w obrazie. Stosując transformację afiniczną do jednorodnie zniekształconego obrazu można skorygować wiele zniekształceń perspektywicznych przez transformację pomiarów z układu idealnego do rzeczywistego. Stosuje się to np. w nawigacji satelitarnej, gdzie wymagane są mapy geometrycznie dokładne.

Transformacja afiniczna jest ważną klasą liniowych transformacji geometrycznych obejmujących kombinacje operatorów translacji, obrotu, odbicia i skalowania. Ogólną własnością tej transformacji jest zachowanie linii prostych jako prostych i równoległych jako równoległych. Jest ona zdefiniowana przez podanie trzech punktów, na których można zbudować równoległobok, czyli odpowiada mapowaniu dowolnego równoległoboku (zazwyczaj jest to kwadrat) w inny dowolny równoległobok. Rysunek 3.7 pokazuje przykłady transformacji afinicznej – zachowującej proporcje wzdłuż osi. Na ogół transformacja ta zmienia rozmiar obrazu, część informacji może być utracona a część obrazu wynikowego może nie posiadać danych źródłowych (jest to "pusty" czyli zazwyczaj czarny obszar).



Rysunek 3.7: Przykłady transformacji afinicznej: a) obraz przed transformacją; b) i c) – transformacja obrazu

Ogólne równanie transformacji afinicznej przedstawia równanie (3.11):

$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = A \times \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + B. \quad (3.11)$$

Definiując tylko macierz B uzyskuje się translację:

$$A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}; B = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}. \quad (3.12)$$

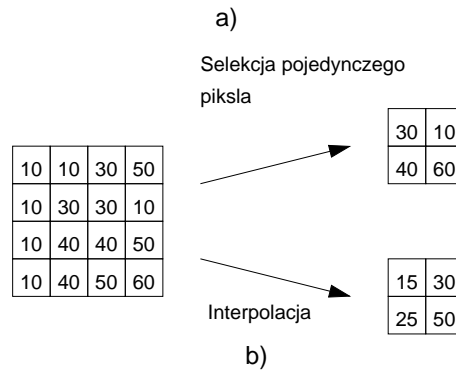
Obrót definiuje się jako przekształcenie afiniczne:

$$A = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}; B = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (3.13)$$

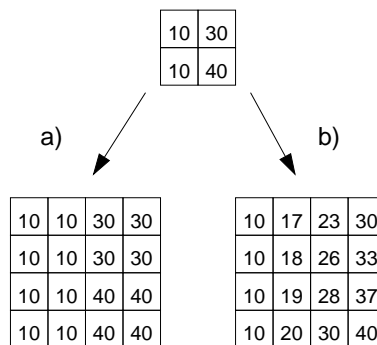
Skalowanie

Skalowanie jest specyficznym przypadkiem transformacji afinicznej. Metoda zmniejszania obrazu, zwana *subsamplingiem*, polega na nadaniu pikselowi wartości jednego z wielu pikseli wejściowych, lub interpolacji wartości piksla na podstawie wartości pikseli wejściowych. Powiększanie uzyskuje się przez replikację lub interpolację pikseli. Skalowania używa się do zmiany wizualizacji obrazu lub jego części, w celu zmiany ilości informacji przechowywanej w obrazie lub jako element przetwarzania niskopoziomowego dla operatora wymagającego obrazu o określonych rozmiarach.

Dwa algorytmy *subsamplingu* przedstawia rysunek 3.8. Wybór pojedynczego piksla jest metodą szybką, jednak wybrany piksel nie musi być reprezentatywny dla grupy, którą zastępuje. Interpolacja jest wolniejsza, a wartość piksla wynikowego jest średnią statystyczną sąsiadów zastępowanych tym pikslem. Obraz można powiększyć stosując albo replikację pikseli, albo interpolację. Rysunek 3.9 przedstawia obie metody.



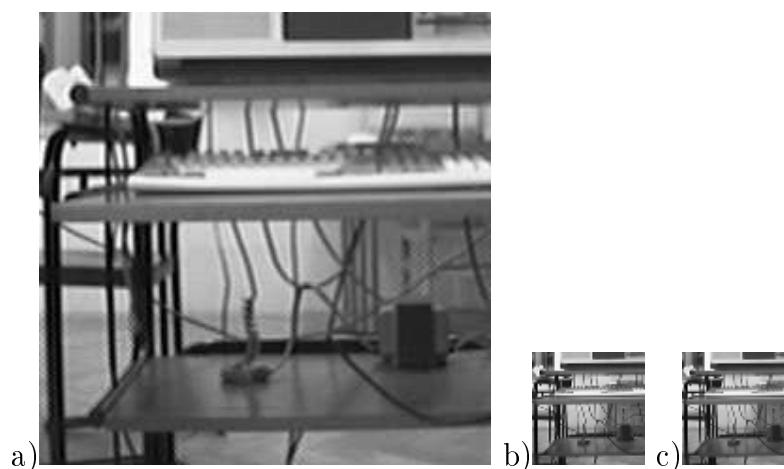
Rysunek 3.8: Schemat działania algorytmów *subsamplingu*: a) wybór pojedynczego piksla; b) interpolacja



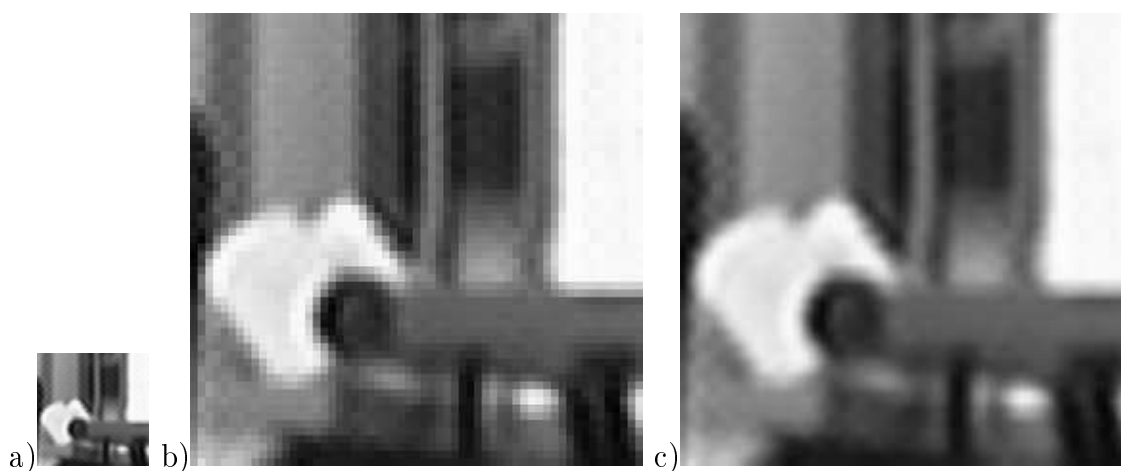
Rysunek 3.9: Schemat działania algorytmów powiększania: a) replikacja pikseli; b) interpolacja

Przy zmniejszaniu obrazu może dojść do utraty informacji. Minimalna częstotliwość próbkowania, przy której informacja nie jest tracona, to tzw. częstotliwość Nyquista. Pozwala ona na odtworzenie szczegółów o wielkości równej długości fali tej częstotliwości. W obrazie oznacza to, że najmniejsze dostrzegalne szczegóły mogą mieć rozmiar 1 piksla. Z kolei przy powiększaniu brak interpolacji powoduje "ziarnistość" obrazu. Rysunek 3.10 pokazuje, efekty zastępowania i interpolacji przy zmniejszaniu obrazu. Widać na nim, że różnica jest mała.

Na rysunku 3.11 porównano metody powiększania obrazu. Przewaga jakości interpolacji nad replikacją jest wyraźnie widoczna, choć z drugiej strony replikacja pikseli jest znacznie szybsza.



Rysunek 3.10: Zmniejszanie obrazu: a) obraz źródłowy; b) bez interpolacji; c) z interpolacją



Rysunek 3.11: Zwiększanie obrazu: a) obraz źródłowy; b) bez interpolacji; c) z interpolacją

3.4.2 Transformacja biliniowa

Klasa transformacji biliniowych jest nadklasą transformacji afinicznych. Transformację biliniową (zwaną też dwuliniową) określają cztery punkty. Jest ona mapowaniem dowolnego czworokąta wypukłego (najczęściej kwadratu) w inny dowolny czworokąt wypukły. Równania transformacji dwuliniowej są następujące:

$$x' = a_{00}x + a_{01}y + a_{02}xy + b_0; \quad (3.14)$$

$$y' = a_{10}x + a_{11}y + a_{12}xy + b_1. \quad (3.15)$$

Znajomość odwzorowanych współrzędnych wszystkich 4 wierzchołków dostarcza 8 niezależnych zmiennych, na podstawie których można ten układ rozwiązać. W transformacji jest 8 nieznanymi współczynnikami i wyznacza się je na podstawie układu 8 niezależnych równań. W

zapisie macierzowym oznaczając macierz odpowiednio ułożonych współrzędnych przed transformacją jako \underline{X} , wektor współrzędnych po transformacji jako \underline{X}' , zaś macierz transformacji jako \underline{A} otrzymuje się wzory transformacji jako (3.16) i (3.17).

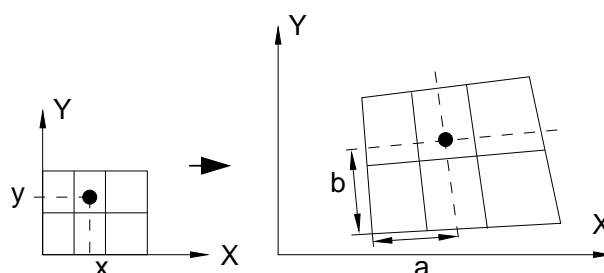
$$\underline{A} \bullet \underline{X} = \underline{X}' \quad (3.16)$$

$$\begin{bmatrix} a_{00} & a_{01} & a_{02} & b_0 \\ a_{10} & a_{11} & a_{12} & b_1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ x_1 & x_2 & x_3 & x_4 \\ y_1 & y_2 & y_3 & y_4 \\ x_1y_1 & x_2y_2 & x_3y_3 & x_4y_4 \end{bmatrix} = \begin{bmatrix} x'_1 & x'_2 & x'_3 & x'_4 \\ y'_1 & y'_2 & y'_3 & y'_4 \end{bmatrix} \quad (3.17)$$

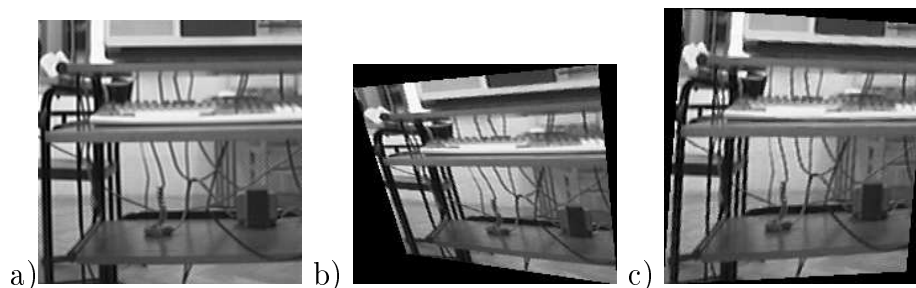
Jeżeli macierz \underline{X} jest kwadratowa, a żadne 3 punkty nie są współliniowe, istnieje macierz odwrotna \underline{X}^{-1} . Macierz \underline{A} wyznacza się według wzoru (3.18):

$$(\underline{A} \bullet \underline{X}) \bullet \underline{X}^{-1} = \underline{X}' \bullet \underline{X}^{-1} \Rightarrow \underline{A} = \underline{X}' \bullet \underline{X}^{-1}. \quad (3.18)$$

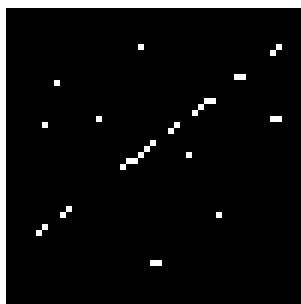
Schemat transformacji geometrycznej przedstawia rysunek 3.12. Transformacja nadaje się do usuwania zniekształceń np. perspektywicznych. W rozdziale 4 opisuję zastosowanie transformacji dwuliniowej do usuwania ogólnych zniekształceń geometrycznych obrazu. Przykład transformacji dwuliniowej przedstawia rysunek 3.13. W transformacji biliniowej część obrazu może być zmniejszona, a jednocześnie inna część - powiększona. Do poprawy jakości obrazu stosuje się jednocześnie *subsampling* i interpolację obrazu.



Rysunek 3.12: Schemat przekształcenia dwuliniowego



Rysunek 3.13: Przekształcenie dwuliniowe: a) obraz wyjściowy; b) i c) – przykłady przekształcenia



Rysunek 3.14: Obraz z linią i punktami zakłócającymi

3.5 Transformata Hough'a

Transformata Hougha [48], [49], [29], [30] jest narzędziem umożliwiającym detekcję w obrazie linii prostych, a w rozszerzonej postaci – dowolnych kształtów oraz kształtów "abstrakcyjnych" nie będących rzeczywistymi obiektami.

W podstawowej wersji transformata pozwala uzyskać informację liniach prostych w obrazie. Obraz musi być po obróbce wstępnej i detekcji krawędzi detektorem np Laplace'a, Marra-Hildretha lub innym, generującym linie cienkie, o grubości jednego piksela. Możliwe jest stosowanie transformaty do obrazów w których krawędzie wykrywa się filtrem Sobela (lub innym dającym szerokie linie), lecz wynik transformaty jest wówczas bardzo niedokładny.

Przykładowy obraz po detekcji krawędzi przedstawia rysunek 3.14. Transformata jest wyznaczana w przestrzeni współczynników linii prostych o równaniach postaci:

$$x \cos \alpha + y \sin \alpha = c, \quad (3.19)$$

gdzie α jest kątem nachylenia prostej względem osi OX , zaś c jest odległością tej prostej od początku układu współrzędnych. W literaturze spotyka się wzór (3.19) również z zamienionymi funkcjami kątów, w postaci

$$x \sin \alpha + y \cos \alpha = c. \quad (3.20)$$

Podczas wyznaczania transformaty, dla każdego punktu o współrzędnych (x, y) wyznacza się równania linii, które mogą przez niego przebiegać i zaznacza się je w układzie współrzędnych c, α . Zaznaczenie polega na zwiększeniu o 1 odpowiedniego elementu tablicy liczników – punkt "głosuje" na rodzinę wszystkich prostych przechodzących przez niego. Proste generuje się z rozdzielczością ustaloną wcześniej, według której dobrano rozmiary tablicy. Dla każdej wartości α wyznacza się odległość c linii o kącie nachylenia α przechodzącej przez punkt $p = (x, y)$ według równania (3.19).

Na przykładowym obrazie z rysunku 3.14 jest kilkanaście punktów współliniowych oraz kilka dodatkowych, nie tworzące wspólnie z żadnym trzecim linii prostej. Transformatę Hougha tych punktów przedstawia rysunek 3.15, przy czym kolejne poziomy jasności odpowiadają ilości zwiększeń danego licznika. Każdemu z punktów na rysunku 3.14 odpowiada jedna linia sinusoidalna o małym progu jasności na rysunku 3.15. Widać, że linie odpowiadające punktom współliniowym przecinają się w jednym punkcie, dając wyraźny jasny punkt, podczas gdy linie odpowiadające pozostałym punktom nie dają tak charakterystycznego sygnału. Wartość komórki w punkcie (α, c) odpowiada ilości punktów głosujących na tę linię, co na rysunku 3.15 jest reprezentowane przez jasność piksli.

W praktyce tak prosta transformata Hougha jest obarczona błędem dokładności, a skupienie często nie jest tak dobre jak na rysunku 3.15. Dodatkowo możliwe jest wygenerowanie



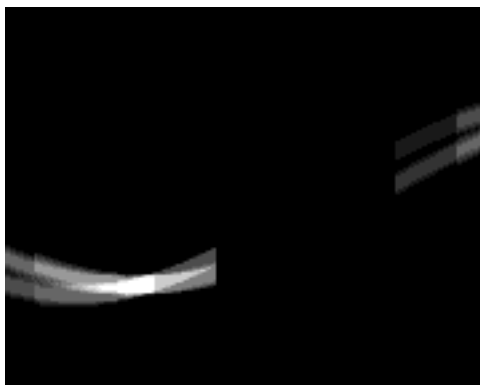
Rysunek 3.15: Obraz transformaty Hougha obrazu z rysunku 3.14

wyraźnego sygnału dla wielu punktów przypadkowo współliniowych, ale nie przylegających do siebie - np jeżeli w obrazie jest dużo krótkich linii pionowych na tej samej wysokości, jeżeli ilość linii krótkich jest porównywalna z ich długością, wówczas można otrzymać sygnał dla linii poziomych, których w rzeczywistości nie ma. Problemy te zostały omówione m.in w [80].

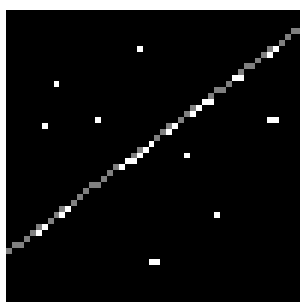
W celu uniknięcia rozproszenia punktu przecięcia na 2 sąsiednie komórki układu współrzędnych α, c , każdy punkt głosuje na długość c wynikającą z równania (3.19), a także na c o 1 jednostkę mniejszą i większą. W efekcie otrzymuje się sygnał w transformacie bardziej rozmyty, ale za to z gwarancją, że wartość transformaty nie jest podzielona między 2 wartości c lub α . Problem przypadkowych linii rozwiązano w ten sposób, że punkt może głosować na linię nachyloną pod kątem α wtedy i tylko wtedy, jeżeli razem ze swoimi 2 sąsiadami jest w stanie utworzyć linię nachyloną pod kątem α . Przykładowy wykres transformaty dla punktów z rysunku 3.14 znajduje się na rysunku 3.16. Ponieważ wartości c z równania (3.19) są identyczne dla $\alpha_2\alpha + \pi$, transformatę można wyznaczać dla $\alpha \in \langle 0; \pi \rangle$. Zakres wartości c mieści się w granicach $\langle -\sqrt{width^2 + height^2}; \sqrt{width^2 + height^2} \rangle$, gdzie $width$ jest szerokością, a $height$ wysokością obrazu w pikslach. Dla bezpieczeństwa dodają 10 z każdej strony przedziału. Na rysunku 3.16 widać wyraźnie, że linie sinusoidalne nie przebiegają przez cały zakres α , a jedynie przez $3/8$ całego przedziału $\langle 0; \pi \rangle$ (każda z linii złożonych z trzech punktów może głosować na linie położone w $3/8$ kąta pełnego). W porównaniu z rysunkiem 3.15 nie ma sygnałów pojedynczych piksli, ponieważ nie mają one sąsiadów z którymi mogłyby tworzyć linie.

Po ustaleniu, że na rysunku jest jakaś linia – przez wybór najjaśniejszego punktu czyli najdłuższej linii – należy ustalić które punkty należą do tej linii (tzn głosowały na nią) i usunąć ich głosy z transformaty. W pozostałej części transformaty można znaleźć kolejną najdłuższą linię i ustalić jej równanie. Proces ten powtarza się dopóki znajdowane linie są nie krótsze niż ustalona wartość progowa lub do usunięcia wszystkich punktów z transformaty.

Efektom rozmycia transformaty jest trudność dokładnego ustalenia równania linii. Dlatego po stwierdzeniu, że na rysunku istnieje najdłuższa linia o długości większej niż progowa, podczas usuwania sygnału należących do wykrytej linii równanie wyznaczam metodą regresji liniowej biorąc za podstawę współrzędne wszystkich punktów głosujących – zob. rozdział 7.3.3. Efekt końcowy – wyznaczoną linię prostą – przedstawia rysunek 3.17. Na rysunku źródłowym 3.14 transformata Hougha połączona z regresją liniową znalazła tylko jedną linię, pokrywającą się z linią faktycznie istniejącą.



Rysunek 3.16: Obraz rozszerzonej transformaty Hougha obrazu z rysunku 3.14



Rysunek 3.17: Linia wykrzyta przy pomocy transformaty Hougha i regresji liniowej

Rozdział 4

Kalibracja kamery

Obraz uzyskany z kamery jest źródłem wielu informacji, w szczególności danych pomiarowych. Aby obraz mógł być traktowany jako wiarygodne źródło danych pomiarowych, należy przeprowadzić kalibrację kamery, podobnie jak kalibruje się inne urządzenia pomiarowe. Dodatkowo, obraz pochodzący praktycznie z każdej kamery jest zniekształcony, i kalibracja powinna prowadzić do usunięcia niepożądanych zniekształceń.

Nie istnieje jedna metoda kalibracji kamery. Różnią się one ilością wyznaczanych parametrów, złożonością obliczeń, ilością obrazów wejściowych oraz ilością wartości, które należy wyznaczyć w inny sposób, np. mierząc odległość.

4.1 Parametry kamery

Parametry niezbędne do wykonania pomiarów dzieli się na wewnętrzne i zewnętrzne. Parametry wewnętrzne są specyficzne dla kamery i można je określić podczas kalibracji, zaś parametry zewnętrzne zależą np. od położenia kamery.

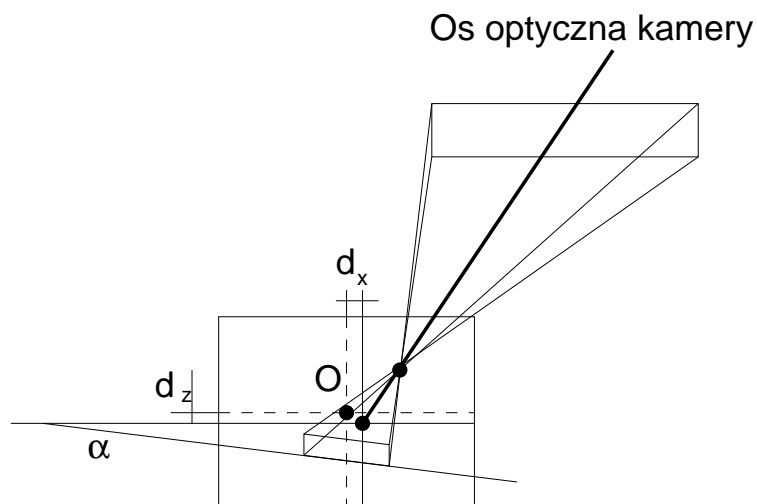
Do wewnętrznych parametrów kamery można zaliczyć: odległość ogniskową f , rozmiar fizyczny obrazu, położenie rzutu osi optycznej na obrazie, kąt obrotu obrazu w płaszczyźnie obrazu, odchylenie osi optycznej kamery od osi prostopadłej do obrazu, zniekształcenia nieliniowe spowodowane zastosowaniem szerokokątnego obiektywu. Schemat odwzorowania obiektu w obrazie z kilkoma parametrami wewnętrznymi przedstawia rysunek 4.1.

Ogniskowa obiektywu jest podstawowym parametrem umożliwiającym wyznaczenie relacji między odległością na obrazie i odległością w świecie rzeczywistym. Oznacza się ją symbolem f . Pojęcia ogniskowej oraz ogniska kamery są znane z optyki. Znaczenie ogniskowej przy wyznaczaniu odległości do znanego obiektu przedstawia rysunek 4.2.

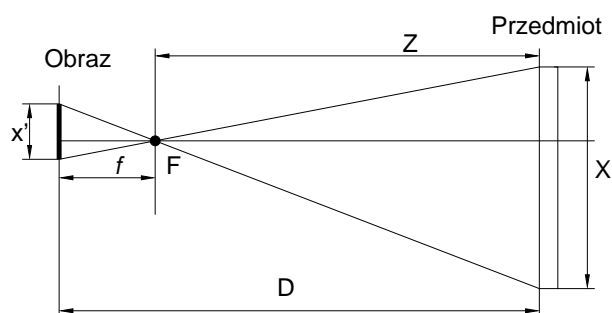
Znając **rozmiar fizyczny obrazu** X , ogniskową f oraz rozmiar obrazowy x' przedmiotu, odległość Z między obiektywem a przedmiotem wyznacza się wg wzoru:

$$D = f + Z = f + f \frac{X}{x'} \quad (4.1)$$

Zniekształcenia nieliniowe radialne są spowodowane rzutowaniem sferycznej powierzchni soczewki na prostokątną matrycę. Chociaż teoretycznie są obecne zawsze, w praktyce trzeba brać je pod uwagę jedynie dla soczewek o krótkich ogniskowych. Inne rodzaje zniekształceń mogą być spowodowane budową układu optycznego kamery. Przykład zniekształceń nieliniowych przedstawia rysunek 4.4. Zniekształcenia nieliniowe opisuje się na podstawie znajomości ogniskowej kamery, lub na podstawie pomiaru zniekształceń obiektów o ustalonym kształcie.



Rysunek 4.1: Schemat powstawania rzutu przedmiotu na obrazie z wyodrębnieniem wewnętrznych parametrów kamery: rzutu odległości d_x i d_z osi optycznej od środka obrazu oraz kąta obrotu α



Rysunek 4.2: Parametry kamery - ogniskowa (f), ognisko (F) i związek pomiędzy rozmiarem na obrazie oraz rozmiarem rzeczywistym

Oprócz parametrów charakteryzujących kamerę ważne jest ustalenie położenia i orientacji kamery w układzie współrzędnych robota. Położenie opisuje się trzema współrzędnymi X, Y, Z, zaś orientację dwoma kątami ϕ (odchylenie od poziomu) i θ .

Położenie rzutu osi optycznej na obrazie nie zawsze jest w środku obrazu. Często elementy światłoczułe, np. kamer z czujnikiem CCD, są na tyle małe, że nawet bardzo mała niedokładność montażu powoduje znaczne przesunięcie miejsca rzutu. Na rys. 4.1 rzut osi optycznej jest oddalony od środka obrazu o wektor d_x, d_z .

4.2 Klasyfikacja metod kalibracji kamery

Krótki przegląd i klasyfikację strategii kalibracji można znaleźć w [88]. Poniżej przedstawiono definicje i klasyfikację podejść do kalibracji zaczerpnięte z tego artykułu uzupełnione przykładami konkretnych metod.

Metody kalibracji kamery różnią się następującymi czynnikami: *złożonością modelu*, *klasyfikacją parametrów modelu*, *miarą dopasowania modelu* i *techniką rozwiązania numerycznego*. Model systemu można opisać równaniem

$$F(\mu) = 0, \quad (4.2)$$

gdzie μ jest wektorem parametrów składającym się z wszystkich punktów obrazu, punktów świata rzeczywistego oraz parametrów kamery. Z wektorem parametrów μ można skojarzyć wektorową zmienną losową U . Niech u będzie wektorem pomiarów U , zaś ϵ – błędem tego pomiaru (tzn. $\nu = u - \epsilon$). Przy założeniu że U ma rozkład normalny z wartością średnią μ i macierzą kowariancji W^{-1} (tzn. $U \sim N(\mu, W^{-1})$), łatwo można pokazać że estymacja maksymalnego podobieństwa $\hat{\mu}$ jest dana rozwiązaniem równania:

$$\min_{\mu} \epsilon^T W \epsilon \text{ należy do } F(u - \epsilon) = 0. \quad (4.3)$$

W ten sposób problem kalibracji można przedstawić, z pewnym uogólnieniem, jako optymalizację nieliniową.

4.2.1 Złożoność modelu

Model systemu

System definiuje się jako wszystkie urządzenia wymagające kalibracji, łącznie ze wszystkimi znanymi ograniczeniami. Typowymi urządzeniami są kamery i projektory. Możliwe warianty obejmują układy wielu kamer, pojedynczą kamerę umieszczoną w wielu różnych miejscach i pary kamera/projektor (w systemach światła strukturalnego). Ograniczenia i założenia są metodą wprowadzania dodatkowej informacji do modelu.

Model kamery Najprostszy model kamery polega na projekcji perspektywicznej ze świata zewnętrznego na matrycę kamery. *Równania współliniowe*, przyjęte dla kamery CCD, opisują tę relację:

$$F_x = s_x(x - x_0) - fU/W = 0; \quad (4.4)$$

$$F_y = s_y(y - y_0) - fV/W = 0, \quad (4.5)$$

gdzie: $P_c = (UVW)^T = R(P^w - P_0)$ – współrzędne punktu sceny względem kamery (w mm),
 P^w – współrzędne punktu sceny w świecie (w mm),
 P_0 – współrzędne środka perspektywy w świecie,
 R – macierz rotacji między układami współrzędnych świata i kamery,
 $p^i = (xy)^T$ – współrzędne obrazowe punktu po projekcji (w pikslach),
 $p_0 = (x_0, y_0)^T$ – współrzędne obrazowe środka optycznego (w pikslach),
 f – efektywna długość ogniskowej kamer (w mm),
 $s = (s_x, s_y)^T$ – współczynniki skalujące (w mm/piksel).

Rzeczywiste kamery są obciążone systematycznymi błędami takimi jak zniekształcenia soczewki i przypadkowymi efektami takimi jak szum sensora. Niektóre błędy systematyczne można włączyć do modelu przez dodanie wyrażeń korygujących Δ_x , Δ_y , do równań (4.4) i (4.4). Przykładowo, zniekształcenia radialne i rozpraszające soczewki można modelować poprawkami opisanymi równaniami (4.6) i (4.7):

$$\Delta_x = \bar{x}(K_1r^2 + K_2r^4 + \dots) + (P_1(r^2 + 2\bar{x}^2) + 2P_2\bar{x}\bar{y})(1 + P_3r^2 + \dots), \quad (4.6)$$

$$\Delta_y = \bar{y}(K_1r^2 + K_2r^4 + \dots) + (2P_1\bar{x}\bar{y} + P_2(r^2 + 2\bar{y}^2))(1 + P_3r^2 + \dots), \quad (4.7)$$

gdzie $\bar{x} = s_x(x - x_0)$; $\bar{y} = s_y(y - y_0)$; $r = \sqrt{\bar{x}^2 + \bar{y}^2}$, a parametry K_i i P_i modelują odpowiednio zniekształcenia radialne i rozpraszające.

Można modelować wiele innych parametrów fizycznych. W wąskiej klasie aplikacji może być konieczne modelowanie zależności zniekształceń od odległości. Nawet po zamodelowaniu tych efektów ciągle możliwe są systematyczne różnice między danymi i modelem. Takie zniekształcenia można modelować dodając wyrażenie korekcyjne, np. w postaci wielomianu. Nie jest konieczne zrozumienie i modelowanie wszystkich występujących zakłóceń i efektów. Bardzo ważne jest żeby nie wprowadzić zbyt wielu parametrów do modelu.

Wiele obrazów z jednej kamery i wielu kamer Załóżmy, że kamerę umieszczano w m miejscach i z każdego pobrano obraz. Parametry kamery można podzielić na zmieniające się wraz ze zmianą obrazu i nie zmieniające się. Przykładowo, położenie kamery zmienia się, ale długość ogniskowej nie (przy założeniu, że nastawy soczewki się nie zmieniają).

Założmy, że jest kilka kamer z których każda dostarcza jeden obraz. W wielu aplikacjach powoduje to proste podwojenie modelu pojedynczej kamery ponieważ wszystkie parametry można rozpatrywać niezależnie. Jednak w aplikacjach wysokiej precyzji, gdzie punkty świata traktuje się jak pomiary (nie jak stałe), każda kamera musi zgodzić się z estymacją położenia każdego punktu świata. W ten sposób punkty świata stają się wspólnymi parametrami i modele kamer nie mogą być dłużej traktowane jako niezależne.

4.2.2 Pozostałe parametry

Parametry modelu można ogólnie sklasyfikować na podstawie ilości dostępnych o nich informacji. *Stała* jest całkowicie zdeterminowana, *wartość swobodna* jest całkowicie nieograniczona, a *mierzony parametr* ma jakąś określoną *a priori* estymatę. Można wyróżnić wiele klasyfikacji parametrów w modelu na podstawie ich wariancji w W^{-1} lub wag w W . Waga stałej jest duża, zaś wartości swobodnej mała lub równa zero. W ten sposób można wprowadzić do modelu wszystkie parametry w spójny sposób. To podejście jest skrajnie ogólne i nadaje się do wielu różnych sytuacji.

Funkcja $\epsilon^T W \epsilon$, w równaniu (4.3) jest miarą dopasowania modelu i powstała przy założeniu że szumy i błędy pomiarów są Gaussowskie. Jest jasne, że miara modelu dopasowania może być zastąpiona inną funkcją. Inne możliwości obejmują błąd maksymalny i błąd całkowity. Jednak suma kwadratów ma przewagę analityczną.

Problem ograniczonej optymalizacji przedstawiony na stronie 45 można rozwiązać stosując sekwencyjne programowanie kwadratowe (SQP). Ponieważ funkcja $\epsilon^T W \epsilon$ jest już kwadratowa, i jest spełnionych wiele dobrze zdefiniowanych założeń, można zastosować prostszą procedurę, znaną jako poprawki najmniejszych kwadratów (LSA). SQP i LSA są blisko związanymi procedurami iteracyjnymi rozwiązującymi problem programowania kwadratowego (QP) w każdym kroku. Dla specyficznej kamery możliwe jest zastosowanie prostszego i mniej złożonego obliczeniowo algorytmu.

4.2.3 Przegląd publikacji dotyczących kalibracji kamery

Wszystkie opisane przeze mnie prace dotyczą modelu pojedynczej kamery. Nie są to jedyne opracowania, lecz publikacje prezentujące typowe zastosowanie danej metody. Uporządkowano je według stopnia komplikacji.

Najprostsze metody zakładają, że błąd montażu jest pomijalny, i jedynym elementem jaki należy korygować jest radialne zniekształcenie obrazu spowodowane krótkoogniskowym obiektywem. W metodach tych wyznacza się zniekształcenia na podstawie znajomości ogniskowej kamery, jak w pracach [11], [73], [64]. Korekcję obrazu wykonuje się przez model, wykonany na podstawie jedynie znajomości ogniskowej i teoretycznie wyznaczonego na tej podstawie kąta widzenia. W pracach tych stosuje się odwzorowanie wycinka sfery na prostokąt funkcjami wielomianowymi i sferycznymi.

Bardziej złożony model zniekształceń opisano w pracach [85] i [86]. Artykuły te przedstawiają metodę korekcji zniekształceń nieliniowych kamer szerokokątnych, wprowadzających największe zniekształcenia. Metoda opiera się na matematycznym modelu zniekształceń. Autorzy wprowadzają 3 rodzaje zniekształceń: przesunięcie środka obrazu, zniekształcenia radialne oraz zniekształcenia decentralizujące. Wszystkie źródła błędów dla danego punktu opisano pięcioma równaniami. Proponowana metoda korekcji polega na wyznaczeniu współczynników tych równań metodą najmniejszych kwadratów. Do kalibracji używa się sztucznie generowanych obrazów z kilkoma liniami prostymi. W zależności od poziomu szumu (0-2 piksele) średni błąd położenia piksla w obrazie wynosi teoretycznie od 0 do 2 pikseli. Kalibracja kamery zajmuje ok 20 minut na komputerze PC z procesorem Pentium II 300MHz. Na podstawie wyznaczonych parametrów sporządza się tablicę przekształceń dla obrazu.

Praca [55] opisuje metodę kalibracji kamery dla systemu gry w piłkę nożną robotów. Metoda usuwa zniekształcenia nieliniowe dla kamer szerokokątnych, korygując regularne zniekształcenia radialne. Dodatkowo system kompensuje odchylenie osi kamery od kierunku prostopadłego do widzianej płaszczyzny (do gry robotów w piłkę). Zastosowano równania perspektywy, którą ustala się wykrywając linie boiska i transformując je do układu "prostopadłego". Proponowana korekcja barw polega na usunięciu wyznaczonego eksperymentalnie tła (dla obserwacji szarej powierzchni) oraz skalowaniu barw przez wyznaczone doświadczalnie współczynniki korekcji (inne dla każdej barwy i intensywności).

Bardziej złożoną technikę kalibracji opisano w [84]. Opisana metoda dokonuje kalibracji kamery na podstawie punktów odnajdywanych w różnych obrazach tego samego przedmiotu, wykonanych z różnych miejsc. Nie jest istotna znajomość położenia kamery w trakcie wykonywania zdjęć czy przemieszczenia kamery między kolejnymi zdjęciami. Metoda pozwala na kali-

brację na podstawie 2 lub 3 obrazów. Dla 2 obrazów liniowymi równaniami można skalibrować kamerę mając 8 punktów, zaś przy 3 obrazach - 7 punktów. Większa ilość odnalezionych punktów na wielu obrazach pozwala na dokładne wyznaczenie parametrów metodą najmniejszych kwadratów. Wartości wyznaczone podczas kalibracji są na podstawie złożonych wzorów, a ich sens fizyczny jest trudny do określenia.

Artykuł [27] opisuje metodę kalibracji zniekształceń radialnych kamer. Metoda pozwala na wyznaczenie parametrów zniekształcenia na podstawie znalezionych linii w obrazie. Dla wyznaczonych i połączonych linii wyznacza się stopień zniekształcenia, w szczególności krzywiznę. Na podstawie pomierzonych krzywizn kilku linii wyznacza się parametry zniekształcenia radialnego. Obraz nie musi być wcześniej przygotowany, ale istotne jest aby znajdowały się na nim dostatecznie długie obrazy linii w rzeczywistości prostych.

Bardzo złożoną procedurę kalibracji podano w pracy [92], wydanej przez firmę Microsoft. Artykuł przedstawia technikę kalibracji kamery przy pomocy wcześniej ustalonej płaskiej siatki o znanych rozmiarach. Wyznacza się parametry kamery na podstawie co najmniej dwóch obrazów wykonanych z różnych miejsc. W efekcie parametry kamery opisuje się sześcioma wartościami, których wymiar fizyczny jest trudny do określenia. Do artykułu jest dołączony program do kalibracji, który można pobrać ze stron internetowych firmy Microsoft.

4.3 Zastosowana metoda kalibracji

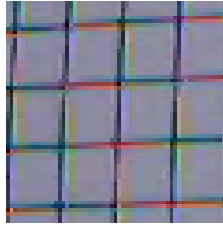
Opracowana metoda pozwala na wyznaczenie parametrów wewnętrznych kamery z dużą dokładnością. Do kalibracji wymagany jest następujący sprzęt: statyw do zamocowania kamery z możliwością pomiaru odległości oraz płynną regulacją odległości kamery od podłoża oraz płaskiej, nieruchomej powierzchni z regularną siatką prostokątną (może być wydrukowana czarna siatka na białym tle). Koniecznym warunkiem jej stosowalności są:

- dobra ostrość obrazu;
- niezmiennosc parametrów układu optycznego kamery pomiędzy kalibracją i akwizycją obrazu roboczego.

Pierwszy warunek oznacza, że dopuszczalne są zniekształcenia kształtów, ale nie może być znacznych zniekształceń ostrości. Wynika to z faktu korzystania z obrazu wzorcowego uzyskanego z kamery jako mapy zniekształceń, co pociąga za sobą wymaganie ostrości takiego obrazu. Tolerancja nie jest dokładnie określona, z praktyki wynika że praktycznie każda kamera szerokokątna ma różne ogniska dla różnych barw i następuje rozszczepienie linii czarnej na barwne, co przedstawia rysunek 4.3. Rozszczepienie barw powodujące zwiększenie szerokości linii do kilku pikseli jest szkodliwe, jednak jeżeli główna część linii nie jest szersza niż 2 piksele, wówczas można uważać że obraz po korekcji będzie wystarczająco dokładny. W stosowanej obróbce wstępnej nie jest wykonywana korekcja ostrości obrazu.

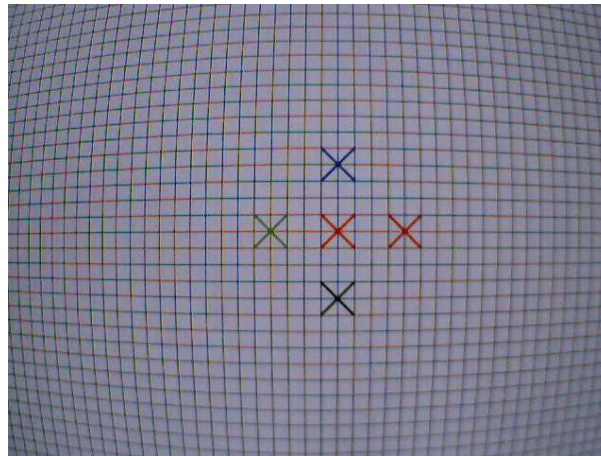
Drugi warunek oznacza, że metoda daje najlepsze wyniki, jeżeli obrazy robocze są uzyskiwane przy takich samych nastawach układu optycznego, przy jakich uzyskano obraz mapy. Wynika on z konieczności pracy przy tych samych parametrach pomiarowych, które były podczas kalibracji, i wprowadzania przez układ optyczny kamery zniekształceń zależnych od nastaw, np. optymalnej odległości, przesłony i regulacji ostrości.

Obraz wzorcowy jest obrazem płaskiej siatki prostokątnej (lub lepiej kwadratowej). W konkretnym przypadku stosowana była kartka papieru w kratkę. Obraz takiej kartki jest zniek-



Rysunek 4.3: Rozszczepienie barw w obrazie czarnej linii na białym tle spowodowane przez optykę kamery

sztalony (Rys 4.4). W zastosowanej kartce wyróżniono dodatkowo miejsce, w którym powinien znaleźć się rzut osi optycznej kamery w obrazie.



Rysunek 4.4: Obraz kartki w kwadratową kratkę ze zniekształceniami nieliniowymi

4.3.1 Przygotowanie do kalibracji

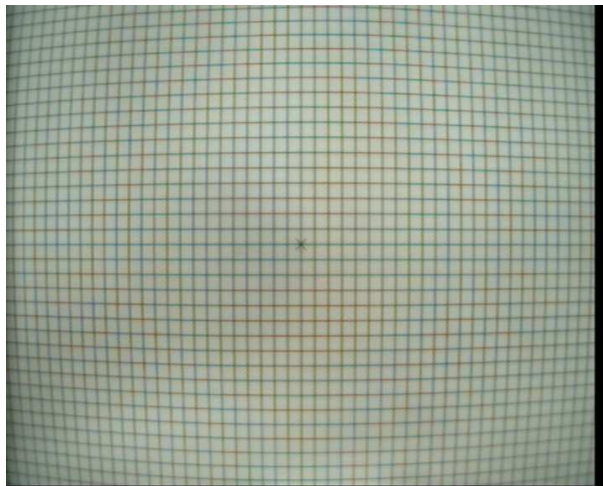
Ponieważ dokładność kalibracji wpływa na dokładność wszystkich późniejszych pomiarów, należy przeprowadzić ją ze szczególną starannością. Kamerę należy zamocować nieruchomo tak, aby powierzchnia kartki była równoległa do brzożu obiektywu. Równoległość najlepiej sprawdzić przysuwając obiektyw do kartki na bardzo małą odległość (mniej niż 1 mm) i ustawiając kąt nachylenia kamery. Ważne jest, aby linie nie zniekształcone przebiegające w pobliżu środka obrazu były dokładnie poziome i pionowe.

Sama siatka powinna być regularna i kwadratowa. Ze względu na różną jasność obrazu w różnych miejscach (obraz jest jaśniejszy na środku) najlepsza jest krata czarna na białym tle, co zapewnia kontrast wystarczający do automatycznej kalibracji. Odległość między węzłami siatki powinna być określona bardzo precyzyjnie. Podczas badań wynosiła 5 mm w pionie i w poziomie. Podczas akwizycji obrazu kalibrującego należy pamiętać, że płaszczyzna tej kartki będzie po kalibracji **reprezentować wszystkie płaszczyzny prostopadłe do osi optycznej kamery**. Dlatego należy jak najdokładniej ustawić kartkę prostopadle do osi optycznej kamery.

4.3.2 Znajdowanie położenia osi optycznej w obrazie

Położenie osi optycznej w obrazie nie zawsze jest w środku obrazu. Stosowana w końcowej fazie pracy kamera charakteryzuje się bardzo dużym przesunięciem położenia osi optycznej względem środka obrazu.

W celu wyznaczenia położenia osi optycznej należy zapewnić sobie możliwość oglądania obrazu z kamery. Punkt na który skierowana jest oś optyczna jest jedynym elementem obrazu nie zmieniającym położenia przy zmianie odległości między kartką z siatką a kamerą. Pozostałe punkty przybliżają się do osi optycznej podczas zwiększania tej odległości, i oddalają podczas jej zmniejszania, jak pokazuje to rysunek 4.6. Po rozpoznaniu tego miejsca należy tak umieścić kartkę, aby oś optyczna była skierowana na przecięcie węzłów. Na rysunku 4.4 oś optyczna jest skierowana na środkowy, czerwony znak "X". Nadal ważny jest poziomy i pionowy przebieg linii nie zniekształconych. Po odpowiednim umieszczeniu kartki należy zapisać obraz do pliku, z zapamiętaniem odległości między kamerą a kartką. Dobrze jest wykonać serię zdjęć dla różnych odległości. Siatka musi wypełniać cały obraz. Obraz nie może mieć dodatkowych znaczników, dlatego należy je usunąć po zapisaniu obrazu kalibrującego (przy pomocy programu graficznego). Możliwe jest też zastosowanie słabiej zaakcentowanego znacznika osi optycznej, który pokazuje rysunek 4.5. Na rysunku tym widoczny jest czarny obszar, który należy usunąć przed dalszą obróbką, zmniejszając wymiaru obrazu.



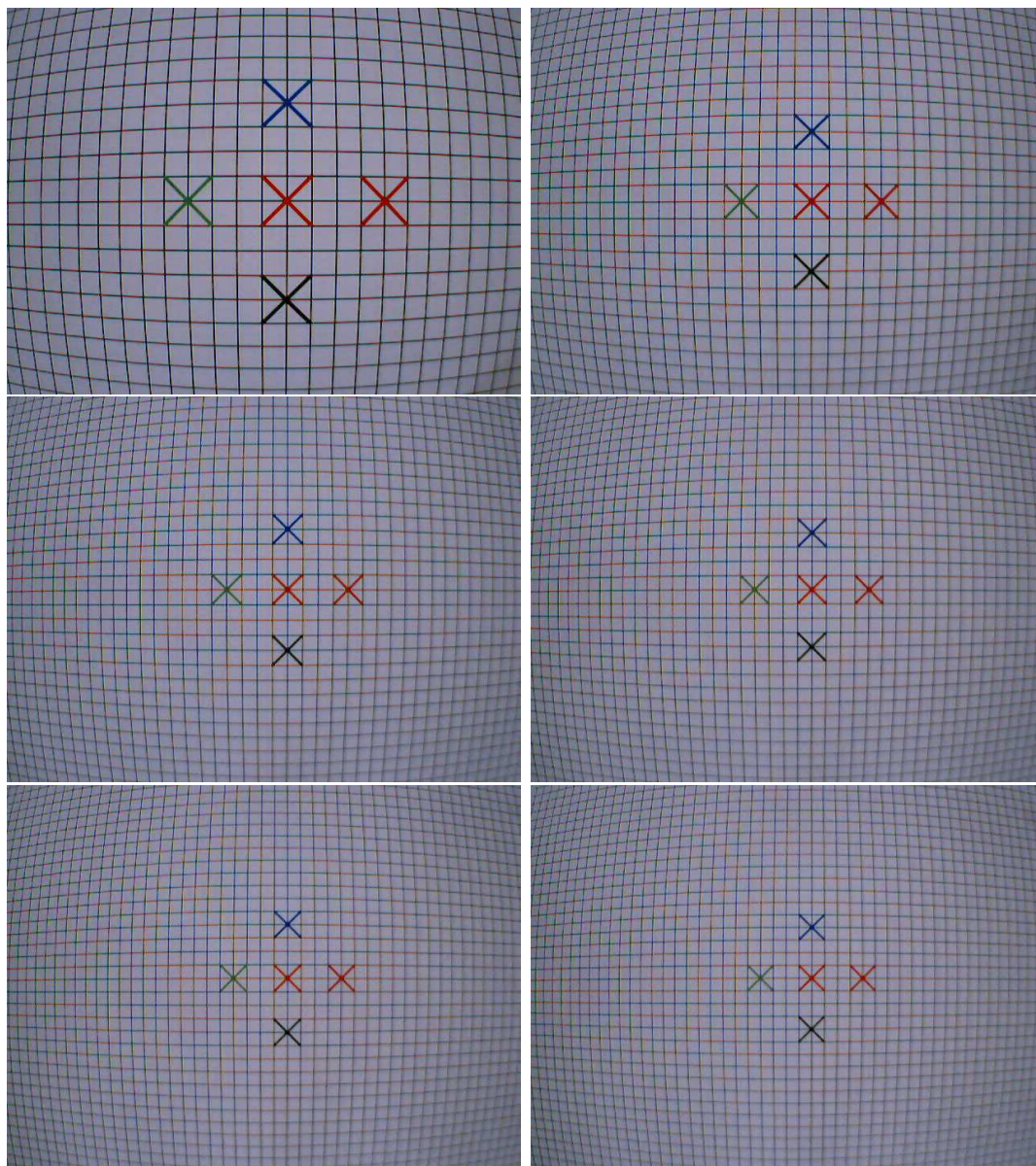
Rysunek 4.5: Obraz kartki w kwadratową kratkę z słabo zaakcentowanym położeniem osi optycznej

Automatyczne wyznaczenie przebiegu linii

Dla uzyskanego obrazu należy utworzyć mapę przebiegu linii, które w rzeczywistości są proste, i w związku z tym w teoretycznie uzyskanym obrazie powinny też być proste. Przebieg linii wyznacza się w następujących etapach:

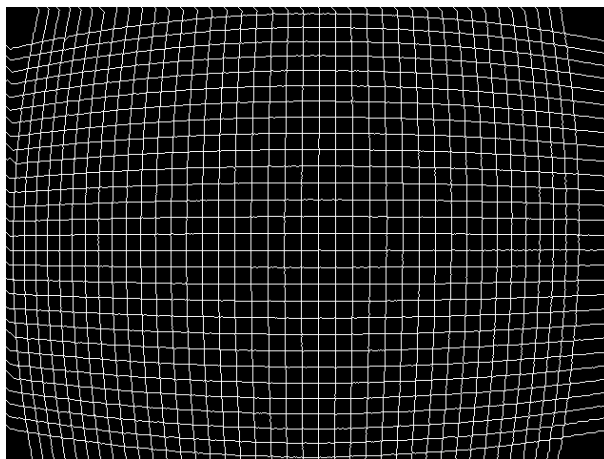
- wyznaczenie przebiegu linii poziomych i pionowych;
- analiza przebiegu linii z uwzględnieniem odległości i kierunku.

Wyznaczenie linii poziomych i pionowych wykonuje się oddzielnie, stosując algorytm Sobela o rozmiarze maski 3x3 (zob. rozdz. 3.3.1). Detekcję pojedynczych linii wykonuje się na



Rysunek 4.6: Poszukiwanie osi optycznej w serii obrazów: miejsce w które "patrzy" kamera jest jedynym, które nie zmienia położenia przy zbliżaniu i oddalaniu kamery.

podstawie maksimum jasności i minimum zmian kierunku linii, oddzielnie dla linii pionowych i poziomych. Dodatkowymi parametrami algorytmu są: minimalna odległość między liniami (eliminuje krótkie linie wynikające np. z nierównomierności oświetlenia), szerokość otoczenia w którym bada się ciągłość linii i stałość kierunku, minimalna jasność linii. Obraz linii wyznaczonych dla obrazu wzorcowego z rysunku 4.4 prezentuje rysunek 4.7. Ostatecznie danymi wyjściowymi tego etapu obróbki jest tablica punktów przecięć linii poziomych i pionowych, definiujących zniekształconą siatkę czworokątów.

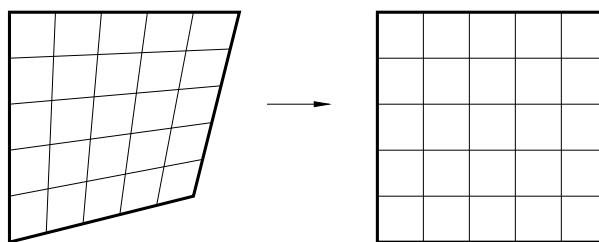


Rysunek 4.7: Linie siatki dla obrazu wzorcowego z rysunku 4.4

Wyznaczenie zniekształceń nieliniowych

Mapę zniekształceń nieliniowych wyznacza się porównując obraz rzeczywisty z teoretycznie oczekiwanym (bez zniekształceń). Obraz bez zniekształceń powinien przedstawiać w tym wypadku kwadratową (lub prostokątną) siatkę. Korekcja obrazu polega na znalezieniu odwzorowania obrazu zniekształconego w obraz bez zniekształceń. Zastosowane odwzorowanie polega na odwzorowaniu każdego widocznego quasi-czworokąta obrazu w kwadrat. Pierwszy etap, opisany w rozdziale 4.3.2, polega na wyróżnieniu na zniekształconym obrazie wzorcowym węzłów siatki. Następnie każdy czworokąt o znanych współrzędnych wierzchołków jest przekształcany w kwadrat z zastosowaniem przekształcenia dwuliniowego (biliniowego), opisanego w rozdziale 3.4.2. Zasadę tego przekształcenia przedstawia rysunek 4.8. Zakłada się przy tym, że odległości między węzłami są na tyle małe, że zniekształcone oczka siatki można traktować jak czworokąty – wówczas błąd wprowadzany przez przekształcenie jest minimalny, i wynosi co najwyżej 1 piksel. Podczas korekcji pomijany jest obszar przy brzegach składający się z częściowo widocznych oczek siatki. Ponieważ za rozmiar kratki w obrazie po korekcji przyjmuje się rozmiar największej kratki przed korekcją (rozmiar ten zmniejsza się wraz z oddalaniem się od środka obrazu), a na obrazie po korekcji wszystkie kratki mają jednakowe rozmiary, obraz po korekcji jest większy niż przed korekcją. Rysunek Rys. 4.9.a przedstawia fragment obrazu rzeczywistego ze zniekształceniami, zaś rysunek Rys. 4.9.b ten sam obraz po korekcji.

Mapę zniekształceń wykonuje się raz dla konkretnej kamery i określonej rozdzielczości obrazu. Wyznaczenie przekształcenia wszystkich pikseli jest znacząco czasochłonne. Oprócz mapy warto w celu przyspieszenia obliczeń utworzyć tablicę przekształceń dla każdego piksela, co pozwala na ok. 200-krotne przyspieszenie późniejszych przekształceń.



Rysunek 4.8: Zasada przekształcenia dwuliniowego



a)

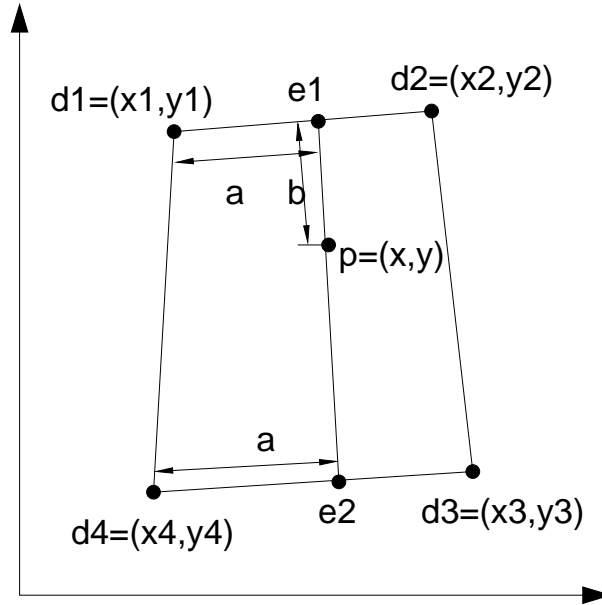


b)

Rysunek 4.9: Korekcja zniekształceń nieliniowych: a) obraz przed korekcją; b) ten sam obraz po korekcji

Wyznaczenie współrzędnych w przekształceniu dwuliniowym

Na Rys. 4.10 pokazano punkty d_1, d_2, d_3, d_4 definiujące czworokąt, definicję punktu p przez współrzędne wynikowe a i b oraz punkty pośrednie e_1 i e_2 .



Rysunek 4.10: Wyznaczenie współrzędnych punktu w obrazie wejściowym na podstawie znajomości współrzędnych w obrazie wyjściowym

Przekształcenie dwuliniowe korygujące pojedynczy prostokąt polega na wyznaczeniu parametrów a i b na podstawie współrzędnych punktów d_1, d_2, d_3, d_4 i p dla każdego piksla obrazu. W rzeczywistości stosowane jest przekształcenie odwrotne, tzn. znane są parametry a i b dla każdego piksla obrazu wynikowego, a konieczne jest wyznaczenie odpowiadających im współrzędnych x i y w obrazie wejściowym. Wyznaczenie tych współrzędnych przebiega następująco [90] :

- Wyznaczenie punktów e_1 i e_2 położonych na bokach czworokąta proporcjonalnie do wartości parametru a ;
- Wyznaczenie punktu p leżącego na linii e_1 - e_2 proporcjonalnie do wartości b .

Współrzędne punktów e_1 i e_2 wyznacza się według wzorów:

$$e_1(a) = d_1 + a(d_2 - d_1), \quad (4.8)$$

$$e_2(a) = d_4 + a(d_3 - d_4). \quad (4.9)$$

Rzeczywiste współrzędne wyznacza się wstawiając współrzędne punktów $d_1 \div d_4$ do powyższych wzorów. Współrzędne punktu p wyznacza się według wzoru:

$$p(a, b) = e_1 + b(e_2 - e_1) = d_1 + a(d_2 - d_1) + b(d_4 - d_1) + ab(d_1 - d_2 + d_3 - d_4). \quad (4.10)$$

Wartości parametrów a i b odpowiadają współrzędnym piksli w wynikowym prostokącie. Przy założeniu, że wynikowy prostokąt ma rozmiary $A \times B$ piksli, a przyjmuje wartości z zakresu $\frac{0}{A} \div \frac{A-1}{A}$, zaś $b - \frac{0}{B} \div \frac{B-1}{B}$. Przyjęty sposób wyznaczania parametrów powoduje, że zawsze są one dodatnie i mniejsze niż 1.

Dokładność przekształcenia dwuliniowego

Błędy w obrazie wynikowym są wprowadzane przez sam mechanizm korekcji. Mogą też być spowodowane błędem ustawienia siatki wzorcowej względem kamery. Mechanizm korekcji wymaga wyznaczenia węzłów zniekształconej siatki. Są one wyznaczane z dokładnością do jednego piksla i taki może być w założeniu maksymalny błąd określenia położenia węzła siatki. Poniżej zamieszczono analizę wpływu błędu wyznaczenia współrzędnych węzłów siatki na współrzędne punktu obrazu wejściowego znając współrzędne (a, b) punktu w obrazie wynikowym. Współrzędna x wyraża się wzorem:

$$x(a, b) = x_1 + a(x_2 - d_1 + b(x_4 - d_1) + ab(x_1 - d_2 + d_3 - d_4)). \quad (4.11)$$

Po przekształceniu wzór ten można zapisać następująco:

$$x(a, b) = x_1(1 - a - b + ab) + x_2(a - ab) + x_3ab + x_4(b - ab). \quad (4.12)$$

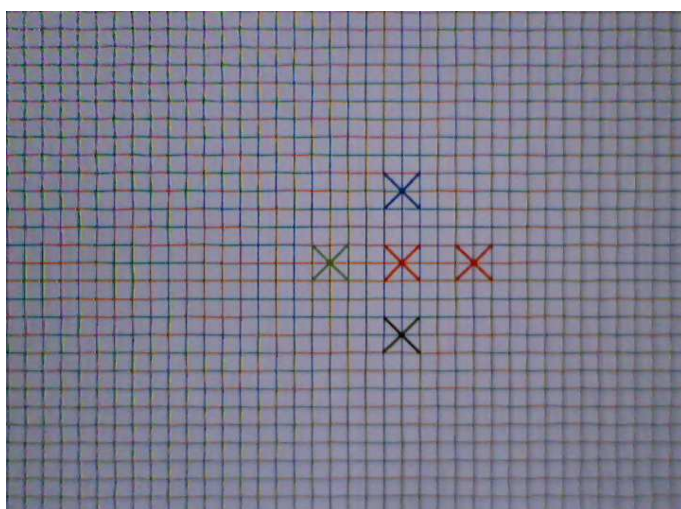
Wszystkie czynniki przy współrzędnych są dodatnie i mniejsze niż 1. Największy błąd występuje, gdy błędy wszystkich współrzędnych mają tę samą wartość, czyli +1 albo -1. Maksymalnie zakłócona wartość x wyraża się wzorem:

$$\begin{aligned} x_{max}(a, b) &= \\ (x_1 + 1)(1 - a - b + ab) + (x_2 + 1)(a - ab) + (x_3 + 1)ab + (x_4 + 1)(b - ab) &= \\ = x(a, b) + (1 - a - b - ab + a - ab + ab + b - b) &= \\ = x(a, b) + 1. & \quad (4.13) \end{aligned}$$

Analogicznie wyraża się błąd współrzędnej Y . Ze wzoru (4.13) wynika, że maksymalny błąd współrzędnych wszystkich punktów wewnątrz czworokąta wynosi 1 piksel i występuje, gdy wszystkie wierzchołki czworokąta są przesunięte w tę samą stronę. Przyjmując, że linia wyznaczona całkowitymi współrzędnymi piksli leży najbliżej rzeczywistej linii w obrazie, współrzędne punktów są tylko zaokrąglone. Błąd wynikający z zaokrąglenia do wartości całkowitej nie jest większy niż $\frac{1}{2}$ piksela. Jako, że korekcja jest odwzorowaniem powiększającym, błąd wprowadzony przez zaokrąglenie współrzędnych przy wyznaczaniu przekształcenia jest nie większy niż piksel. Można go pominąć, jeżeli piksel wynikowy jest wyznaczany jako średnia ważona otoczenia piksela źródłowego. Wówczas ogólny błąd wynosi maksymalny wynosi $\frac{1}{2}$ piksela i jest to maksymalna wielkość dodatkowego zniekształcenia, jakie może wprowadzić korekcja. Dokładność tę można poprawić stosując analizę obrazu wzorcowego z dokładnością subpikslową. Wyznaczenie siatki z obrazu oraz korekcję zniekształceń opisano w [45].

Wyznaczenie położenia rzutu osi optycznej

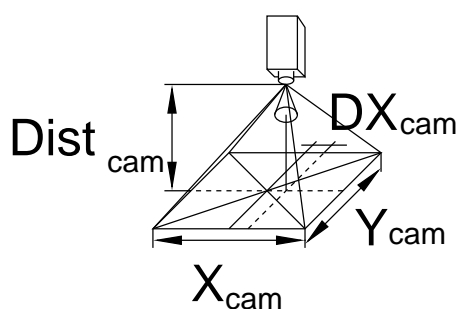
Jak wspomniano w rozdziale 4.3.2, obraz wzorcowy powinien być tak wykonany, żeby rzut osi optycznej na obrazie znajdował się w węźle siatki. Wówczas współrzędne tego punktu są łatwe do wyznaczenia po korekcji na podstawie tablicy wierzchołków do transformacji. Jeżeli zaś znane są współrzędne tego punktu w obrazie wzorcowym, ale nie leży on w węźle siatki, wówczas wyznaczenie jego współrzędnych w obrazie po korekcji wymaga rozwiązania układu równań przekształcenia biliniowego podanego w rozdziale 3.4.2. Jest to układ 8 lub 9 równań liniowych, którego szczegółową analizę pominięto ze względu na możliwość odpowiedniego sporządzenia obrazu wzorcowego. Przykład położenia rzutu osi optycznej w skorygowanym obrazie 4.4 przedstawia rysunek 4.11. Bardzo wyraźnie widoczna jest asymetria kamery, kierunek osi optycznej jest różny od kierunku od środka obrazu.



Rysunek 4.11: Obraz wzorcowy z zaznaczonym rzutem osi optycznej po korekcji zniekształceń

Kamera jako urządzenie pomiarowe

Podczas kalibracji generowany jest obraz bez zniekształceń nieliniowych. Obraz ten odpowiada rzutowaniu przestrzeni na płaszczyznę prostopadłą do powierzchni czołowej kamery. Pomiary rozmiarów są możliwe dzięki wyznaczeniu rozmiarów poziomego i pionowego obszaru znajdującego się w określonej odległości od kamery i wypełniającego dokładnie cały ekran. Rozmiary, oznaczane w dalszej części pracy jako X_{cam} i Y_{cam} , te można bardzo prosto wyznaczyć na podstawie liczby kraterki siatki w obrazie wzorcowym po korekcji i znajomości ich rozmiaru. Rozmiar oznaczony $Dist_{cam}$ należy dokładnie zmierzyć podczas kalibracji. Wymienione trzy wielkości i znajomość położenia osi optycznej na obrazie są wszystkimi parametrami pomiarowymi. Rysunek 4.12 przedstawia wielkości wyznaczone podczas kalibracji. Dla zastosowanej kamery wyznaczone parametry wynoszą odpowiednio: $Dist_{cam}=160\text{mm}$, $X_{cam}=190\text{mm}$, $Y_{cam}=140\text{mm}$.



Rysunek 4.12: Podstawowe parametry wyznaczone podczas kalibracji kamery

4.3.3 Zautomatyzowana kalibracja kamery

W celu łatwiejszej kalibracji można zautomatyzować etapy kalibracji polegające na przetwarzaniu obrazów. Algorytm automatycznej kalibracji został zaimplementowany w języku Visual C++. Danymi wejściowymi algorytmu są:

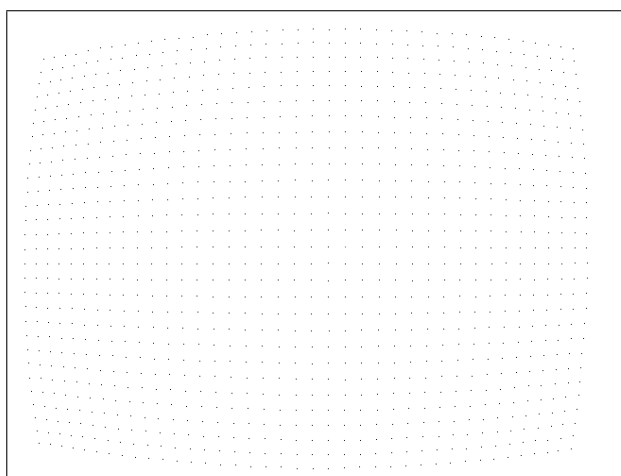
- obraz prostokątnej siatki wykonany przy dokładnie prostopadłym ustawieniu osi kamery względem płaszczyzny siatki,
- współrzędne obrazowe punktu rzutowania osi optycznej, wyznaczone wcześniej dla podanego obrazu. Punkt ten **musi** znajdować się na przecięciu linii,
- odległość między poziomymi i pionowymi liniami siatki oraz odległość między kamerą i siatką wyrażone w tych samych jednostkach.

Obraz musi być wykonany przy dobrej ostrości. Dla uzyskania jak najlepszej dokładności korekcji, w obrazie powinno być jak najwięcej całych elementów siatki. Oznacza to, że siatka powinna znajdować się w najdalszej odległości od kamery, przy której elementy siatki są dobrze różniane w procesie automatycznej detekcji i wypełniają cały obraz.

Podczas detekcji linii poszukuje się linii złożonych z najciemniejszych punktów. Linie te muszą spełniać następujące warunki: być ciągle, mieć prawie niezmienny kierunek w małym obszarze, leżeć od siebie w odległości nie bliższej niż określona. Okno z ustawieniami programu przed wykryciem linii przedstawia rysunek 4.13. Efektem działania tego algorytmu jest obraz siatki, którego przykład znajduje się na rysunku 4.7. Oprócz obrazu generowana jest tablica współrzędnych przecięć linii, co umożliwia wyznaczenie rozmiarów obrazu wynikowego oraz współczynników transformat dwuliniowych do korekcji. Ponieważ poszukiwanie punktów w obrazie wejściowym przez analizę linii jest czasochłonne, generowany jest obraz z mapą punktów w których przecinają się linie. Ten obraz można załadować przy uruchomieniu programu, a znalezienie wg niego współrzędnych punktów przecięć jest bardzo szybkie. Przykład takiej siatki przedstawia rysunek 4.14. Informacja o punkcie w który jest skierowana oś optyczna jest zachowana przez kolor. Wszystkie wierzchołki są czarne, tylko ten na osi optycznej jest czerwony. Zaletami tego podejścia są: większa szybkość analizy obrazu w porównaniu z obrazem źródłowym, przechowanie informacji o położeniu osi optycznej oraz bardzo duży stopień kompresji obrazu metodą bezstratną. Przykładowy obraz w formacie PNG o rozdzielczości 640x480 pikseli zajmuje niecałe 2 KB. Algorytm analizujący siatkę potrafi wydostać tę informację z obrazu. Metodę automatycznej kalibracji opisano też w [46].



Rysunek 4.13: Konfiguracja wykrywania linii siatki w programie



Rysunek 4.14: Siatka wierzchołków przechowująca informację o parametrach kamery

Rozdział 5

Detekcja i reprezentacja znaczników w literaturze

5.1 Momenty

Momenty stosuje się w wielu aspektach obróbki obrazów, od rozpoznawania wzorców do kodowania obrazu. Zastosowane do obrazu opisują jego zawartość z uwzględnieniem osi współrzędnych. Są tak zaprojektowane, żeby uwzględniały zarówno globalne właściwości obrazu, jak i geometryczne szczegóły. Przy założeniu ciągłości obraz może być traktowany jako dwuwymiarowa dystrybucja gęstości funkcji ciągłej dwu zmiennych $f(x, y)$. Moment funkcji dwuwymiarowej ciągłej definiuje się ogólnie jako:

$$M_{jk} = \int \int_{\zeta} \psi_{jk}(x, y) f(x, y) dx dy; j, k = 0, 1, 2, \dots + \infty \quad (5.1)$$

Bazową funkcją jest ψ_{jk} . Moment jest ważonym opisem $f(x, y)$ na płaszczyźnie ζ . Dyskretna postać momentu wyraża się wzorem:

$$M_{jk} = \sum_x \sum_y \psi_{jk}(x, y) P_{xy}, \quad (5.2)$$

gdzie $P(x, y)$ oznacza wartość jasności piksela o współrzędnych (x, y) .

5.1.1 Momenty geometryczne (kartezjańskie)

Stosowane do identyfikacji obiektów momenty geometryczne są wartościami wyznaczanymi dla funkcji dwu zmiennych. Dla funkcji ciągłej dwóch zmiennych moment geometryczny rzędu j, k definiuje się wg wzoru (5.3):

$$M_{j,k} = \int_{x=-\infty}^{+\infty} \int_{y=-\infty}^{+\infty} f(x, y) x^j y^k dx dy. \quad (5.3)$$

Momenty funkcji dyskretniej dwu zmiennych wyznacza się według wzoru (5.4):

$$M_{j,k} = \sum_{x=0}^X \sum_{y=0}^Y f(x, y) x^j y^k. \quad (5.4)$$

W praktyce momenty wyznacza się dla obszarów o skończonych rozmiarach, jakimi są obrazy. Dodatkowo, zazwyczaj przed identyfikacją wyodrębnia się obszar który należy identyfikować, w którym wartość funkcji jasności $f(x, y)$ jest stała. Oznaczając rozmiary obszaru przez X i Y , otrzymuje się:

$$M_{j,k} = \sum_{x=0}^X \sum_{y=0}^Y x^j y^k. \quad (5.5)$$

Bardziej praktyczne zastosowanie znajdują dyskretne momenty centralne, wyznaczone wg wzoru (5.6) i znormalizowane momenty centralne, wyznaczone wg wzoru (5.7):

$$\mu_{j,k} = \sum_{x=0}^X \sum_{y=0}^Y (x - \bar{x})^j (y - \bar{y})^k, \quad (5.6)$$

$$\eta_{j,k} = \frac{\mu_{j,k}}{(XY \mu_{0,0})^{\frac{j+k}{2}}}. \quad (5.7)$$

Symbole \bar{x} i \bar{y} oznaczają współrzędne środka ciężkości. Wyznacza się je na podstawie momentów: $\bar{x} = \frac{M_{1,0}}{M_{0,0}}$, $\bar{y} = \frac{M_{0,1}}{M_{0,0}}$.

Wartości momentów wyznaczone wg wzoru (5.5) do rzędu N , gdzie $N = j + k$, wystarczą do wyznaczenia wszystkich momentów centralnych i znormalizowanych momentów centralnych do rzędu N włącznie. Nie jest konieczne wyznaczanie momentów centralnych za pomocą wzoru (5.6). Przedstawiają to wzory (5.8) i (5.9).

$$\mu_{0,1} = \sum_{x=0}^X \sum_{y=0}^Y (x - \bar{x})^0 (y - \bar{y}) = \sum_{x=0}^X \sum_{y=0}^Y \left(y - \frac{\sum_{x=0}^X \sum_{y=0}^Y y}{\sum_{x=0}^X \sum_{y=0}^Y 1} \right) = \sum_{x=0}^X \sum_{y=0}^Y y - \sum_{x=0}^X \sum_{y=0}^Y y = 0. \quad (5.8)$$

$$\begin{aligned} \mu_{2,0} &= \sum_{x=0}^X \sum_{x=0}^Y (x - \bar{x})^2 = \sum_{x=0}^X \sum_{x=0}^Y (x^2 - 2x\bar{x} + \bar{x}^2) = \\ &= \sum_{x=0}^X \sum_{x=0}^Y x^2 - 2 \sum_{x=0}^X \sum_{x=0}^Y x\bar{x} + \sum_{x=0}^X \sum_{x=0}^Y \bar{x}^2 = M_{2,0} - 2M_{1,0}\bar{x} + M_{0,0}\bar{x}^2 \end{aligned} \quad (5.9)$$

Li w pracy doktorskiej [58] opisuje wnioskowanie o kształcie i symetrii obiektu na podstawie momentów znormalizowanych centralnych niskich rzędów ($J + k \leq 3$). Wartości momentów $\mu_{0,1}$ i $\mu_{1,0}$ są zawsze równe 0. Moment $\eta_{1,1}$ ma wartość 0, jeżeli kształt ma poziomą lub pionową oś symetrii. Momenty trzeciego rzędu określają niesymetryczność (*ang. skewness*) obiektu. Dla obiektu symetrycznego względem osi x , $\eta_{3,0} = 0$ i $\eta_{1,2} = 0$. Z kolei dla kształtu symetrycznego względem osi y , $\eta_{0,3} = 0$ i $\eta_{2,1} < 0$.

Momenty oparte na bazie $x^p y^q$ nie są ortogonalne, ponieważ baza nie jest ortogonalna. Powiązanie rośnie wraz ze wzrostem rzędu momentu, co powoduje że momenty wyższych rzędów opisują bardzo silnie skorelowane opisy. Efektem tego jest duże znaczenie małych zmian wartości momentów, przez co wymagana jest bardzo duża precyzja obliczeń. Stąd momenty te są bardzo podatne na błędy.

Istnieją jednakże też momenty ortogonalne, uzyskane z ortogonalnego zbioru bazowego. Ich przewagą nad momentami centralnymi jest wymagana mniejsza dokładność wartości dla równie dokładnego opisu kształtu. Do tego typu momentów zalicza się momenty Legendre'a i Zernike'a.

5.1.2 Niezmienniki

Momenty centralne znormalizowane pozwalają na identyfikację obiektów według kształtów niezależnie od położenia i skali. Nie są jednak odporne na zmiany orientacji, tzn wartości te zmieniają się jeżeli obiekt zostanie obrócony. Są one podstawą do wyznaczenia wielu niezmienników przez różnych autorów.

Niezmienniki Hu

Najstarsze znane niezmienniki zostały opisane przez Hu w [38]. Ich definicje są następujące:

$$H1 = \eta_{2,0} + \eta_{0,2}, \quad (5.10)$$

$$H2 = (\eta_{2,0} - \eta_{0,2})^2 + 4 * (\eta_{1,1})^2, \quad (5.11)$$

$$H3 = (\eta_{3,0} - 3\eta_{1,2})^2 + (3\eta_{2,1} - \eta_{0,3})^2, \quad (5.12)$$

$$H4 = (\eta_{3,0} + \eta_{1,2})^2 + (\eta_{2,1} + \eta_{0,3})^2, \quad (5.13)$$

$$H5 = (\eta_{3,0} - 3\eta_{1,2})(\eta_{3,0} + \eta_{1,2})[(\eta_{3,0} + \eta_{1,2})^2 - 3(\eta_{2,1} + \eta_{0,3})^2] + \\ + (3\eta_{2,1} - \eta_{0,3})(\eta_{2,1} + \eta_{0,3})[3(\eta_{3,0} + \eta_{1,2})^2 - (\eta_{2,1} + \eta_{0,3})^2], \quad (5.14)$$

$$H6 = (\eta_{2,0} + \eta_{0,2})[(\eta_{3,0} + \eta_{1,2})^2 - 3(\eta_{2,1} + \eta_{0,3})^2] + 4\eta_{1,1}(\eta_{3,0} + \eta_{1,2})(\eta_{2,1} + \eta_{0,3}), \quad (5.15)$$

$$H7 = (3\eta_{2,1} - \eta_{0,3})(\eta_{2,1} + \eta_{0,3})[(\eta_{3,0} + \eta_{1,2})^2 - 3(\eta_{2,1} + \eta_{0,3})^2] + \\ - \eta_{3,0} - 3\eta_{1,2})(\eta_{3,0} + \eta_{1,2})[3(\eta_{3,0} + \eta_{1,2})^2 - (\eta_{2,1} + \eta_{0,3})^2]. \quad (5.16)$$

Niezmienniki te są niewrażliwe na obrót obiektu. Pozwalają zidentyfikować obiekt niezależnie od jego położenia w obrazie, skali oraz rotacji. Niezmienniki te nie stanowią kompletnego opisu obiektu, jednak można też wyprowadzić niezmienniki wyższych rzędów.

5.1.3 Niezmienniki przekształcenia afinicznego

Niezmienniki przekształcenia afinicznego (*Affine Moment Invariants*, AMI) zostały opracowane w roku 1994 [33] na bazie momentów geometrycznych. Kontynuacja tej tematyki znajduje się w kolejnych pracach J. Flussera, m.in. [35], [36], [34], [94].

Niezmienniki AMI zostały wyprowadzone z teorii algebraicznych niezmienników. Pełne wyprowadzenie znajduje się w [32]. AMI są niezmiennicze przy *ogólnej transformacji afinicznej*:

$$u = a_0 + a_1x + a_2y, \quad (5.17)$$

$$v = b_0 + b_1x + b_2y, \quad (5.18)$$

gdzie (x, y) oraz (u, v) oznaczają współrzędne punktu odpowiednio przed i po transformacji.

Pierwsze niezmienniki definiuje się wzorami:

$$I_1 = \frac{1}{\mu_{00}^4}(\mu_{20}\mu_{02} - \mu_{11}^2), \quad (5.19)$$

$$I_2 = \frac{1}{\mu_{00}^{10}}(\mu_{30}^2\mu_{03}^2 - 6\mu_{30}\mu_{21}\mu_{12}\mu_{03} + 4\mu_{30}\mu_{12}^3 + 4\mu_{03}\mu_{21}^3 - 3\mu_{21}^2\mu_{12}^2), \quad (5.20)$$

$$I_3 = \frac{1}{\mu_{00}}(\mu_{20}(\mu_{21}\mu_{03} - \mu_{12}^2) - \mu_{11}(\mu_{30}\mu_{03} - \mu_{21}\mu_{12}) + \mu_{02}(\mu_{30}\mu_{12} - \mu_{21}^2)), \quad (5.21)$$

$$I_4 = \frac{1}{\mu_{00}^3} (\mu_{20}^3 \mu_{03}^2 - 6\mu_{20}^2 \mu_{11} \mu_{12} \mu_{03} - 6\mu_{20}^2 \mu_{02} \mu_{21} \mu_{03} + 9\mu_{20}^2 \mu_{02} \mu_{12}^2 + 12\mu_{20} \mu_{11}^2 \mu_{21} \mu_{03} + 6\mu_{20} \mu_{11} \mu_{02} \mu_{30} \mu_{03} - 18\mu_{20} \mu_{11} \mu_{02} \mu_{21} \mu_{12} - 8\mu_{11}^3 \mu_{30} \mu_{03} - 6\mu_{20} \mu_{02}^2 \mu_{30} \mu_{12} + 9\mu_{20} \mu_{02}^2 \mu_{21}^2 + 12\mu_{11}^2 \mu_{02} \mu_{30} \mu_{12} - 6\mu_{11} \mu_{02}^2 \mu_{30} \mu_{21} + \mu_{02}^3 \mu_{30}^2), \quad (5.22)$$

$$I_5 = \frac{1}{\mu_{00}^6} (\mu_{40} \mu_{04} - 4\mu_{31} \mu_{13} + 3\mu_{22}^3), \quad (5.23)$$

$$I_6 = \frac{1}{\mu_{00}^9} (\mu_{40} \mu_{04} \mu_{22} + 2\mu_{31} \mu_{22} \mu_{13} - \mu_{40} \mu_{13}^2 - \mu_{04} \mu_{31}^2 - \mu_{22}^3). \quad (5.24)$$

Warto zauważyć, że niezmienniki wyznacza się z momentów centralnych nieznormalizowanych. J. Flusser i jego współpracownicy udowadniają, że niezmienniki te można z powodzeniem stosować do rozpoznawania liter pisma ręcznego. Bardziej zaawansowane niezmienniki, opisane w [35], [36], [34], [94], pozwalają na rozpoznanie obiektów w obrazach zaszumionych, rozmytych działaniem filtru Gaussa oraz nieostrych z powodu np. ruchu kamery podczas akwizycji zdjęcia. Artykuły te definiują momenty zespolone, analogiczne do rzeczywistych. Dla funkcji ciągłej dwu zmiennych $f(x, y)$ momenty zespolone zwykle wyznacza się wg wzoru:

$$c_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x + iy)^p (x - iy)^q f(x, y) dx dy \quad (5.25)$$

Artykuł [36] zawiera dowód twierdzenia, że zespolone centralne znormalizowane momenty obrazu są odporne na obrót, przesunięcie i zniekształcenia liniowe (spowodowane ruchem, *ang. blur*). Artykuł [34] zawiera dowód niewrażliwości tych momentów na obrót, zniekształcenia splotem z funkcją centrosymetryczną oraz metodę uniezależnienia momentów od zmian kontrastu przez normalizację.

5.2 Cechy niezmiennicze względem skali (SIFT)

Cechy niezmiennicze względem skali (*Scale Invariant Feature Transform*, **SIFT**) pozwalają określić punkty stałe otoczenia nie zmieniające się przy zmianie skali, odległości i kąta patrzenia [62], [63]. Metoda została opracowana w celu identyfikacji elementów lokalizujących niezależnie od przekształcenia skalującego, obrotu, odpornego na szum i zniekształcenia obrazu.

W metodzie stosuje się przetwarzanie etapowe. W pierwszym etapie wyznacza się położenie kluczy w przestrzeni skali – czyli określa się skalę kluczy. Odporność na zmianę skali osiągnięto przez rozmycie obrazu i wyznaczanie parametrów dla różnej skali. Wyodrębnianie obiektów z obrazu opiera się na wyznaczeniu maksimum i minimum funkcji Gaussa w przestrzeni obrazu. W celu uniezależnienia się od skali obrazu stosuje się piramidę obrazów, w której kolejne poziomy powstają przez przeskalowanie poprzedniego ze współczynnikiem 1,5. Dla zachowania wysokich częstotliwości pierwszy obraz jest powiększony ze współczynnikiem 2. Następnie obraz dolny (A) mnożony jest przez funkcję Gaussa z $\sigma = 2$, a górny (B) przez funkcję Gaussa ze współczynnikiem $\sigma = \sqrt{2}$. Maksima i minima różnicy obrazów $A - B$ odpowiadają kluczom.

Dla każdego piksla obrazu A wyznacza się wartość i orientację gradientu. Uzyskane wartości są progowane wartością 0,1 maksymalnej wartości gradientu. Dla większej niezależności gradientu od zmian oświetlenia kierunek gradientu wyznacza się również na podstawie otoczenia z lokalnego histogramu gradientów. Histogramy tworzy się przy użyciu okna Gaussa z σ 3 razy większym niż obrazu A . Każdemu kluczowi przypisuje się kanoniczną orientację, co umożliwia

wykrywanie obrotu obrazu. Do uzyskania niezależności względem skali oraz orientacji kluczy każdy obiekt jest przechowywany jako 160 próbek obrazu.

Metoda wykrywa 70% do 90% powtarzalnych kluczy. Obrazy wejściowe były poddawane zmianom kontrastu, obrotowi o 20° , skalowaniu, rozciąganiu, ściskaniu, dodaniu 10% szumu oraz wszystkim wymienionym transformacjom.

Obiekty wykrywane tą metodą są bardzo małe, stanowią one fragmenty większych obiektów. Przedmioty należy opisywać jako przestrzennie uporządkowany zestaw obiektów, co przedstawia rysunek 5.1. Widoczne kwadraty oznaczają wielkości znalezionych i zidentyfikowanych obiektów, zaś kąt nachylenia reprezentuje orientację obiektu względem orientacji bazowej. Zaznaczono też kontury zidentyfikowanych obiektów. Opis obiektu za pomocą cech SIFT zastosowano do lokalizacji w połączeniu ze stereowizją [79], oraz zrealizowano SLAM stosując tylko te cechy [78], [76]. W ostatnich latach podejmuje się próby zastosowania tej techniki dla kamery dookólnej.



Rysunek 5.1: Przykład detekcji fragmentów przedmiotów metodą SIFT oraz identyfikacji przedmiotów

5.3 Detekcja znaczników

Wykrywanie znaczników w obrazie jest możliwe w obszarach o dużym gradiencie lub charakterystycznych cechach takich jak jasność, barwa lub sąsiedztwo barw. Bardziej zaawansowane metody wyznaczają cechy niezależne od skali czy jasności średniej obrazu. Do cech umożliwiających wyznaczanie znaczników należą:

- barwa segmentu (zob też rozdział 7.1),
- kształt opisywany różnymi sposobami (zob też rozdz 7.2),
- gęstość i ułożenie krawędzi,
- tekstura,
- inne elementy - np **SIFT** - Scale Invariant Feature Transform (zob 5.2).

Przy lokalizacji znacznikowej możliwe jest albo selektywne rozpoznawanie pojedynczych znaczników, których jest w otoczeniu względnie mało, albo rozpoznawanie wielu obiektów przy identyfikacji wspartej metodami statystycznymi, np. filtrami Bayesa (rozdz. 2.4). Do pierwszego typu znaczników należą np. znaczniki identyfikowane przez kształt (opisane w rozdziale 5.3.2), także stosowane w proponowanej metodzie, opisane w rozdziale 7.2. Do znaczników występujących bardziej "masowo" należą znaczniki identyfikowane przez krawędzie, barwę (rozdział 5.3.1) i inne cechy. Możliwe jest też stosowanie mieszanych metod detekcji i identyfikacji znaczników.

5.3.1 Detekcja na podstawie barwy

Barwę znacznika można definiować jako pojedynczy obszar jednobarwny lub jako zbiór obszarów barwnych o określonych barwach i konfiguracji. W przypadku pojedynczej barwy detekcja sprowadza się do wykrycia pikseli o barwie mieszczącej się w określonym zakresie, stosując model HSV lub $r'g'$. W przypadku wielu obiektów barwnych można stosować np. szachownicę barwną lub dwa obszary barwne sąsiadujące ze sobą.

W wielu pracach, np. [54], [56] stosuje się "znormalizowany" model RGB. Modyfikacje powodują zmianę własności modelu, pozwalając na oddzielne określenie barwy i jasności. W pracy [54] przeprowadza się normalizację barwy, przy czym kolor jest opisywany dwoma parametrami:

$$r' = \frac{R}{R + G + B} \quad (5.26)$$

$$g' = \frac{G}{R + G + B} \quad (5.27)$$

gdzie R, G i B oznaczają składowe czerwoną, zieloną i niebieską, zaś r' i g' barwy znormalizowane. W rzeczywistości barwa znormalizowana posiada składowe odpowiadające odcieniowi i nasyceniu w modelu HSV/HSL, zaś informacja o jasności jest tracona bezpowrotnie. Taki model umożliwia wykrywanie np. koloru skóry. Wadą tego systemu jest konieczność ponownej kalibracji systemu po zmianie oświetlenia. Po normalizacji koloru traci się informację o jasności. Podejście takie pozwala wykryć dany kolor niezależnie od jego jasności. Ta sama informacja, jaka pozostaje w składowych r oraz g znormalizowanej barwy, jest tożsama z nasyceniem i odcieniem barw w modelach HSV czy HSL.

Rozpoznawanie na podstawie barwy stosuje się szeroko w grach robotów w piłkę nożną z podwieszoną kamerą. W tych systemach, w których ważna jest szybkość obróbki, obiekty (gracze, piłka, elementy boiska) identyfikuje się wyłącznie na podstawie barwy. W typowych przypadkach obiekty poruszające się mają nasycone barwy kontrastujące ze sobą, boisko (tło) jest zielone, zaś identyfikowane granice boiska są białe.

Pola znacznika w postaci barwnej szachownicy ([91]) mają charakterystyczne, nasycone barwy. Kolory wykrywa się w przestrzeni znormalizowanych barw bez jasności $r'g'$, opisanej w rozdziale 3.1.5. Potencjalny znacznik wykrywa się sporządzając histogramy barw lokalny i globalny. Po wykryciu znacznika aktualizuje się wzorce barw według ostatnio wykrytych, w celu odporności algorytmu na zmiany oświetlenia.

5.3.2 Detekcja na podstawie kształtu

Kształt obiektu opisuje się na podstawie jego konturu lub wyznaczonego obszaru. Dla konturu znanych jest wiele metod identyfikacji obiektu, w tym odległości brzegu od środka [9], ogólną transformatę Hougha dla konturów [57], [49].

Schemat detekcji kształtów za pomocą transformaty Hougha opisano w [30], [29], a skrócony opis zamieszczono w rozdziale 3.5. W pracach [40], [70] opisano opracowane metody detekcji obiektów za pomocą tej transformaty. Praca [70] podaje sposób rozróżnienia obiektów w transformacie.

Dla obszaru można stosować metody korelacyjne lub oparte o momenty i niezmienniki. Metody korelacyjne są jednymi z najstarszych metod. W pracach [65], [60], [15], [4] opisano wykrywanie w ten sposób małych znaczników. Dopasowanie dużych obiektów do wzorca w ten sposób jest trudniejsze, jednak i ten problem rozwiązano przechowując dane o otoczeniu widocznym z różnych miejsc [3]. Lokalizacja odbywa się przez porównanie wyglądu otoczenia z zapamiętanym wcześniej widokiem otoczenia, pomijając wyznaczanie konkretnych znaczników. Podczas nawigacji wyznacza się korelację obrazu z kamery z dwoma wzorcowymi obrazami: uzyskanym z punktu najbliższego estymowanemu położeniu oraz z następnego, w pobliżu którego robot powinien się znaleźć, jeżeli będzie poprawnie realizował wytyczoną ścieżkę.

Detekcja znaczników na podstawie momentów i niezmienników jest trudna ze względu na zniekształcenia perspektywiczne. Jan Flusser [35] [36] [34] [94] przedstawił metodę rozpoznawania znaczników za pomocą specjalnie wyznaczonych niezmienników oraz momentów zespolonych. Jest to metoda dobra do rozpoznawania liter lub rozróżniania obiektów o kształtach różniących się znacznie.

Znacznik zbudowany z dwu sąsiadujących obiektów ([52]) umieszcza się na tle dwu białych kartek. Całość jest rozmieszczona według określonej geometrii. Detekcja następuje na podstawie momentów Zernike'a połączonego z topologicznym rozkładem podobnych obiektów w otoczeniu.

Rozdział 6

Charakterystyka proponowanej metody lokalizacji

6.1 Wprowadzenie

Proponowana metoda w całości opiera się na analizie obrazu. Wyznaczenie położenia odbywa się według schematu przedstawionego na rysunku 6.1. Obraz wejściowy uzyskiwany podczas akwizycji może pochodzić z pojedynczej kamery. Akceptowane są obrazy czarno-białe, w odcieniach szarości oraz barwne.

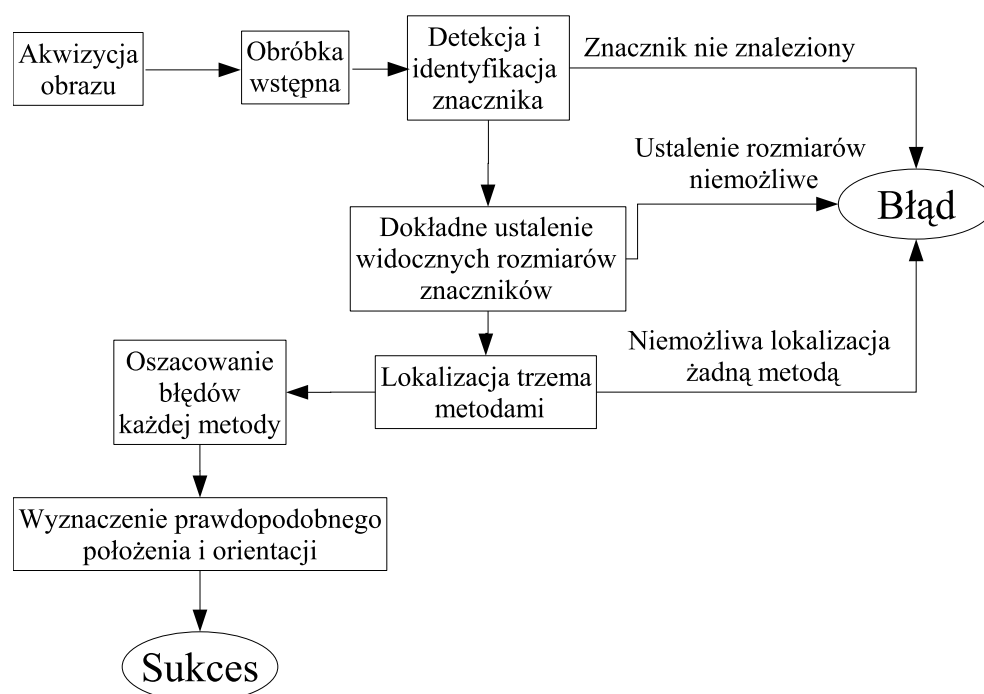
Obraz uzyskany podczas akwizycji powinien być ostry i nie rozmyty. Ponieważ w większości kamer stosuje się przeplot, a uzyskany obraz składa się z 2 lub 6 obrazów uzyskanych w różnym czasie, konieczne jest zatrzymanie robota na czas akwizycji obrazu. Jeżeli robot nie może się zatrzymać, jego prędkość powinna być na tyle mała, żeby znacznik nie był rozmyty. Po akwizycji obrazu robot może kontynuować jazdę. Nie jest konieczne zatrzymanie robota, jeżeli kamera pozwala uzyskać obraz ostry i nie rozmyty podczas ruchu.

Korekcja obrazu usuwa zniekształcenia nieliniowe obrazu i wymaga uprzedniej kalibracji kamery. Stosowaną procedurę kalibracji oraz korekcji obrazu opisano w rozdziale 4. Podczas kalibracji uzyskuje się obraz wzorca przy roboczych ustawieniach kamery. Do korekcji obrazów podczas lokalizacji wystarczy jeden obraz wzorcowy. Na podstawie obrazu wzorcowego uzyskuje się mapę zniekształceń, według której koryguje się wszystkie kolejne uzyskane obrazy.

Znacznikami, do których można stosować przedstawioną w niniejszej pracy metodę lokalizacji, są obiekty płaskie o charakterystycznych barwie i kształcie. Barwa pozwala na wstępną selekcję potencjalnych znaczników. Na podstawie kształtu i barwy obiekt może być zidentyfikowany jako znacznik. Do lokalizacji wykorzystywane są rozmiary i położenie prostokąta opisanego na znaczniku, dlatego najbardziej odpowiednie są znaczniki o obrysie prostokątnym. Do wyznaczenia położenia robota konieczny jest jeden w całości widoczny znacznik. Ze względu na ograniczoną dokładność pomiaru oraz możliwość zasłonięcia znacznika wskazane jest przygotowanie większej ilości znaczników, o różnych barwach i kształtach.

Ekstrakcja obiektów odbywa się przez wyodrębnienie obszarów o charakterystycznej barwie, a następnie identyfikację obiektów o charakterystycznych kształtach. Stosowaną metodę detekcji barwy i kształtu obiektów opisano w rozdziale 7. Mechanizm ekstrakcji zależy od tego, czy obraz jest barwny czy w odcieniach szarości. Dla obrazu w odcieniach szarości jest to progowanie z jednym lub dwoma progami. Mechanizm wyodrębniania obszarów z obrazu barwnego przez filtrowanie barw według odcienia, nasycenia i jasności opisano w rozdziale 7.1. Po wyodrębnieniu obszarów barwnych wyznacza się ich cechy charakterystyczne – wartości

momentów centralnych znormalizowanych (opisanych w rozdziale 5.1.1) oraz niektórych niektórych niezmienników H_u i AMI (rozdział 5.1.2). Metodę identyfikacji kształtów oraz wyniki eksperymentów zamieszczono w rozdziale 7.2. Obiekt, który ma poszukiwaną barwę oraz pasuje do poszukiwanego kształtu zostaje uznany za znacznik. Po identyfikacji znacznika wyznacza się jego dokładne wymiary obrazowe na podstawie wyznaczenia prostokąta opisanego na znaczniku. Prostokąt wykrywa się za pomocą detektora krawędzi Marra–Hildretha (rozdział 3.3.2), a dokładne współrzędne narożników za pomocą transformaty Hougha i regresji liniowej (rozdział 7.3.3). Wyznaczanie prostokąta opisanego na znaczniku opisano w całości w rozdziale 7.3. Proponowana metoda pozwala na prawidłowe wyznaczenie narożników prostokąta jeżeli znacznik jest częściowo zasłonięty lub nie wpisuje się dokładnie w prostokąt, w szczególności jeżeli narożniki znacznika nie leżą w narożnikach prostokąta. Procedurę i przykład wyznaczenia położenia zasłoniętego narożnika zamieszczono w rozdziałach 7.3.2 ÷ 7.3.4. Przykładowe stosowane kształty znaczników przedstawia rysunek 7.27 na stronie 96.

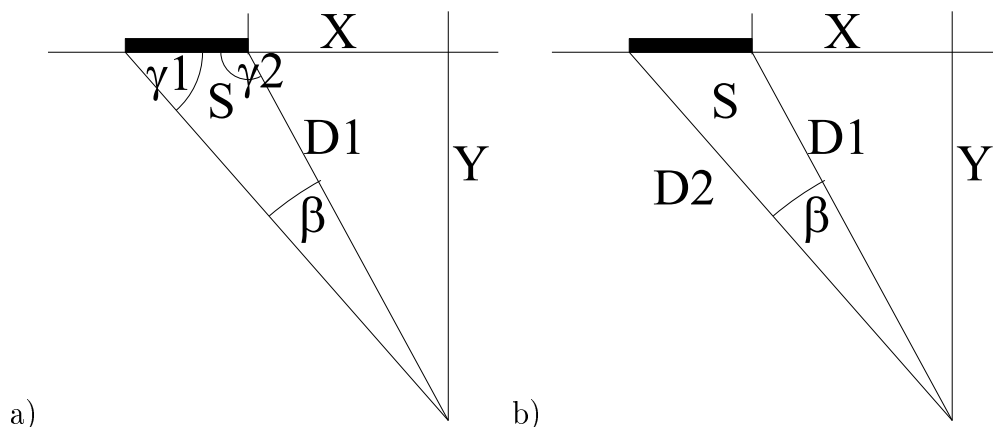


Rysunek 6.1: Schemat pojedynczej operacji wyznaczania położenia robota

Po wyznaczeniu rozmiarów znacznika następuje obliczenie położenia i orientacji kamery (i robota na którym jest zamontowana) względem znacznika. Pierwszymi możliwymi do wyznaczenia wartościami są azymuty do pionowych krawędzi (odległości kątowe między krawędziami a osią optyczną kamery), co opisano w rozdziale 8.1. Znajomość azymutu do krawędzi pozwala wyznaczyć szerokość kątową znacznika β , co opisano w rozdziale 8.3. W połączeniu ze znajomością rzeczywistej wysokości krawędzi znacznika oraz wysokości w obrazie liczonej jako długość krawędzi prostokąta pozwala na wyznaczenie odległości do krawędzi znacznika, co opisano szczegółowo w rozdziale 8.2. Schematycznie lokalizację tą metodą przedstawia rysunek 6.2a. Wartością znaną jest rzeczywista szerokość znacznika S . Znajomość odległości do jednej z krawędzi D_1 , odległości między krawędziami S oraz kąta β pozwala wyznaczyć pozostałe kąty na podstawie twierdzenia sinusów i twierdzenia cosinusów. Te wartości pozwalają wyznaczyć współrzędne kamery X i Y oraz jej orientację względem znacznika α . Szczegółowo tę metodę wyznaczania położenia opisano w rozdziale 8.4. Ponieważ do wyznaczenia położenia wystar-

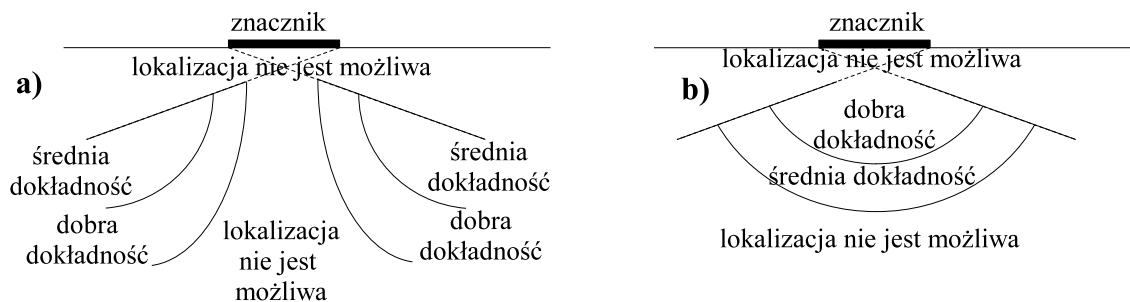
czy odległość do jednej z pionowych krawędzi znacznika, metodę tę można traktować jak dwie metody, różniące się tylko krawędzią do której mierzy się odległość.

Do wyznaczenia położenia i orientacji robota względem znacznika wystarczy znajomość 3 długości boków trójkąta. Są to odległości obu krawędzi znacznika od robota oraz odległość między krawędziami znacznika. Poglądowo przedstawiono tę sytuację na rysunku 6.2b. Metodę lokalizacji na podstawie samych odległości przedstawiono w rozdziale 8.5.

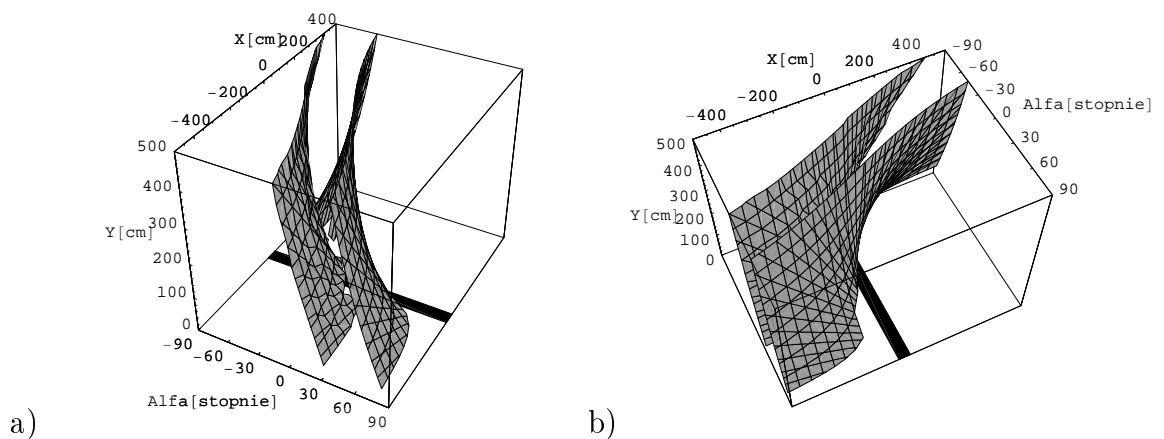


Rysunek 6.2: Schemat wyznaczenia położenia robota: a) na podstawie kąta widzenia znacznika oraz odległości do jednej krawędzi znacznika; b) na podstawie odległości do obu krawędzi znacznika

Obie metody wzajemnie się uzupełniają, pozwalając na wyznaczenie położenia w zwartym obszarze, w którym znacznik jest dobrze widoczny. Poglądowo strefy dokładności obu metod pokazano na rysunku 6.3. W rozdziałach 8.4.1 oraz 8.5.1 przedstawiono metodę oszacowania błędu pomiaru. Błąd ten można wyznaczyć w czasie rzeczywistym, na podstawie wyznaczonych rozmiarów obrazowych znacznika oraz wyznaczonego położenia i orientacji robota.



Rysunek 6.3: Schemat zakresów dokładności dla metod lokalizacji: a) z jedną odległością (rozdział 8.4) b) z dwoma odległościami (rozdział 8.5)



Rysunek 6.4: Kształt strefy w której jest możliwa lokalizacja, tzn. w której znacznik jest widoczny w całości

6.2 Zakres lokalizacji

W rozdziałach 8.4 i 8.5 zamieszczono wykresy obrazujące błąd lokalizacji przy różnych zakłóceniach oraz strefy określonej dokładności lokalizacji. Dla każdej z metod zamieszczono oddzielne wykresy dla współrzędnych X i Y oraz orientacji robota oznaczonej α . Wszystkie wykresy zostały sporządzone dla tych samych parametrów kamery i znacznika. Parametry kamery (wyznaczone dla jednej z używanych kamer) wynoszą:

- Odległość między kamerą a siatką kalibrującą $Dist_{cm} = 16,0\text{cm}$,
- Rozmiar poziomy obrazu po kalibracji $DX_{cm} = 19,0\text{cm}$, $DX_p = 722\text{px}$,
- Rozmiar pionowy obrazu po kalibracji $DY_{cm} = 28,0\text{cm}$, $DY_p = 532\text{px}$,
- Położenie osi optycznej w obrazie $DLX_p = 418$.

Znaczenie i wyznaczanie tych parametrów omówiono przy omawianiu kalibracji kamery w rozdziale 4. Przyjęte rozmiary znacznika to 70 cm wysokości i 50 cm szerokości.

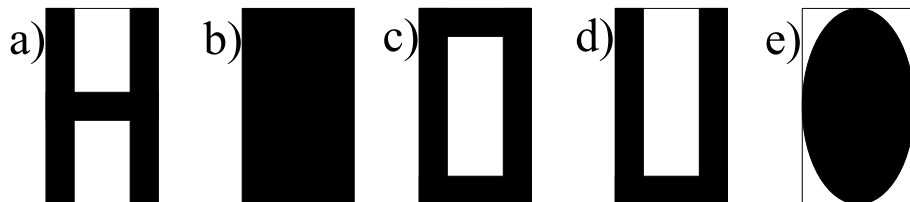
Wykresy przedstawiające błąd są sporządzone dla kąta patrzenia wynoszącego 0° (choć nie dla każdych współrzędnych lokalizacja jest możliwa). Wykresy przedstawiające płaszczyzny o określonej dokładności lokalizacji są sporządzone podobnie jak rysunek 6.4, w rzucie rysunku 6.4.a. Przedstawia on zakres współrzędnych w których jest możliwa lokalizacja, tzn. znacznik o rozmiarach 50x70 cm jest widoczny w całości. Strefa ta znajduje się między dwoma płaszczyznami. Dla współrzędnej $Y \leq 100$ granice tej strefy są nierówne, powyżej tej odległości można w uproszczeniu przyjąć, że granice są płaszczyznami równoległymi. Widoczna asymetria między stroną lewą i prawą znacznika wynika z niesymetrycznego położenia znacznika względem współrzędnej X oraz niesymetrycznego położenia osi optycznej kamery w obrazie. Położenie znacznika oznaczono czarnym prostokątem o współrzędnych $Y = 0$, $-50 \leq X \leq 0$, widocznym jako czarny pas. Wykresy te czytać można w następujący sposób: dla określonej orientacji α wyznacza się przekrój przez wykres i otrzymuje się granice strefy na płaszczyźnie $(X, Y) \in \mathbb{R}^2$. Wszystkie wykresy dokładności sporządzono dla zakresów $-500\text{ cm} \leq X \leq 500\text{ cm}$, $0\text{ cm} \leq Y \leq 500\text{ cm}$, $-90^\circ \leq \alpha \leq 90^\circ$.

6.3 Stosowane znaczniki

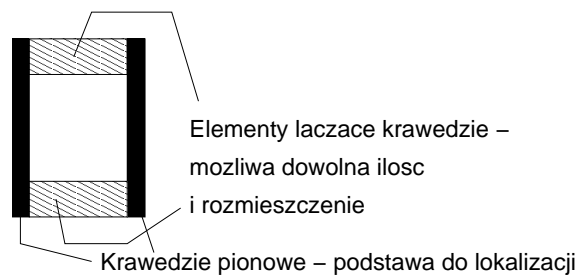
Aby pojedynczy znacznik wystarczył do lokalizacji, musi zawierać odpowiednią ilość informacji. W tradycyjnych metodach stosuje się wiele znaczników (np w triangulacji) lub wiele sensorów (np w stereowizji). Aby możliwa była lokalizacja na podstawie jednego znacznika przy pomocy jednego sensora, znacznik musi zawierać co najmniej tyle informacji, ile trzy znaczniki w triangulacji. Proponowana metoda lokalizacji opiera się na następujących założeniach:

- znaczniki są płaskie,
- każdy znacznik posiada dwie zewnętrzne pionowe krawędzie możliwe do identyfikacji i zmierzenia w obrazie,
- znane są dokładne rozmiary znacznika.

Metoda lokalizacji opisana w rozdziałach 8.4 wymaga, żeby w znaczniku można było wyznaczyć wysokość "w linii prostej", czyli żeby najwyższy punkt znajdował się nad najniższym. Wymagania te spełniają różne kształty, a sam znacznik nie musi być spójnym obszarem. Metoda ta wymaga, żeby można było określić widoczną szerokość znacznika. Warunki te spełniają wszystkie znaczniki z rysunku 6.5. Metoda lokalizacji opisana w rozdziale 8.5 wymaga, żeby w znaczniku dało się znaleźć dwie pionowe krawędzie oddalone od siebie oraz odległość między tymi krawędziami. Wysokości krawędzi nie muszą być jednakowe. Z kształtów zamieszczonych na rysunku 6.5 jedynie kształt z rysunku 6.5.e nie spełnia tego warunku. W trakcie eksperymentów stosowano znaczniki o kształtach zamieszczonych na rysunku 6.5. Podstawowym kształtem znacznika odpowiednim dla proponowanej metody lokalizacji jest obszar ograniczony dwoma pionowymi krawędziami o znanych wysokościach, oraz o znanej odległości między tymi krawędziami. Schemat znacznika przedstawia rysunek 6.6.



Rysunek 6.5: Kształty znaczników stosowanych w pomiarach

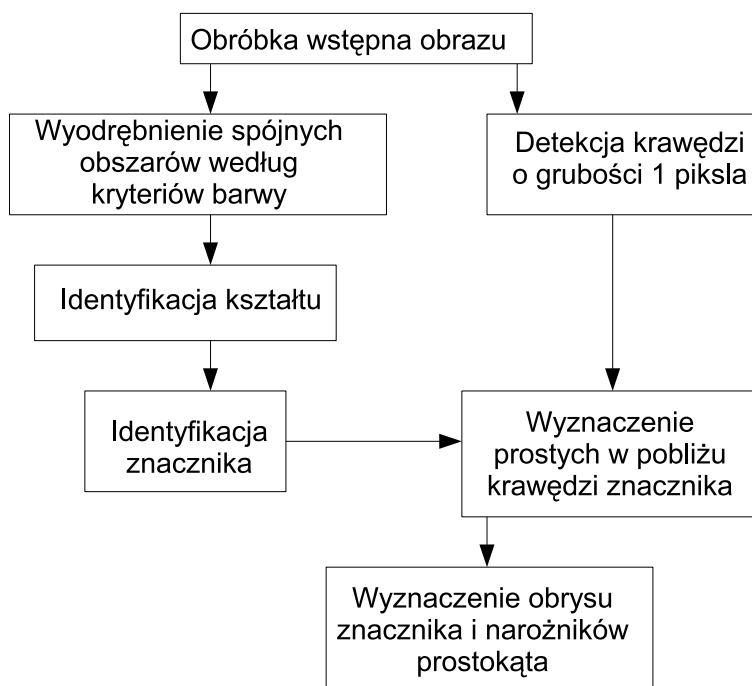


Rysunek 6.6: Schemat budowy znacznika do proponowanej metody lokalizacji

Wykrywanie znaczników może być dziś wykonane wieloma metodami. Do omawianych metod jako znaczniki nadają się obiekty płaskie, jednospójne i względnie duże. Istotną cechą jest możliwość określenia wysokości pionowej krawędzi. W obiekcie z rysunku 6.5.e takiej wysokości nie ma.

Podczas eksperymentów stosowano różne techniki wykrywania znaczników. Detekcja znaczników jest wieloetapowa. Schemat detekcji przedstawia rysunek 6.7. W pierwszym etapie wykonuje się obróbkę wstępną, m.in. usuwającą zniekształcenia nieliniowe. Następnie wyodrębnia się obszary odpowiadające barwom znaczników, oddzielnie dla każdej sprawdzanej barwy. Stosowaną metodę detekcji obszarów o określonej barwie opisano w rozdziale 7.1. Następnie sprawdza się kształt wyodrębnionych obszarów na podstawie momentów centralnych, niezmienników H_u oraz AMI (rozdział 7.2). Jeżeli wyodrębniony obszar barwny ma oczekiwany kształt, zostaje uznany za znacznik. Możliwe jest stosowanie dodatkowych kryteriów identyfikacji (np sąsiedztwo barw), jeżeli w otoczeniu jest więcej obiektów o tych samych kształtach i barwie.

Równoległe z detekcją znaczników wykonuje się detekcję krawędzi o grubości 1 piksla metodą Marra–Hildretha (rozdz. 3.3.2, 7.3.1). Po wykryciu linii pozostawia się tylko te, które przebiegają w pobliżu krawędzi znacznika, co opisano w rozdziale 7.3.2. Z krawędzi wykrytych w obrazie znajdujących się w pobliżu krawędzi znacznika wyznacza się równania prostych (rozdz 7.3.3), a z nich wyznacza się prostokątny obrys znacznika (rozdz 7.3.4).



Rysunek 6.7: Schemat detekcji znaczników

Rozdział 7

Zastosowana detekcja znaczników

7.1 Detekcja obiektów o określonej barwie

W niniejszym rozdziale opisano technikę wykrywania obszarów o jednorodnej barwie. Obszary te mogą być naturalnymi znacznikami. Przykładowe obiekty jakie można wyodrębnić to np. plakaty (znacznik "naturalny"), niektóre meble (szafa) i drzwi o charakterystycznej barwie.

Opracowana metoda pozwala na łatwe znalezienie obiektu o barwie charakterystycznej w jego otoczeniu, w dużym stopniu niezależnie od zmian różnic w oświetleniu spowodowanych częściowym zacienieniem czy różnicami w odległości przedmiotów od źródła światła [44].

W większości codziennych zastosowań wizyjnych kolor reprezentuje się w postaci modelu RGB. Kolor jest w nim opisywany jasnością trzech sumujących się składowych barw. Model ten jest bardzo popularny w cyfrowej reprezentacji obrazu, ponieważ stosuje się go do reprezentacji barw na ekranie kineskopu. Inne ważne modele barw to CMY (CMYK), HSV i CIE. Szerzej modele barw przedstawiono w rozdziale 3.1.

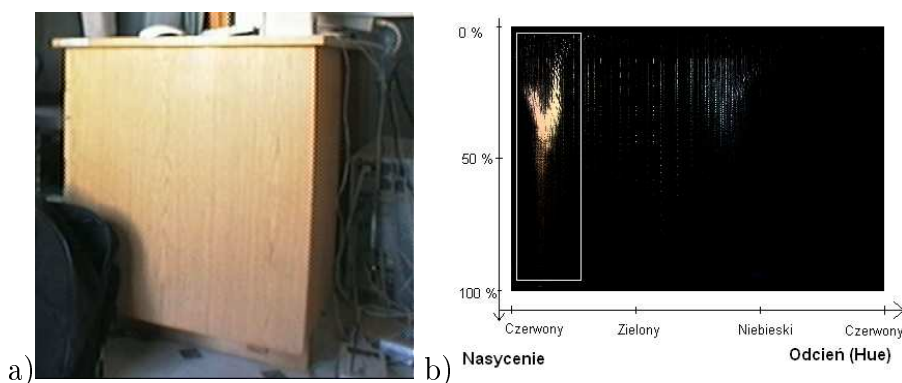
7.1.1 Wybór barw wzorcowych

Proponowana metoda powinna znajdować w obrazie obszary łatwo wykrywalne ze względu na charakterystyczną barwę. Metoda wyszukiwania powinna spełniać następujące założenia:

- umożliwiać wykrywanie obiektu o dowolnej barwie mieszczącej się w określonym zakresie;
- charakteryzować się odpornością na zmiany oświetlenia;
- działać w miarę możliwości szybko.

Powyższe warunki spełniają modele barw rozdzielające jasność od barwy. Ze względu na możliwość zakłóceń i utratę informacji w modelu r'g' zdecydowano się na model HSV. Za obszar jednokolorowy uważany jest obszar o odcieniu mieszczącym się nie dalej niż w określonej odległości od wyznaczonego, zaś nasycenie i jasność mogą być zmienne w określonym zakresie. Wynika to z faktu, że różne fragmenty tej samej powierzchni mogą być różnie oświetlone, zacienione lub mieć różną jasność wynikającą np z różnic odległości części obiektu od ogniska kamery. Intensywność oświetlenia wpływa przede wszystkim na jasność i nasycenie barwy, a na odcień jedynie wtedy, jeżeli barwa oświetlenia znacząco odbiega od białej. Przykład nierównomiernie oświetlonego obiektu, którego odcień jest wszędzie jednakowy, pokazuje rys 7.1. Rysunek 7.1.a pokazuje zdjęcie obiektu, zaś 7.1.b - histogram barw, którego osiami są

nasylenie (*Saturation*) w pionie i odcień (*Hue*) w poziomie. Na osi poziomej histogramu jest odcień, zaś na osi pionowej – nasylenie. Najmniej nasycone barwy znajdują się na górze, zaś najbardziej na dole obrazu. W zaznaczonej ramką części histogramu mieszczą się wszystkie punkty obiektu (biurka). Z histogramu wynika, że nawet obiekty postrzegane jako barwne mają w rzeczywistości barwy słabo nasycone, tzn nasylenie tylko niewielkiej części pikseli przekracza 50%.



Rysunek 7.1: Histogram barw: Obiekt nierównomiernie oświetlony i histogram obrazu w modelu HSV

Zakres barw

Podczas sporządzania histogramu wyznacza się nasylenie i odcień każdego punktu obrazu. Ilość punktów o tych samych parametrach przekłada się na jasność punktu we współrzędnych odcień i nasylenie.

Problemem podczas sporządzania histogramu jest konwersja niektórych ciemnych barw RGB na mocno nasycone barwy w modelu HSV. Dla uniknięcia fałszywych sygnałów, pochodzących głównie od obiektów ciemnych, podczas sporządzania histogramu stosuje się progi nasylenia zależne od jasności [7], co pozwala wyeliminować bardzo ciemne punkty o pozornie "dużym" nasyleniu. Przykładowo, punkt praktycznie czarny ($R=4, G=1, B=1$) mógłby zostać zinterpretowany jako punkt czerwony o nasyleniu 60%. Filtr ten jednocześnie eliminuje możliwość identyfikacji punktów czarnych i szarych. Działanie filtru polega na odrzucaniu pikseli o zbyt małej jasności dla danego nasylenia. Tabela 3.3 (strona 32, rozdz. 3.1.5) zawiera warunki, jakie muszą spełniać jasność i nasylenie, aby odcień punktu nie był nieokreślony i punkt był uwzględniany przy sporządzaniu histogramu.

Uczenie barw charakterystycznych

Możliwe jest automatyczne ustalenie i zapamiętanie przez system kolorów rozpoznawanych obiektów. Polega ono na rozpoznaniu we wzorcowym obrazie kolorów o dużym nasyleniu, kontrastujących z innymi kolorami. Szczegółowo opisano metodę wyznaczania barw w [43]. W tym celu wykonuje się następujące kroki:

- sporządzenie histogramu HSV w małej skali;
- znalezienie obszarów o nasyleniu większym od progowego, zajmujących wąskie pasmo odcienia;

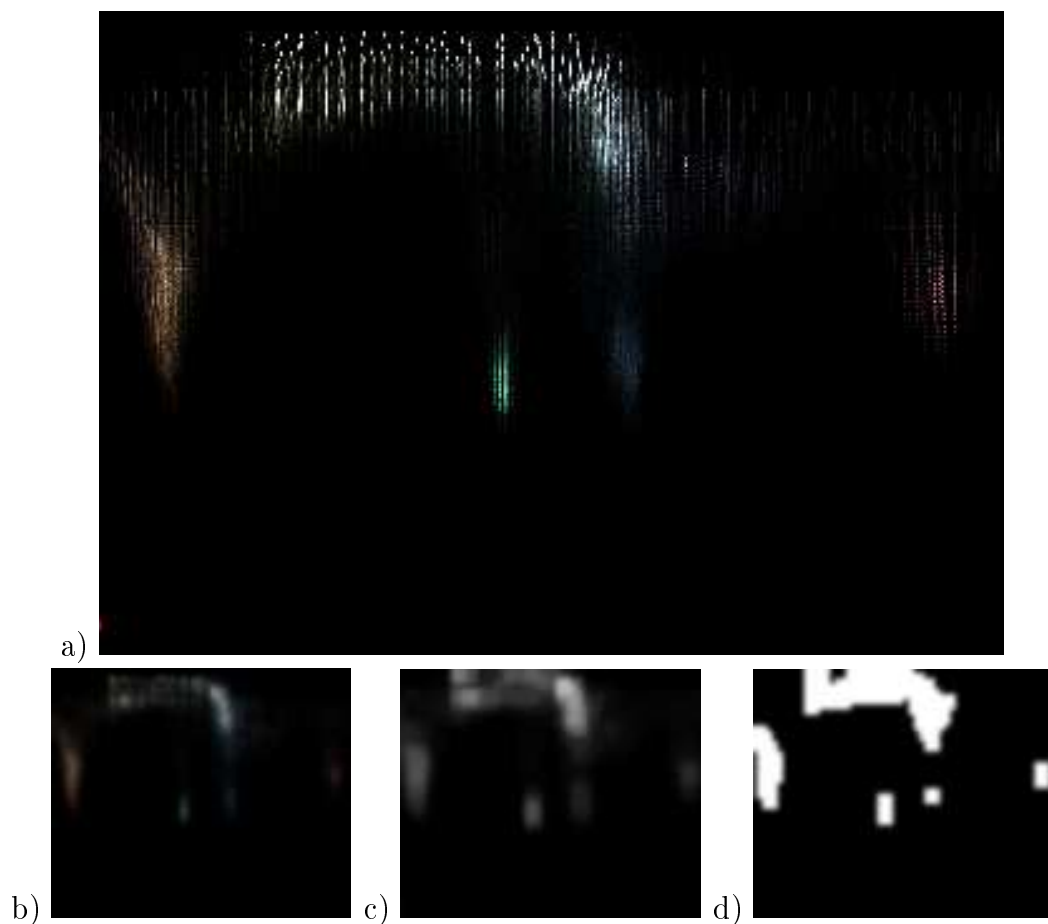
- określenie zakresu parametrów barw odpowiadających tym odcieniom.

Sporządzany jest histogram dwuwymiarowy na płaszczyźnie HS, której współrzędnymi są odcień i nasycenie koloru. Rysunek z obiektami barwnymi, stanowiący dane wejściowe algorytmu uczenia barw, przedstawia rysunek 7.2. Postać histogramu w kolejnych fazach detekcji barw przedstawia rysunek 7.3. W histogram zlicza się punkty obrazu o nasyceniu i barwie należących do wyznaczonych przedziałów. Podstawą do automatycznego wyboru danej barwy jako wzorca jest obecność na obrazie odpowiednio dużej ilości piksli o zbliżonym odcieniu i nasyceniu większym od progowego, co daje efekt jasnego obszaru w histogramie, co przedstawia rysunek 7.3.a. Aby łatwo można było wyznaczyć takie obszary, histogram wykonuje się w małej rozdzielczości - większa spowodowałaby rozproszenie punktów i w efekcie w histogramie nie byłoby dużych ilości piksli o określonej barwie. Przykład takiego histogramu przedstawia rysunek 7.3.b. Piksele o zbliżonych parametrach odpowiadają wówczas temu samemu punktowi histogramu. Zaimplementowany algorytm wykonuje histogram w rozdzielczości 60 odcieni i 50 stopni nasycenia. Uzyskany histogram jest poddawany działaniu filtra maksymalnego (rys 7.3.c, a następnie progowaniu. Obiekty o barwach charakterystycznych są widoczne w postaci niewielkich obszarów w obszarze dużego nasycenia barw. Na rysunku 7.3.d barwy charakterystyczne widoczne są jako małe obszary w dolnej części (nasycenie barw rośnie z góry na dół).



Rysunek 7.2: Rysunek wejściowy dla algorytmu uczenia barw

Określenie barw na podstawie histogramu wymaga wyznaczenia progowej ilości piksli oraz zakresu nasycenia i jasności. Wartości te dobrano eksperymentalnie. Wyznaczenie progu nasycenia wymaga odcięcia barw nasyconych od barwy białej oświetlonej lub odbijającej barwny obiekt. Rysunek 7.4 przedstawia typowe histogramy HS dla obrazów barwnych. Można na nich dość łatwo oddzielić barwy nasycone od barw słabo nasyconych (szarej, czarnej i białej). Rysunek 7.4.a przedstawia histogramy przed progowaniem. Kolejne wiersze przedstawiają te same histogramy po progowaniu, które powinno dokonać najlepszego podziału na barwy. Progi



Rysunek 7.3: Wybór barw do rozpoznawania na podstawie histogramu: a) Histogram w dużej rozdzielczości; b) ten sam histogram w małej rozdzielczości; c) poddany działaniu filtra maksymalnego; d) po progowaniu - widać 4 charakterystyczne kolory o odpowiednim nasyceniu

ustalono na wartościach: 20 (rys 7.4.b), 30 (rys 7.4.c), 40 (rys 7.4.d), 50 (rys 7.4.e) i 60 (rys 7.4.f). Mniejsza wartość progu powoduje zauważenie większej ilości barw (obszary jasne w histogramie mają większą powierzchnię), ale jednocześnie powodują połączenie kilku barw w jedną oraz łączenie barwy nasyconej z nienasyconą o tym samym odcieniu. Optymalny wydaje się podział wg wartości progu 40 – co przedstawia rysunek 7.4.d. W oprogramowaniu można ustalić wartość progu, domyślna ustalona na podstawie badań wynosi 40 i jest odpowiednia dla większości przypadków. Ponieważ nierównomierne oświetlenie może powodować że odcień obiektu jest rozciągnięty na pewien zakres, należy uwzględnić fakt że na brzegach zakresu ilości pikseli są mniejsze niż próg. Dlatego zastosowano dwa progi: główny, wyznaczający charakterystyczne barwy, oraz dodatkowy - domyślnie 10. Obszar wyznaczany przez dodatkowy próg jest brany pod uwagę, jeżeli przynajmniej niektóre wartości obszaru przekraczają wartość podstawowego progu.

Rysunek 7.5 przedstawia próg oddzielający skupiska obszarów białych od barwnych. Pozioma linia w histogramie znajduje się na wysokości 20% nasycenia. Obszary znajdujące się całkowicie w obszarze "barwnym", tj. o nasyceniu powyżej progowej wartości 20% stanowią barwy wzorcowe.

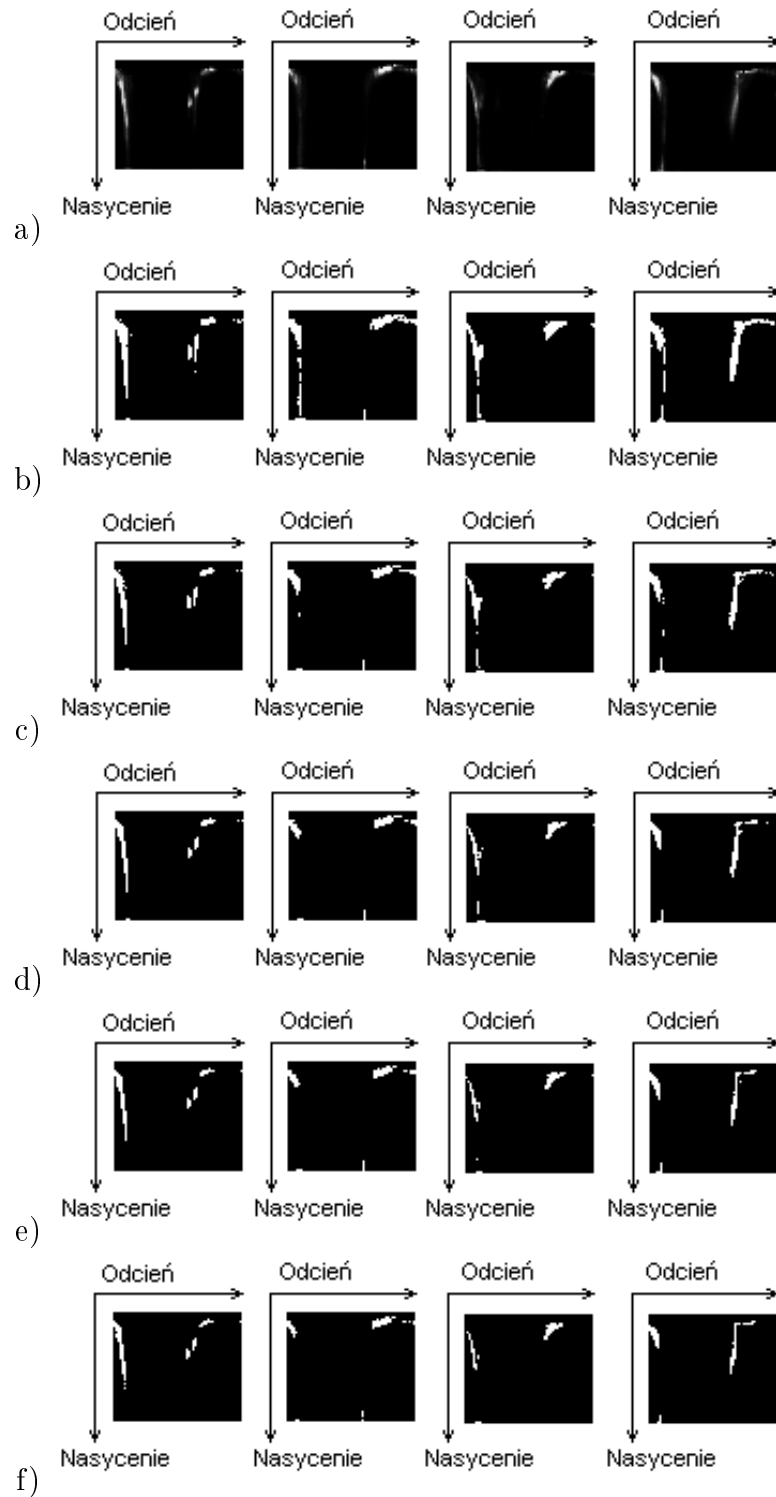
Przeprowadzono eksperymenty z kolorową kamerą robota mobilnego B-14 oraz z kamerą stacjonarną USB. Zestaw kamera i *frame grabber* robota oraz kamera USB pozwalają uzyskać obraz w maksymalnej rozdzielczości 640 na 480 pikseli, przy rozdzielczościach barwy 15 i 24 bity. Przeprowadzono eksperymenty z rozpoznawaniem obiektów po uczeniu systemu w zmiennym oświetleniu i z zastosowaniem kolorów wzorcowych generowanych automatycznie.

Uczenie systemu przeprowadzono dla wybranych obiektów barwnych. Standardowo jako barwy charakterystyczne algorytm wybiera barwy obiektów dużych. Typowa barwa to barwa "drewniana", pośrednia między brązową i pomarańczową - jest to kolor mebli i drewnianych podłóg. Bardzo duże znaczenie ma ustawienie kamery względem obiektu - obiekty o wzorcowych barwach muszą zajmować odpowiednio duży obszar obrazu, aby nie zostały pominięte. Minimalna powierzchnia to ok. 1000 pikseli, w przypadku obrazów o małym rozproszeniu barw (tzn. małej ilości barw) nawet 30000 - 40000 pikseli - czyli obiekt wzorcowy powinien mieć w obrazie rozmiary 30x30, a czasem 200x200 pikseli żeby był zauważony przez mechanizm automatycznej detekcji barw. Skrajny przypadek przedstawiono na rysunku 7.6, gdzie barwa płyty CD znajdującej się w czarno-białym otoczeniu była zauważana jako wzorcowa dopiero przy wielkości obiektu rzędu 200 na 200 pikseli (obiekt zajmuje ok. 10% powierzchni obrazu).

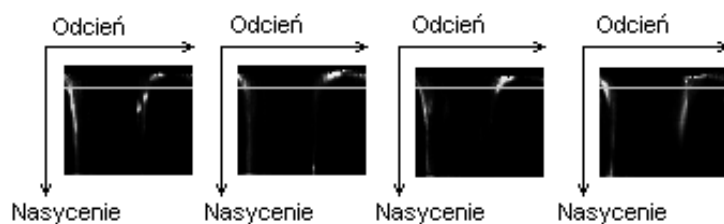
Zdefiniowany kolor określa się w granicach nasycenia, odcienia i jasności. Podczas uczenia definiuje się dopuszczalne minimalne nasycenie, w którym powinien znaleźć się obszar po progowaniu (20%-30%), minimalną (20-30) i maksymalną (200-255) jasność oraz wartość progu (30-60). Modyfikując te wartości można dostroić optymalnie system do wzorcowego zdjęcia w celu wykrycia wszystkich charakterystycznych barw.

Uczenie barw na podstawie wielu obrazów

Wykrycie określonej barwy wymaga aby obiekty tej barwy zajmowały sporą część obrazu. W przypadku rozmieszczenia interesujących obiektów w różnych miejscach uzyskanie jednego obrazu z wszystkimi obiektami barwnymi może być niewykonalne. Rozszerzeniem metody detekcji barw w jednym obrazie jest wyznaczanie barw na podstawie serii obrazów. W pierwszym etapie detekcji następuje wykrycie barw dla każdego obrazu z osobna algorytmem przedstawionym w rozdziale 7.1.1. W drugim etapie następuje scalanie barw podobnych, powodujących powstanie barw o zakresie odcienia nie przekraczającym 60° (w skali 360°). Podczas scalania



Rysunek 7.4: Różne progowe ilości piksli dla histogramów HS: a) przed progowaniem; b) ÷ f) - z progami 20 ÷ 60



Rysunek 7.5: Próg oddzielający barwy nasycone od odcieni szarości



Rysunek 7.6: Minimalna wielkość obiektu barwnego w otoczeniu czarno-białym podczas uczenia barw

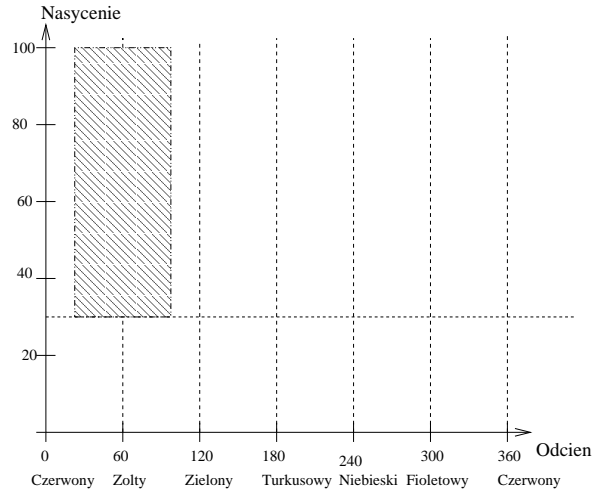
przyjęto następujące reguły:

- jeżeli jedna barwa znajduje się całkowicie w innej, jest pomijana,
- jeżeli dwie barwy mają ten sam przedział odcienia (lub jeden zawiera się w drugim) i różne zakresy nasycenia lub jasności, w wyniku scalenia powstaje jeden przedział o większym z zakresów odcienia, nasycenia i jasności,
- barwy o częściowo zachodzących odcieniach można łączyć, rozszerzając zakres nasycenia i jasności do największej występującej wartości. Jeżeli szerokość odcienia przekracza 60° , tworzy się odpowiednią liczbę oddzielnych barw o zakresie odcienia 60° , zachodzących na siebie co najmniej 20° (mniej jeżeli dzielony przedział nie jest dostatecznie szeroki).

Wybór barw wzorcowych bez uczenia

Uzyskanie barw wzorcowych z obrazu wymaga, żeby odpowiednie obiekty zajmowały przynajmniej 5-10% powierzchni obrazu dla rozdzielczości 648 na 480 piksli. Jest to uzasadnione, ponieważ wiarygodne wykrycie obiektu nie jest możliwe jeżeli obiekt zajmuje mało miejsca w obrazie (może to być zakłócenie). W niektórych sytuacjach próba nauczania systemu barw wzorcowych nie daje dobrych wyników. Dzieje się tak np gdy poszukiwane obiekty są małe na obrazach uczących. Możliwe jest określenie barw wykrywanych bez uczenia. Wówczas każdy z kolorów zajmuje równy fragment osi odcieni, zaś każdy kolor zajmuje dopuszczalny zakres nasycenia. Gdy kolorów wzorcowych jest 6, wówczas każdy ze zbiorów odpowiada jednemu z

6 odcieni (czerwony, żółty, zielony, turkusowy, niebieski, fioletowy) o dużej szerokości zakresu. Zasadę podziału przedstawiono na rysunku 7.7.



Rysunek 7.7: Automatycznie generowane kolory wzorcowe

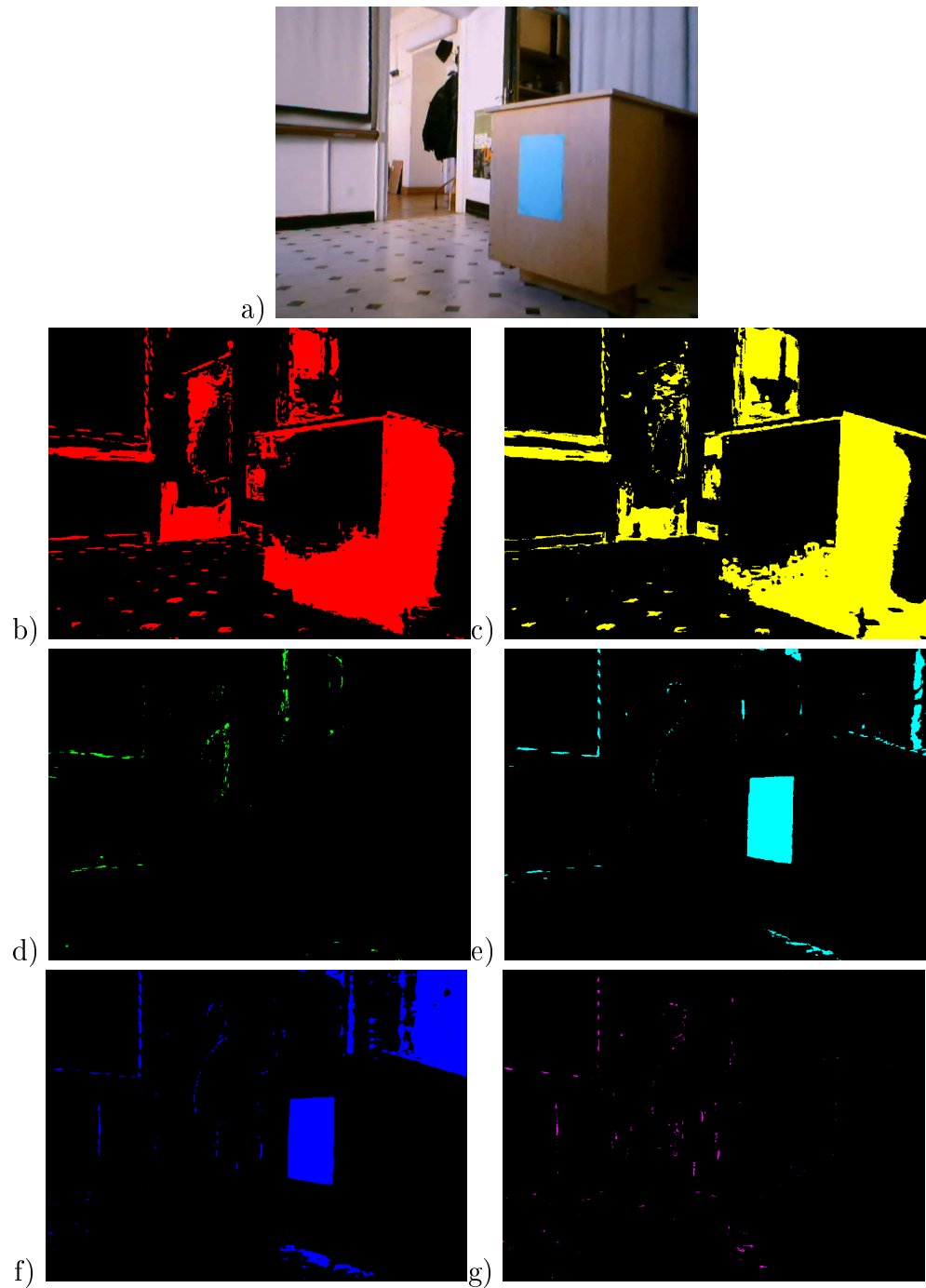
Przeprowadzono eksperyment z wykrywaniem barw wyznaczonych automatycznie przez równomierny podział skali odcieni. Dla 6 barw uzyskano wyniki znacznie różniące się od uzyskanych z uczącym się systemem. Główne różnice są następujące:

- Wykrywane są obiekty o różnych kolorach, w ogólności wszystkie barwne;
- Wiele wykrytych obszarów nie odpowiada konkretnym obiektom, są przypadkowe;
- Obszary o barwach pośrednich między wzorcowymi są różnie klasyfikowane zależnie od oświetlenia.

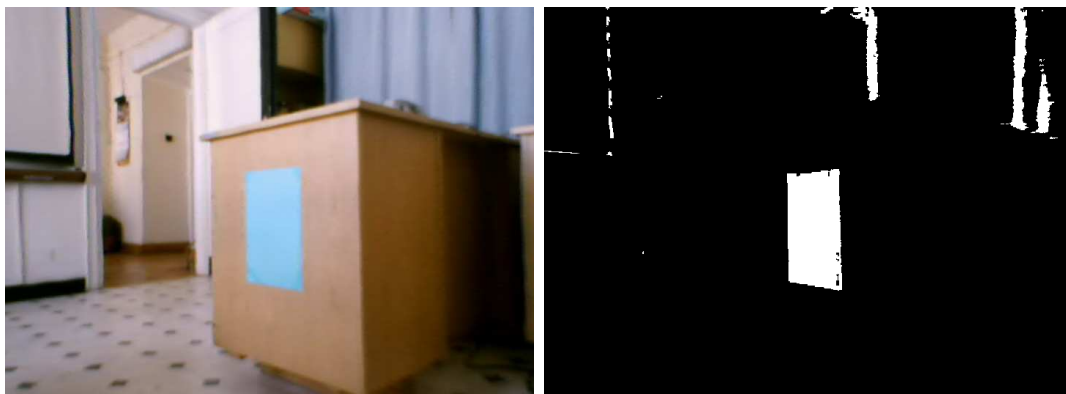
Przykładowy obraz oraz wykrywane na nim elementy przedstawia rysunek 7.8. Ze względu na dużą ilość barw i elementów, obszary wykrywane dla każdej barwy przedstawiono na oddzielnym obrazie. Rysunek 7.8.a przedstawia obraz wejściowy, 7.8.b – obszary widziane jako czerwone, 7.8.c – żółte, 7.8.d – zielone, 7.8.e – turkusowe, 7.8.f – niebieskie i 7.8.g – fioletowe. Bardzo dobrze widoczne jest duplikowanie obiektów w różnych barwach, oraz wykrywanie dużej ilości przypadkowych obszarów. Oprócz wielu niewłaściwych obiektów znaleziono znaczny w barwach turkusowej i niebieskiej.

7.1.2 Znajdowanie obszarów o charakterystycznej barwie

Najprostszy i najwolniejszy algorytm znajdowania obszarów polega na sprawdzeniu każdego piksela czy mieści się on w zakresie parametrów HSV oznaczających jedną z poszukiwanych barw. Metoda ta daje dobre rezultaty jeżeli poszukiwane obiekty nie są duże, a w obrazie nie ma dużego szumu barwnego. Przykład działania tego algorytmu w otoczeniu dobrze oświetlonym przedstawia rysunek 7.9. W eksperymencie tym poszukuje się obszaru zielono-niebieskiego, o zakresie odcienia 180–220. Podczas progowania wszystkich pikseli obrazu zostają zauważone wszystkie obszary - na prawym obrazie widoczne jako jasne plamy.



Rysunek 7.8: Wykrywanie obszarów według barw zdefiniowanych automatycznie: a) – obraz wejściowy; b) ÷ g) – znalezione obszary barwne.

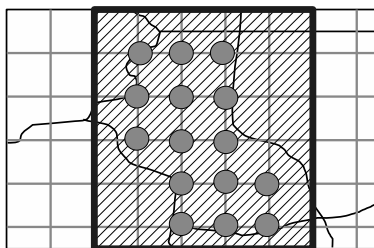


Rysunek 7.9: Przykład wykrywania wszystkich piksli o określonej barwie w otoczeniu dobrze oświetlonym



Rysunek 7.10: Przykład wykrywania wszystkich piksli o określonej barwie w obrazie z zakłóceniami

Jak widać na rysunku 7.10, obszary czarne pod wpływem oświetlenia i odbłasków mogą wydawać się barwne i zostają wówczas zaliczone do poszukiwanej klasy. Dla obrazów z zakłóceniami barw lepsze wyniki daje wstępne odrzucenie małych obszarów, które są wykrywane jako barwne. Jest to uzasadnione, ponieważ poszukiwany znacznik musi być względnie duży (musi mieć co najmniej kilkadziesiąt pikseli wysokości i szerokości) żeby zadziałał algorytm detekcji linii i żeby można było zidentyfikować go i wystarczająco dokładnie ustalić jego kształt i rozmiary obrazowe. W dodatku źródłem zakłóceń może być padające jaskrawe światło (np słoneczne) które czasem nadaje zabarwienie obiektom czarnym i białym. W opisanej poniżej metodzie zakłada się, że wykrywany obiekt musi być dość duży (zajmować powierzchnię większą niż 400 pikseli), aby można było go jednoznacznie identyfikować. Przy odstępach punktów kontrolnych równych 10 (tzn., że sprawdzany jest co dziesiąty piksel w pionie i w poziomie) faktycznie sprawdzany jest 1% powierzchni obrazu. Pozwala to szybsze znalezienie poszukiwanych kolorów mimo wyznaczania wartości nasycenia i odcienia pikseli. Dodatkowo, ponieważ poszukiwane obiekty mają określoną minimalną powierzchnię (400 pikseli), odrzucane są obiekty mniejsze niż 400 pikseli, czyli 4 piksele znalezione z odstępem co 10. Schemat metody szybkiego wykrywania obszarów pokazuje rysunek 7.11. Cienkie nierówne linie oznaczają granice obszarów o różnych barwach. Linie przerywane - linie siatki, w której węzłach sprawdza się kolor pikseli. Gruba ramka z zakreskowanym obszarem oznacza region, w którym w całości znajduje się poszukiwany kolor. Rysunek 7.12 przedstawia wyodrębnione obszary na obrazie z rysunku 7.10.



Rysunek 7.11: Szybkie wykrywanie obszarów o jednakowym kolorze



Rysunek 7.12: Przykład przyspieszonego wykrywania obszarów barwnych na obrazie z zakłóceniami

Czas wykrywania obszarów zależy od dwu czynników: ilości barw wzorcowych, do których powinien być wprost proporcjonalny oraz ilości punktów i obszarów każdego poszukiwanego

koloru. Zależność czasu przetwarzania od drugiego czynnika jest skomplikowana, wpływ mają tutaj:

- Ilość znalezionych pikseli danego koloru wzorcowego;
- Czas scalania pojedynczych oczek siatki w obraz, który zależy od kształtu obszaru;
- Czas łączenia obszarów rozłącznych przylegających do siebie o tej samej długości boku przylegającego;
- Czas wykrywania i pomijania usuwania izolowanych pikseli i obszarów składających się z mniejszej liczby pikseli niż progowa (w badaniach 4).

Czasy wykrywania obszarów podaje tabela 7.1. Pozycje, w których znajdowano 3 kolory dotyczą barw nauczonych, zaś 6 kolorów zostało wygenerowanych automatycznie. W przypadku 6 kolorów wyznaczonych przez równomierny podział skali odcieni duży jest udział przypadkowych obszarów, wydłużających czas przetwarzania. Pomiar wykonano w programie napisanym w języku Visual C++ v6.0 Professional, na komputerze PC z procesorem AMD K6 350MHz, działającym pod kontrolą systemu operacyjnego Windows NT 4.0. Najdłuższy z czasów nie przekracza 40 ms, co pozwala na przetwarzanie nawet na stosunkowo wolnym komputerze 25 klatek na sekundę, szczególnie jeżeli system szuka nauczonych wcześniej kolorów.

Ilość barw wzorcowych	Ilość wykrytych obszarów	Czas szukania [ms]
3	5	9
3	9	9
3	9	11
6	17	27
6	19	21
6	27	38
6	28	23
6	36	28

Tablica 7.1: Czasy wykrywania obszarów barwnych przy różnych parametrach i ilości barw wzorcowych

7.2 Detekcja charakterystycznych kształtów

Niniejszy rozdział przedstawia metodę detekcji znaczników o charakterystycznych kształtach z zastosowaniem momentów i niezmienników. Odpowiednie definicje znajdują się w rozdziale 5.1.

Detekcja kształtów obiektów odbywa się za pomocą momentów geometrycznych i niezmienników. Momenty centralne są identyczne dla wszystkich obiektów o tych samych kształcie i wielkości niezależnie od położenia. Z kolei momenty centralne znormalizowane są niewrażliwe na zmianę rozmiaru (skalowanie), również jeżeli skala pozioma nie jest identyczna z pionową. Dlatego nadają się do identyfikacji obiektu widzianego z różnych odległości.

Zgodnie z teorią nie są wrażliwe na obrót wokół osi prostopadłej do płaszczyzny obiektu niezmienniki H_u [38] i niezmienniki przekształcenia afinicznego [33], opisane w rozdziale 5.1.

Najbardziej korzystne do detekcji obiektów wydają się niezmienniki, ze względu na największą ilość dopuszczalnych stopni swobody przy określaniu położenia i orientacji znacznika względem kamery. Poniżej przedstawiono doświadczalnie wyznaczanie wartości dla kilku kształtów znaczników: liter U, H, O oraz pełnego prostokąta.

Momenty są funkcjami zależnymi od kształtu i pozwalają na identyfikację. Liczba momentów, jakie należy wyznaczyć w celu opisanego kształtu, jest teoretycznie nieograniczona. W praktyce już kilka momentów najniższych rzędów pozwala na rozróżnienie i identyfikację wielu kształtów. Dodatkowo, we wzorach momentów wyższych rzędów występują wielomiany dużych stopni, przez co wartości te są bardzo wrażliwe na drobne niedokładności, powstające np. przy zmianie wielkości obiektu. Praktycznie w implementacji stosowano momenty, których suma rzędów nie przekracza 3: $\eta_{0,0}, \dots, \eta_{3,0}, \eta_{2,1}, \eta_{1,2}, \eta_{0,3}$. Wyznaczenie momentów wyższych rzędów nie jest bardziej czasochłonne niż wyznaczenie momentów niższych rzędów, jednak uzyskiwane wartości dla tego samego obiektu zbyt różnią się od siebie. Tabele 7.2, 7.3 i 7.4 przedstawiają wartości momentów centralnych znormalizowanych, niezmienników H_u oraz AMI dla obiektów widzianych z różnych miejsc i odległości.

Znacznik	$\mu_{0,0}$	$\mu_{0,2}$	$\mu_{0,3}$	$\mu_{1,1}$	$\mu_{1,2}$	$\mu_{2,0}$	$\mu_{2,1}$	$\mu_{3,0}$
U - minima	0,34	0,028	-0,0088	-0,018	-0,011	0,1	-0,0037	-0,086
U - maksima	0,46	0,1	0,007	0,0084	-0,0057	0,33	0,012	-0,012
H - minima	0,31	0,021	-0,0097	-0,0055	-0,0023	0,084	-0,0035	0,00070
H - maksima	0,48	0,098	0,0065	0,013	0,00072	0,28	0,016	0,036
O - minima	0,10	0,032	-0,044	-0,025	-0,09	0,030	-0,024	-0,015
O - maksima	0,57	0,22	0,07	0,016	0,005	0,27	0,012	-0,0013

Tablica 7.2: Porównanie momentów centralnych znormalizowanych dla różnych kształtów

Znacznik	H_u [1]	H_u [2]	H_u [3]	H_u [4]	H_u [5]	H_u [6]
U - minima	0,20	0,00035	1,37E-10	3,22E-10	-4,68-07	1,90241E-06
U - maksima	0,36	0,089	3,3E-06	3,5E-07	5,8E-05	0,0026
H - minima	0,14	2,3E-06	3,1E-12	3,6E-16	-2,4E-09	-1,9E-06
H - maksima	0,30	0,065	3,5E-06	2,0E-07	3,1E-06	0,00027
O - minima	0,072	0,00078	1,0E-12	2,1E-14	-4,1E-09	-1,1E-06
O - maksima	0,29	0,045	6,1E-06	1,1E-05	5,3E-05	0,0002

Tablica 7.3: Wartości niezmienników H_u dla różnych kształtów

Z przeprowadzonych pomiarów wynikają następujące wnioski:

- najbardziej stabilne są momenty centralne niskich rzędów (gdy suma rzędów nie przekracza 3),
- najbardziej stabilne są niezmienniki H_u wyznaczone na podstawie niskich potęg momentów niskich rzędów,
- najbardziej stabilne są AMI również wyznaczone na podstawie niskich potęg momentów niskich rzędów.

Znacznik	AMI [1]	AMI [2]	AMI [3]	AMI [4]	AMI [5]	AMI [6]
U - minima	0,044	1,6E-06	-2,6E+21	0,000035	0	0
U - maksima	0,077	1,5E-05	3,3E+20	0,00024	0	0
H - minima	0,035	-6,4E-06	-2,8E+21	3,2E-07	0	0
H - maksima	0,058	3,9E-09	-1,4E+17	0,00022	0	0
O - minima	0,041	-2,6E-05	-1,0E+27	2,7E-07	0	0
O - maksima	0,10	8,0E-06	1,8E+19	0,0089	0	0

Tablica 7.4: Wartości niezmienników AMI dla różnych kształtów

Wartości momentów centralnych wyższych rzędów (widać to już dla $\mu_{3,0}$ i $\mu_{0,3}$) dążą do zera, przy czym nawet 64-bitowa dokładność obliczeń nie ma na to wpływu. Dlatego pominięto te wartości. Z niezmienników H_u jedynie dwa pierwsze są na tyle różne od zera, żeby można było brać ich wartości pod uwagę. Podobnie jest z AMI, gdzie tylko pierwszy niezmiennik wykazuje stabilność, a pozostałe dążą do zera lub nieskończoności. Z przeprowadzonych eksperymentów wynika, że większość niezmienników nie nadaje się do zastosowań związanych z transformacją perspektywiczną. W przypadku momentów wyższych rzędów i wyznaczanych na ich podstawie niezmienników poważnym źródłem błędów jest niestabilność samych momentów, a dokładniej ich duża wrażliwość na małe zmiany obrazu. Z tego powodu do identyfikacji obiektów stosowano:

- momenty centralne znormalizowane η_{jk} , których suma rzędów $j + k$ nie przekracza 3,
- dwa pierwsze niezmienniki H_u (oznaczane w tabeli 7.3 $H_u[1]$ i $H_u[2]$) (wartości pozostałych są zbyt bliskie zeru lub zbyt różne przy widoku tego samego kształtu z różnych miejsc),
- pierwszy niezmiennik AMI.

Przed wyznaczeniem momentów obszar jest określony funkcją binarną, tzn jest czarny albo biały. Momenty są funkcjami obszaru, zmieniającymi się przy zmianie kształtu. Powstaje pytanie, czy wpływ zmian kształtu i wielkości obiektu pod wpływem zmiany odległości i kąta widzenia obiektu może być pominięty lub jakoś uwzględniony przy identyfikacji. Wpływ ten na pewno nie może być pominięty. Jednak wraz ze zmianą kąta patrzenia i wielkości obrazowej obiektu wartości momentów zmieniają się płynnie w określonych granicach. Dlatego przyjęto, że kształt można zapamiętać jako zestaw zakresów momentów, definiujący 11-wymiarowy "hiperprostopadłościan" w 11-wymiarowej przestrzeni. Osie tej przestrzeni są zdefiniowane przez:

- 8 znormalizowanych momentów centralnych $\eta_{00} - \eta_{30}$, z wykluczeniem momentów η_{01} i η_{10} , które zawsze mają wartość 0,
- 2 niezmienniki H_u ,
- jeden niezmiennik AMI.

Teoretyczne wyznaczenie krańcowych wartości wymaga złożonej analizy, która w dodatku nie uwzględniałaby błędów kształtu wprowadzanych przez dyskretyzację obrazu.

W eksperymentach krańcowe wartości uzyskano na podstawie wyznaczania momentów dla obiektu widocznego z różnych miejsc jako najbardziej skrajne wartości, przesunięte w kierunku

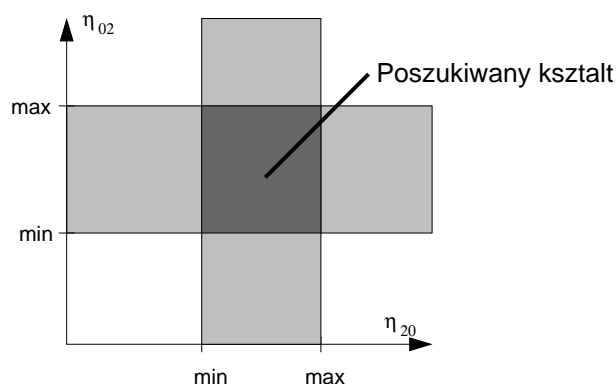
zwiększenia wartości bezwzględnej o kilka procent. Wymagało to napisania odpowiedniego oprogramowania, zapisującego wyznaczone momenty obszaru. Dodatkowo, wartości momentów i niezmienników dla każdego znacznika wymagają serii pomiarów, polegających na wyznaczaniu tych wartości dla znacznika widocznego z różnych miejsc i odległości. Ponieważ trudno jest stwierdzić a priori, dla których widoków obiektu momenty i niezmienniki przyjmują wartości skrajne, zastosowano podejście adaptacyjne. W pierwszej fazie wyznaczano wartości dla wielu różnych położzeń i orientacji kamery względem obiektu. Podczas prac na lokalizacją w przypadku braku identyfikacji widocznego znacznika następowała odpowiednia korekcja przedziałów.

Tabela 7.5 przedstawia wartości 8 momentów dla znacznika w kształcie litery O widzianego z różnych miejsc i pod różnymi kątami. Wiersze "Min" i "Max" zawierają odpowiednio dolną i górną granicę zakresu wyznaczoną na podstawie większej ilości obserwacji (nie wszystkie są zamieszczone w tabeli).

LP	X	Y	η_{00}	η_{02}	η_{03}	η_{11}	η_{12}	η_{20}	η_{21}	η_{30}
1	-100	200	0,497	0,0885	0,0011	-0,0023	-0,00033	0,16	0,0011	-0,009
2	0	200	0,461	0,0725	-0,0006	-0,0085	-0,0006	0,181	0,00243	-0,013
3	80	220	0,531	0,0878	0,0014	-0,0017	-0,00054	0,154	0,000346	-0,0059
4	-60	240	0,468	0,0812	0,0039	0,0040	-0,00009	0,168	-0,0002	-0,0076
5	100	240	0,489	0,0844	0,0049	0,002	0,00009	0,156 7	-0,0012	-0,0030
Min			0,10	0,032	-0,044	-0,025	-0,09	0,030	-0,024	-0,015
Max			0,57	0,22	0,07	0,016	0,005	0,27	0,012	-0,0013

Tablica 7.5: Wartości momentów dla znacznika "O" widzianego z różnych miejsc

Aby obszar z wyznaczonymi wartościami momentów został uznany za obiekt o znanym kształcie, przeprowadza się klasyfikację według odległości Hamminga. Obszar zostaje uznany za obiekt o znanym kształcie, jeżeli wszystkie wartości brane pod uwagę (momenty i niezmienniki) mieszczą się w przedziałach wartości identyfikujących dany kształt. Schematycznie dla dwóch współrzędnych przedstawia to rysunek 7.13. Rysunki 7.14 - 7.16 przedstawiają znacznik poprawnie identyfikowany widziany z różnych miejsc. Fragmenty rysunków po prawej stronie, na których jest tylko znacznik, zostały znalezione i wydodrębnione automatycznie przez program poszukujący znaczników na podstawie 10 momentów centralnych znormalizowanych $\eta_{j,k}$, gdzie $j + k \leq 3$, dwóch niezmienników H_u oraz jednego niezmiennika AMI.



Rysunek 7.13: Schemat klasyfikacji obszaru miarą odległości Hamminga dla momentów η_{02} i η_{20}



Rysunek 7.14: Wykrywany znacznik w kształcie litery "O" - widok 1



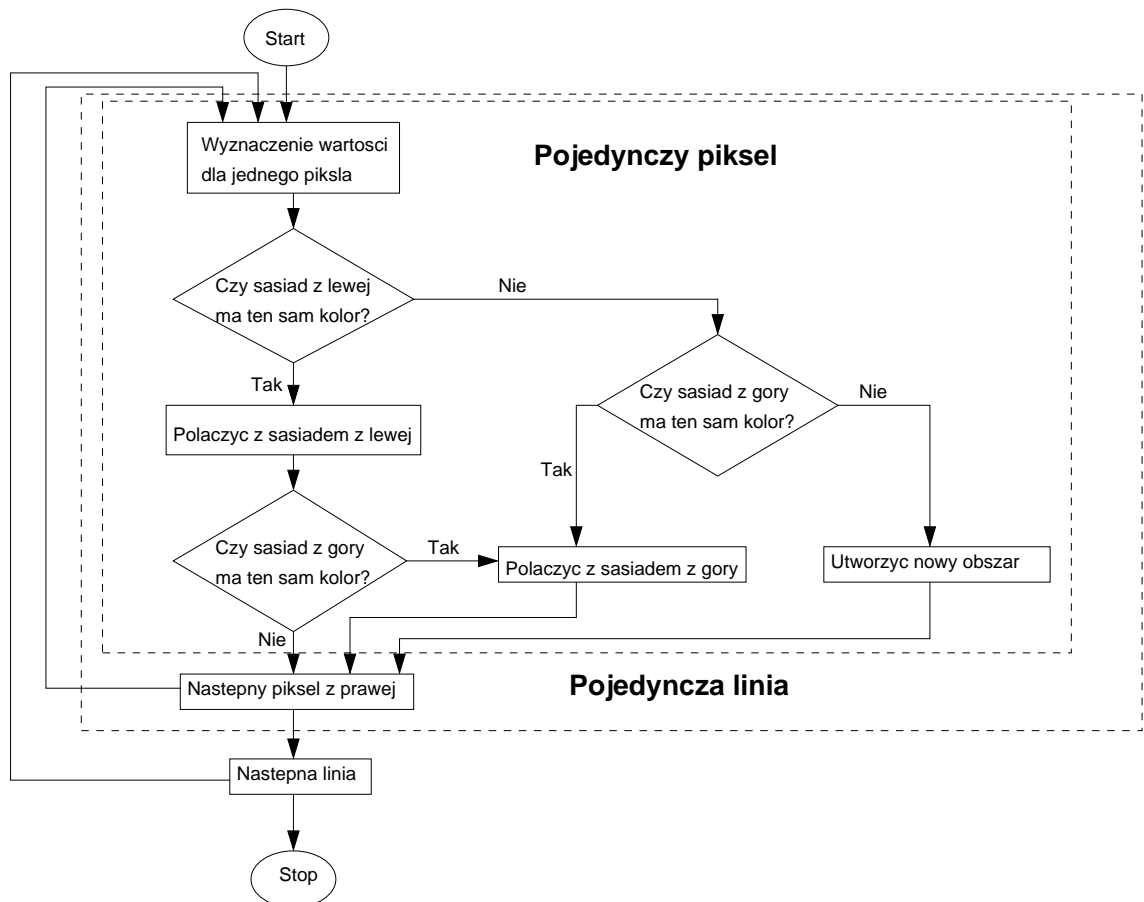
Rysunek 7.15: Wykrywany znacznik w kształcie litery "O" - widok 2



Rysunek 7.16: Wykrywany znacznik w kształcie litery "O" - widok 3

Podczas analizy obrazu najczęściej znajduje się wiele obszarów o określonej barwie, które są przekształcane w obszary białe na czarnym tle. Dlatego wyznaczanie momentów przeprowadza się dla wszystkich obiektów znajdujących się w obrazie, zarówno białych jak i czarnych. Stosowano szybki algorytm wyznaczający te momenty skanujący obraz liniami i wykonujący wszystkie niezbędne obliczenia w jednym przebiegu. Schemat algorytmu przedstawia rysunek 7.17. Obszary są przechowywane w pamięci komputera jako obiekty dynamiczne, co pozwala na przechowanie ich dowolnej liczby oraz łączenie ich ze sobą. Podczas skanowania pojedynczej linii najczęściej jest znajdowanych wiele obszarów o tej samej barwie oddzielonych obszarami o innej barwie. Przy skanowaniu kolejnych linii obszary czasem zachodzi konieczność połączenia rozdzielnych obszarów w jeden spójny, co dokonuje się przez zsumowanie momentów i odpowiednie określenie łącznego rozmiaru obszaru. Dzięki przechowywaniu obszarów jako obiektów dynamicznych ich łączenie nie stanowi problemu.

Podczas skanowania obrazu wyznacza się jedynie momenty geometryczne "zwykłe", zależne od współrzędnych i wyrażające się wzorem (5.5). Dla każdego piksla jest to iloczyn jego współrzędnych podniesionych do odpowiednich potęg. Po wyznaczeniu wszystkich obszarów i ich momentów "zwykłych" wyznacza się na ich podstawie momenty centralne i znormalizowane (zob wzory (5.8) i (5.9)). Oprócz momentów i niezmienników wyznacza się rozmiary liniowe każdego obszaru, jego położenie oraz zapamiętuje informację czy dany obszar dotyka brzegu obrazu. Ta ostatnia informacja ma znaczenie przy ustalaniu czy dany obiekt jest widoczny w całości.



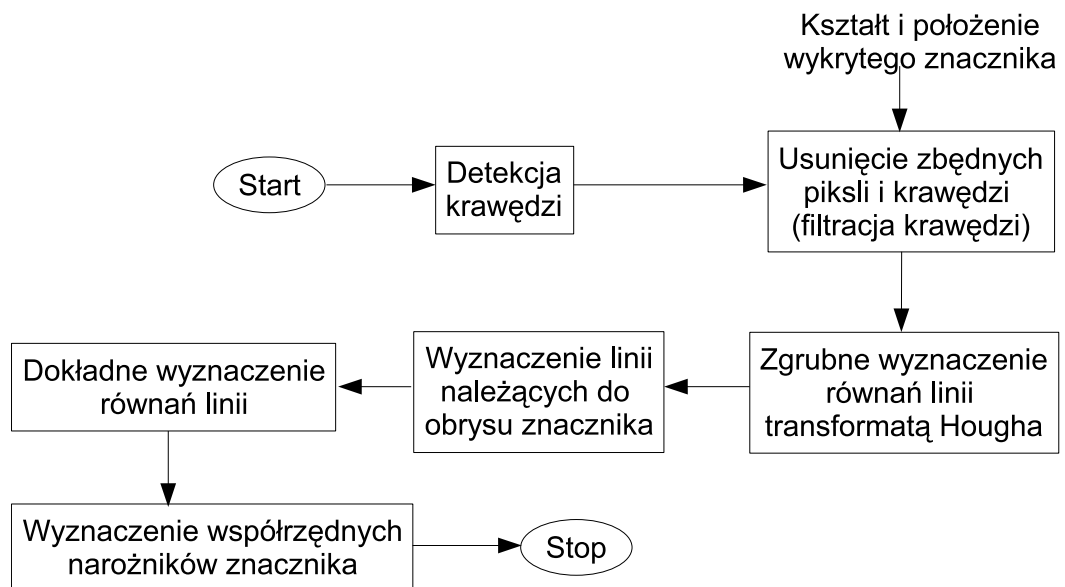
Rysunek 7.17: Schemat algorytmu wyznaczającego momenty geometryczne wszystkich obszarów w obrazie podczas jednego skanowania obrazu

7.3 Pomiar prostokąta opisanego na znaczniku

Techniki opisane w rozdziałach 7.1 i 7.2 są przeznaczone do odnalezienia obiektów charakterystycznych w obrazach. Po wykryciu i identyfikacji znacznika konieczne jest wyznaczenie jego rozmiarów w obrazie.

Niniejszy rozdział prezentuje zastosowaną metodę detekcji i określania rozmiarów obiektów prostokątnych. Danymi wejściowymi algorytmu są: obraz wejściowy po wykryciu znaczników za pomocą momentów i niezmienników (zob. rozdział 7.2) oraz sam obraz wejściowy po korekcji zniekształceń nieliniowych.

Schemat blokowy algorytmu rozpoznawania obrysu prostokątnego znajduje się na rysunku 7.18. Przykładowy obraz wejściowy z danymi znajduje się na rysunku 7.19. Poszukiwanym obiektem jest jasnoniebieska kartka umieszczona na boku biurka.



Rysunek 7.18: Schemat blokowy algorytmu detekcji prostokątnego obrysu znaczników



Rysunek 7.19: Przykładowy obraz wejściowy z poszukiwanym obiektem prostokątnym



Rysunek 7.20: Obraz wejściowy poddany działaniu detektora Marra-Hildretha: a) cały obraz; b) fragment odpowiadający znalezionemu obszarowi barwnemu

7.3.1 Detekcja linii w obrazie

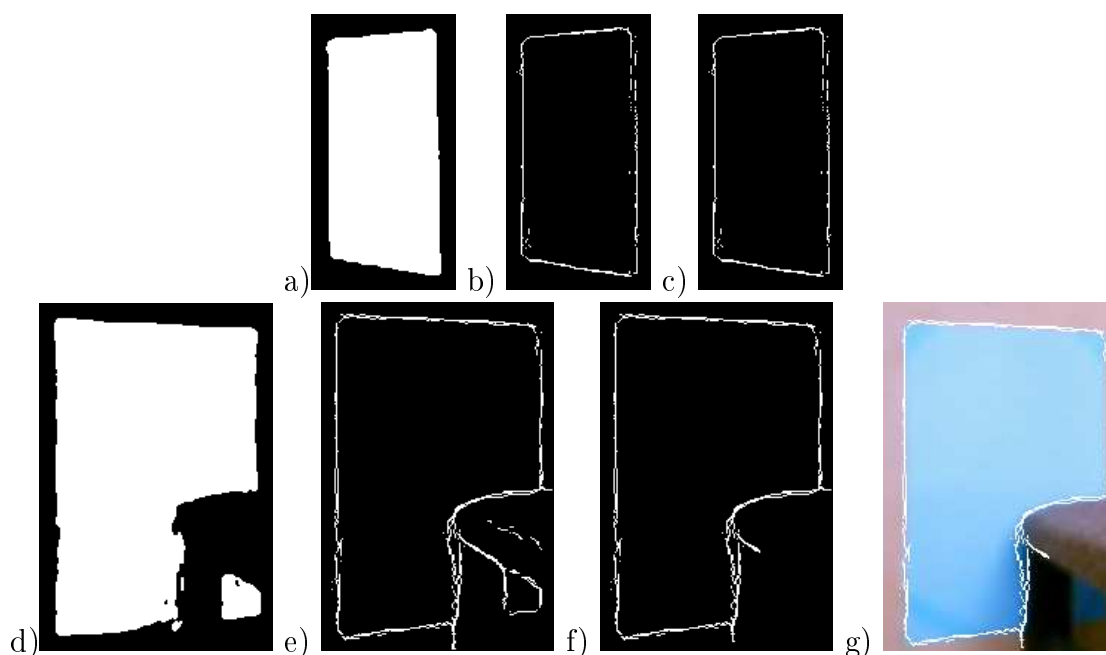
Danymi wejściowymi etapu jest obraz po korekcji zniekształceń i filtracji Gaussa. Powinien to być ten sam obraz, na którym poszukuje się obszarów o określonych barwach. Dzięki temu można dopasować znalezione obszary do krawędzi.

Detekcję linii wykonuje się algorytmem Sobela z filtracją linii przez drugą pochodną (algorytm Marra-Hildretha [66]). Metodę tę opisano w rozdziale 3.3. W pierwszej fazie stosuje się klasyczny algorytm Sobela, wykrywający wartości gradientów jasności. W drugiej fazie algorytm znajduje maksima i odrzuca pozostałe piksele. Jeżeli obraz wejściowy był barwny, wówczas po wykryciu krawędzi następuje łączenie płaszczyzn w jeden obraz. Na końcu algorytm wykonuje progowanie krawędzi. Jeżeli znalezione obszary są znacznie mniejsze niż wielkość obrazu, wówczas korzystniejsze czasowo jest wyszukiwanie krawędzi tylko w fragmentach obrazu, w których znajdują się znalezione obszary. Rysunek 7.20 przedstawia działanie detektora na obrazie po korekcji zniekształceń i filtracji Gaussa. Na rysunku 7.20.a przedstawiono wykryte krawędzie w całym obrazie, zaś na rysunku 7.20.b – poddany działaniu detektora fragment obrazu, na podstawie wcześniej znalezionego obszaru o charakterystycznej barwie. Poszukiwanie krawędzi w fragmentach obrazu pozwala skrócić czas obliczeń, co ma duże znaczenie przy transformacji Hough'a (zob rozdz. 3.5). Przy wyborze fragmentu obrazu do detekcji krawędzi należy pamiętać, że stosuje się detektor Sobela (zmniejszający obraz na brzegach o wielkość maski -1) oraz wyznacza się pochodne cyfrowe 1 i 2 stopnia, co zmniejsza użytkowe rozmiary obrazu o 4 piksele. W detektorze zastosowano operator Sobela o rozmiarze 3x3 piksele oraz progi o wartości 15 w skali 0 ÷ 255. Ten detektor powoduje zmniejszenie rozmiarów poziomego i pionowego obrazu o 6 pikseli i dlatego wybierając fragment należy uwzględnić obszar większy o 3 piksele z każdej strony. W eksperymentach stosowano też operator Sobela o rozmiarze 5x5 pikseli i wartości progowej 25.

7.3.2 Filtracja wykrytych krawędzi

W opisywanym etapie następuje filtrowanie pikseli krawędzi wykrytych w obrazie. Filtrem jest odległość danego punktu od wyznaczonego obszaru, znalezione przez progowanie barwy lub

jasności. Piksele krawędzi, które są przez filtr usuwane, leżą dalej od brzegu obszaru niż określona wartość graniczna d_b . Wszystkie punkty leżące wewnątrz obszaru lub poza nim w odległości nie większej niż d_b są zachowywane bez zmian. W implementacji stosuje się najpierw rozrost znalezionej obszaru o d_b pikseli w każdym kierunku, a podczas filtrowania wystarczy wybrać te punkty krawędzi, które leżą w powiększonym obszarze. Rysunek 7.21 przedstawia wykryte krawędzie po filtracji czyli usunięciu punktów nie należących do obszaru. W przypadku obiektu nie zasłoniętego krawędzie na zbiory punktów krawędzi 7.21.b i 7.21.c są identyczne i nie widać efektu filtrowania. W przypadku obiektu częściowo zasłoniętego efekt działania filtru czyli usunięcie pikseli krawędzi nie należących do znalezionej obszaru pokazuje rysunek 7.21.d, 7.21.e i 7.21.f [44]. Rysunek 7.21.g przedstawia złożenie pozostawionych krawędzi i obszaru barwnego.



Rysunek 7.21: Filtracja krawędzi na podstawie sąsiedztwa obiektu: a), d) obszary barwne; b), e) odpowiadające im krawędzie przed filtracją; c), f) krawędzie po filtracji; g) złożenie krawędzi i obszaru na jednym rysunku

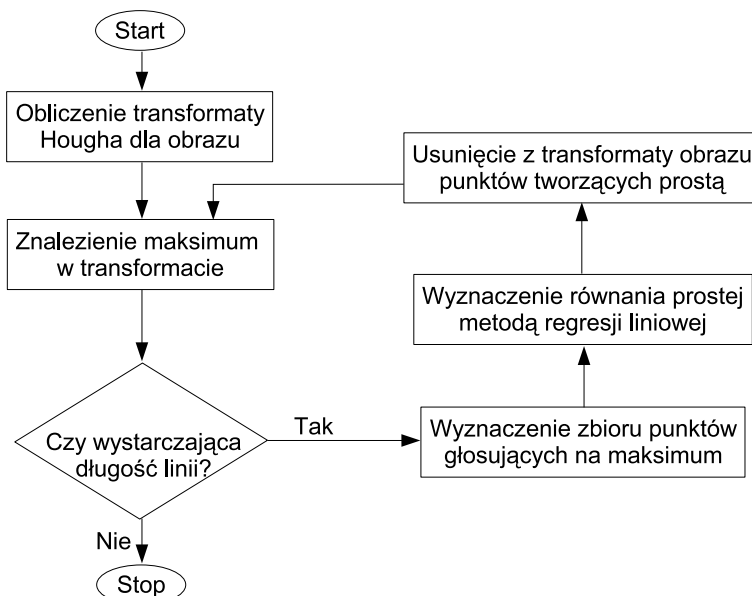
7.3.3 Wyznaczenie równań linii

Po przefiltrowaniu wykrytych krawędzi pozostaje zbiór punktów związanych z określonym obszarem barwnym. Detekcja krawędzi transformatą Hough'a pozwala na przybliżone określenie równań linii. Jest to metoda pozwalająca na wykrycie najdłuższych linii w obrazie, niekoniecznie ciągłych, co pozwala wykryć krawędzie z przerwami. Metoda ta jest wrażliwa na punkty nie należące do linii, leżące np. poza krawędzią obiektu, ale w przybliżeniu na przedłużeniu linii. Filtracja ograniczająca zbiór punktów do punktów leżących w pobliżu krawędzi obszaru (opisana w poprzednim punkcie 7.3.2) pozwala wyeliminowanie mogących całkowicie zafałszować wyznaczany przebieg linii. Takie samo zadanie spełnia wyodrębnienie z obrazu fragmentu otaczającego znaleziony obszar i wyznaczanie krawędzi tylko w tym fragmencie. Wyznaczenie równań krawędzi przebiega w dwóch etapach:

1. Przybliżone wyznaczenie przebiegu linii transformatą Hough'a;

2. Dokładne wyznaczenie równań metodą regresji liniowej.

Algorytm wyznaczania równań linii przedstawiono na rysunku 7.22.



Rysunek 7.22: Schemat algorytmu wyznaczania równań linii prostych w obrazie

Detekcja transformatą Hough'a

Transformata Hough'a, opisana w rozdziale 3.5, pozwala na detekcję równań linii o prostych wyrażonych równaniem $x \cos \alpha + y \sin \alpha = c$. Poszukiwanie równania sprowadza się do znalezienia maksimum w transformacie. Ponieważ obliczenia te mają na celu wyznaczenie przybliżonych wartości współczynników, zastosowano rozdzielczość kątową 1 stopień i liniową 1 piksel. W typowych warunkach fragmenty obrazu (dane wejściowe transformaty) nie są duże, więc obliczenia nie są bardzo czasochłonne.

Przy wykonywaniu transformaty sprawdza się, czy piksel może należeć do prostej wraz z sąsiadami. Piksel zostaje uznany za leżący na prostej z sąsiadami, jeżeli w odległości dwóch pikseli w każdą stronę znajdują się co najmniej dwa inne, tworzące z nim odcinek prostej. W celu sprawdzenia tego wyznacza się 8 kierunków (poziomy, pionowy, dwa pośrednie o nachyleniu $\pm 45^\circ$ i 4 pośrednie o nachyleniach $\pm 22,5^\circ$ i $\pm 67,5^\circ$). Jeżeli piksel należy do odcinka prostej o odpowiednim nachyleniu, wówczas wykonuje się transformację, ale tylko dla współczynników α odpowiadających spodziewanemu nachyleniu prostej.

W celu uniknięcia podziału głosów na jedną linię między dwa sąsiednie punkty transformaty, czyli dwie sąsiednie wartości współczynnika, każdy piksel obrazu wejściowego głosuje na 3 wartości współczynnika c . Dla skrócenia obliczeń, współczynniki α mogą przyjmować wartości z zakresu $0 \div \pi$. Rysunek 7.23 przedstawia wejściowy fragment obrazu oraz znormalizowaną transformację Hougha tego fragmentu. Normalizacja polega na proporcjonalnym przeskalowaniu wartości transformaty do przedziału $0 \div 255$. Linie pionowe (o nachyleniu 0° względem osi OY) są widoczne na górze i na dole transformaty, zaś poziome - na linii poziomej przechodzącej przez środek obrazu. Z rysunku 7.23.b wynika, że w obrazie są 3 linie dokładnie pionowe oraz 3 linie o małym nachyleniu, zbliżone do poziomym.



Rysunek 7.23: Transformata Hougha fragmentu obrazu z zakłóconym prostokątem: a) obraz wejściowy; b) obraz transformaty Hougha tego obrazu

Dokładne wyznaczenie równań linii

Wyznaczanie transformaty Hough'a jest czasochłonne, a jej złożoność obliczeniowa rośnie wraz ze wzrostem rozdzielczości kątowej i wielkości obrazu. Dlatego stosuje się ją tylko jako wstępne narzędzie do wyznaczenia przybliżonych równań linii w obrazie. Następnie wyznacza się dokładne równania linii na podstawie regresji liniowej. Wynikiem tego etapu jest zbiór równań prostych przebiegających w znalezionym fragmencie i w jego pobliżu.

Regresja liniowa jest to metoda wyznaczania równania linii wyznaczonej przez zbiór punktów o znanych współrzędnych. Linia wyznaczona tą metodą leży w minimalnej odległości średniokwadratowej od wszystkich punktów. Podstawowe parametry linii: współczynniki a , b wyznacza się według wzorów:

$$a = \frac{n \sum xy - (\sum x)(\sum y)}{n \sum x^2 - (\sum x)^2}, \quad (7.1)$$

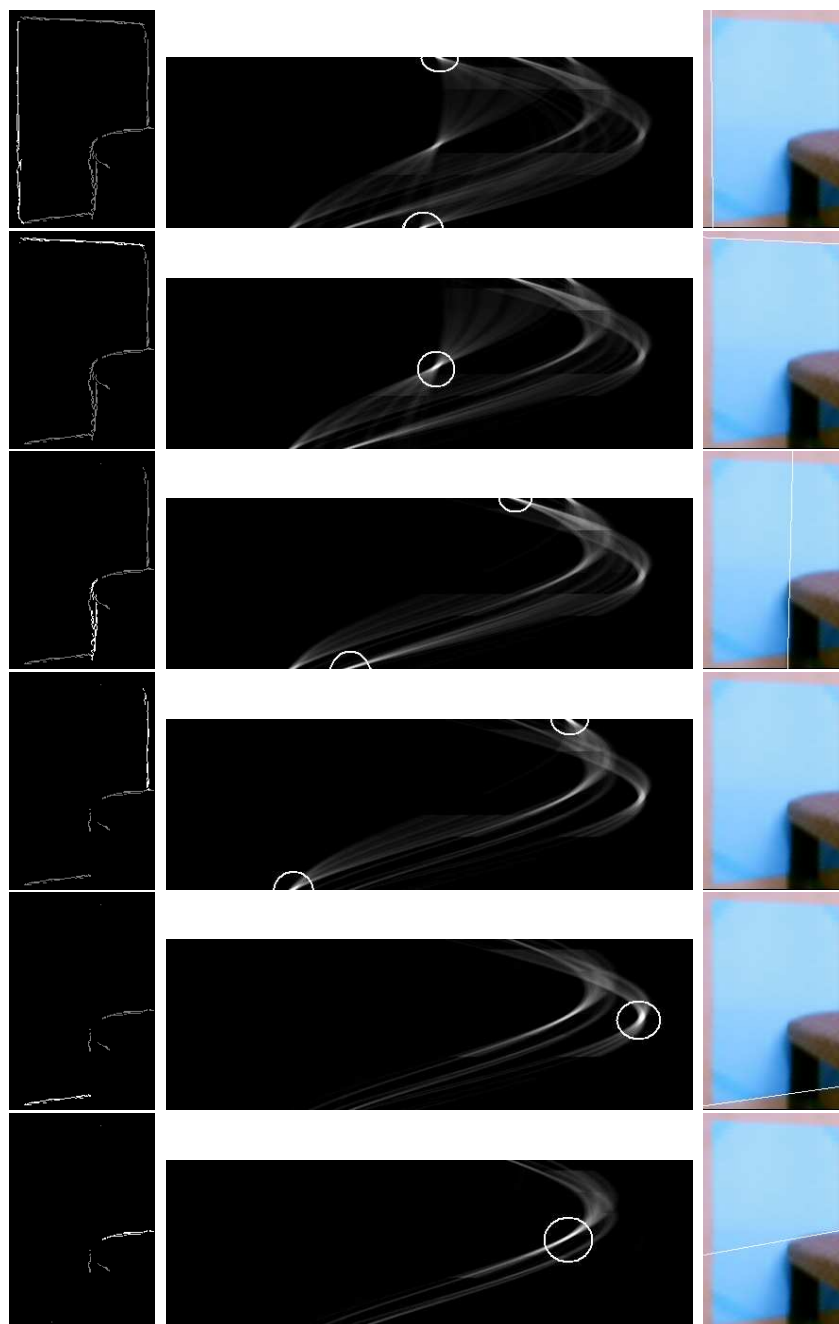
$$b = \frac{n \sum x^2 \sum y - \sum xy \sum x}{n \sum x^2 - (\sum x)^2}. \quad (7.2)$$

Punkty, które wyznaczają prostą muszą na niej leżeć wraz z otoczeniem, pomija się w ten sposób wpływ innych blisko położonych linii i bardzo krótkich fragmentów. Dla każdego punktu sprawdza się czy w odległości do 2 pikseli w kierunku wyznaczonym przez α znajdują się co najmniej 2 inne piksele współliniowe i tylko takie punkty zalicza się do zbioru punktów znalezionej prostej. Może to powodować efekt uboczny pozostawienia w transformacie obrazów pojedynczych pikseli.

W implementacji algorytmu pokazanego na rysunku 7.22 poszukuje się 16 najdłuższych linii prostych. Dla każdej znalezionej prostej zapamiętuje się jej długość i współczynniki c i α . Rysunek 7.24 pokazuje znormalizowane obrazy transformaty oraz wykrywane linie w kolejnych cyklach algorytmu. W lewej kolumnie znajdują się piksele wykryte przez detektor krawędzi i nie zaliczone jeszcze do żadnej prostej. Piksele o kolorze białym są zaliczone do najdłuższej linii wyznaczonej na podstawie transformaty, piksele szare należą do innych linii. W środkowej kolumnie jest obraz transformaty Hough'a (znormalizowany do przedziału $0 \div 255$). W miarę usuwania najdłuższych linii zwiększa się jasność pozostałych na skutek przyjętego odwzorowania, w którym najjaśniejszy punkt zawsze przyjmuje wartość 255. W prawej kolumnie zamieszczono prostą wykreśloną na podstawie regresji liniowej z białych punktów w kolumnie lewej. Prosta jest naniesiona na fragment źródłowego obrazu. Dobre dopasowanie prostych do rzeczywistych

krawędzi jest wyraźne w pierwszych wierszach. W miarę zmniejszania się długości linii ich dopasowanie do obrazu jest coraz gorsze. Poniżej progu ok. 20 punktów wykrywane proste stają się zbiorem przypadkowych pikseli, niekoniecznie odpowiadających rzeczywistym krawędziom w obrazie.

Wyznaczenie równania prostej metodą regresji daje wynik w postaci $y = ax + b$. Żeby uwzględnić także linie pionowe, proste opisuje się równaniami typu $cy = ax + b$, gdzie c może przyjmować tylko wartości 0 i 1. Wynikiem działania algorytmu wyznaczającego równania linii jest zbiór prostych opisanych trzema współczynnikami i uszeregowanych według długości.



Rysunek 7.24: Linie wykrywane w kolejnych cyklach algorytmu detekcji linii. W kolumnach od lewej: wejściowe piksele krawędzi z wyróżnioną najdłuższą linią; wejściowy obraz transformaty Hougha z wyróżnioną najdłuższą linią; wykryta linia na tle obrazu

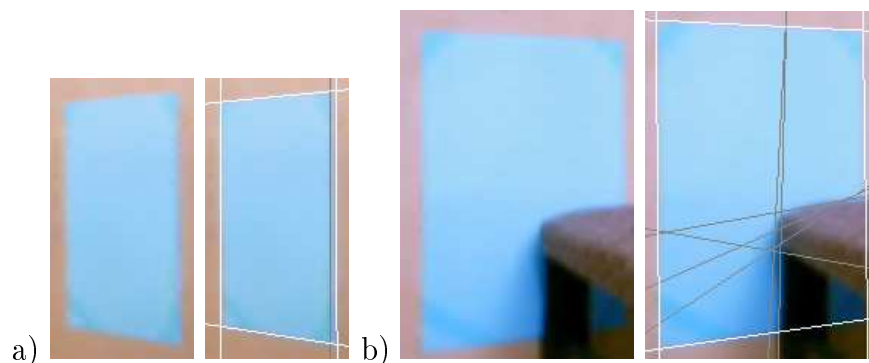
7.3.4 Określenie zewnętrznych krawędzi znacznika

Poprzednie etapy pozwalają na ustalenie zbioru linii prostych znajdujących się w sąsiedztwie obiektu. Po wyznaczeniu równań linii kolejnym etapem jest ustalenie które z nich są krawędziami obiektu. To jest zależne od kształtu obiektu a także od tego czy jest widoczny w całości czy częściowo zasłonięty. W pierwszym przybliżeniu można przyjąć, że obszar jest prostokątny i może być częściowo zasłonięty, tak jak na rysunkach przykładowych 7.20, 7.23 i 7.24. Warunkami uznania obszaru za prostokątny obiekt są:

1. obecność linii dokładnie pionowych lub prawie dokładnie pionowych ograniczających obszar z lewej i z prawej strony;
2. obecność linii ograniczających obszar z góry i z dołu, przy czym linie te nie mogą być odchylone od poziomu o więcej niż 45° .

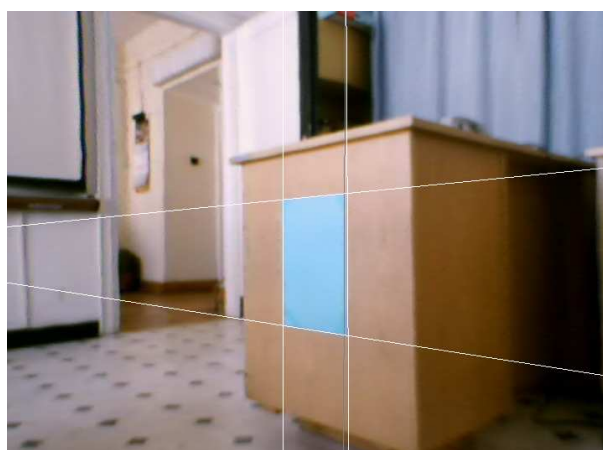
Różne możliwe zasłonięcia powodują różne widoczne kształty, wówczas wykrywane są więcej niż 4 linie proste. W takiej sytuacji wybiera się najdłuższą linię obecną z każdej z 4 stron. Jeżeli nie da się znaleźć linii z każdej strony, wówczas oznacza to że obiekt uznany za znacznik nie ma obrysu prostokątnego lub jest zbyt zniekształcony (albo zasłonięty).

Jeżeli linii ograniczających jest więcej niż jedna, za krawędź obszaru uznaje się najdłuższą z nich. Takie podejście umożliwia prawidłowe rozpoznawanie obiektów, których krawędzie są częściowo zasłonięte. Stopień zasłonięcia zależy od konturu obiektu zasłaniającego, jeżeli tworzą go proste wówczas nie może być zasłonięte więcej niż 40% dowolnej krawędzi (co pozostawia krawędź zewnętrzną jako najdłuższą z każdej strony), w przeciwnym wypadku jeżeli obiekt znaleziony jest duży zasłonięte może być więcej niż 50% długości brzegów obiektu (jeżeli kształt zasłaniający nie wprowadza dodatkowych prostych, np jest elipsą). W prezentowanym przykładzie na rysunkach 7.20, 7.23 7.24 i 7.25.b pomimo częściowego zasłonięcia prostokąta obiekt został prawidłowo rozpoznany jako prostokąt i zmierzony. Przy innym kształcie znacznika i jego znacznym zasłonięciu może wystąpić problem z identyfikacją za pomocą momentów i niezmienników metodami opisanym w rozdziale 7.2. Rysunek 7.26 przedstawia cały rysunek źródłowy 7.19 ze znalezionym znacznikiem prostokątnym z rysunku 7.25.a.



Rysunek 7.25: Wyodrębnienie poszukiwanego obszaru i wyznaczenie jego krawędzi: a) prostokąt bez zasłonięcia; b) prostokąt częściowo zasłonięty

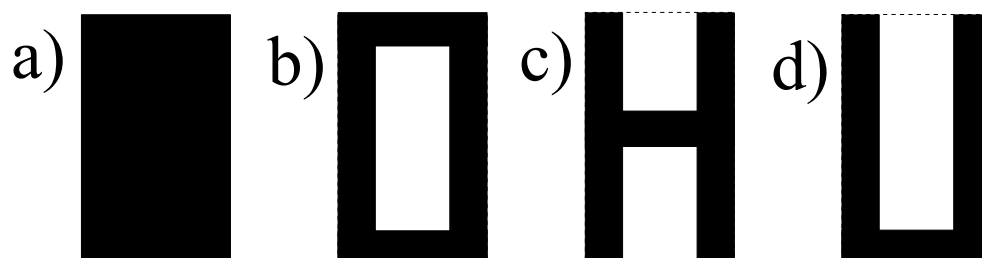
Pomiar rozmiarów obiektu (znacznika) prostokątnego odbywa się przez wyznaczenie punktów przecięć boków czworokąta, które zawierają się w 4 wyznaczonych prostych o równaniach wyznaczonych metodą regresji liniowej. Tak wyznaczone współrzędne narożników są bardzo dokładne (pod warunkiem że ilość wykrytych punktów krawędzi składających się na każdą



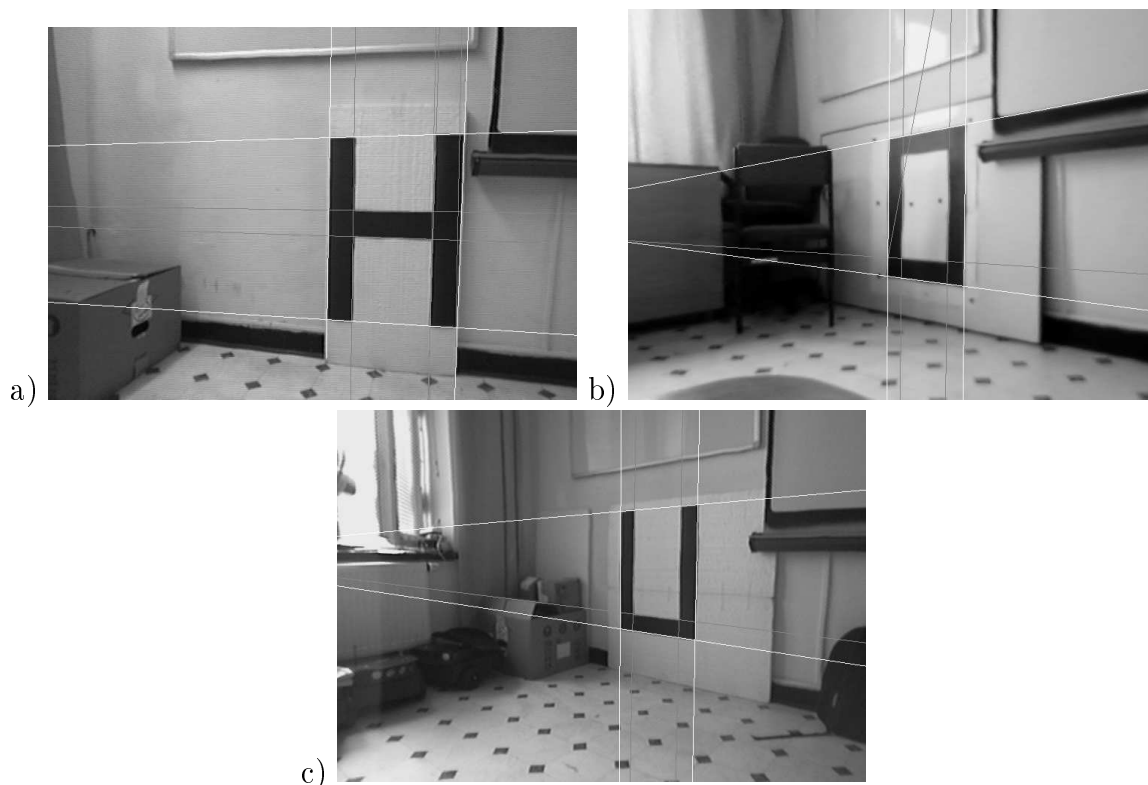
Rysunek 7.26: Obszar wyodrębniony z obrazu

prostą przekracza 20), ponieważ regresja liniowa zapewnia minimalizację błędu średniokwadratowego. W ten sposób uzyskano niezależność od wysokoczęstotliwościowych zakłóceń zawsze występujących na granicy obszarów o różnej jasności i barwie. W rozdziale 7.3.5 przedstawiono wyniki doświadczalne.

Uogólnienie pomiaru rozmiarów prostokąta do pomiaru rozmiarów obiektów (znaczników) o innych kształtach jest możliwe przynajmniej dla niektórych kształtów. Kształt inny niż prostokąt musi się dobrze wpisywać w kontur prostokątny, tzn. brzegi prostokąta opisanego na obiekcie muszą być styczne z najbardziej zewnętrznymi brzegami obszaru. Odchylenia od kształtu prostokątnego mogą być "do wewnątrz obszaru" (wcięcia), ale nie na zewnątrz (wypukłości). Przykładowe kształty spełniające te warunki prezentuje się na rysunku 7.27. Im dłuższe są krawędzie styczne do brzegów prostokąta, tym większa jest dokładność wyznaczanych współrzędnych. W kształtach innych niż pełny prostokąt znajdują się więcej niż 4 proste należące do znacznika. Jeżeli linie brzegowe styczne do opisanego prostokąta są krótkie, może to powodować błędną identyfikację brzegów z powodu możliwości wystąpienia przypadkowych krótkich linii zakłócających. Po wykryciu wszystkich linii dłuższych niż progowa, następuje ich klasyfikacja. Jeżeli linia przebiega dalej od środka obiektu niż 3 piksele, przydziela się ją do jednej z 4 stron, zaś linie bliższe środkowi są uznawane za nie należące do brzegów obiektu. Po znalezieniu pierwszej linii z danej strony jest uznawana za brzegową. Jeżeli zostanie znaleziona kolejna i jest położona dalej niż poprzednia "brzegowa", zostaje uznana za brzegową jeżeli jest nie krótsza niż 60% ostatniej prostej "brzegowej". Przykłady znaczników wykrytych i zmierzonych opisywaną metodą przedstawiają rysunki 7.27.a, 7.27.b, 7.27.c i 7.27.d.

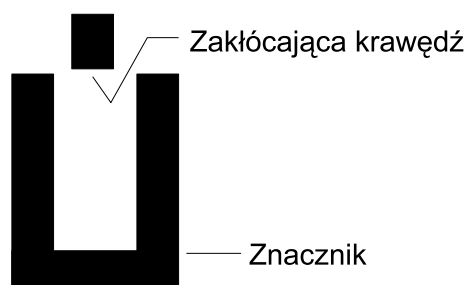


Rysunek 7.27: Przykładowe kształty znaczników dających się zmierzyć opisywaną metodą

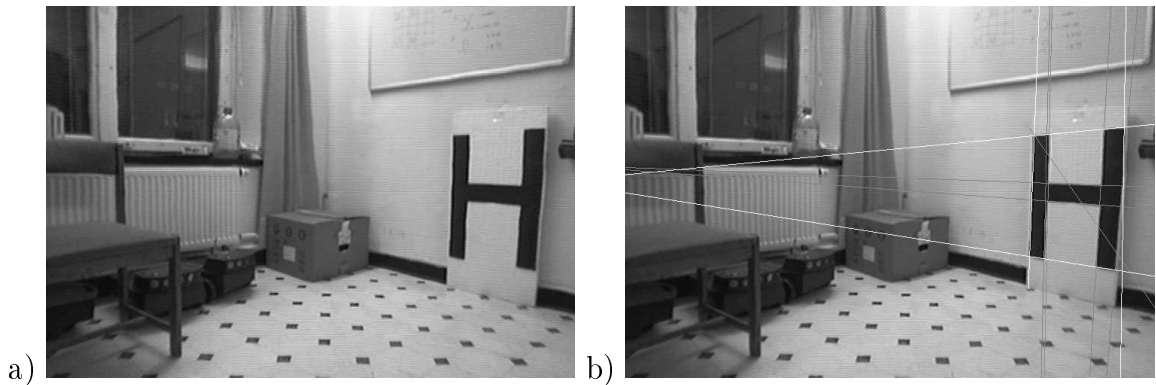


Rysunek 7.28: Wykrywanie i pomiar rozmiarów znaczników wpisanych w prostokąt: a) litera H, b) ramka, c) litera U

W przypadku gdy zewnętrzny kształt znacznika odbiega od prostokątnego (na rysunku 7.27.c i 7.27.d, w pobliżu brzegów prostokąta opisanego na znaczniku nie powinny znajdować się inne obiekty widoczne przy wyodrębnianiu obiektu podstawowego. Przykład tak zakłóconego znacznika przedstawia rysunek 7.29. Nie każde zakłócenie w obszarze prostokąta opisanego na znaczniku może zniekształcić równania prostych brzegowych i w związku z tym współrzędne narożników. Zakłócający wpływ na równania linii mogą mieć **tylko krawędzie "obcego" obiektu położone w pobliżu brzegu prostokąta opisanego na znaczniku i równoległe do tej krawędzi**. Przykład zakłóceń kształtu pokazuje rysunek 7.30 - są one spowodowane krawędzią i cieniem blisko prawej strony znacznika. W przedstawianym konkretnym przypadku (obraz jest w odcieniach szarości) rozwiązaniem problemu może być progowanie jasności oparte na średniej lokalnej jasności, opisane w [53].



Rysunek 7.29: Schemat znacznika z zakłóceniem, które może zniekształcić pomiary



Rysunek 7.30: Niedokładne określenie rozmiaru znacznika spowodowane nałożeniem zniekształceń oraz sąsiedztwem innych obiektów (po prawej stronie)

Wartość	Minimum	Maksimum	Średnia wartość	Wariancja
X0	297,3	297,6	297,35	0,01
Y0	294,4	294,9	294,54	0,02
X1	358,4	358,7	358,61	0
Y1	297,3	297,4	297,46	0,01
X2	294,8	295,2	295,17	0,01
Y2	395,3	396,2	395,67	0,04
X3	358,0	358,3	358,15	0
Y3	403,7	404,3	403,91	0,03

Tablica 7.6: Własności statystyczne wyznaczania współrzędnych narożników znacznika dla 1000 próbek.

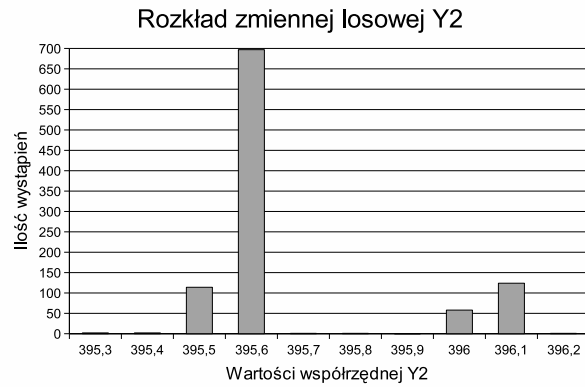
7.3.5 Dokładność wykrywania krawędzi

Ze względu na uśrednianie danych, zmienna losowa wartości ma małą wariancję. Tabela 7.6 przedstawia wyniki uzyskane podczas jednego z testów. Test polegał na wyznaczeniu współrzędnych znacznika 1000 razy przy nieruchomej kamerze.

Rysunek 7.31 przedstawia rozkład zmiennej losowej Y2, która w omawianym przykładzie ma największą wariancję.

Z rozkładów zamieszczonych w tabeli 7.6 i na rysunku 7.31 wynika, że wariancja współrzędnych traktowanych jako zmienne losowe jest mała. Można przyjąć, że nie popełnia się dużego błędu przyjmując że wartość uzyskana z jednego pomiaru jest wartością dokładną. Dla zmiennej Y2 prawdopodobieństwo uzyskania wartości 395,6 wynosi 0,7. Wartości skrajne uzyskane podczas testu różnią się bardzo mało, mniej niż 1 piksel. Błąd wynikający z tego podejścia jest mniejszy niż błąd wprowadzany przez korekcję geometrii, wynoszący 1 piksel (zob rozdz. 4.3.2).

Przy zaokrągleniu współrzędnych do 1 piksela można uzyskać co najwyżej 2 wartości współrzędnych. Rzeczywista rozdzielczość metody pomiaru współrzędnych przez regresję wynosi mniej niż 1 piksel. Dowodem jest zmniejszenie minimalnego błędu lokalizacji opartej na opisanych pomiarach znacznika z 15 cm przy zaokrągleniu do 1 piksela do 2 cm przy zachowaniu części ułamkowej, co opisano w rozdziałach 8 i 9. Zmiana błędu lokalizacji przy zachowaniu



Rysunek 7.31: Rozkład zmiennej losowej Y2, jako najbardziej rozproszona dystrybucja

części ułamkowej współrzędnych wynosi prawie jeden rząd, ale zależność błędów lokalizacji od błędu współrzędnych nie jest liniowa. Można przyjąć, że podając wynik z dokładnością do 0,1 piksla ostatnia cyfra jest znacząca, a może nawet dokładna.

Rozdział 8

Proponowana metoda lokalizacji

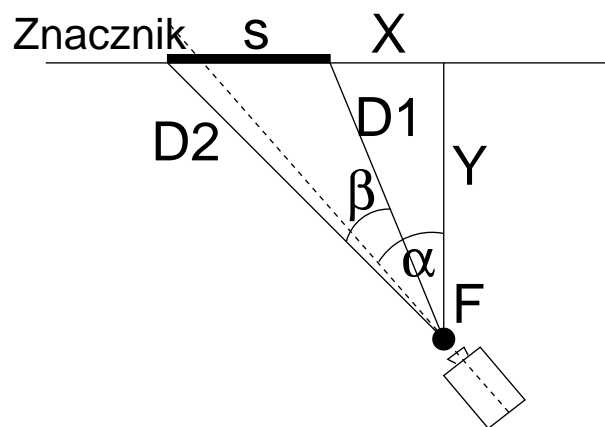
W niniejszym rozdziale opisano technikę wyznaczania położenia kamery (zamontowanej na robocie) względem znacznika. Znacznik musi być wykrywalny metodami opisanymi w rozdziale 7. Warunek ten spełnia znacznik jednobarwny i o zewnętrznym kształcie prostokątnym lub dobrze wpisującym się w prostokątną obwiednię (zob. rozdział 7.3). W dalszej części pracy przyjęto dla uproszczenia, że znacznik jest prostokątem. W przypadku innych kształtów znaczników słowo "prostokąt" odnosi się do prostokątnej obwiedni znacznika. Lokalizacja na podstawie znacznika prostokątnego wymaga uprzedniej detekcji i pomiarów znacznika. W niniejszym rozdziale przyjęto, że znacznik został znaleziony w obrazie, zidentyfikowany i zmierzony za pomocą metod opisanych w rozdziale 7. Lokalizacja względem znacznika wymaga uprzedniego wyznaczenia długości zewnętrznych krawędzi pionowych znacznika, azymutu tych krawędzi względem osi optycznej kamery i odległości do tych krawędzi od ogniska kamery.

Rzut z góry znacznika widzianego przez kamerę przedstawia rysunek 8.1. Wszystkie wyznaczone wzory umożliwiające wyznaczenie położenia mają początek w schemacie 8.1. Elementy stałe układu to znacznik i jego znana szerokość (S). Elementy mierzone to:

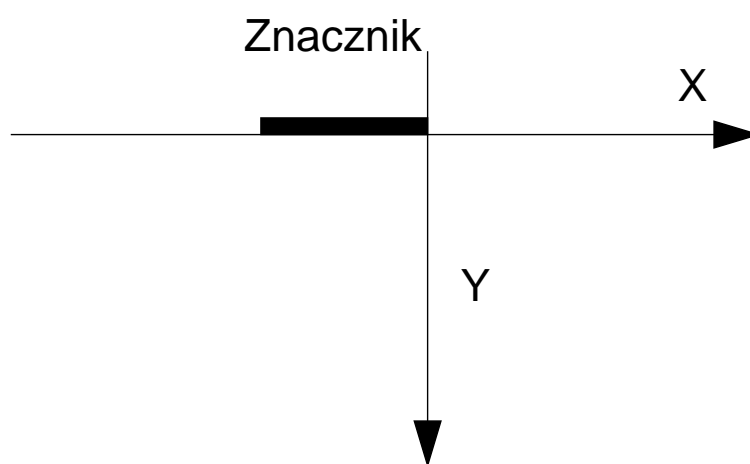
- odległość od ogniska kamery do prawego brzegu znacznika, oznaczana dalej symbolem $D1$,
- odległość od ogniska kamery do lewego brzegu znacznika $D2$,
- kąt widzenia szerokości znacznika β ,
- odległość od ogniska do płaszczyzny znacznika Y ,
- odległość od prawego brzegu znacznika do punktu najbliższego do ogniska na płaszczyźnie znacznika X ,
- kąt między osią optyczną kamery a kierunkiem prostopadłym do płaszczyzny znacznika α .

We wszystkich obliczeniach pominięto ogniskową kamery f . Pozwala to na uproszczenie obliczeń i przedstawienie przejrzystszych wyrażeń opisujących współrzędne, które i tak są dość skomplikowane. W eksperymentach stosowano kamery o bardzo krótkich ogniskowych, wynoszących kilka mm.

We wszystkich dotyczących lokalizacji stosowano układ współrzędnych związany ze znacznikiem, zorientowany jak na rysunku 8.2. Może on być układem globalnym, lub przesuniętym względem układu globalnego lokalnym układem znacznika.



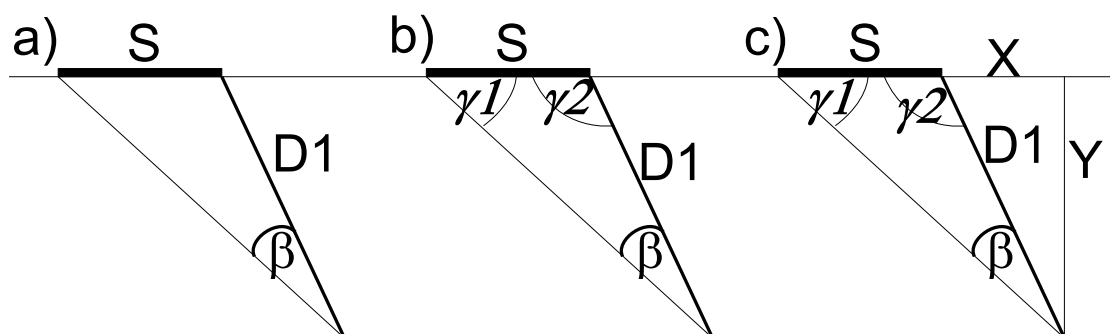
Rysunek 8.1: Elementy geometryczne układu znacznika i kamery



Rysunek 8.2: układ współrzędnych związany ze znacznikiem

Ponieważ znacznik został zidentyfikowany, znane są zarówno jego rzeczywiste rozmiary (w jednostkach metrycznych), jak i rozmiary obrazowe – szerokość obrazowa oraz wysokości prawej i lewej krawędzi. Znajomość wymienionych trzech wielkości mierzonych oraz rzeczywistej szerokości znacznika i wysokości obu krawędzi pozwala na wyznaczenie bezpośrednio odległości $D1$ i $D2$ oraz kąta β . Położenie kamery względem znacznika, opisywane przez X , Y i α , można wyznaczyć dwoma metodami.

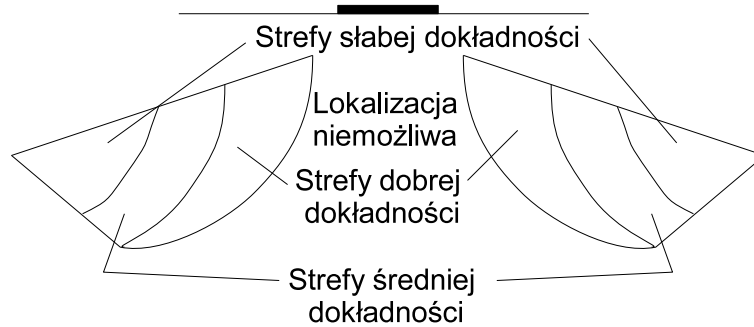
W pierwszej metodzie pozę robota wyznacza się na podstawie znanej szerokości znacznika, zmierzonej odległości do jednej z pionowych krawędzi znacznika oraz obserwowanej szerokości kątowej znacznika, co opisano w rozdziale 8.4. Schematycznie etapy pozycjonowania przedstawia rysunek 8.3. Szerokość znacznika S jest znana przed pomiarem. Na podstawie zmierzonej odległości $D1$ oraz kąta β i twierdzenia sinusów wyznacza się wartość kąta $\gamma1$. Przedtem trzeba ustalić czy jest to kąt ostry czy rozwarty, do tego wystarczy porównać ze sobą odległości do obu krawędzi znacznika. Następnie wyznacza się kąt $\gamma2$ i znane są wszystkie kąty, co pozwala na ustalenie pozycji robota. Metoda jest dość dokładna, jednak ma duże martwe pole w obszarze przed znacznikiem. W martwej strefie możliwe jest wyznaczenie odległości między robotem a krawędzią znacznika, ale nie jest możliwe wyznaczenie odległości oznaczonej jako X na rysunku 8.2. Obecność martwego pola wynika z praktycznie jednakowej szerokości kątowej znacznika obserwowanego z różnych miejsc przy samej odległości od krawędzi. Ponieważ w metodzie potrzeba odległości do jednej z dwu pionowych krawędzi znacznika, można wyznaczyć pozę robota na podstawie odległości do lewej lub prawej krawędzi znacznika. Rozkład stref dla tej metody przedstawia rysunek 8.4.



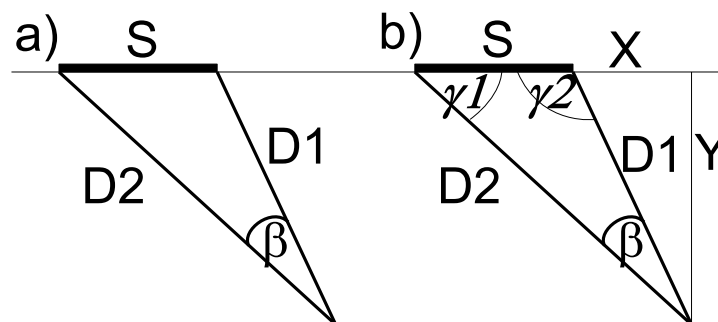
Rysunek 8.3: Schemat etapów wyznaczania pozycji robota metodą opartą o pomiar jednej odległości: a) wielkości mierzone bezpośrednio ($D1$, β) i znane (S); b) wielkości wyznaczone pośrednio ($\gamma1$, $\gamma2$); c) pozycja robota (X i Y)

W drugiej z metod pozycjonowania wyznacza się pozę robota na podstawie znanej szerokości znacznika oraz zmierzonych odległości do obu krawędzi znacznika, co opisano w rozdziale 8.5. Metoda jest mniej dokładna niż poprzednia, jednak nie ma martwej strefy przed znacznikiem. Schematycznie etapy pozycjonowania przedstawia rysunek 8.5. Znając wielkości $D1$, $D2$ i S można wyznaczyć wszystkie kąty, w szczególności kąt $\gamma2$. Przy wyznaczaniu współrzędnych X i Y znajomość wartości tego kąta nie jest potrzebna, ponieważ można wyprowadzić wzory dla X i Y zależny tylko od $D1$, $D2$ i S (zob rozdz. 8.5). Metoda jest mniej dokładna od poprzedniej, i ma mniejszą strefę przydatności – zob. rys. 8.6.

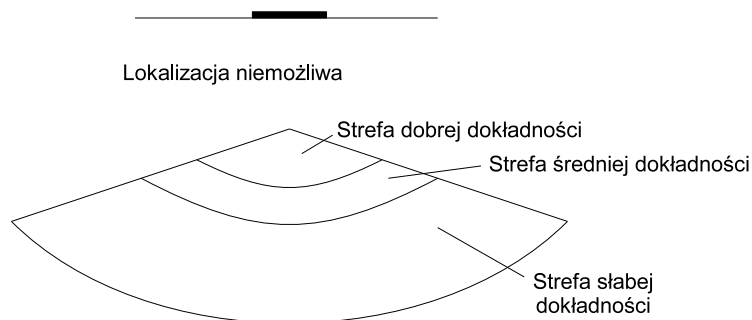
Różnice w dokładności metod pozycjonowania wynikają z ilości przekształceń mierzonych wielkości. Pierwszą wyznaczaną wielkością są azymuty krawędzi $\beta1$ i $\beta2$, na ich podstawie oraz widocznych wysokości można ustalić odległości $D1$ i $D2$. W metodzie pierwszej (z jedną



Rysunek 8.4: Strefy dokładności lokalizacji metodą z jedną odległością



Rysunek 8.5: Schemat etapów wyznaczania pozycji robota metodą opartą o pomiar dwu odległości: a) wielkości mierzone bezpośrednio ($D1$, $D2$, β) i znane (S); b) pozycja robota (X i Y)



Rysunek 8.6: Strefy dokładności lokalizacji metodą z dwoma odległościami

odległością do krawędzi) wyznacza się bezpośrednio 2 wielkości (β_1 i β), oraz jedną wartość zależną $D1$. W metodzie drugiej (z dwoma odległościami) wyznacza się dwie wielkości bezpośrednio (β_1 i β_2), oraz dwie wielkości zależne $D1$ i $D2$.

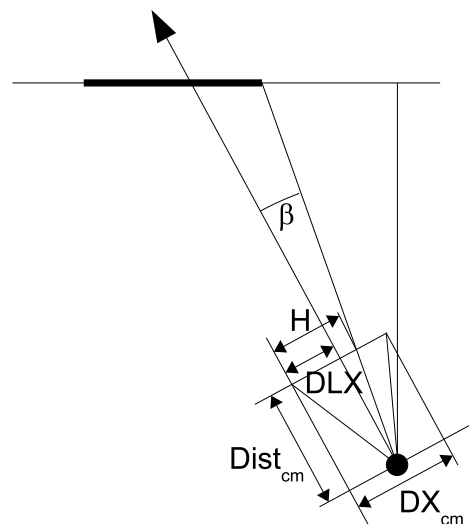
Strefy przydatności obu metod uzupełniają się umożliwiając dokładną lokalizację w całym obszarze przed znacznikiem, pod warunkiem że kąt między osią kamery a płaszczyzną znacznika nie jest zbyt ostry, tzn. jest powyżej 15° .

8.1 Wyznaczenie azymutu do obiektu

Wyznaczenie azymutu jest podstawą do wyznaczenia odległości kamery od obiektu lub jego części. Azymut wyznacza się na podstawie współrzędnych ekranowych obiektu oraz parametrów wyznaczanych podczas kalibracji kamery w rozdziale 4. Rzut z góry prezentuje rysunek 8.7. Znaczenie parametrów jest identyczne jak na rysunku 4.12:

- $Dist_{cm}$ oznacza odległość od ogniska kamery do siatki kalibrującej w jednostkach metrycznych,
- DX_{cm} oznacza szerokość pola widzenia siatki kalibrującej po kalibracji w jednostkach metrycznych,
- DLX_p oznacza współrzędną X kierunku osi optycznej w pikselach po korekcji obrazu, zaś DLX_{cm} odległość w jednostkach metrycznych między brzegiem pierwszej widocznej kratki z lewej strony (po korekcji obrazu) a punktem w który jest skierowana oś optyczna kamery. Schematycznie przedstawiono układ współrzędnych na rysunku 8.2. Oznaczenia te stosowano we wszystkich wzorach tego rozdziału.

Kierunek osi optycznej kamery



Rysunek 8.7: Parametry niezbędne do wyznaczenia azymutu do obiektu

Dodatkowo, H_p oznacza współrzędną obrazową X obiektu do którego wyznacza się azymut (w pikselach). Azymut oznaczono symbolem β . Symbolem H_{cm} oznaczono odległość między brzegiem pierwszej widocznej kratki z lewej (na obrazie kalibrującym) a obiektem gdyby został

naniesiony na siatkę kalibrującą (i byłby tak samo widoczny jak podczas pomiaru). Podstawowy wzór jest następujący:

$$\beta = \arctan \frac{(H_{cm} - DLX_{cm}) \text{ w jednostkach metrycznych}}{Dist_{cm}}. \quad (8.1)$$

Licznik równania (8.1) można wyznaczyć ze wzoru (8.2):

$$\frac{H_{cm} - DLX_{cm}}{DX_{cm}} = \frac{H_p - DLX_p}{DX_p}, \quad (8.2)$$

gdzie DX_p oznacza szerokość obrazu po korekcji w pikslach. Ostatecznie azymut wyraża się wzorem:

$$\beta = \arctan \frac{(H_p - DLX_p) * DX_{cm}}{DX_p * Dist_{cm}}. \quad (8.3)$$

Poniżej przedstawiono obliczenia wykazujące stabilność funkcji azymutu, przede wszystkim jej ciągłość. Jednocześnie pokazano wrażliwość funkcji wyznaczającej azymut na małe błędy (różnice) wartości wejściowych. Azymut wyznacza się według równania (8.3). Stabilność sprawdzono wyznaczając pochodne wg wszystkich wartości. Jako zmienne traktuje się wartości mierzone oraz stałe wyznaczone podczas kalibracji kamery. Równania (8.4) ÷ (8.7) opisują pochodne cząstkowe względem zmiennych równania (8.3), odpowiednio względem wartości H oraz parametrów kamery: DLX_{cm} , DX_{cm} i $Dist_{cm}$. Mianowniki wszystkich wyrażeń są identyczne. Dzielenie przez 0 może wystąpić tylko w wypadku, gdyby wartość 0 przyjął DX_p (rozmiar poziomy obrazu), DX_{cm} czyli szerokość obszaru podczas kalibracji lub $Dist_{cm}$, czyli odległość między ogniskiem kamery a siatką podczas kalibracji. Wszystkie te wielkości są zawsze dodatnie, więc funkcja wyznaczająca azymut jest ciągła. W niekorzystnym przypadku może wystąpić $H_p - DLX_p = 0$, ale nie spowoduje to dzielenia przez 0. Wartości liczbowe pochodnych powstały z podstawienia do wzorów (8.4) ÷ (8.7) wartości wyznaczonych dla kamery stosowanej w badaniach: $DX_p = 722px$, $DX_{cm} = 19cm$, $Dist_{cm} = 16cm$, $DLX_p = 418px$. Wartości liczbowe wyraźnie wskazują na małą wrażliwość wyników na wahania i błędy danych wejściowych oraz parametrów kalibracji. Najbardziej znaczące (o jednakowej wadze) są błędy wyznaczenia położenia obiektu H_p oraz położenia osi optycznej w obrazie DLX_p , zaś kilkanaście razy mniejsze znaczenie mają błędy wyznaczenia odległości $Dist_{cm}$ i DX_{cm} .

$$\frac{\partial \beta}{\partial H_p} = \frac{Dist_{cm} DX_{cm} DX_p}{(H_p - DLX_p)^2 DX_{cm}^2 + Dist_{cm}^2 DX_p^2} = \frac{608}{369664 + (H_p - DLX_p)^2} \quad (8.4)$$

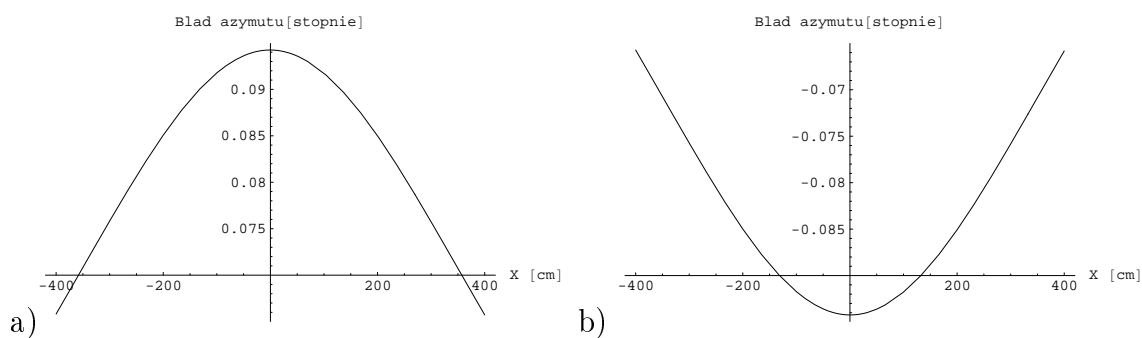
$$\frac{\partial \beta}{\partial DLX_p} = -\frac{Dist_{cm} DX_{cm} DX_p}{(H_p - DLX_p)^2 DX_{cm}^2 + Dist_{cm}^2 DX_p^2} = -\frac{608}{369664 + (H_p - DLX_p)^2} \quad (8.5)$$

$$\frac{\partial \beta}{\partial DX_{cm}} = \frac{Dist_{cm} (H_p - DLX_p) X_{Img}}{(H_p - DLX_p)^2 DX_{cm}^2 + Dist_{cm}^2 X_{Img}^2} = \frac{32(X - DLX_{cm})}{369664 + (H_p - DLX_p)^2} \quad (8.6)$$

$$\frac{\partial \beta}{\partial Dist_{cm}} = \frac{DX_{cm} (H_p - DLX_p) DX_p}{(H_p - DLX_p)^2 DX_{cm}^2 + Dist_{cm}^2 DX_p^2} = -\frac{38(X - DLX_{cm})}{369664 + (H_p - DLX_p)^2} \quad (8.7)$$

Jakość pomiaru określono jego wrażliwością na błąd wartości wejściowej. Wykresy błędu kąta dla przyjętych błędów wejściowych przedstawia rysunek 8.8. Z danych pomiarowych zamieszczonych w rozdziale 7.3 wynika, że błąd wartości wejściowej jest mniejszy niż 1 piksel.

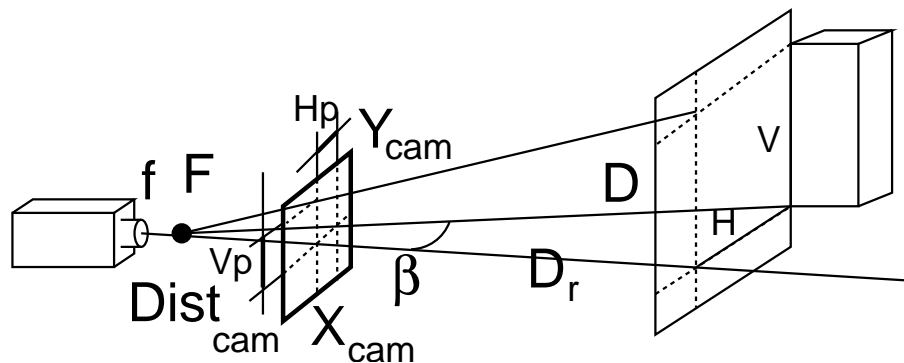
Zakładam, że wynosi on dla wykresu 8.8.a 1 piksel i dla 8.8.b -1 piksel. Z wykresu wynika, że wartość błędu i rozdzielczość pozioma są dużo mniejsze niż 1 stopień, a nawet $0,1$ stopnia, co jest oczywiste jeżeli porówna się poziomą rozdzielczość liniową obrazu (720 pikseli po korekcji) z kątem widzenia kamery wnoszącym ok 50° . Dla różnych kamer wartości parametrów modelu (zob rozdz 4) są różne, jednak zawsze rozdzielczość kątowna i błąd azymutu są znacznie mniejsze niż 1° .



Rysunek 8.8: Zależność błędu wyznaczenia azymutu od błędu wejściowych współrzędnych wynoszącego: a) $+1$ piksel, b) -1 piksel

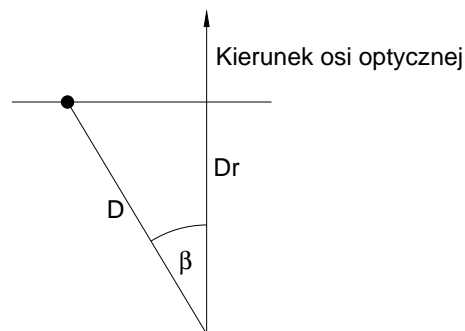
8.2 Wyznaczenie odległości do obiektu o znanej wysokości

Pomiar odległości odbywa się na podstawie porównania znanej wysokości obiektu z wysokością obrazową. Na obrazie mierzy się widoczną wysokość obiektu oraz jego odległość od punktu przecięcia osi optycznej z obrazem. Na rysunku 8.9 przedstawiono oznaczenie symboliczne wielkości mierzonych oraz rzut widoku obiektu na płaszczyznę obrazu kalibrującego. Znaną wysokość obiektu oznaczono jako V , zaś jego odległość w płaszczyźnie poziomej od osi optycznej kamery jako H . Wielkości te są wyrażone w metrycznych jednostkach miary odległości.



Rysunek 8.9: Wyznaczanie odległości do obiektu o znanej wysokości

W rzucie z góry, zamieszczonym na rysunku 8.10, pokazano, że wysokość obrazowa obiektu V_p zależy od odległości między ogniskiem kamery a płaszczyzną prostopadłą do osi optycznej zawierającą widoczny obiekt. Jest to spowodowane rzutowaniem obiektów z tej samej odległości na jedną płaszczyznę podczas korekcji obrazu (zob. rozdział 4). Zależność tę uwzględniono przy wyprowadzeniu (wzoru 8.12).



Rysunek 8.10: Wyznaczanie odległości do obiektu: rzut z góry

Na wszystkich wykresach w niniejszym rozdziale przyjęto następujące rozmiary znacznika: wysokość $V = 70\text{cm}$, szerokość $H = 50\text{cm}$. Są to rozmiary jednego ze znaczników stosowanych w eksperymentach. Z kolei na obrazie zmierzona odległość od osi optycznej oznaczono $H_p - DLX_p$, zaś zmierzona wysokość – V_p . Obie te wartości wyrażone są w pikselach. Z danych uzyskanych podczas kalibracji kamery można uzyskać odpowiadające im wartości H_{cm} i V_{cm} wyrażone w jednostkach metrycznych. Są to wysokość obiektu i jego położenie w poziomie jakie miałby obiekt naniesiony na siatkę kalibrującą, przy założeniu że obiekt na siatce kalibrującej wygląda

tak samo widoczny jak podczas pomiaru. Oznaczając rozdzielczość poziomą i pionową obrazu jako DX_p i DY_p i rozważając rzut obrazu na płaszczyznę otrzymuje się następujące równania:

$$H_m = \frac{H_p - DLX_p}{DX_p} DX_{cm}, \quad (8.8)$$

$$V_m = \frac{V_p}{DY_p} DY_{cm}. \quad (8.9)$$

Następujące równania wyznaczają odległość D obiektu od ogniska kamery:

$$D_r = Dist_{cm} * \frac{V}{V_m} = Dist_{cm} \frac{DY_p * V}{V_p * DY_{cm}}, \quad (8.10)$$

$$\beta = \arctan \frac{H_m}{Dist_{cm}} = \arctan \frac{(H_p - DLX_p) DX_{cm}}{DX_p Dist_{cm}}, \quad (8.11)$$

$$D = \frac{D_r}{\cos \beta} = D_r \sqrt{1 + \tan^2 \beta} = Dist_{cm} \frac{DY_p * V}{V_p * DY_{cm}} \sqrt{1 + \frac{(H_p - DLX_{cm})^2 * DX_{cm}^2}{DX_p^2 Dist_{cm}^2}} = \frac{DY_p * V}{DX_p * V_p * DY_{cm}} \sqrt{DX_p^2 * Dist_{cm}^2 + (H_p - DLX_{cm})^2 * DX_{cm}^2}. \quad (8.12)$$

8.2.1 Dokładność wyznaczenia odległości

Wrażliwość odległości na błąd określają pochodne cząstkowe. Pochodne cząstkowe odległości względem parametrów kamery wyrażają równania (8.13) ÷ (8.16). Pochodne cząstkowe względem wysokości obrazowej obiektu i odległości poziomej obiektu od rzutu osi optycznej wyrażają wzory (8.17) i (8.18). Ze względu na skomplikowane równania nie przeprowadzono pełnej analizy tych funkcji, lecz jedynie podano analizę ciągłości i oszacowania dokładności dla określonych warunków.

Wszystkie pochodne wyrażone wzorami (8.13) ÷ (8.16) są ciągłe, po warunkiem że wyrażenie pod pierwiastkiem jest nie ujemne, a mianowniki większe niż 0. Warunek nieujemności jest spełniony, jeżeli $Dist_{cm}^2 DX_p^2 \geq 0$. Warunek ten jest spełniony zawsze, ponieważ są to dwie stałe dodatnie: odległość przy kalibracji $Dist_{cm} > 0$ i rozmiar obrazu w pikselach DX_p . Dla współrzędnej $H_p = DLX_p$ drugi składnik po pierwiastkiem jest zerowy, a dla innego H_p dodatni. W mianownikach wszystkich pochodnych występują zmienna V_p i stałe dodatnie DX_p (rozmiar pionowy widzianego obiektu i rozdzielczość pozioma obrazu) oraz wyrażenie obecne we wzorach (8.4) ÷ (8.7) na azymut. Nie powodują one możliwości wystąpienia dzielenia przez 0. We wzorze (8.12) jest dzielenie przez $\cos \beta$, i później się ono upraszcza. Teoretycznie dzielenie przez 0 może wystąpić gdy $\cos \beta = 0 \Rightarrow \beta = \pm \frac{\pi}{2}$. Żadna kamera (jeżeli nie zastosuje się np zwierciadła parabolicznego) nie pozwala widzieć obiektu znajdującego się w kierunku pod kątem prostym do osi optycznej. Można założyć, że $\cos \beta \neq 0$.

$$\frac{\partial D}{\partial DX_{cm}} = \frac{V DX_{cm} (H_p - DLX_{cm})^2 DY_p}{DY_{cm} V_p DX_p \sqrt{Dist_{cm}^2 DX_p^2 + DX_{cm}^2 (H_p - DLX_p)^2}} = \frac{70}{19 V_p \sqrt{133448704 + DX_{cm}^2 (H_p - 418)^2}} \quad (8.13)$$

$$\frac{\partial D}{\partial DY_{cm}} = \frac{V DY_p}{DY_{cm}^2 V_p DX_p} \sqrt{Dist_{cm}^2 DX_p^2 + DX_{cm}^2 (H_p - DLX_p)^2} =$$

$$-980 \frac{\sqrt{544388 + H_p(H_p - 836)}}{V_p DY_{cm}^2} \quad (8.14)$$

$$\frac{\partial D}{\partial Dist_{cm}} = \frac{Dist_{cm} V DX_p Y_p}{V_p DY_{cm} \sqrt{Dist_{cm}^2 DX_p^2 + DX_{cm}^2 (H_p - DLX_p)^2}} =$$

$$\frac{101080 Dist_{cm}}{V_p \sqrt{1444 Dist_{cm}^2 + (H_p - 418)^2}} \quad (8.15)$$

$$\frac{\partial D}{\partial DLX_{cm}} = -\frac{V DX_{cm}^2 (H_p - DLX_{cm}) DY_p}{V_p DY_{cm} DX_p \sqrt{Dist_{cm}^2 DX_p^2 + DX_{cm}^2 (H_p - DLX_p)^2}} =$$

$$-\frac{70(H_p - DLX_{cm})}{V_p \sqrt{369664 + (H_p - DLX_{cm})^2}} \quad (8.16)$$

$$\frac{\partial D}{\partial H_p} = \frac{V DX_{cm}^2 (H_p - DLX_p) DY_p}{V_p DY_{cm} DX_p \sqrt{Dist_{cm}^2 DX_p^2 + DX_{cm}^2 (H_p - DLX_p)^2}} =$$

$$\frac{70(H_p - 418)}{V_p \sqrt{544388 + H_p(H_p - 418)}} \quad (8.17)$$

$$\frac{\partial D}{\partial V_p} = -\frac{V DY_p}{DY_{cm} V_p^2 DX_p} \sqrt{Dist_{cm}^2 DX_p^2 + DX_{cm}^2 (H_p - DLX_p)^2} =$$

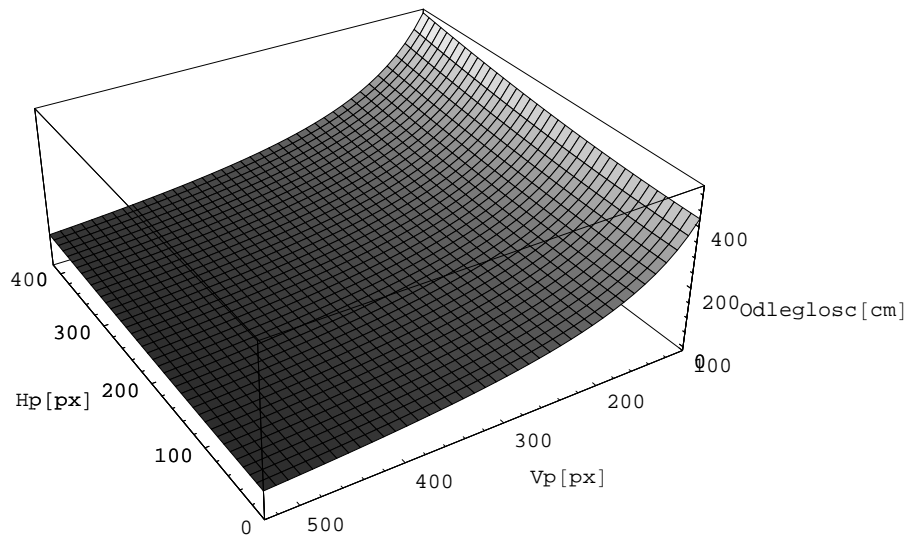
$$-\frac{70}{V_p^2} \sqrt{544388 + H_p(H_p - 836)} \quad (8.18)$$

Dokładność wyznaczenia odległości jest ograniczona skończoną rozdzielczością poziomą i pionową obrazu. Dokładność wyznaczono licząc różnicę wartości funkcji przy zmianie współrzędnych obrazowych o 1 piksel.

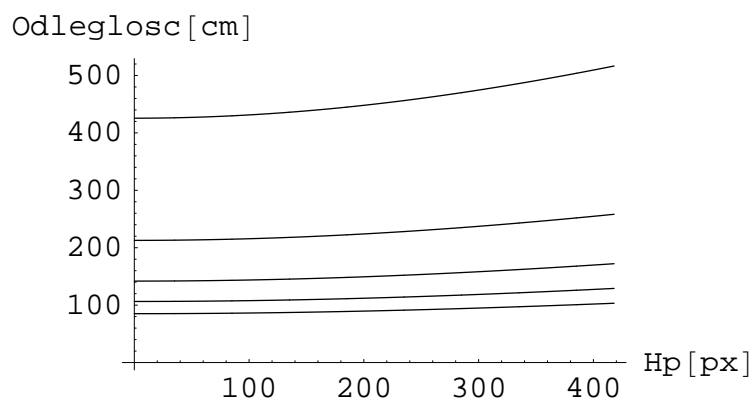
Wartości $Dist_{cm} = 16,0\text{cm}$, $DX_{cm} = 19,0\text{cm}$, $DY_{cm} = 28,0\text{cm}$, $DX_p = 722$, $DY_p = 532$, $DLX_p = 418$ wyznaczono podczas kalibracji kamery metodą opisaną w rozdziale 4.

Odległość D i dokładność jej wyznaczenia zależą od kąta β , dlatego poniższe wykresy przedstawiają zależność zmiany D od V_p i β . Rysunek 8.11 przedstawia zależność wyznaczonej odległości D od rozmiaru obrazowego V_p i odległości obrazowej od osi optycznej H_p . Przekroje tego wykresu wzdłuż osi wysokości obrazowej V_p i odległości poziomej od osi optycznej H_p są przedstawione na rysunkach 8.12 i 8.13. Rysunek 8.12 przedstawia zależność D od obrazowej odległości poziomej obiektu od osi optycznej H_p dla wybranych wartości wysokości obrazowej V_p : 100, 200, 300, 400 i 500 pikseli. Zależność ta jest rosnąca szybciej niż liniowo. Wpływ azymutu na zmierzoną odległość nie jest bardzo duży. Odległość jest tylko elementem umożliwiającym wyznaczenie położenia kamery i pominięcie azymutu w modelu powodowałoby znaczne błędy

pomiarów. Rysunek 8.13 przedstawia zależność odległości D od wysokości obrazowej obiektu V_p dla wybranych wartości odległości od osi optycznej H_p : 0, 100, 200, 300 i 400 piksli. Zależność jest w przybliżeniu hiperboliczna. Z wykresu wynika, że dla zastosowanej kamery i znacznika o wysokości 70 cm zakres odległości w których znacznik jest dobrze widoczny i można tę odległość mierzyć wynosi 5 metrów. Skala wykresów kończy się na wartości wysokości obrazowej $V_p = 100px$, ponieważ mniejsze rozmiary obrazowe znacznika powodują gorszą dokładność pomiarów i mogą utrudniać rozpoznanie znacznika.

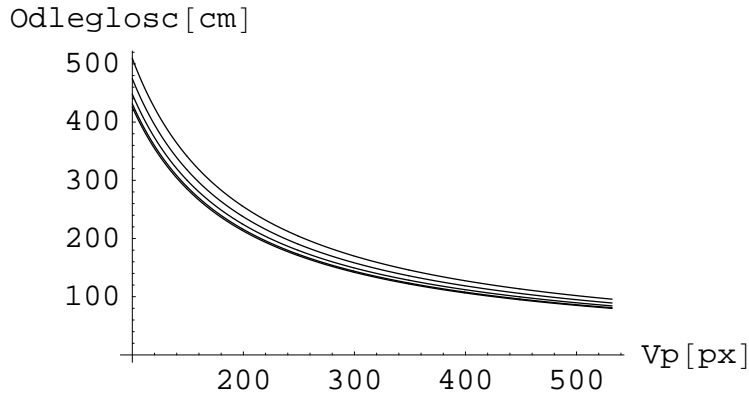


Rysunek 8.11: Zależność wyznaczonej odległości D od wysokości obrazowej V_p i odległości od osi optycznej H_p



Rysunek 8.12: Zależność wyznaczonej odległości D odległości od osi optycznej H_p

Zdolność rozdzielczą wyznaczenia odległości określono jako maksymalny błąd pomiaru przy najbardziej niekorzystnej kombinacji zakłóceń wszystkich wartości wyznaczanych z obrazu.



Rysunek 8.13: Zależność wyznaczonej odległości D od wysokości obrazowej V_p

Przyjęty błąd wyznaczenia położenia brzegu obiektu w obrazie wynosi jeden piksel. Zdolność rozdzielczą wyznaczono przy zmianie odległości poziomej H_p , związanej z azymutem, o 1 piksel oraz wysokości V_p o dwa piksele (ponieważ V_p wyznaczają współrzędne dwóch punktów). Rysunek 8.14 przedstawia różnicę odległości D przy zmianie wysokości obrazowej V_p o **a)** jeden piksel i **b)** dwa piksele dla wartości H_p wynoszących kolejno 0, 100, 200, 300 i 400 pikseli. Rysunek 8.15 przedstawia różnicę odległości D przy odległości obrazowej od osi optycznej H_p o jeden piksel. Funkcje te zostały wyznaczone jako różnica wartości funkcji $D(H_p, V_p)$. Przykład dla $V_p = 100$ i błędzie 2 pikseli przedstawia równanie (8.19):

$$\Delta D(H_p)|_{V_p=100, \Delta V_p=2} = D(H_p, V_p = 102) - D(H_p, V_p = 100) \quad (8.19)$$

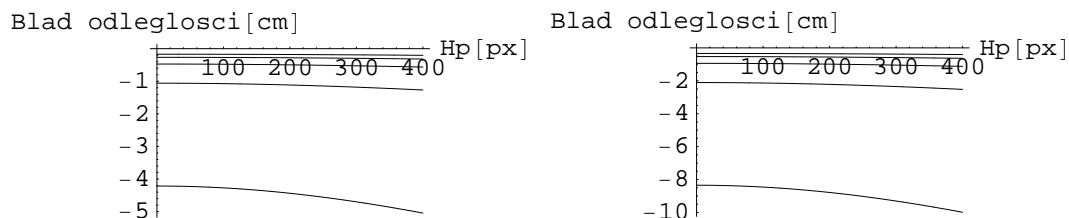
Z wykresów tych wynikają następujące wnioski:

- wpływ błędzie odległości obrazowej od osi optycznej H_p jest kilkanaście razy mniejszy niż taki sam błąd wysokości obrazowej V_p ;
- błąd odległości spowodowany błędem wysokości obrazowej V_p rośnie proporcjonalnie do błędzie V_p ;
- najmniejszy błąd występuje gdy obiekt znajduje się na osi optycznej kamery.

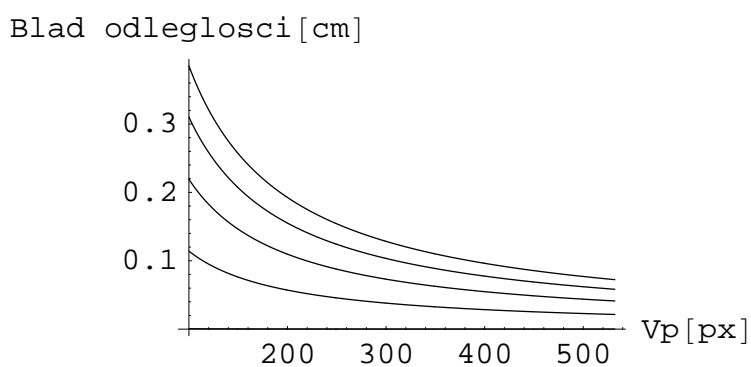
Dla znacznika o wysokości 70 cm oczekiwane błędzie maksymalne pomiaru odległości nie powinny być większe niż 10 cm (w zakresie 1-5 metrów). W przypadku znacznika o mniejszej wysokości, proporcjonalnie zmniejsza się użytkowy zakres odległości.

W bardziej złożonym, ale mającym znaczenie praktyczne przypadku, interesująca jest dokładność wyznaczenia odległości zależna od położenia kamery w lokalnym i globalnym układzie współrzędnych i przy różnych kierunkach patrzenia kamery. W celu wyznaczenia tych zależności wprowadzono funkcje pomocnicze:

- rzeczywista wysokość obrazowa $Vl_p = Vl_p(\beta_n, D_n)$,
- rzeczywista odległość obrazowa obiektu od osi optycznej $Hl_p = Hl_p(\beta_n, D_n)$.



Rysunek 8.14: Zależności zmiany wartości (czyli możliwego błędu) wyznaczonej odległości D od odległości od osi optycznej H_p przy zmianie V_p o **a)** 1 piksel i **b)** 2 piksele. Wartości wysokości obrazowych wynoszą 100, 200, 300, 400 i 500 piksli



Rysunek 8.15: Zależność błędu odległości D od wysokości obrazowej V_p przy zmianie H_p o jeden piksel dla Odległości poziomych H_p wynoszących 100, 200, 300, 400 piksli

Funkcje te oznaczają oczekiwane położenie i wysokość obiektu w obrazie przy wyznaczonych a priori: odległości D – oznaczanej jako D_n oraz β – oznaczanej jako β_n dla lokalnego układu współrzędnych biegunowych robota. Wartości te są wyznaczone jako liczby rzeczywiste, z pominięciem skończonego rozmiaru piksla. Wartość Hl_p zależy w lokalnym układzie współrzędnych tylko od kąta β i parametrów kamery, zaś Vl_p od odległości D , kąta β i parametrów kamery:

$$Hl_p = \frac{Dist_{cm}DX_p \tan \beta_n}{DX_{cm}} \quad (8.20)$$

$$Vl_p = \frac{Dist_{cm}DY_p V}{DY_{cm}D_n \cos \beta_n} \quad (8.21)$$

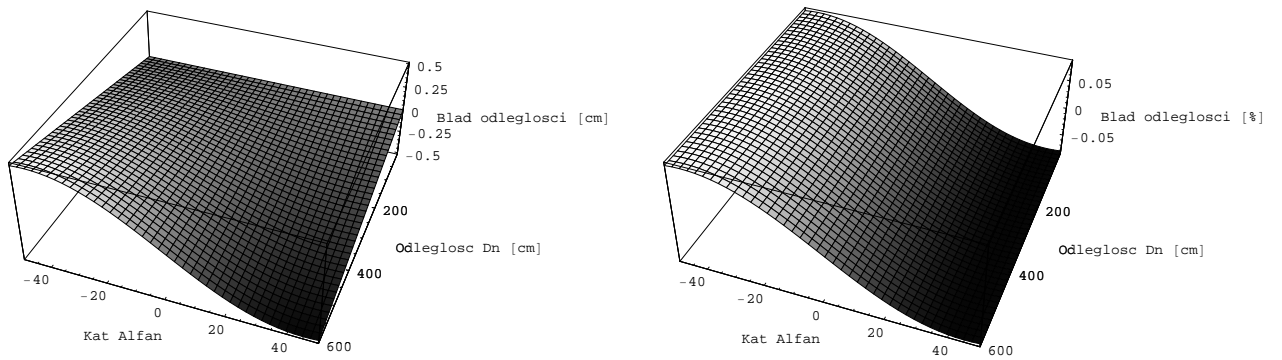
Rzeczywiste wartości współrzędnych i rozmiarów są liczbami całkowitymi jeżeli liczy się piksele, lub rzeczywistymi jeżeli wyznaczono je na podstawie równań prostych. Błąd maksymalny współrzędnych zależy od przyjętego maksymalnego błędu położenia i rozmiaru. Wyznaczenie tego błędu opisano w rozdziale 7.3 – zwykle jest on mniejszy niż 1 piksel. Błąd pomiaru dla określonych X , Y , α wyznacza się następująco:

- Wyznaczenie D_n i β_n na podstawie X , Y i α ,
- Wyznaczenie Hl_p i Vl_p ,
- Zniekształcenie Hl_p i/lub Vl_p o określoną wartość, wynikającą z przyjętego maksymalnego błędu współrzędnych,
- Wyznaczenie X , Y i α dla zniekształconych Hl_p i Vl_p i porównanie z wartościami początkowymi.

Opisany schemat wyliczenia maksymalnego błędu pomiaru zastosowano także dla metod wyznaczania położenia opisanych w rozdziałach 8.4 i 8.5. Poniższe wykresy sporządzono dla przesunięć (maksymalnych błędów) współrzędnych:

- w poziomie (H_p) o 1 piksel - rysunek 8.16;
- w pionie (V_p) o 1 piksel - rysunek 8.17;
- w pionie (V_p) o 2 piksele - rysunek 8.18;
- jednocześnie w poziomie (H_p) o 1 piksel i w pionie (V_p) o 2 piksele - rysunek 8.19.

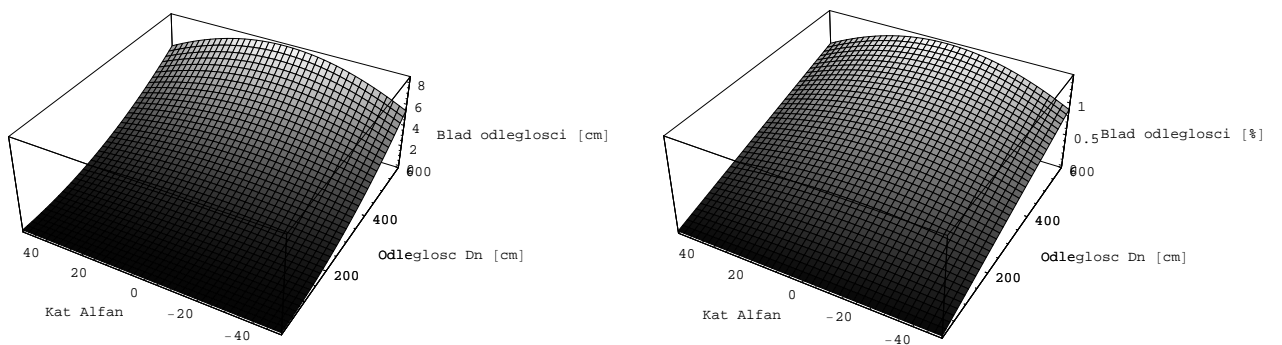
Z przedstawionych wykresów wynika, że decydujące znaczenie dla dokładności ma błąd wysokości obrazowej V_p . Błąd wyznaczonej odległości D jest większy jeżeli oś optyczna kamery jest skierowana na obiekt, ponieważ zmiana wysokości obrazowej powoduje pozorne przesunięcie obiektu wzdłuż osi odchyłonej w poziomie od mierzonej odległości o kąt β .



a)

b)

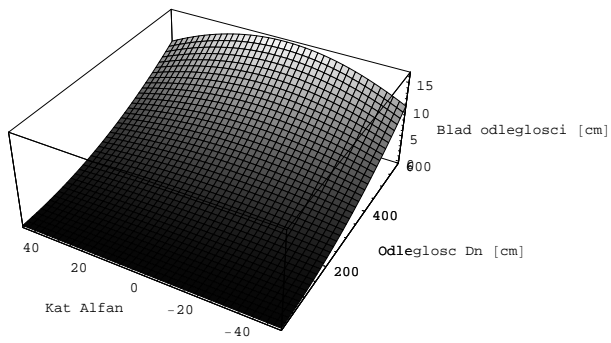
Rysunek 8.16: Błąd wyznaczenia odległości w układzie lokalnym przy przesunięciu Hl_p o 1 piksel a) w cm b) procentowy



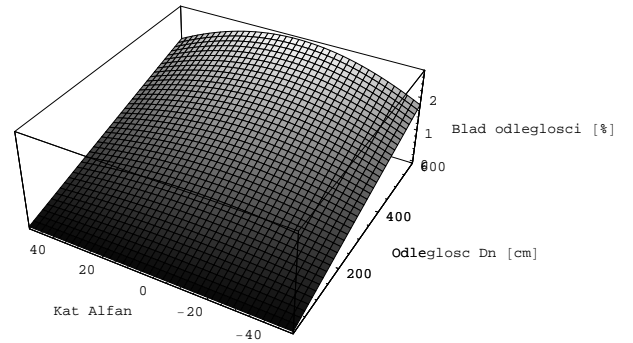
a)

b)

Rysunek 8.17: Błąd wyznaczenia odległości w układzie lokalnym przy przesunięciu Vl_p o 1 piksel a) w cm b) procentowy

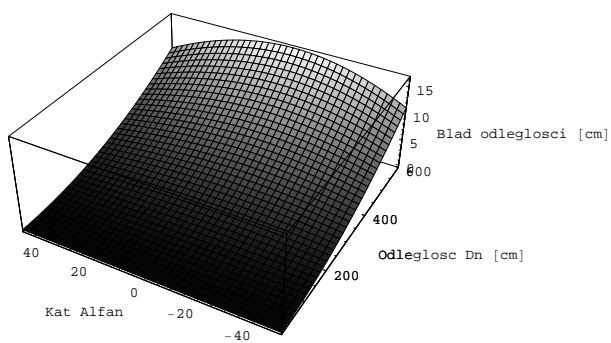


a)

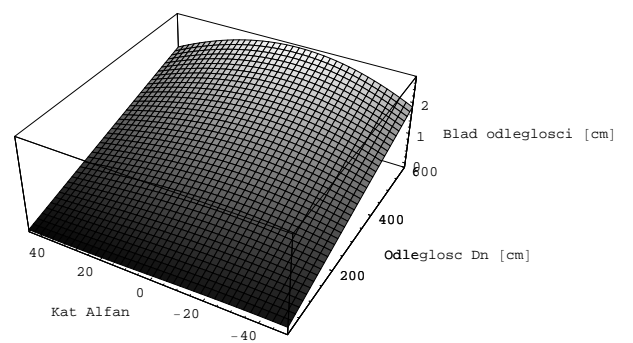


b)

Rysunek 8.18: Błąd wyznaczenia odległości w układzie lokalnym przy przesunięciu Vl_p o 2 piksele a) w cm b) procentowy



a)



b)

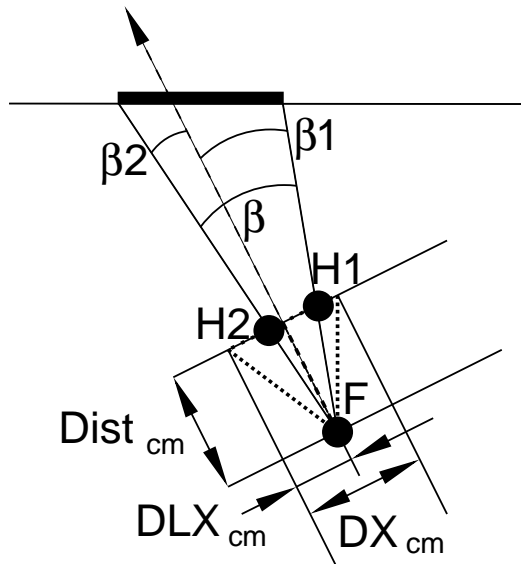
Rysunek 8.19: Błąd wyznaczenia odległości w układzie lokalnym przy przesunięciu Hl_p o 1 piksel a) w cm b) procentowy i Vl_p o 2 piksele

8.3 Wyznaczenie kąta widzenia znacznika

Kąt ten wyznacza się przez analizę rzutu z góry sceny z kamerą obserwującą znacznik, przedstawionego na rysunku 8.3. Elementy wyróżnione są wyznaczone podczas kalibracji kamery:

- odległość od ogniska kamery do płaszczyzny siatki kalibrującej Dist_{cm} ,
- szerokość obszaru widzianego po korekcji obrazu, wyrażona w jednostkach metrycznych DX_{cm} i w pikslach DX_{p} .
- odległość pozioma rzutu osi optycznej od lewego brzegu obrazu wyrażona w jednostkach metrycznych DLX_{cm} i w pikslach DLX_{p} .

Elementami mierzonymi bezpośrednio w obrazie są współrzędne brzegów znacznika oznaczone jako H1 i H2.



Rysunek 8.20: Wyznaczanie kąta β widzenia znacznika

Możliwe jest dwojaki sposób postępowania: wyznaczenie długości boków w trójkącie FH1H2 i następnie wyznaczenie kąta β korzystając z twierdzenia cosinusów lub wyznaczenie kątów składowych β_1 i β_2 , a następnie kąta β jako ich sumy. Poniżej przedstawiono oba przypadki oraz szczegółowe wyprowadzenie wzoru na kąt końcowy. Wzór końcowy jest identyczny niezależnie od metody wyznaczania go. Ponieważ rzeczywiste wartości kąta β nie przekraczają 30° , (jeżeli kamera jest bliżej znacznika, jego wysokość zazwyczaj przekracza wysokość obrazu), lepiej stosować obliczenia wartości pośrednich oparte na funkcji arcus tangens.

Wprowadzono następujące oznaczenia boków trójkąta FH1H2:

$$H1_{cm} = \frac{H1_p * DX_{cm}}{X_p},$$

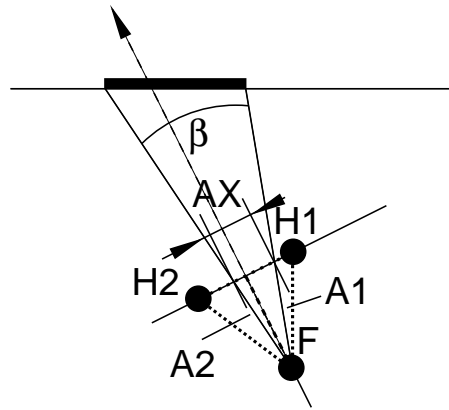
$$H2_{cm} = \frac{H2_p * DX_{cm}}{X_p},$$

$$AX = H1_{cm} - H2_{cm} = \frac{(H1_p - H2_p) * DX_{cm}}{DX_p},$$

$$A1 = \sqrt{Dist_{cm}^2 + (H1_{cm} - DLX_{cm})^2},$$

$$A2 = \sqrt{Dist_{cm}^2 + (H2_{cm} - DLX_{cm})^2},$$

gdzie $H1_p$ i $H2_p$ oznaczają współrzędne poziome brzegów znacznika wyrażone w pikselach, zaś $H1_{cm}$ i $H2_{cm}$ te same współrzędne rzutowane na siatkę kalibrującą wyrażone w jednostkach metrycznych. $A1$ i $A2$ to odległości od ogniska kamery do punktów przecięcia płaszczyzny kalibrującej z prostymi łączącymi ognisko z brzegami znacznika, zaś AX – odległość między punktami przecięcia. Schematycznie przedstawiono to na rysunku 8.21.



Rysunek 8.21: Schemat oznaczeń boków trójkąta FH1H2

Prawdziwe są następujące zależności (na podstawie tw. cosinusów):

$$AX^2 = A1^2 + A2^2 - 2A1A2 \cos \beta, \quad (8.22)$$

$$\beta = \arccos \frac{A1^2 + A2^2 - AX^2}{2A1A2}. \quad (8.23)$$

Dla drugiego wariantu wprowadzono tymczasowe oznaczenia:

$$B1 = H1_{cm} - DLX_{cm} = \frac{(H1_p - DLX_p) * DX_{cm}}{DX_p},$$

$$B2 = H2_{cm} - DLX_{cm} = \frac{(H2_p - DLX_p) * DX_{cm}}{DX_p}.$$

Prawdziwe są następujące zależności:

$$\tan \beta_1 = \frac{B1}{Dist_{cm}}, \quad (8.24)$$

$$\tan \beta_2 = \frac{B2}{Dist_{cm}}, \quad (8.25)$$

$$\beta_1 = \arctan \frac{B1}{Dist_{cm}}, \quad (8.26)$$

$$\beta_2 = \arctan \frac{B2}{Dist_{cm}}, \quad (8.27)$$

$$\beta = \beta_1 - \beta_2. \quad (8.28)$$

Ostatecznie zależność kąta β od współrzędnych obiektu $H1$, $H2$ oraz położenia osi optycznej w obrazie opisuje wzór (8.29):

$$\beta = \arctan \frac{DX_{cm}(H1_p - DLX_p)}{Dist_{cm}DX_p} - \arctan \frac{DX_{cm}(H2_p - DLX_p)}{Dist_{cm}DX_p} \quad (8.29)$$

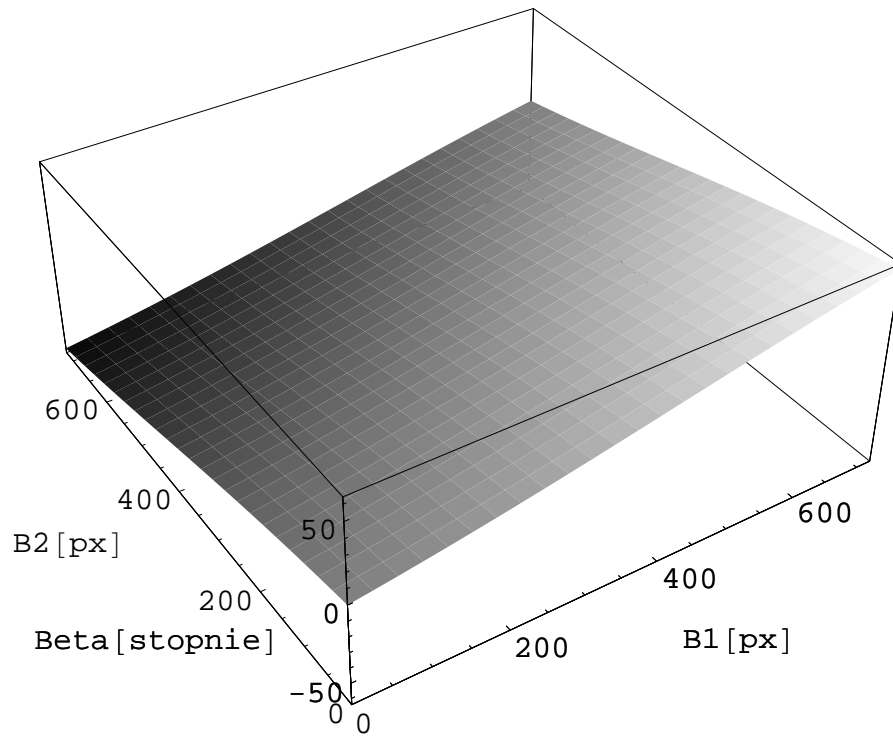
8.3.1 Dokładność wyznaczenia kąta

Oczywistym ograniczeniem dokładności pomiaru jest skończona rozdzielczość obrazu w pikslach. Dlatego analiza dokładności obu metod powinna wyznaczyć zmianę kąta β przy zmianie o jeden piksel wartości $H1$ i $H2$ dla obu metod.

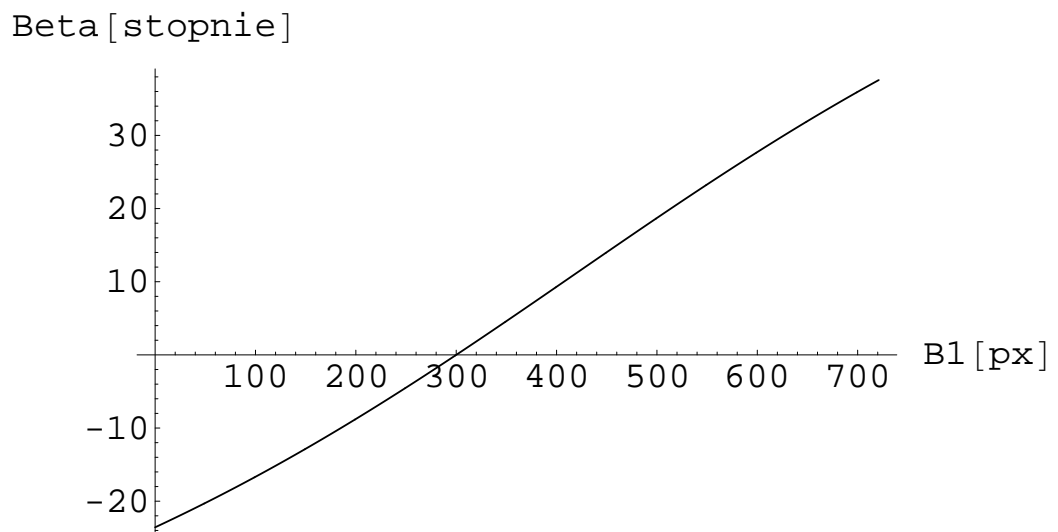
Wartości $Dist_{cm} = 16,0\text{cm}$, $X_{cm} = 19,0\text{cm}$, $X_p = 722$, $DLX_p = 418$ zostały wyznaczone podczas kalibracji kamery (rozdział 4). Dla tych parametrów kamery kąt β wyznacza się według równania (8.30):

$$\beta = \frac{180}{\pi}(\arctan 0,00164931(X1_p - 418) - \arctan 0,00164931(X2_p - 418)) \quad (8.30)$$

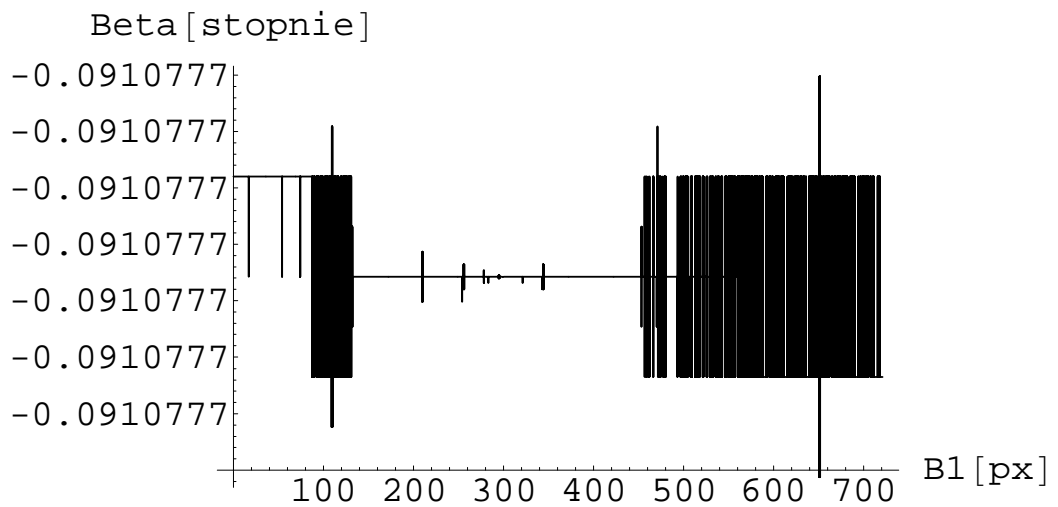
Wykres zamieszczony na rysunku 8.22 przedstawia zależność kąta β od wartości $B1$ i $B2$ wyrażonych w pikslach. Wykres 8.23 przedstawia przekrój przez wykres dla $B2=300$ piksli. Jak widać, wykres przypomina funkcję tangens. W zakresie kątów mierzonych w rzeczywistości ($<30^\circ$) i gdy mierzony obiekt znajduje się w pobliżu osi optycznej wykres jest prawie liniowy. Oznacza to, że kąt można wyznaczyć z określoną rozdzielczością prawie niezależnie od jego wartości. Różnicę funkcji $\beta(X1_p + 1) - \beta(X1_p)$ przedstawiono na wykresie 8.24. Z ostatniego wykresu wynika, że przy zmianie współrzędnej o jeden piksel dla małych kątów zmiana kąta jest praktycznie stała i dla używanej kamery wynosi $0,091^\circ/px$. Najlepszą rozdzielczość kątową i dokładność wyznaczenia kąta uzyskuje się, gdy obiekt znajduje się w pobliżu osi optycznej kamery. W zakresie współrzędnych poziomych 200-500 czyli w pobliżu osi optycznej kamery na wykresie z rysunku 8.24 błąd wynosi 0, co jest spowodowane skończoną dokładnością i rozdzielczością obliczeń numerycznych.



Rysunek 8.22: Zależność kąta β od poziomych współrzędnych obrazowych obiektu



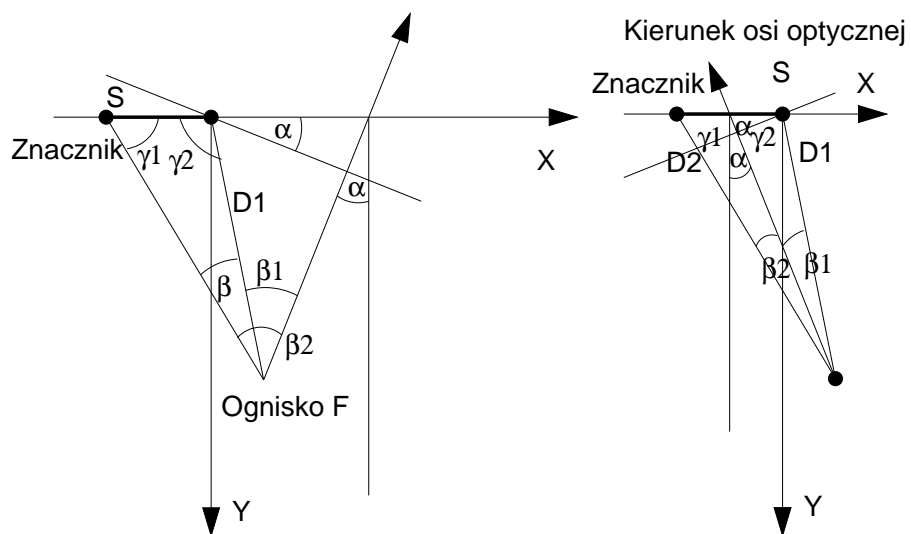
Rysunek 8.23: Wartości kąta β przy współrzędnej obrazowej $X_{2p} = 300$



Rysunek 8.24: Zmiana kąta β przy zmianie współrzędnej obrazowej $X1_p$ o 1 piksel dla $X2_p = 300$

8.4 Wyznaczanie położenia na podstawie jednej odległości

Metoda wyznacza wartości X i Y na podstawie odległości między krawędziami znacznika, odległości między kamerą a jedną z krawędzi znacznika oraz kąta widzenia znacznika. Oznaczenia odległości $D1$, kątów $\beta1$ i $\beta2$ oraz dodatkowych kątów $\gamma1$ i $\gamma2$ pokazuje rysunek 8.25. Prawdziwe są następujące zależności:



Rysunek 8.25: Kąty i odległości wyznaczone podczas lokalizacji opisywaną metodą

$$X = -D1 \cos \gamma_2, \quad (8.31)$$

$$Y = D1 \sin \gamma_2. \quad (8.32)$$

Dla sytuacji z lewej strony rysunku 8.25 prawdziwe jest równanie:

$$\frac{\pi}{2} - \alpha - \beta_2 + \gamma_1 = \pi. \quad (8.33)$$

Wzór (8.33) jest prawdziwy gdy kąty β_1 i β_2 są ostre. Jeżeli, tak jak na lewej części rysunku 8.25, $\alpha > 0$ i $\beta_2 < 0$, wszystkie wartości kątów są dodatnie po podstawieniu do wzoru (8.33). Jeżeli $\alpha < 0$ i $\beta_2 < 0$, wówczas wyrażenie $\frac{\pi}{2} - \alpha$ ze wzoru (8.33) jest kątem rozwartym i wszystkie kąty są dodatnie (tę sytuację przedstawiono na prawej części rysunku 8.25). Zatem:

$$\alpha = \gamma_1 - \frac{\pi}{2} - \beta_2. \quad (8.34)$$

Pozostaje wyznaczyć wartości kątów γ_1 i γ_2 . Kąt γ_2 wyznacza się znając β i γ_1 :

$$\gamma_2 = \pi - \gamma_1 - \beta, \quad (8.35)$$

ponieważ kąty te są kątami trójkąta. Kąt γ_1 wyznacza się z twierdzenia sinusów:

$$\frac{D1}{\sin \gamma_1} = \frac{S}{\sin \beta}, \quad (8.36)$$

zaś kąt β jako różnicę kątów:

$$\beta = \beta_1 - \beta_2. \quad (8.37)$$

Ponieważ twierdzenie sinusów nie pozwala określić, czy kąt jest ostry czy rozwarty, konieczne jest uzyskanie tej informacji inną drogą. Najprostsza jest wykorzystanie twierdzenia cosinusów i wyznaczenie znaku cosinusa kąta γ_1 :

$$D1^2 = D2^2 + S^2 - 2 \cos(\gamma_1), \quad (8.38)$$

Z czego wynika, że:

$$\text{sgn}(\cos(\gamma_1)) = \text{sgn}(D2^2 + S^2 - D1^2) \quad (8.39)$$

Końcowe wzory dla kąta $\gamma_1 \leq \frac{\pi}{2}$, co jest równoznaczne ze spełnieniem nierówności $D1^2 \leq D2^2 + S^2$, mają następującą postać:

$$\gamma_1 = \arcsin\left(\frac{D1 \sin(\beta_1 - \beta_2)}{S}\right), \quad (8.40)$$

$$\gamma_2 = \pi - \arcsin\left(\frac{D1 \sin(\beta_1 - \beta_2)}{S}\right) - \beta_1 + \beta_2, \quad (8.41)$$

$$X = D1 \cos\left(\arcsin\left(\frac{D1 \sin(\beta_1 - \beta_2)}{S}\right) + \beta_1 - \beta_2\right), \quad (8.42)$$

$$Y = D1 \sin\left(\arcsin\left(\frac{D1 \sin(\beta_1 - \beta_2)}{S}\right) + \beta_1 - \beta_2\right), \quad (8.43)$$

$$\alpha = \arcsin\left(\frac{D1 \sin(\beta_1 - \beta_2)}{S}\right) - \frac{\pi}{2} - \beta_2. \quad (8.44)$$

zaś dla kąta $\gamma_1 > \frac{\pi}{2}$:

$$\gamma_1 = \pi - \arcsin\left(\frac{D_1 \sin(\beta_1 - \beta_2)}{S}\right), \quad (8.45)$$

$$\gamma_2 = \arcsin\left(\frac{D_1 \sin(\beta_1 - \beta_2)}{S}\right) - \beta_1 + \beta_2, \quad (8.46)$$

$$X = -D_1 \cos\left(\arcsin\left(\frac{D_1 \sin(\beta_1 - \beta_2)}{S}\right) - \beta_1 + \beta_2\right), \quad (8.47)$$

$$Y = D_1 \sin\left(\arcsin\left(\frac{D_1 \sin(\beta_1 - \beta_2)}{S}\right) - \beta_1 + \beta_2\right), \quad (8.48)$$

$$\alpha = \frac{\pi}{2} - \arcsin\left(\frac{D_1 \sin(\beta_1 - \beta_2)}{S}\right) - \beta_2, \quad (8.49)$$

$$(8.50)$$

8.4.1 Dokładność wyznaczenia położenia

Matematyczne wyznaczenie dokładności pomiaru wymaga podobnego postępowania jak w przypadku kąta β i odległości D :

- wyznaczenia położenia i wysokości krawędzi znacznika w obrazie,
- wyznaczenia położenia ogniska i orientacji osi optycznej kamery na podstawie położenia i rozmiarów obrazowych znacznika,
- badania zmian współrzędnych przy zmianie obrazowych wielkości i położenia krawędzi znaczników o wartość błędu.

Wyznaczenie położenia i wielkości znacznika w obrazie wykonuje się, podobnie jak wyznaczenie położenia i rozmiaru pojedynczej pionowej krawędzi, przez wyznaczenie odległości i kąta β dla każdej krawędzi. Zależności odległości D_{1r} i D_{2r} oraz kątów β_{1r} i β_{2r} od współrzędnych robota X_r i Y_r oraz orientacji α_r wyznaczają wzory (8.51) ÷ (8.54):

$$D_{1r} = \sqrt{X_r^2 + Y_r^2}; \quad (8.51)$$

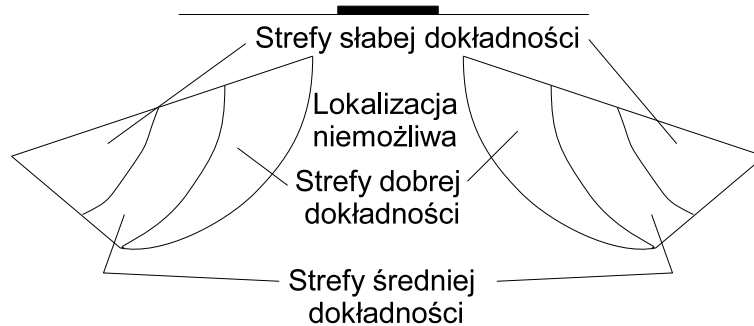
$$D_{2r} = \sqrt{(X_r + S)^2 + Y_r^2}; \quad (8.52)$$

$$\beta_{1r} = \alpha_r - \arctan\left(\frac{X_r}{Y_r}\right); \quad (8.53)$$

$$\beta_{2r} = \alpha_r - \arctan\left(\frac{X_r + S}{Y_r}\right). \quad (8.54)$$

Błąd wyznaczony na podstawie równań zależy od błędu wyznaczenia współrzędnych narożników znacznika. Błędu statystycznego nie wyznaczano. Poglądowo rozkład stref dobrej, średniej i złej lokalizacji oraz braku możliwości lokalizacji przedstawia rysunek 8.26.

Do teoretycznego wyznaczenia dokładności pomiarów przyjęto, że każda współrzędna narożnika może różnić się o 1 piksel od wartości prawidłowej. Wartość ta jest równa maksymalnemu błędowi położenia piksli po korekcji obrazu, wyznaczonemu w rozdziale 4.3.2. Dodatkowy błąd, wprowadzany przez niedokładność wyznaczenia współrzędnych narożników pominięto, ponieważ wartości te wyznacza się na podstawie dużej ilości danych, co powoduje że wyznaczona



Rysunek 8.26: Strefy dokładności lokalizacji metodą z jedną odległością

wartość średnia jest bardzo bliska prawdziwej. Poniżej umieszczono wykresy obrazujące błąd w centymetrach i w procentach współrzędnych. Procentowy błąd współrzędnych jest wyliczany wg równań (8.56), (8.55) i (8.57):

$$\Delta X\% = 100\% * \left| \frac{X - X_r}{X_r} \right| \quad (8.55)$$

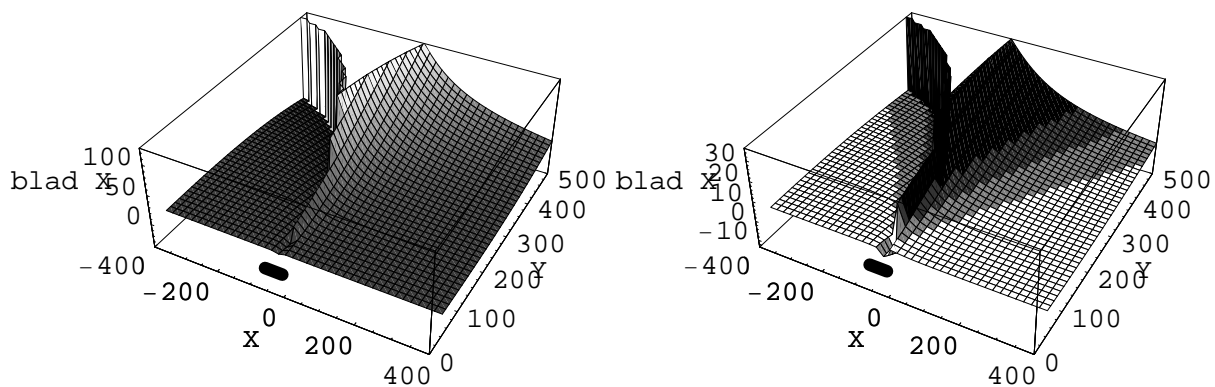
$$\Delta X_N\% = 100\% * \left| \frac{X - X_r}{D1} \right| \quad (8.56)$$

$$\Delta Y\% = 100\% * \left| \frac{Y - Y_r}{Y_r} \right| \quad (8.57)$$

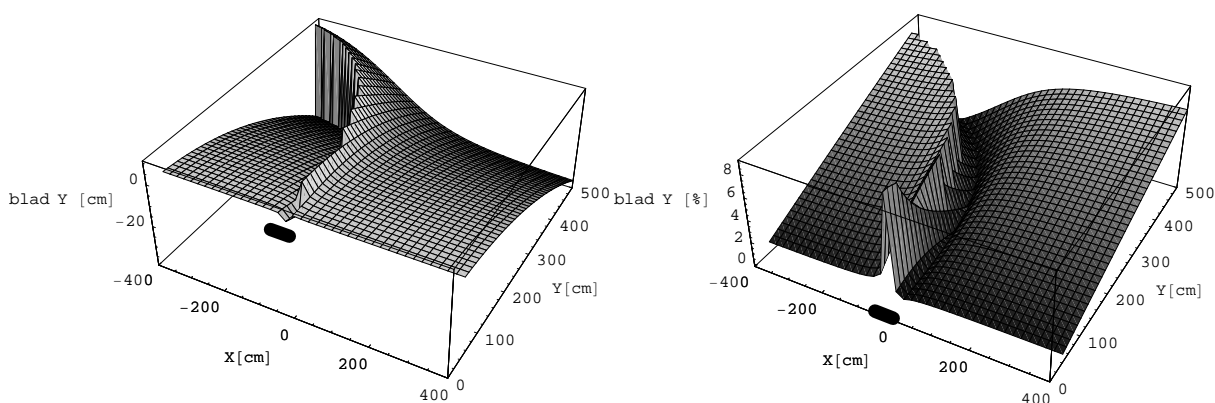
Wyznaczanie błędu procentowego współrzędnej X według odległości od znacznika według wzoru (8.56) może wydawać się, jednak należy pamiętać że położenie wartości 0 współrzędnej X jest umowne. Dla odróżnienia błąd wyrażony równaniem (8.55) nazwano zgodnie z definicją błędem procentowym, zaś błąd wyrażony wzorem (8.56) błędem znormalizowanym. Wyznaczenie błędu procentowego określenia kąta pominięto, wyznaczono tylko błąd w stopniach. Powodem jest brak punktu odniesienia – jeżeli by liczyć błąd orientacji względem odległości katowej od umownego kierunku to nie wiadomo jaki kierunek wybrać, a jeżeli względem np. wartości kąta pełnego to wystarczy liniowe przeskalowanie wartości.

Poniższe wykresy przedstawiają błędy współrzędnych X, Y i kąta α dla ustalonego błędu współrzędnych narożników znacznika. Wszystkie wykresy sporządzono dla wartości kąta $\alpha = 0$. We wszystkich wykresach użyto następujących oznaczeń błędów współrzędnych: dHp1 oznacza błąd poziomy lewej krawędzi, dHp2 – prawej, dVp1 – błąd wysokości lewej krawędzi, dVp2 – prawej. Wielkości tych błędów są wyrażone w pikselach. Na wykresach z błędem wyrażonym w procentach (po prawej stronie) kolor biały oznacza błąd poniżej 2%, jasnoszary – od 2% do 5%, ciemnoszary – od 5% do 10%, czarny – powyżej 10%. Położenie znacznika o szerokości 50 cm oznaczono czarną poziomą linią.

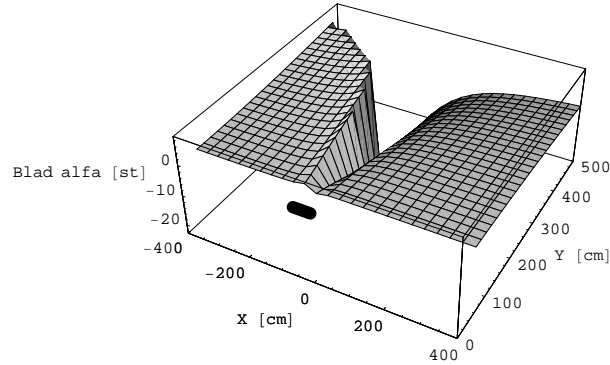
Charakterystyczną anomalią jest duży błąd lub nieokreśloność funkcji w obszarze "przed znacznikiem", tzn. dla wartości X bliskich 0. Anomalia ta istnieje w praktyce podczas prób lokalizacji, co przedstawiono w rozdziale 9.1. Anomalia ta wynika z wyznaczania kąta γ_1 pomiędzy płaszczyzną znacznika a kierunkiem do kamery przy pomocy funkcji arcus sinus. W praktyce podczas wyznaczania γ_1 jako wartości pośredniej często argument funkcji jest poza dozwolonym zakresem (czyli poza przedziałem $<-1, 1>$), jeżeli kamera znajduje się przed znacznikiem. Jeżeli wartości składowe argumentu tej funkcji są zakłócone, może on przyjąć



Rysunek 8.27: Błąd współrzędnej X wyrażony w cm i znormalizowany w % dla przyjętych błędów $dHp1=-1$, $dHp2=1$, $dVp1=2$, $dVp2=-2$



Rysunek 8.28: Błąd współrzędnej Y wyrażony w cm i w % dla przyjętych błędów $dHp1=-1$, $dHp2=1$, $dVp1=2$, $dVp2=-2$



Rysunek 8.29: Błąd orientacji α X wyrażony stopniach dla przyjętych błędów $dHp1=-1$, $dHp2=1$, $dVp1=2$, $dVp2=-2$

wartość 1 lub większą. W takim przypadku wartości funkcji (a więc i błędu) nie da się wyznaczyć, nie można też wyznaczyć położenia robota. Z wykresów zamieszczonych na rysunkach 8.27-8.29 wynika też, że błąd współrzędnej X jest względnie mały prawie wszędzie poza obszarem na wprost znacznika.

Poniżej zamieszczono analizę wpływu błędu wyznaczenia współrzędnych na zbieżność funkcji. Ze względu na wielkość i złożoność wyrażeń końcowych, przeprowadzono analizę częściową na wzorach pośrednich.

Dla współrzędnej X (gdy $0 \leq \gamma_1 \leq \pi/2$) pochodne wyrażają się wzorami:

$$\frac{\partial X}{\partial D_1} = \cos(\beta_1 - \beta_2 + \gamma_1), \quad (8.58)$$

$$\frac{\partial X}{\partial \beta_1} = \frac{\partial X}{\partial \gamma_1} = -D_1 \sin(\beta_1 - \beta_2 + \gamma_1) \quad (8.59)$$

$$\frac{\partial X}{\partial \beta_2} = D_1 \sin(\beta_1 - \beta_2 + \gamma_1). \quad (8.60)$$

Dla współrzędnej Y pochodne wyrażają się wzorami:

$$\frac{\partial Y}{\partial D_1} = D_1 \sin(\beta_1 - \beta_2 + \gamma_1), \quad (8.61)$$

$$\frac{\partial Y}{\partial \beta_1} = \frac{\partial Y}{\partial \gamma_1} = D_1 \cos(\beta_1 - \beta_2 + \gamma_1), \quad (8.62)$$

$$\frac{\partial Y}{\partial \beta_2} = -D_1 \cos(\beta_1 - \beta_2 + \gamma_1). \quad (8.63)$$

Pochodne kąta α wyrażają się wzorami:

$$\frac{\partial \alpha}{\partial \beta_1} = -1, \quad (8.64)$$

$$\frac{\partial \alpha}{\partial \gamma_1} = 1. \quad (8.65)$$

Ciągłość funkcji kątów (β_1 i β_2) omówiono w rozdziale 8.1, zaś ciągłość funkcji odległości $D1$ – w rozdziale 8.2.1. Wszystkie te funkcje są ciągłe w całej rozważanej przestrzeni. Różnica wyrażeń na X i α dla $x < -\frac{S}{2}$ i $x > -\frac{S}{2}$ powoduje nieciągłość funkcji współrzędnych robota i jego orientacji przed znacznikiem. Pochodne X , Y i α w drugiej części przestrzeni różnią się od pochodnych wyrażonych wzorami (8.58) ÷ (8.65) jedynie znakami przy niektórych funkcjach składowych, co nie ma wpływu na ciągłość w tym zakresie.

Kąt γ_1 wyznaczono według wzoru (8.66):

$$\gamma_1 = \arcsin \left(\frac{D1 \sin(\beta_1 - \beta_2)}{S} \right) = \arcsin \frac{Dist_{cm} DY_p V \sqrt{1 + \frac{X_{1p}^2 DX_{cm}^2}{Dist_{cm}^2 DX_p^2}} (\arctan \frac{X_{1p} DX_{cm}}{Dist_{cm} DX_p} - \arctan \frac{X_{2p} DX_{cm}}{Dist_{cm} DX_p})}{SY_{1p} DY_{cm}}. \quad (8.66)$$

Po podstawieniu odpowiednich parametrów kamery do równania uzyskuje się:

$$\gamma_1 = \arcsin \left(\frac{4256 \sqrt{1 + \frac{H_{1p}^2}{369664}} (\arctan \frac{H_{1p}}{608} - \arctan \frac{H_{2p}}{608})}{5V_{1p}} \right). \quad (8.67)$$

W równaniach (8.66) wyrażenia H_{1p} i H_{2p} oznaczają różnicę między współrzędną X obiektu a współrzędną X rzutu osi optycznej DLX_p . Uproszczenie to wprowadzono dla poprawienia czytelności wzoru. Kolejnym etapem przekształcenia jest podstawienie za wartości współrzędnych wyrażeń reprezentujących zależności współrzędnych obrazowych od rzeczywistych współrzędnych i parametrów kamery. Oznaczając współrzędne według zasady $H_{1p} = HR_{1p} + dH_{1p}$, gdzie HR_{1p} oznacza dokładne położenie krawędzi w obrazie, zaś dH_{1p} jego zaburzenie, otrzymuje się wyrażenie opisane wzorem (8.68):

$$\gamma_1 = \arcsin \left(\frac{4256 \sqrt{1 + \frac{(HR_{1p} + dH_{1p})^2}{369664}} (\arctan \frac{HR_{1p} + dH_{1p}}{608} - \arctan \frac{HR_{2p} + dH_{2p}}{608})}{5(VR_{1p} + dV_{1p})} \right). \quad (8.68)$$

Po podstawieniu za współrzędne wyrażeń z rzeczywistymi współrzędnymi robota w układzie współrzędnych znacznika X_r , Y_r , α_r argument funkcji arcsin przedstawia wzór (8.69):

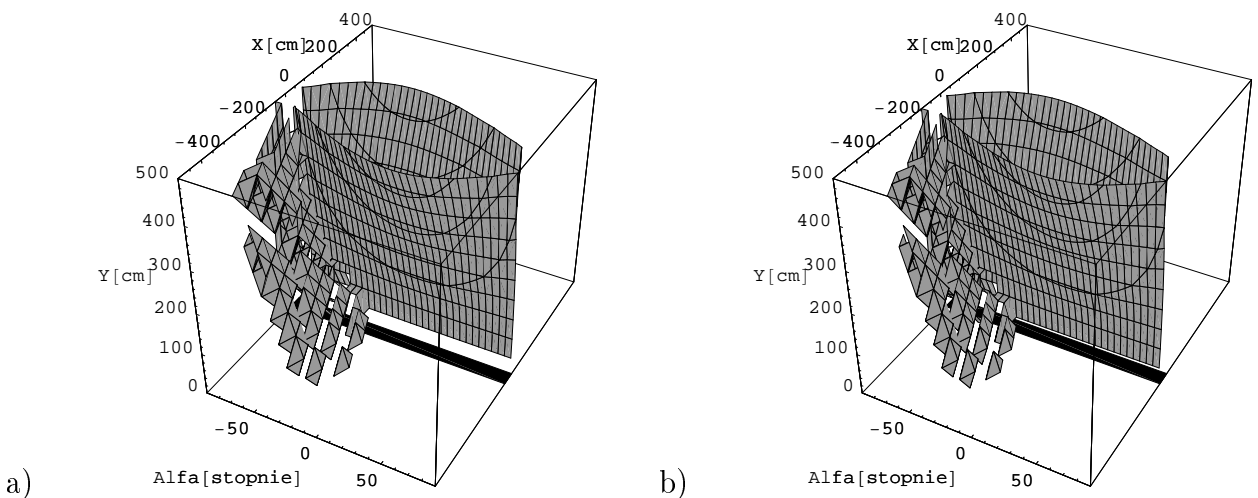
$$\frac{851, 2 \sqrt{1 + \frac{(dH_{1p} + 608 \tan(\alpha_r - \arctan \frac{X_r}{Y_r}))^2}{369664}}}{dV_{1p} + \frac{42560}{\cos(\alpha_r - \arctan \frac{X_r}{Y_r})} \sqrt{X_r^2 + Y_r^2}} * \left[\arctan \left(\frac{dH_{1p}}{608} + \tan(\alpha_r - \arctan \frac{X_r}{Y_r}) \right) - \arctan \left(\frac{dH_{2p}}{608} + \tan(\alpha_r - \arctan \frac{50 + X_r}{Y_r}) \right) \right]. \quad (8.69)$$

Dla sprawdzenia czy rzeczywiście przed znacznikiem nie da się wykonać pomiarów, wybrano przykładowe położenie robota w punkcie $X_r = 0$ cm, $Y_r = 250$ cm, $\alpha_r = 5^\circ$. Następnie wyznacza się wartość wyrażenia wg wzoru (8.69) dla $X_r = 0$, $Y_r = 250$, $\alpha_r = 5^\circ$ i różnych zakłóceń dH_{1p} , dH_{2p} i dV_{1p} . Wyniki obliczeń znajdują się w tabelicy (8.1). Wynika z niej, że dla przyjętego błędu odczytu współrzędnych $\pm 0,5$ piksla mogą wystąpić kombinacje błędów składowych uniemożliwiające wyznaczenie współrzędnych.

$dH1_p$	$dH2_p$	$dV1_p$	argument
0	0	0	0,986978
0,5	0	0	0,991129
-0,5	0	0	0,982827
0,5	0	-0,5	0,995191
-0,5	1	0	0,977109
0,5	-1	-0,5	1,00105

Tablica 8.1: Wpływ zakłóceń odczytu współrzędnych na wartość argumentu funkcji arcus sinus

Wyznaczenie obszaru, w którym przy założonym błędzie współrzędnych nie ma gwarancji otrzymania wyniku, wymaga rozwiązania nierówności przestępnych (uzyskanych z równania (8.69)) z różnymi wartościami $dH1_p$, $dH2_p$ i $dV1_p$, a następnie wyznaczenia części wspólnej rozwiązań. Ponieważ równanie (8.69) jest przestępne, przybliżone rozwiązanie najlepiej jest wyznaczyć wyznaczając wartości X_r , Y_r , α_r i sprawdzając, czy da się w tym miejscu wyznaczyć współrzędne, tzn. czy wartość równania (8.69) jest w zakresie dozwolonych argumentów funkcji arcus sinus. Rysunek 8.30 prezentuje płaszczyzny w przestrzeni (X, Y, α) , na których wartość wyrażenia opisanego wzorem (8.69) wynosi 1. Bardzo wyraźnie widoczna jest strefa przed znacznikiem, w której niezależnie od α określenie położenia może nie być możliwe. Dodatkowo przy obserwacji znacznika pod znacznym kątem również wartość wyrażenia (8.69) wynosi 1, lecz jest to obszar poza zakresem obserwacji ze względu na ograniczone rozmiary obrazu. W obszarze tym przebiega asymptota spowodowana funkcją tangens, dla której argument wynosi $\frac{\pi}{2}$. Rysunek 8.30.a przedstawia płaszczyzny dla błędu współrzędnych 0,5 piksla, zaś 8.30.b dla błędu wynoszącego 1 piksel. Widoczne jest podobieństwo kształtu stref, a przy mniejszym błędzie szerokość martwej strefy jest większa. W obu przypadkach szerokość martwej strefy wzrasta przy wzroście odległości Y od znacznika. Położenie znacznika oznaczono czarnym prostokątem na dole. W rzeczywistości znacznik nie ma na tym wykresie wysokości. Znacznik znajduje się we współrzędnej X w przedziale $<-50, 0>$ i dla współrzędnej Y=0, jednak taka linia byłaby niewidoczna. Dotyczy to wszystkich wykresów obrazujących dokładność w przestrzeni (X, Y, α) .

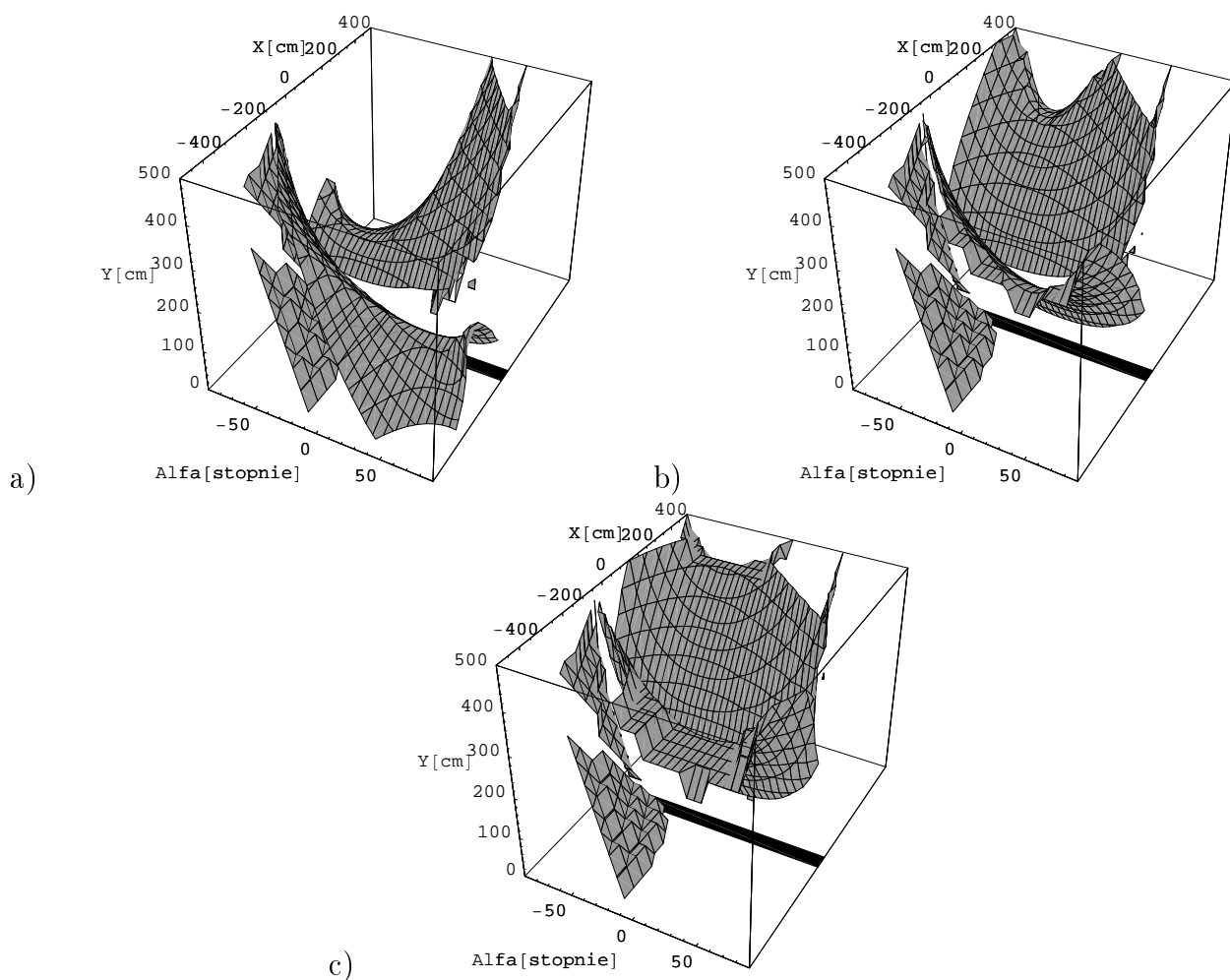


Rysunek 8.30: Przebieg płaszczyzn ograniczających strefę możliwej nieokreśloności położenia: a) dla błędu 0,5 piksla; b) dla błędu 1 piksel

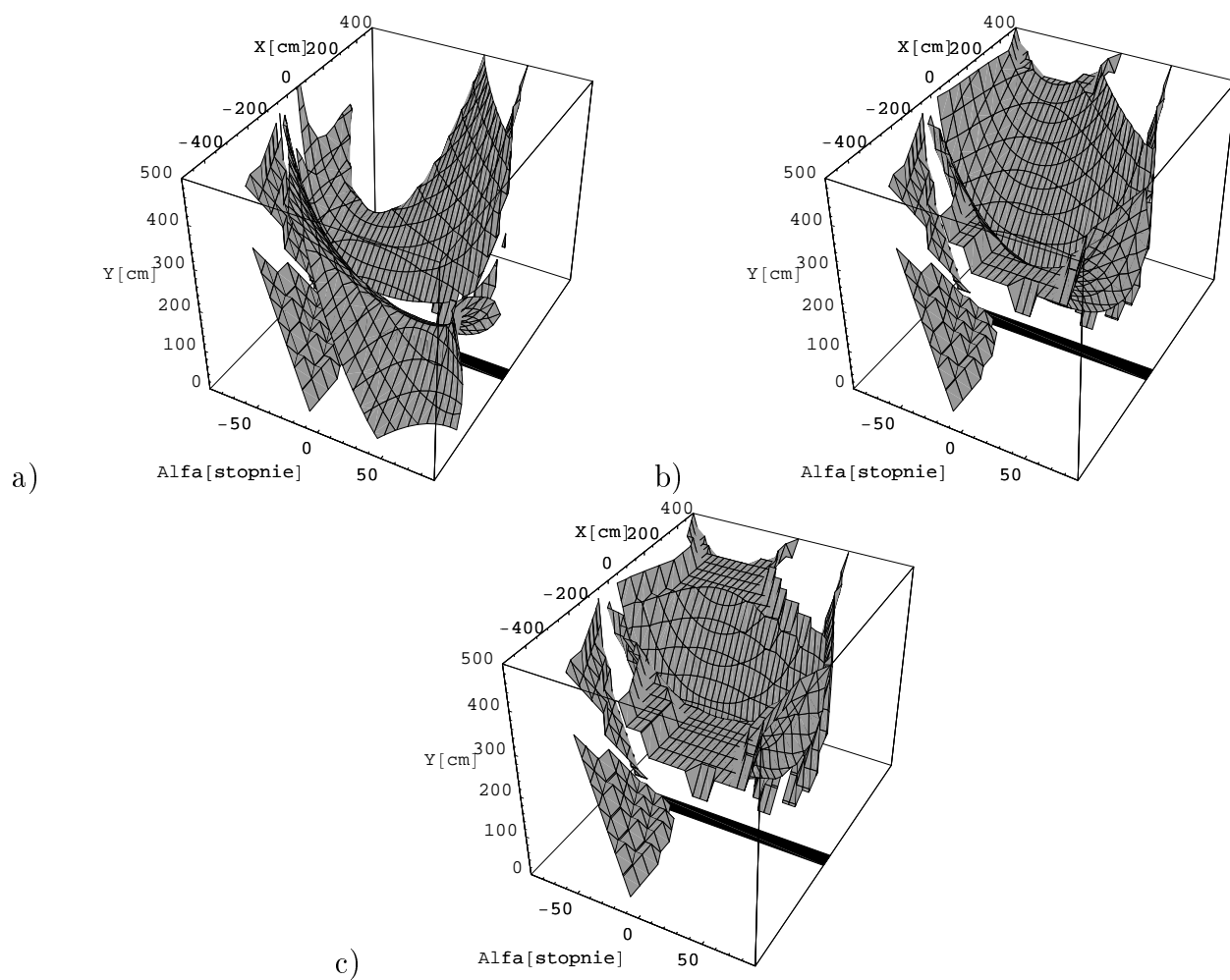
Dokładność współrzędnej X jest najmniejsza spośród wartości X , Y , α . Wykresy 8.31.a, 8.31.b i 8.31.c przedstawiają płaszczyzny w przestrzeni $(X, Y, \alpha) \in \mathfrak{R}^3$, na których maksymalny błąd wyznaczenia współrzędnej X wynosi kolejno 2%, 5% i 10%. Płaszczyzny, na których błąd znornmalizowany wynosi 2%, 5% i 10 % przedstawiono na rysunku 8.32.a, 8.32.b i 8.32.c. Interpretacja tych wykresów jest następująca: błąd pomiaru mniejszy niż określona wartość (2% lub 5%) występuje dla punktów znajdujących się w przestrzeni X, Y, α poniżej płaszczyzny o wartości odpowiednio 2% lub 5%. Widać to przy porównaniu tych wykresów – płaszczyzna o większej wartości błędu jest położona wyżej. Wynika z tego, że im dalej robot jest przesunięty w bok od znacznika, tym większa jest dokładność pomiaru. Tak więc wynik lokalizacji jest dobry, jeżeli znacznik jest widoczny pod kątem ostrym. Częściowo wynika to ze zwiększenia mianownika (którą jest współrzędna X na rysunku 8.31), ale błąd w centymetrach też się zmniejsza. Przed znacznikiem znajduje się strefa martwa, widoczna na wszystkich wykresach. Strefa dobrej dokładności teoretycznie rozciąga się w nieskończoność, co pokazuje rysunek 8.31.d. Widać na nim, że co prawda dokładność 2% osiąga się i dla 10 metrów, ale jest to błąd rzędu kilkudziesięciu cm. W miarę wzrostu odległości dobrą dokładność można uzyskać tylko przy coraz ostrzejszym kącie patrzenia na znacznik. Najlepszą dokładność pomiaru uzyskuje się, gdy kąt między płaszczyzną znacznika a linią od kamery do brzegu znacznika jest $< 60^\circ$ lub $> 120^\circ$. Dokładność ta pogarsza się gdy znacznik jest widziany pod bardzo ostrym kątem, czyli $< 30^\circ$ lub $> 150^\circ$.

Analogicznie, wykresy na rysunku 8.33.a i 8.33.b przedstawiają powierzchnię błędu pomiaru odpowiednio 2% i 5%. Dla współrzędnej Y również istnieje martwa strefa przed znacznikiem. Dodatkowo, dla wszystkich mierzonych wielkości, w strefie w której jest możliwa lokalizacja (zob. rys. 6.4) występuje minimum funkcji ze względu na Y , co oznacza że zasięg dobrej dokładności dla współrzędnych jest najgorszy gdy środek znacznika znajduje się na osi optycznej kamery. Rysunek 8.34 przedstawia wykresy powierzchni, na których błąd orientacji α wynosi odpowiednio: 8.34.a $\div 2^\circ$, 8.34.b $\div 5^\circ$, 8.34.c $\div 10^\circ$. Bardzo wyraźnie widać, że strefa najgorszej dokładności znajduje się przed znacznikiem. Tylko przed znacznikiem dla $Y \geq 2,5\text{m}$ błąd orientacji może przekroczyć 10° . Na rysunku 8.34.a widać, że gdy współrzędna Y osiąga 3-3,5 metra, dokładność orientacji 5° można uzyskać tylko dla dużych (ponad 200 cm) odległości X i ostrego kąta patrzenia.

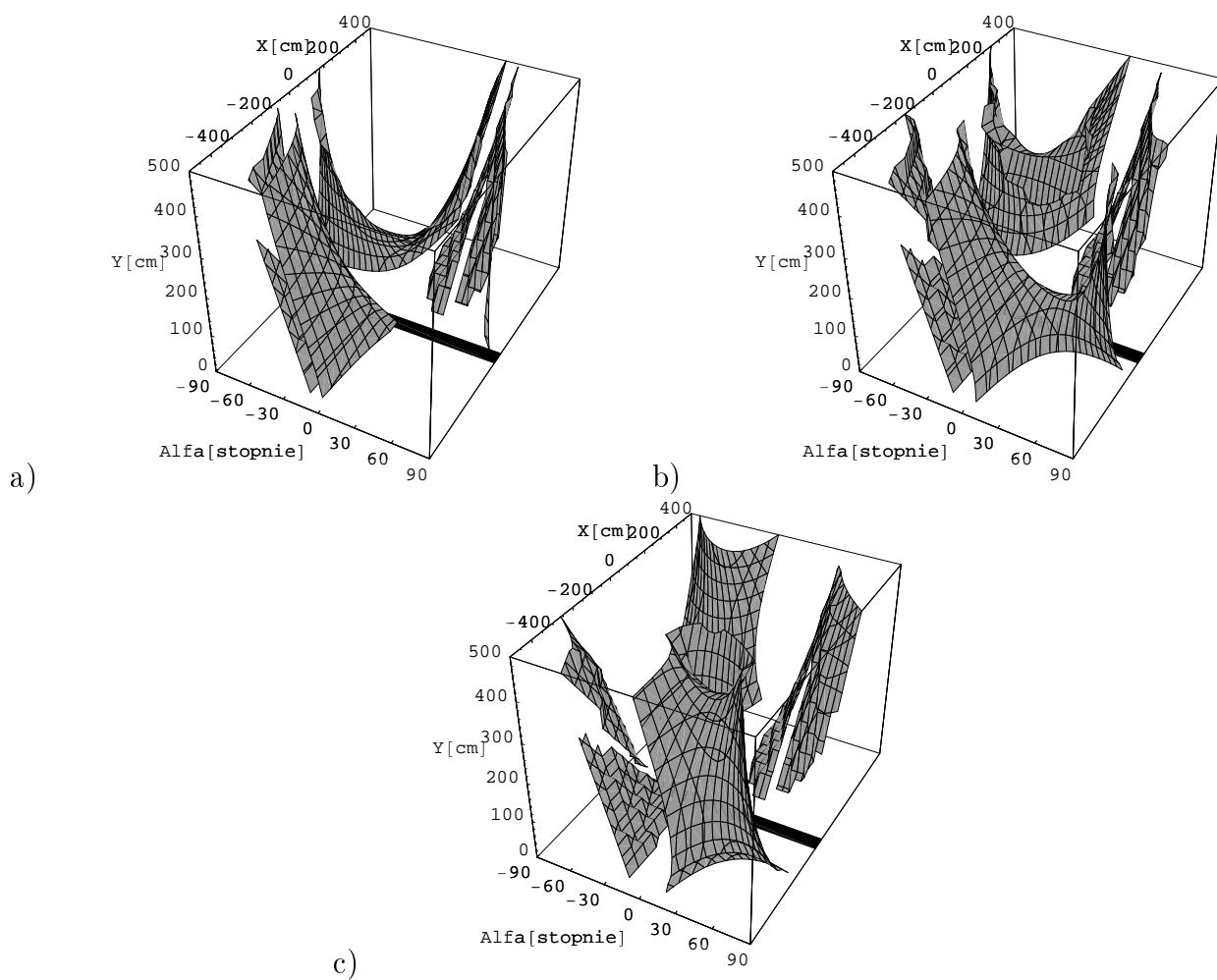
Zamieszczone wykresy odzwierciedlają przebieg matematycznie określonej funkcji dokładności. Funkcja ta różni się od rzeczywistej dokładności tym, że nie uwzględnia ograniczonego pola widzenia kamery.



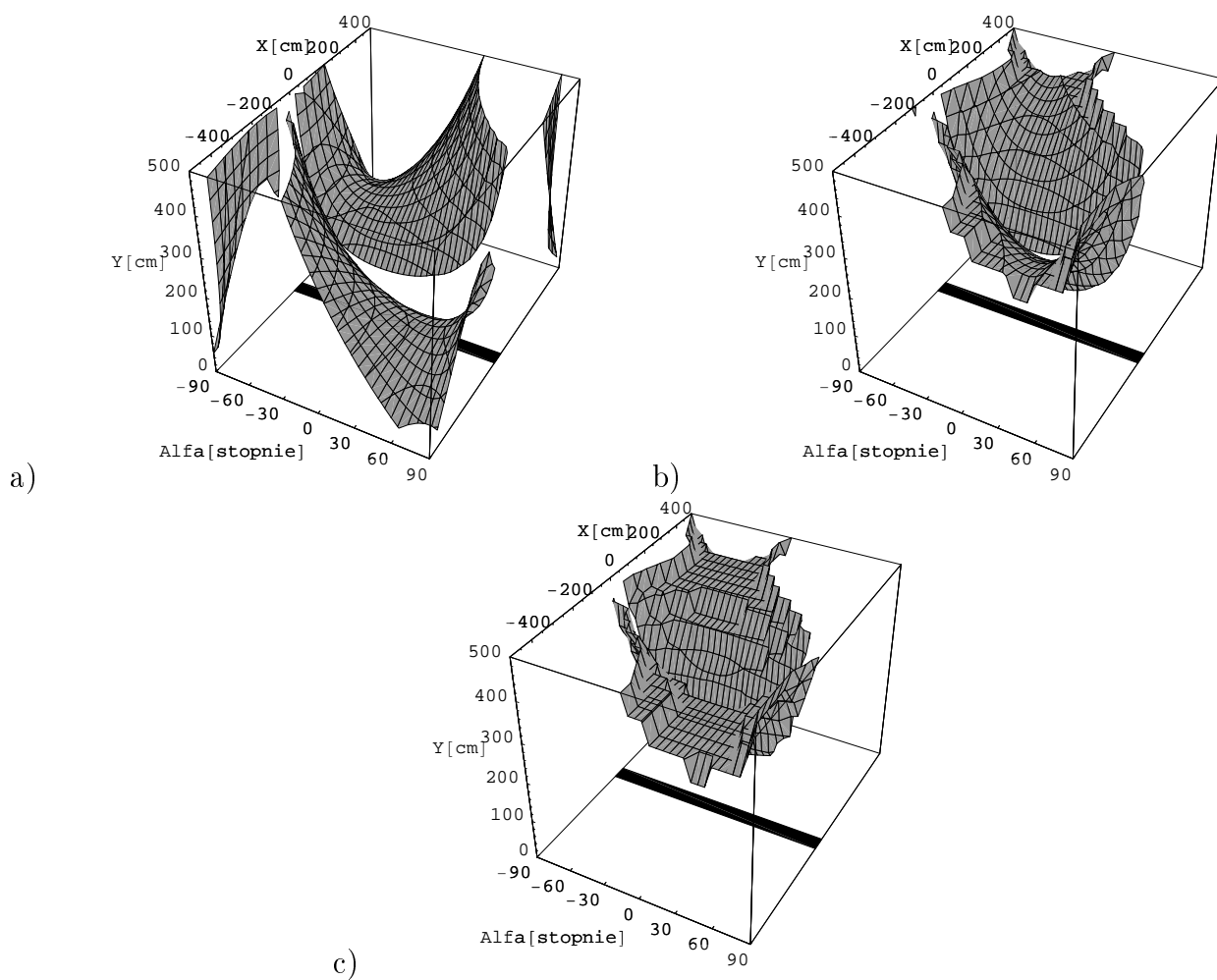
Rysunek 8.31: Powierzchnie w przestrzeni współrzędnych robota, na których maksymalny błąd pomiaru współrzędnej X wynosi: a) 2%; b) 5%; c) 10%; d) 2% dla zasięgu 10m.



Rysunek 8.32: Powierzchnie w przestrzeni współrzędnych robota, na których maksymalny znormalizowany błąd pomiaru współrzędnej X wynosi: a) 2%; b) 5%; c) 10%.



Rysunek 8.33: Powierzchnie w przestrzeni współrzędnych robota, na których maksymalny błąd pomiaru współrzędnej Y wynosi: a) 2%; b) 5%; c) 10%.



Rysunek 8.34: Powierzchnie w przestrzeni współrzędnych robota, na których maksymalny błąd pomiaru orientacji α wynosi: a) 2° ; b) 5° ; c) 10° .

8.5 Wyznaczanie położenia na podstawie dwóch odległości

Metoda wyznacza odległość na podstawie odległości do brzegów znacznika D1 i D2. Metoda jest mniej dokładna niż poprzednia, ale działa najlepiej w obszarze przed znacznikiem, gdzie zawodzi metoda opisana w rozdziale 8.4. Rysunek 8.35 przedstawia rzut z góry podczas wyznaczania położenia. Prawdziwe są zależności:

$$x = D1 \cos(\pi - \gamma1) = -D1 \cos \gamma1, \quad (8.70)$$

$$y = \sqrt{D1^2 - x^2} = \sqrt{D1^2(1 - \cos^2 \gamma1)}. \quad (8.71)$$

Wartość $\cos(\gamma1)$ można łatwo wyznaczyć znając długości wszystkich boków trójkąta korzystając z twierdzenia cosinusów:

$$D2^2 = D1^2 + S^2 - 2SD1 \cos \gamma1, \quad (8.72)$$

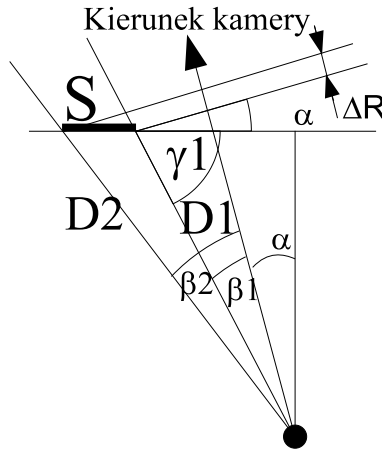
$$\cos \gamma1 = \frac{D1^2 + S^2 - D2^2}{2D1S}. \quad (8.73)$$

Ostatecznie współrzędne ogniska kamery wyznacza się wg wzorów:

$$X = \frac{1}{2S}(D2^2 - D1^2 - S^2), \quad (8.74)$$

$$Y = \frac{1}{2S} \sqrt{(D1 + D2 - S)(D1 + S - D2)(D2 + S - D1)(D1 + D2 + S)}. \quad (8.75)$$

Wyznaczenie kąta α jest łatwe, jak widać z równania (8.77). Sytuację wyjaśnia rysunek 8.35:



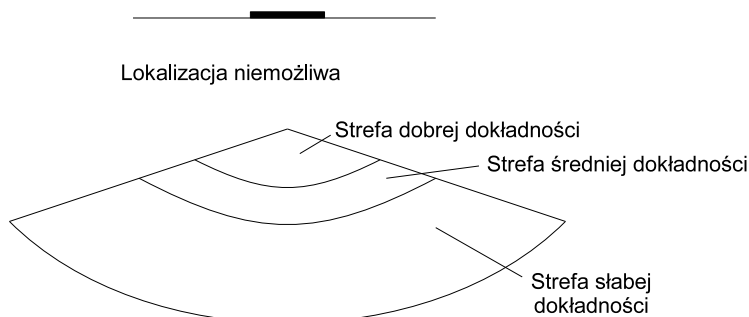
Rysunek 8.35: Wyznaczanie kąta α na podstawie odległości D1, D2 oraz kątów $\beta1$ i $\beta2$

$$\sin \alpha = \frac{\Delta R}{S} = \frac{DR1 - DR2}{S} \quad (8.76)$$

$$\alpha = \arcsin \frac{DR1 - DR2}{S} \quad (8.77)$$

8.5.1 Dokładność wyznaczenia położenia

Poglądowo rozkład stref dobrej, średniej i złej lokalizacji oraz braku możliwości lokalizacji przedstawia rysunek 8.36.



Rysunek 8.36: Strefy dokładności lokalizacji metodą z dwoma odległościami

W celu wyznaczenia ciągłości wyznaczonych współrzędnych na błędy poszczególnych zmierzonych wartości wyznacza się pochodne cząstkowe położenia i orientacji, podobnie jak w rozdziale 8.4.1. Pochodne współrzędnej X wyrażają się wzorami:

$$\frac{\partial X}{\partial D1} = -\frac{D1}{S}, \quad (8.78)$$

$$\frac{\partial X}{\partial D2} = \frac{D2}{S}. \quad (8.79)$$

Pochodne współrzędnej Y wyrażają się wzorami:

$$\frac{\partial Y}{\partial D1} = \frac{D1(-D1^2 + D2^2 + S^2)}{4\sqrt{(D1 + D2 - S)(D1 + S - D2)(D2 + S - D1)(D1 + D2 + S)}} \quad (8.80)$$

$$\frac{\partial Y}{\partial D2} = \frac{D2(D1^2 - D2^2 + S^2)}{4\sqrt{(D1 + D2 - S)(D1 + S - D2)(D2 + S - D1)(D1 + D2 + S)}} \quad (8.81)$$

$$(8.82)$$

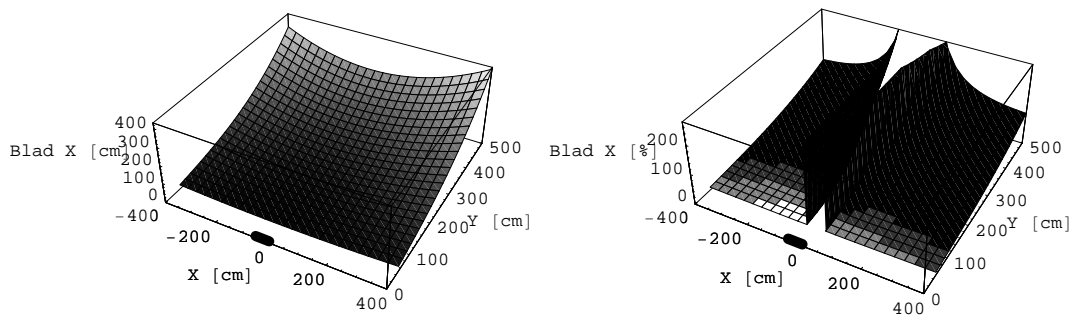
$$(8.83)$$

Wszystkie funkcje kątowe (β_1, β_2) są ciągłe w całej przestrzeni (X, Y, α) , co pokazano w rozdziale 8.1. Funkcje odległości D1 i D2 są ciągłe, co pokazano w rozdziale 8.2.1.

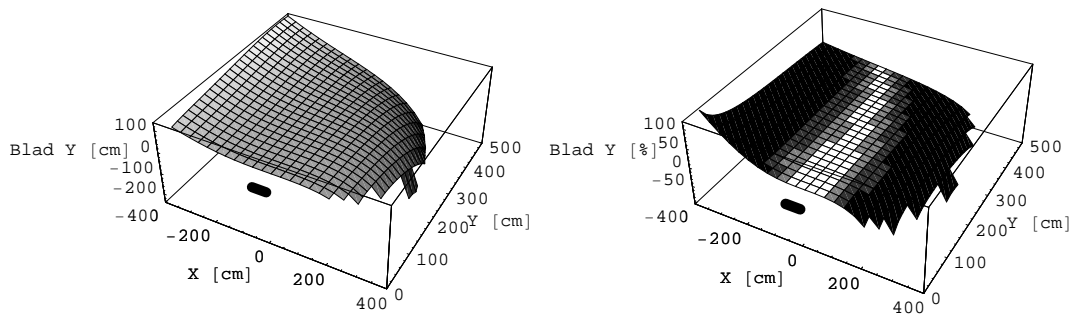
Przy teoretycznym wyznaczaniu błędu pomiaru oparto się na tych samych założeniach co w poprzedniej metodzie:

- błąd wyznaczenia współrzędnej narożnika nie przekracza jednego piksela;
- błąd wyznaczenia wysokości krawędzi nie przekracza dwóch piksli.

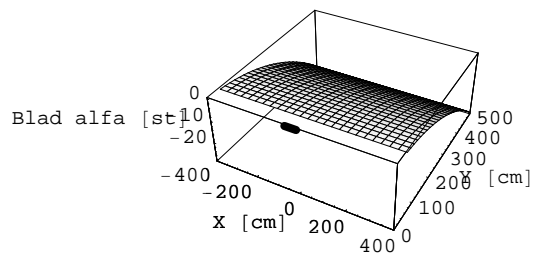
Błąd wyznacza się podobnie jak w rozdziale 8.4, wyznaczając najpierw oczekiwane rozmiary i położenie znacznika w obrazie, wprowadzając określone błędy współrzędnych i porównując wyznaczone współrzędne i kąt z zadanymi. Przykładowe wykresy błędów liniowego i procentowego dla współrzędnych X, Y i kąta α ilustrują rysunki 8.37 – 8.39.



Rysunek 8.37: Błąd liniowy i procentowy wyznaczenia współrzędnej X przy $dHp1=1$, $dVp1=2$, $dHp2=-1$, $dVp2=-2$

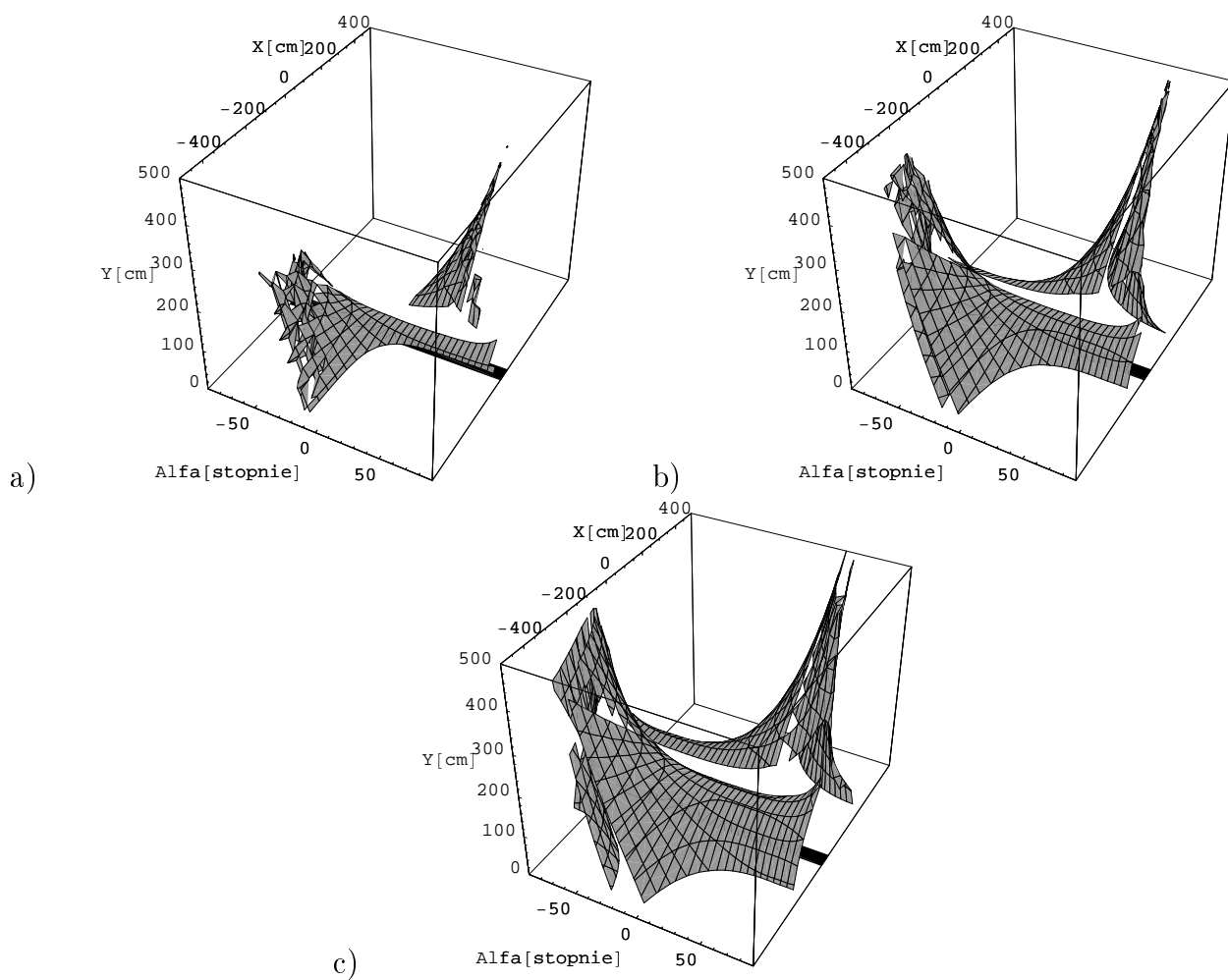


Rysunek 8.38: Błąd liniowy i procentowy wyznaczenia współrzędnej Y przy $dHp1=1$, $dVp1=2$, $dHp2=-1$, $dVp2=-2$

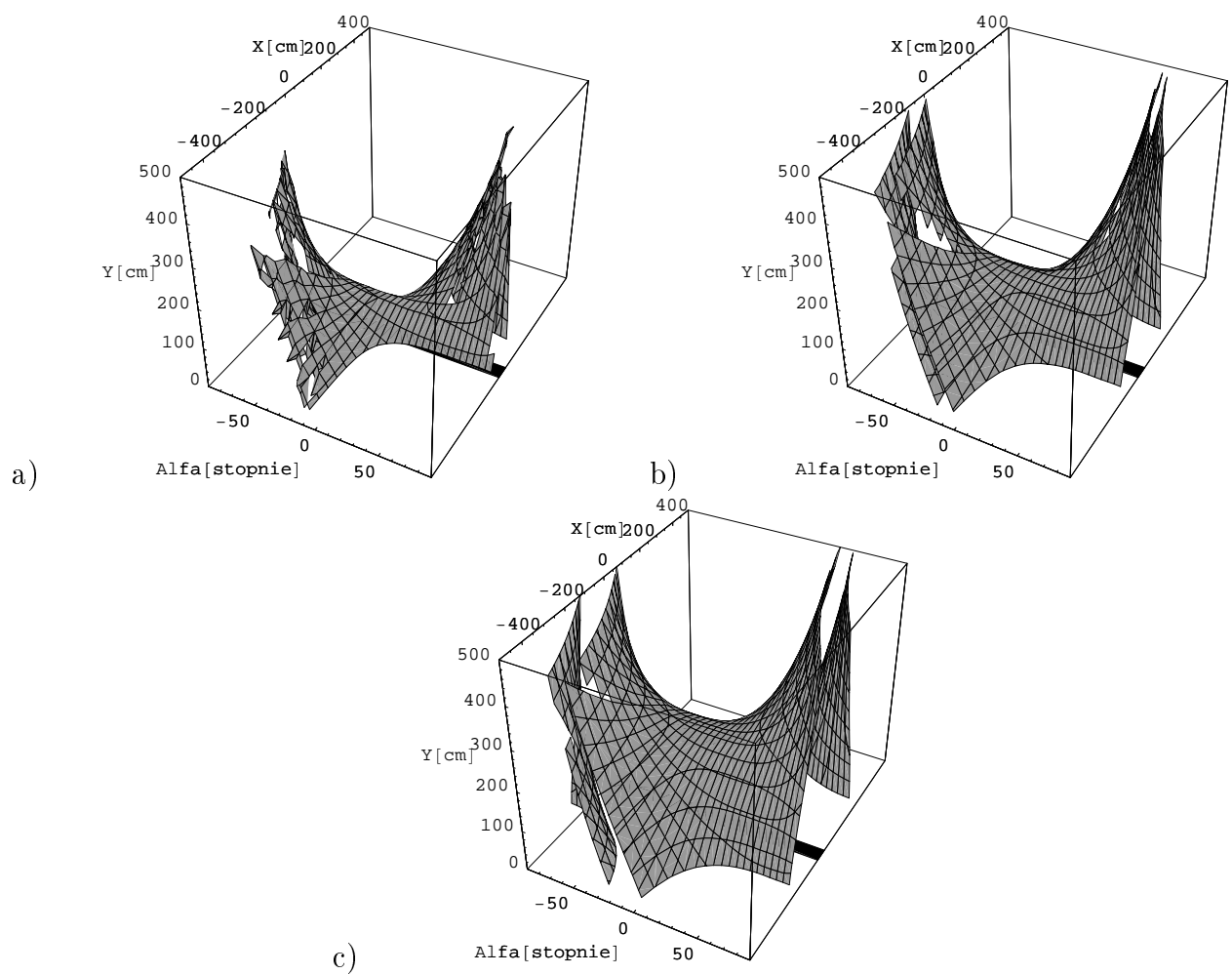


Rysunek 8.39: Błąd wyznaczenia kąta α przy $dHp1=1$, $dVp1=2$, $dHp2=-1$, $dVp2=-2$

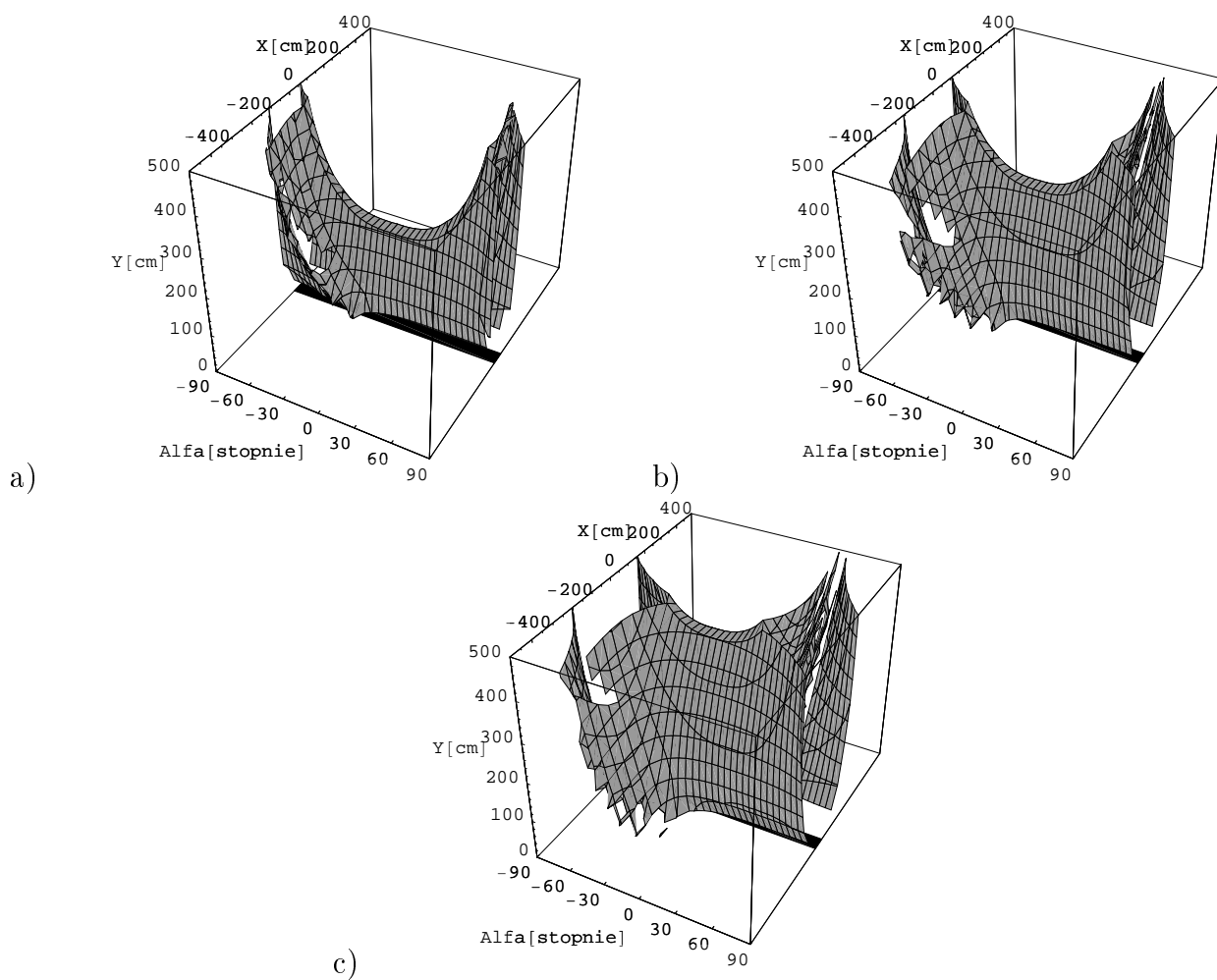
Ponieważ we wzorach (8.74), (8.75) i (8.77) nie ma skomplikowanych wyrażeń trygonometrycznych, funkcja obrazująca błąd pomiaru jest bardziej przewidywalna. Najmniejszy błąd pomiaru jest bezpośrednio przed znacznikiem, w obszarze gdzie znacznik zaczyna być widoczny w całości i zajmuje prawie całą wysokość obrazu (tzn w minimalnej odległości w której mieści się w obrazie). Wraz z oddalaniem się kamery od znacznika (w dowolną stronę) rośnie błąd lokalizacji. Poniższe wykresy przedstawiają zakresy współrzędnych X i Y oraz kąta α , dla których błąd jest mniejszy niż 2%, 5% i 10%. Podobnie jak w rozdziale 8.4, błąd kątów wyznacza się tylko w stopniach. Praktyczny zakres odległości dla znacznika o wysokości 70cm i szerokości 50cm wynosi 1-5m. Wyniki eksperymentów przedstawiono w rozdziale 9.2. Zakres kątów α przy których można uzyskać dobrą dokładność lokalizacji to $\pm 20^\circ$ odchylenia od prostej łączącej kamerę ze znacznikiem. Kształt powierzchni o jednakowej dokładności pomiaru przypomina siodło. Współrzędne X i szczególnie Y można wyznaczyć dokładnie tylko przed znacznikiem, co jest widoczne w ten sposób, że nad czarną linią powierzchnia sięga do największej współrzędnej Y w przekroju wzdłuż osi OX. Najmniejsza dokładność jest dla kąta $\alpha=0$, czyli jeżeli oś optyczna kamery nie jest skierowana prostopadle do płaszczyzny znacznika to można uzyskać dobrą dokładność dla większych odległości Y. Dla orientacji α zakres współrzędnych X i Y (szczególnie X) dla których można uzyskać dobrą dokładność jest nieco większy, lecz również powierzchnia jednakowej dokładności ma podobny kształt do siodła. Pewna niesymetryczność wykresów wynika znów z niesymetrycznego umieszczenia znacznika w układzie współrzędnych oraz z niesymetrycznej budowy kamery.



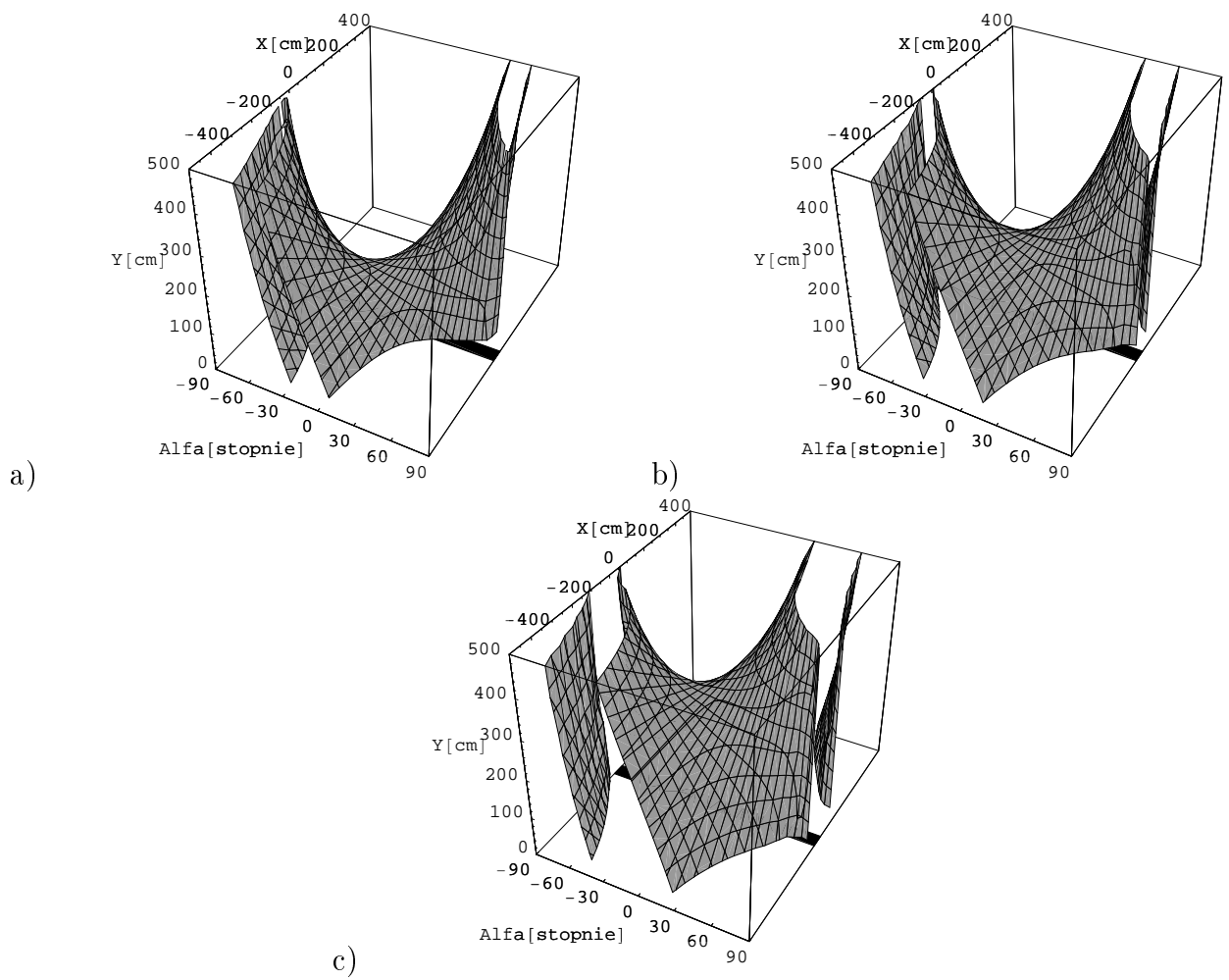
Rysunek 8.40: Powierzchnie w przestrzeni współrzędnych robota, na których maksymalny błąd pomiaru współrzędnej X wynosi: a) 2%; b) 5%; c) 10%.



Rysunek 8.41: Powierzchnie w przestrzeni współrzędnych robota, na których znormalizowany maksymalny błąd pomiaru współrzędnej X wynosi: a) 2%; b) 5%; c) 10%.



Rysunek 8.42: Powierzchnie w przestrzeni współrzędnych robota, na których maksymalny błąd pomiaru współrzędnej Y wynosi: a) 2%; b) 5%; c) 10%.



Rysunek 8.43: Powierzchnie w przestrzeni współrzędnych robota, na których maksymalny błąd pomiaru orientacji α wynosi: a) 2° ; b) 5° ; c) 10° .

Rozdział 9

Pomiary i testy

Przeprowadzono eksperymenty wykazujące wystarczającą dokładność lokalizacji do nawigacji robotem. Wykonano testy dwóch rodzajów: dokładności statycznej oraz test praktyczny. Dokładność statyczną wyznaczono dla znacznika, których był czarnym znakiem o rozmiarach 50x70 cm, a do nawigacji podczas jazdy stosowano barwną kartkę A4 o rozmiarach 21,0x29,7 mm. Eksperymenty potwierdzają, że dla dwu metod wyznaczania położenia istnieją różne strefy dobrej dokładności, które łącznie obejmują cały obszar z którego znacznik jest widoczny jako dostatecznie szeroki, tj. pod kątem większym niż 15° .

Wyniki dla metody z pomiarem jednej odległości i szerokości obiektu znajdują się w rozdziale 9.1. Potwierdzają one dobrą dokładność metody jeżeli znacznik jest widoczny pod kątem ostrym i złą dokładność lub wręcz niemożność lokalizacji, gdy kamera znajduje się przed znacznikiem.

Wyniki metody z wyznaczaniem dwu odległości zamieszczono w rozdziale 9.2. Potwierdzają one, że dokładność metody jest zgodna z intuicyjnymi oczekiwaniami, czyli największą dokładność uzyskuje się gdy kamera znajduje się naprzeciwko znacznika. Statystyczny błąd pomiaru, tzn. różnica między wartościami wyznaczonymi a rzeczywistymi, jest większy niż dla metody z pomiarem jednej odległości.

Ze względu na ograniczenia techniczne wykonano serie po 1000 pomiarów wykonanych z tego samego miejsca bez poruszenia kamery. Z powodu uśredniania danych podczas wyznaczania współrzędnych metodą regresji liniowej różnice między pomiarami z tego samego miejsca są niewielkie. Pełniejszy rozkład statystyczny możnaby uzyskać umieszczając kamerę w tym samym miejscu, z wymuszeniem ruchu kamery między pomiarami.

Test praktyczny polegał porównaniu nawigacji robota na podstawie odometrii i znacznika z nawigacją na podstawie samej odometrii. W trakcie testu wykonano testy UMBMark (zob. rozdz. 2.2.1) z użyciem robotów Pioneer i Navigator. Przebieg i wyniki testu opisano w rozdziale 9.3. Wyniki badań wykazują, podczas nawigacji z kamerą błąd znajomości położenia (na podstawie odometrii) narastał w czasie między kolejnymi wyznaczeniami położenia na podstawie kamery, po czym był sprowadzany do małej wartości.

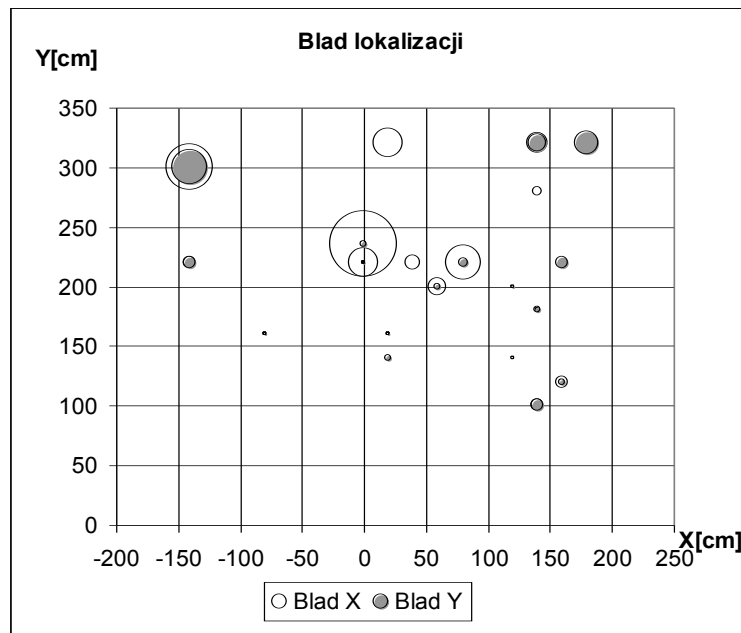
9.1 Wyniki statyczne dla metody z pomiarem widocznej szerokości

Poniżej przedstawiono wyniki pomiarów statycznych dla 23 punktów. Pomiary wykonywano w seriach po 1000 dla każdego punktu. Użyto znacznika w kształcie prostokątnej czarnej ramki o szerokości 50 cm i wysokości 70 cm. Wykres 9.1 przedstawia średni błąd pomiaru współrzędnych

wyrażony w cm. Promienie okręgów odpowiadają błędom pomiaru przy znanym położeniu kamery w środku okręgu. Wykres 9.10 przedstawia procentowy błąd pomiaru. Proporcje średnic okręgów odpowiadają proporcjom błędów procentowego. Największy okrąg odpowiada wartości 11,6%. Błąd procentowy dla współrzędnej X wyznacza się na podstawie wzoru:

$$Err_X\% = 100 * \frac{X_{pom} - X_r}{\sqrt{X_r^2 + Y_r^2}}, \quad (9.1)$$

gdzie X_r i Y_r oznaczają ustalone wcześniej współrzędne kamery, zaś X_{pom} – zmierzoną wartość współrzędnej X. Tabela 9.1 zawiera minimalne, maksymalne i średnie wartości współrzędnej X oraz błędów pomiarowych.



Rysunek 9.1: Wartości średniego błędów pomiaru współrzędnych X i Y w cm dla metody pomiaru odległości i szerokości

Zamieszczone tabele i wykresy potwierdzają anomalię dokładności w pobliżu linii prostopadłej do płaszczyzny znacznika. Dla punktu (-40, 240) znajdującego się przed znacznikiem (znacznik znajduje się w zakresie współrzędnych $X < -50, 0 >$) pomiar był niewykonalny, ponieważ punkt ten leży w martwej strefie metody. Największe błędy występują jeżeli współrzędna X kamery znajduje się w przedziale do 100 cm od brzegów znacznika – tam liniowy błąd pomiaru przekracza 10 a miejscami nawet 40 cm dla współrzędnej X i dochodzi do 20 cm dla współrzędnej Y. Gdy kamera znajduje się w odległości Y nie większej niż 3 metry, a wartość współrzędnej X nie znajduje się w strefie przed znacznikiem, błąd obu współrzędnych nie przekracza 10 cm, a często jest mniejszy niż 5 cm. Wyniki te świadczą o tym, że poza strefą małej dokładności przed znacznikiem i bardzo dużymi odległościami (powyżej 3 metrów) można bardzo dokładnie wyznaczyć położenie kamery.

9.1.1 Porównanie z wartościami oczekiwanymi

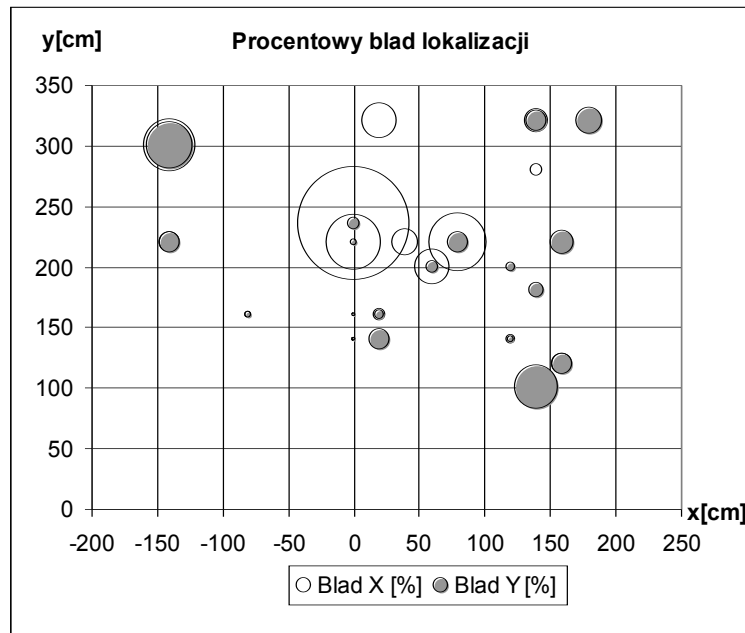
Poniżej zamieszczono porównanie otrzymanych wyników z oczekiwanymi. Podczas wyznaczania błędów oczekiwanego zakładałem, że błąd wyznaczenia współrzędnych obrazowych wynosi

X	Y	Min	Max	Avg	Min_Err	Max_Err	Avg_Err	Avg_Err %
320	180	177,9	187,1	186,7	2,1	-7,1	6,7	1,8
320	140	139,6	225,7	148,8	0,4	-85,7	8,8	2,5
320	20	30,5	31,8	31,6	-10,5	-11,8	11,6	3,6
300	-140	-136,9	-114	-121,7	-3,1	-26,0	18,3	5,6
280	140	142,3	150,9	143,9	-2,3	-10,9	3,9	1,2
240	-40	0	0	0	0	0	0	0
235	0	-36,9	-22,7	-27,3	22,7	36,9	27,3	11,6
220	-140	-135,2	-131,9	-134,4	-4,8	-8,1	5,6	2,1
220	40	32,7	34,5	33,6	5,5	7,3	6,4	2,9
220	0	-17	-10,2	-12,5	10,2	17,0	12,5	5,7
100	140	135,2	135,7	135,3	4,3	4,8	4,7	2,8
120	160	154,4	156	155,3	4,0	5,6	4,7	2,4
140	0	-0,7	-0,2	-0,4	0,2	0,7	0,4	0,3
140	20	19,8	20,4	19,9	0,2	-0,4	0,1	0,1
140	120	117,5	118,1	117,8	1,9	2,5	2,1	1,2
160	0	-0,6	-0,2	-0,5	0,2	0,6	0,5	0,3
160	20	21,2	22,7	22,2	-1,2	-2,7	2,2	1,4
160	-80	-82,1	-79,6	-81,2	-0,4	2,1	1,2	0,7
180	140	138,4	138,4	138,4	1,6	1,6	1,6	0,7
200	60	51,2	58,3	52,5	1,7	8,8	7,5	3,6
200	120	119,5	120,2	119,9	-0,2	0,5	0,1	0,1
220	80	65,3	66,4	66,0	13,6	14,7	14,0	6,0
220	160	156,8	162,4	160,9	-2,4	3,2	0,9	0,3

Tablica 9.1: Graniczne i średnie wartości błędu współrzędnej X dla metody lokalizacji pomiarem odległości i szerokości

X	Y	Min	Max	Avg	Min_Err	Max_Err	Avg_Err	Avg_Err %
320	180	327,8	338,7	329,3	7,8	18,7	9,3	2,9
320	140	278	328,9	327,0	-42,0	8,9	7,0	2,2
320	20	319,8	320,1	319,8	-0,2	0,1	0,2	0,1
300	-140	309,4	316,6	314,4	9,4	16,6	14,4	4,8
280	140	279,1	280,7	279,8	-0,9	0,7	0,2	0,1
240	-40	0	0	0	0	0	0	0
235	0	231,1	232	231,8	-3,9	-3,0	3,2	1,3
220	-140	224,3	227	225,2	4,3	7,0	5,2	2,4
220	40	220,4	220,8	220,5	0,4	0,8	0,6	0,2
220	0	221,2	221,8	221,6	1,2	1,8	1,6	0,7
100	140	104,4	104,8	104,6	4,4	4,8	4,6	4,6
120	160	122	122,7	122,7	2	2,7	2,6	2,1
140	0	139,5	139,7	139,6	-0,5	-0,3	0,4	0,3
140	20	136,6	136,9	136,9	-3,4	-3,1	3,1	2,2
140	120	140,5	140,9	140,8	0,5	0,9	0,8	0,6
160	0	160,5	160,7	160,7	0,5	0,7	0,7	0,4
160	20	161,7	161,9	161,8	1,7	1,9	1,8	1,1
160	-80	158,1	159,9	158,8	-1,9	-0,1	1,2	0,8
180	140	183,1	183,1	183,1	3,1	3,1	3,1	1,7
200	60	196,1	197,7	197,5	-3,9	-2,3	2,5	1,3
200	120	201,5	202,4	202,0	1,5	2,4	2,0	1,0
220	80	224,2	225	224,6	4,2	5,0	4,6	2,1
220	160	224,5	226,9	225,3	4,5	6,9	5,3	2,4

Tablica 9.2: Graniczne i średnie wartości błędu współrzędnej Y dla metody lokalizacji z pomiarem odległości i szerokości

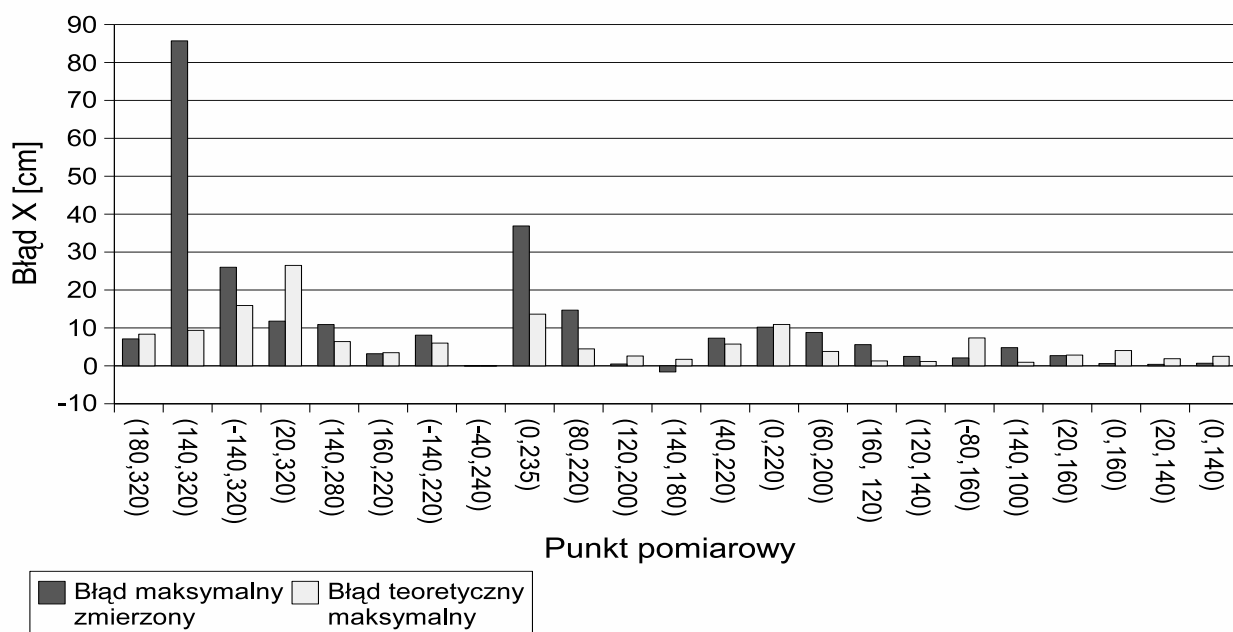


Rysunek 9.2: Wartości średniego błędu pomiaru współrzędnych X i Y (w %) dla metody z pomiarem odległości i szerokości; największy okrąg odpowiada 11,6% błędowi

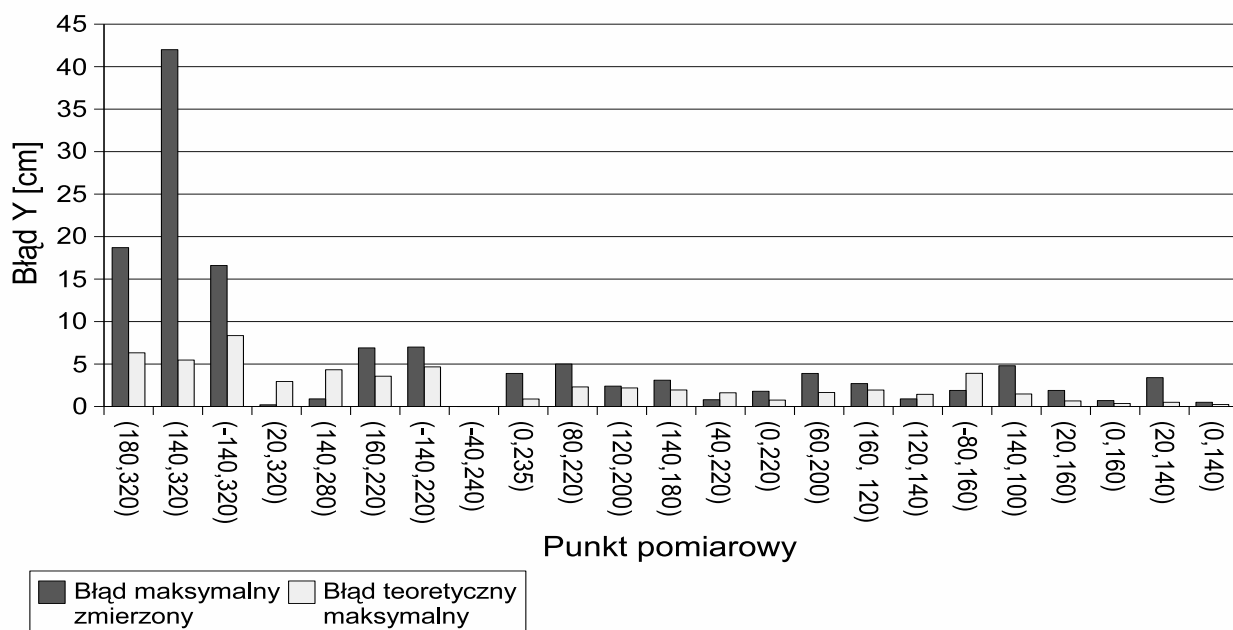
0,5 piksla. Błąd oczekiwany jest maksymalnym błędem jaki wystąpił przy sprawdzeniu wszystkich kombinacji błędów składowych. Wykresy 9.3 i 9.4 przedstawiają porównanie minimalnych i maksymalnych wartości odchylenia od wartości rzeczywistej. Wykresy 9.5 i 9.6 przedstawiają porównanie szerokości przedziałów wartości współrzędnych. Dane z wykresów można interpretować następująco:

1. Kilka punktów pomiarowych jest obarczonych dużym błędem. Są to punkty leżące względnie blisko osi prostopadłej przechodzącej przez środek znacznika. Punkt (-40, 240) leży w martwej strefie i również teoretycznie (metodą opisaną w rozdziale 8.4) nie da się w nim wyznaczyć położenia zastosowaną metodą. Pełną analizę istnienia martwej strefy zamieszczono na stronie 126, zob. też rysunek 8.30.
2. Punkty leżące dalej od osi znacznika pozwalają na dokładniejsze wyznaczenie współrzędnych. Szerokość przedziału (rozrzut wartości) otrzymanych współrzędnych jest przeważnie mniejsza niż oczekiwana na podstawie rozdziału 8.4, jednak nie są to wartości uzyskane w normalnych badaniach statystycznych.
3. Widoczna jest tendencja wzrostu błędu liniowego wraz ze wzrostem odległości od znacznika, szczególnie w kierunku prostopadłym do płaszczyzny znacznika. Błąd procentowy nie rośnie szybko wraz ze wzrostem odległości, przynajmniej dla opisanych przypadków.

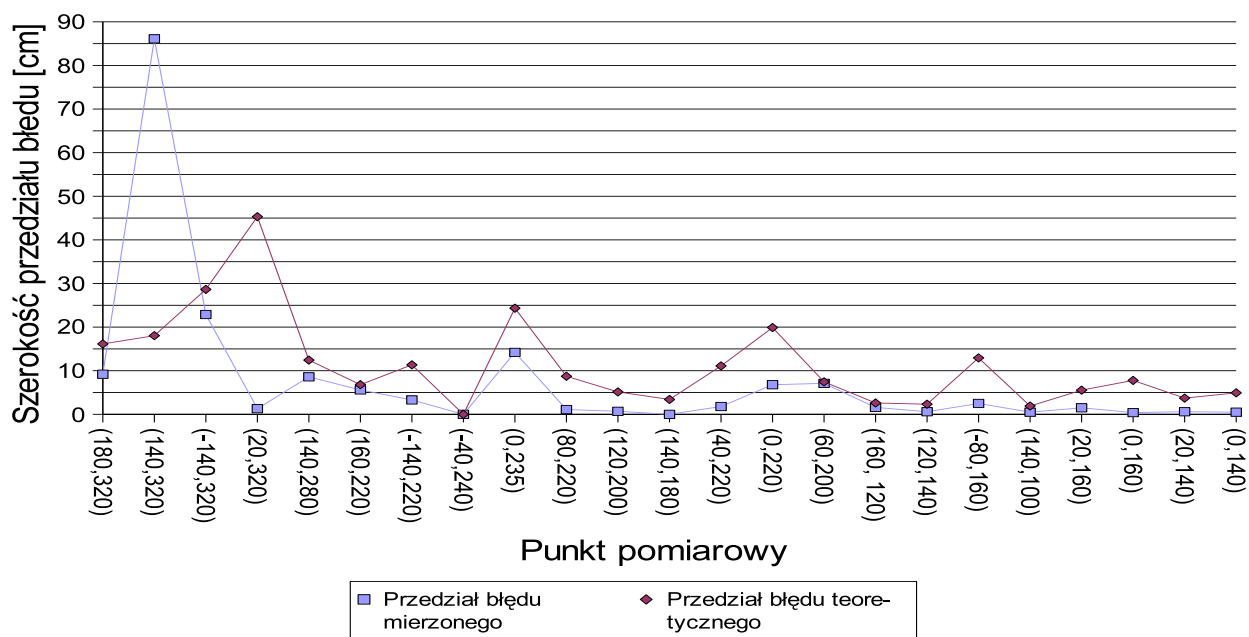
Wykresy na rysunkach 9.5 i 9.6 pokazują, że dla punktu drugiego o współrzędnych (140, 320) rozpiętość błędów jest większa niż przewidywana, zaś dla pozostałych mniejsza. Jednak wartości średnie są dobre (zob. tabele 9.1 i 9.2), więc jest to efekt błędu pojedynczego pomiaru.



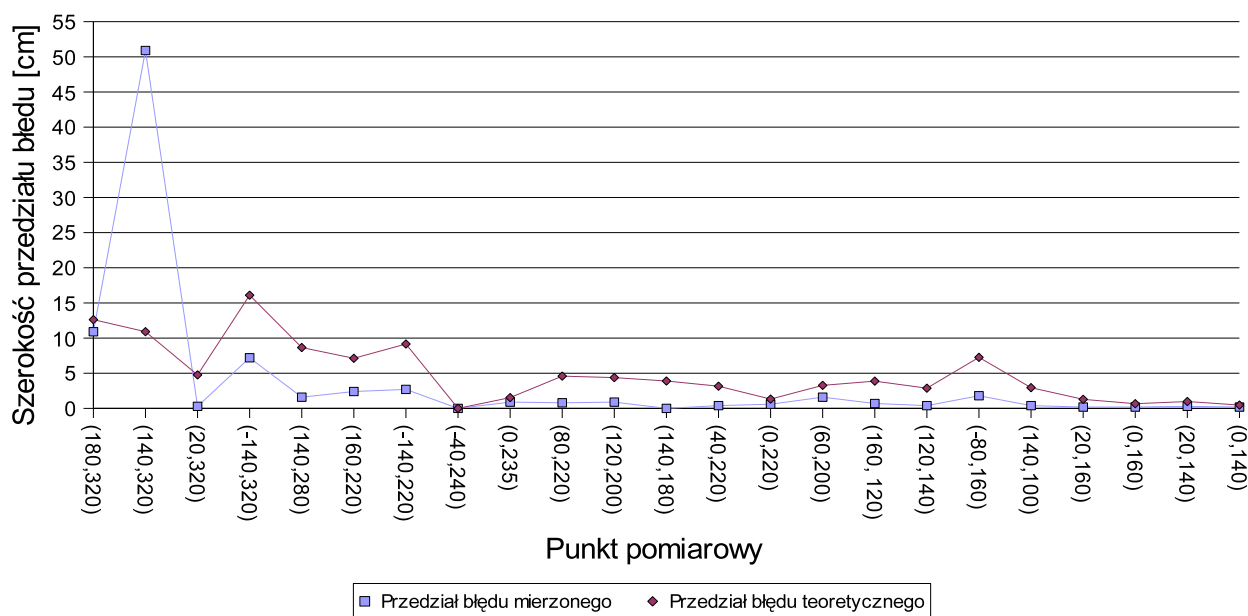
Rysunek 9.3: Porównanie minimalnych i maksymalnych błędów współrzędnej X pomierzonych i wyznaczonych teoretycznie



Rysunek 9.4: Porównanie minimalnych i maksymalnych błędów współrzędnej Y pomierzonych i wyznaczonych teoretycznie



Rysunek 9.5: Porównanie rozpiętości przedziałów błędów współrzędnej X pomierzonych i wyznaczonych teoretycznie



Rysunek 9.6: Porównanie rozpiętości przedziałów błędów współrzędnej Y pomierzonych i wyznaczonych teoretycznie

9.1.2 Własności statystyczne metody

Rozkłady błędów przedstawione w niniejszym rozdziale oraz w rozdziale 9.2.2 dotyczą serii pomiarów wykonanych bez poruszenia kamery. Nie jest to typowy rozkład statystyczny, ze względu na uśrednianie danych wejściowych podczas wyznaczania współrzędnych znacznika.

Źródłem błędów statystycznych jest głównie wpływ zmian oświetlenia na widoczny rozmiar i położenie znacznika w obrazie. Zmiany te są niewielkie, rzędu 1 piksla lub mniejsze. Tabela 9.3 przedstawia podstawowe parametry statystyczne pomiarów: wartości średnie, wariancję i odchylenie standardowe pomiarów obu współrzędnych oraz współczynnik ich wzajemnej korelacji. W kolumnach $XAvg$ i $YAvg$ znajdują się wartości średnie z pomiarów. Standardowe odchylenia od wartości średniej znajdują się w kolumnach $XDev$ i $YDev$, zaś wariancje wartości współrzędnych znajdują się w kolumnach $XVar$ i $YVar$. W kolumnie $Corr$ znajduje się współczynnik korelacji obu pomierzonych wartości współrzędnych. Warto zauważyć, że dla niektórych punktów (140, 180) wszystkie 1000 pomiarów dało ten sam wynik z dokładnością do 0,1 cm. Dla wszystkich przypadków wariancja i odchylenie standardowe od wartości średniej są niewielkie, najczęściej dla każdego punktu dominuje jedna para wartości współrzędnych. Dla większości punktów otrzymane wartości współrzędnych są skorelowane w znacznym stopniu lub nawet całkowicie.

Y	X	XAvg	XDev	XVar	YAvg	YDev	YVar	Corr
320	180	186,7	0,4	0,1	329,3	0,5	0,3	-1,0
320	140	148,8	2,6	6,8	327,0	2,0	4,2	-0,8
320	20	31,6	0,2	0,1	319,8	0,01	0,0001	-0,1
300	-140	-121,7	2,2	4,8	314,4	0,7	0,5	0,9
280	140	143,9	1,9	3,5	279,8	0,4	0,2	-0,6
240	-40	0,0	0,0	0,0	0,0	0,0	0,0	0,0
235	0	-27,3	1,8	3,2	231,8	0,2	0,03	0,4
220	-140	-134,4	0,8	0,6	225,2	0,6	0,3	1,0
220	40	33,6	0,6	0,4	220,5	0,2	0,03	0,8
220	0	-12,5	3,2	10,2	221,6	0,2	0,02	-0,9
100	140	135,3	0,1	0,01	104,6	0,1	0,01	0,6
120	160	155,3	0,2	0,03	122,6	0,2	0,05	1,0
140	0	-0,4	0,05	0,002	139,6	0,1	0,01	-1,0
140	20	19,9	0,0	0,001	136,9	0,0	0,0005	-0,9
140	120	117,8	0,2	0,03	140,8	0,1	0,01	0,6
160	0	-0,5	0,1	0,01	160,7	0,01	0,00004	-0,1
160	20	22,2	0,5	0,2	161,8	0,04	0,002	1,0
160	-80	-81,2	0,8	0,7	158,8	0,6	0,4	1,0
180	140	138,4	0,00003	0,0000000007	183,1	0,0	0,0	-1,0
200	60	52,5	0,2	0,1	197,5	0,1	0,004	-0,5
200	120	119,9	0,2	0,04	202,0	0,2	0,1	-0,2
220	80	66,0	0,04	0,001	224,6	0,2	0,1	-0,1
220	160	160,9	2,2	4,9	225,3	0,9	0,8	-1,0

Tablica 9.3: Parametry statystyczne pomiarów dla metody z pomiarem odległości i szerokości

Poniższe tabele i wykresy przedstawiają szczegółowo następujące sytuacje:

- Punkt o bardzo małej wariancji pomiarów – takich punktów jest wiele, wybrano (120, 200) – zob. tabela 9.4;
- Punkt o dużym błędzie średnim położony naprzeciw znacznika (0, 235);
- Punkt położony w dużej odległości od znacznika o współrzędnych (-140, 300).

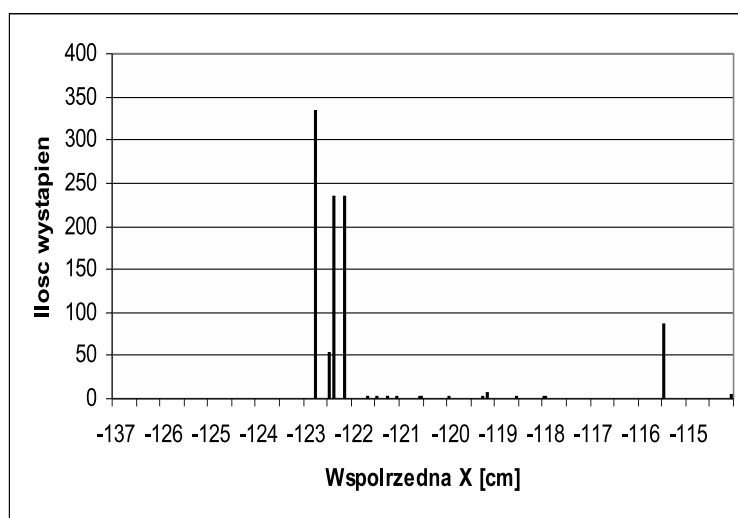
Dla pierwszego przypadku współrzędne są silnie skorelowane. Dla każdej wartości współrzędnej X istnieje jedna (dwie dla $X = 119,9$ cm) główna współrzędna Y , której można oczekiwać przy wystąpieniu określonej współrzędnej X . Jeszcze wyraźniej widać to na przykładzie drugim. Dla każdego z punktów współrzędne przyjmują kilka głównych wartości, nawet jeżeli wszystkich uzyskanych wartości jest bardzo dużo. Wynika to z wyznaczania obu współrzędnych na podstawie tych samych wielkości – odległości od znacznika oraz kąta między osią optyczną kamery i osią prostopadłą do znacznika. Większa wartość jednej współrzędnej powoduje zmniejszenie wartości drugiej współrzędnej, ponieważ obie współrzędne leżą na okręgu którego środkiem jest środek znacznika, a promieniem – odległość od znacznika do kamery.

X [cm]	Y [cm]	ilość wystąpień
119,5	202,0	1
119,5	202,2	2
119,6	202,0	253
119,6	202,2	2
119,7	202,1	1
119,8	202,4	122
119,9	201,5	98
119,9	201,7	1
119,9	202,2	1
119,9	202,4	83
120	202	188
120,1	201,9	212
120,1	202,2	34
120,1	202,3	1
120,2	202,2	1

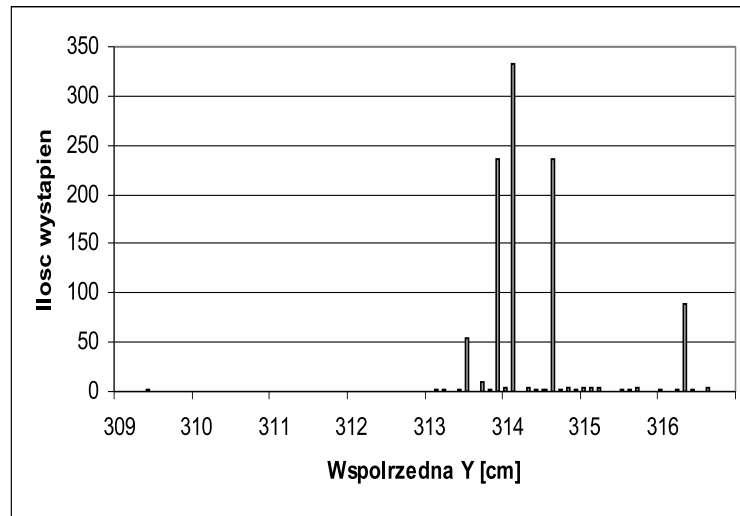
Tablica 9.4: Ilości wystąpień par współrzędnych na 1000 pomiarów w punkcie $X=120$, $Y=200$

X [cm]	Y [cm]	ilość wystąpień
-36,9	231,1	1
-34,8	231,2	1
-35,9	231,2	2
-35,3	231,2	1
-34,9	231,3	1
-34,3	231,4	1
-28,8	231,5	63
-34	231,5	1
-26,1	231,6	1
-26,3	231,6	1
-26,9	231,6	1
-27,8	231,6	1
-32,7	231,6	15
-26,1	231,7	279
-27,8	231,7	1
-31,4	231,7	1
-30,1	231,8	162
-24,7	231,8	2
-25	231,8	1
-24,1	231,8	1
-31,6	231,8	1
-31,3	231,8	1
-25	231,9	12
-23,8	231,9	1
-22,7	231,9	1
-26,6	232	447

Tablica 9.5: Ilości wystąpień par współrzędnych na 1000 pomiarów w punkcie $X=0$, $Y=235$



Rysunek 9.7: Rozkład współrzędnej X dla punktu pomiarowego $X=-140$, $Y=300$ [cm]



Rysunek 9.8: Rozkład współrzędnej Y dla punktu pomiarowego $X=-140$, $Y=300$ [cm]

9.2 Wyniki statyczne dla metody pomiaru dwóch odległości

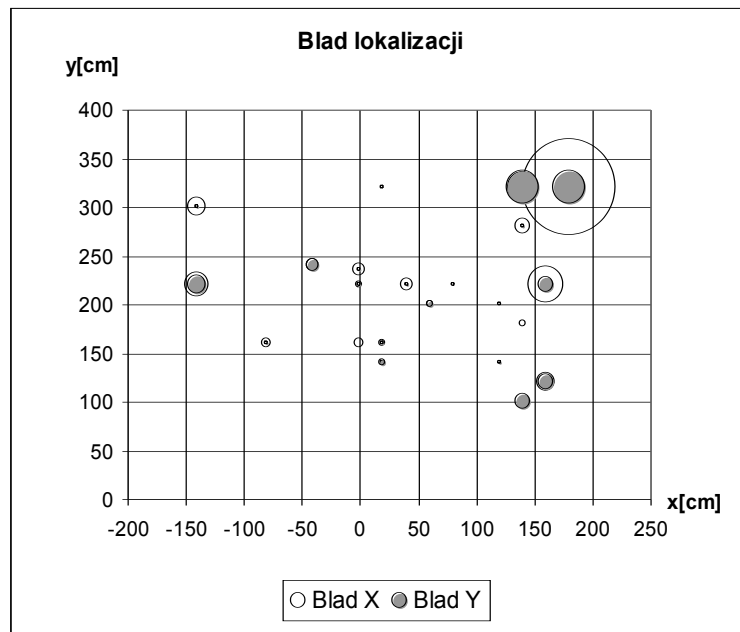
Niniejszy rozdział przedstawia wyniki pomiarów dla metody lokalizacji opisanej w rozdziale 8.5. Poniżej znajdują się wyniki eksperymentów przeprowadzonych, podobnie jak w rozdziale 9.1, dla znacznika w kształcie prostokątnej ramki o szerokości 50 cm i wysokości 70 cm. Wyniki obejmują 23 punkty, w seriach po 1000 pomiarów dla jednego punktu. Kolejno zamieszczono wykresy: błędów współrzędnych w centymetrach 9.9 i w procentach 9.10. Na pierwszym wykresie wielkości okręgów odpowiadają błędom współrzędnych, tzn współrzędna rzeczywista znajduje się w obszarze ograniczonym okręgiem. Na wykresie z błędem procentowym (Rys. 9.10) wielkości okręgów są proporcjonalne do błędu procentowego, ale nie do skali współrzędnych. Największy okrąg symbolizuje błąd 12,5% Rzeczywiste wielkości zamieszczono w tabelach 9.6 i 9.7.

Przedstawione dane pozwalają na wyprowadzenie dwóch wniosków:

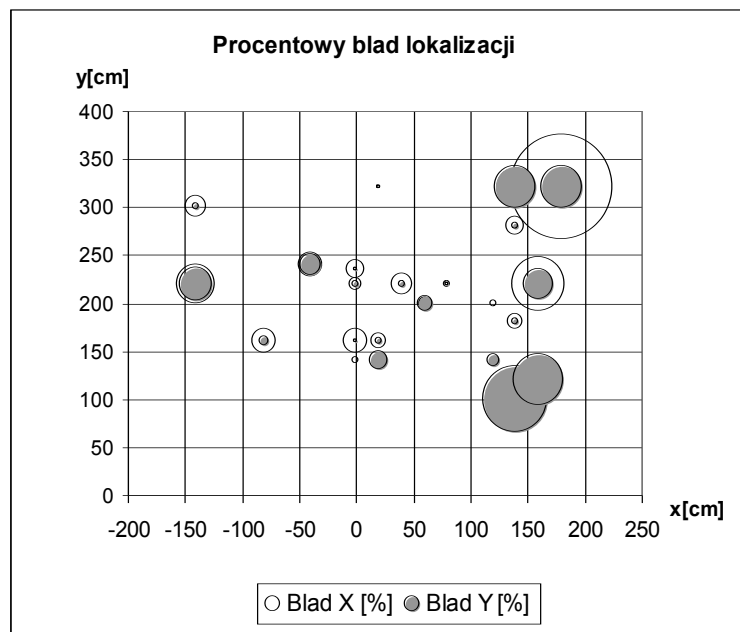
1. Błąd jest najmniejszy w gdy kamera znajduje się w małej odległości od znacznika, blisko linii prostopadłej do środka znacznika,
2. Błąd współrzędnej Y jest znacznie mniejszy niż błąd współrzędnej X.

Błąd współrzędnej rośnie wraz ze wzrostem odległości kamery od znacznika. Wyniki te są zgodne z oczekiwaniami, ponieważ naprzeciw znacznika odległość między znacznikiem i kamerą, a więc i błąd tej odległości, są najmniejsze. Procentowy błąd zwiększa się przy wzroście odległości kamery od znacznika, jeżeli kamera oddala się w kierunku równoległym do płaszczyzny znacznika. Błąd ten nie rośnie znacząco przy wzroście odległości między kamerą i znacznikiem w kierunku prostopadłym do płaszczyzny znacznika.

Błąd średni procentowy współrzędnej Y tylko dla dwóch przypadków znacznie przekracza 5%. Ponieważ nie jest to dla największych odległości, sugeruje to błąd skumulowanie błędów korekcji obrazu lub błąd wyznaczenia położenia rzeczywistego. Średni błąd liniowy współrzędnej Y przekracza 10 cm dla odległości powyżej 3 metrów (przy wysokości znacznika 70 cm i szerokości 50 cm).



Rysunek 9.9: Wartości średniego błędu pomiaru współrzędnych X i Y w cm dla metody z dwoma odległościami



Rysunek 9.10: Wartości średniego błędu pomiaru współrzędnych X i Y (w %) dla metody z dwoma odległościami; największy okrąg odpowiada 12,5% błędowi

X	Y	Min	Max	Avg	Min_Err	Max_Err	Avg_Err	Avg_Err %
180	320	192,7	292,9	225,9	-12,7	-112,9	45,9	12,5
140	320	105,3	159,4	124,7	-19,4	34,7	15,3	4,4
20	320	19,1	22,2	21,8	0,9	-2,2	1,8	0,6
-140	300	-156,0	-126,3	-149,1	-13,7	16,0	9,1	2,7
140	280	137,9	156,0	147,7	2,1	-16,0	7,7	2,5
-40	240	-48,9	-40,6	-46,9	0,6	8,9	6,9	2,9
0	235	-8,0	-3,6	-5,7	3,6	8,0	5,7	2,4
-140	220	-140,6	-125,7	-127,7	0,6	-14,3	12,3	4,7
40	220	41,0	46,9	45,7	-1,0	-6,9	5,7	2,6
0	220	0,4	8,9	3,9	-0,4	-8,9	3,9	1,8
140	100	141,8	145,0	144,1	-1,8	-5,0	4,1	2,4
160	120	149,9	156,5	151,3	3,5	10,1	8,7	4,4
0	140	0,8	1,9	1,1	-0,8	-1,9	1,1	0,8
20	140	21,2	22,2	21,5	-1,2	-2,2	1,5	1,1
120	140	120,2	123,1	121,3	-0,2	-3,1	1,3	0,7
0	160	4,4	4,7	4,6	-4,4	-4,7	4,6	2,9
20	160	22,9	24,0	23,3	-2,9	-4,0	3,3	2,0
-80	160	-75,9	-73,8	-74,6	-4,1	-6,2	5,4	3,0
140	180	144,1	144,4	144,1	-4,1	-4,4	4,1	1,8
60	200	54,4	58,5	56,7	1,5	5,6	3,3	1,6
120	200	118,4	125,4	122,3	1,6	-5,4	2,3	1,0
80	220	79,3	84,4	82,2	0,7	-4,4	2,2	1,0
160	220	177,0	180,2	178,3	-17	-20,2	18,3	6,7

Tablica 9.6: Graniczne i średnie wartości błędu współrzędnej X dla metody lokalizacji z dwoma odległościami

X	Y	Min	Max	Avg	Min_Err	Max_Err	Avg_Err	Avg_Err %
180	320	240,3	326,1	303,6	-79,7	6,1	16,4	5,1
140	320	318,4	343,3	336,8	-1,6	23,3	16,8	5,3
20	320	320,6	321,0	320,6	0,6	1,0	0,6	0,2
-140	300	298,3	313,8	302,3	-1,7	13,8	2,3	0,7
140	280	273,8	285,0	277,8	-6,2	5,0	2,2	0,8
-40	240	232,9	233,9	233,2	-7,1	-6,1	6,8	2,8
0	235	233,0	233,8	233,4	-2,0	-1,2	1,6	0,7
-140	220	221,7	229,7	229	1,7	9,7	9,0	4,1
40	220	218,2	219,5	218,4	-1,8	-0,5	2,0	0,7
0	220	221,4	222,4	221,9	1,4	2,4	1,6	0,9
140	100	90,6	96,3	92,0	-9,4	-3,7	8,0	8,0
160	120	119,3	129,8	127,6	-0,7	9,8	7,5	6,3
0	140	139,5	139,7	139,6	-0,5	-0,3	0,4	0,3
20	140	136,3	136,7	136,6	-3,7	-3,3	3,4	2,4
120	140	136,0	138,7	137,8	-4,0	-1,3	2,2	1,6
0	160	160,5	160,6	160,6	-0,5	0,6	0,6	0,4
20	160	161,3	161,8	161,6	1,3	1,8	1,6	1,0
-80	160	161,7	162,1	161,9	1,7	2,1	1,9	1,2
140	180	178,5	178,7	178,7	-1,5	-1,3	1,3	0,7
60	200	195,8	196,8	196,3	-4,2	-3,2	3,7	1,8
120	200	198,1	202,9	200,6	-1,9	2,9	0,6	0,3
80	220	218,1	220,4	219,2	-1,9	0,4	0,8	0,4
160	220	209,9	212,2	211,8	-10,1	-7,8	8,2	3,7

Tablica 9.7: Graniczne i średnie wartości błędu współrzędnej Y dla metody lokalizacji z dwoma odległościami

Błąd liniowy współrzędnej X rośnie wraz ze wzrostem odległości kamery od osi znacznika, słabiej przy wzroście odległości Y. Dla odległości Y=200 cm średni błąd X wynosi 2-3 cm na wprost znacznika, i 5-20 cm w odległości X wynoszącej około 150 cm od znacznika.

9.2.1 Porównanie z wartościami oczekiwanymi

Poniższe dane zawierają porównanie zmierzonych błędów lokalizacji metodą opisaną w rozdziale 8.5 z wartościami oczekiwanymi wyznaczonymi teoretycznie. Wartości teoretyczne były wyliczone przy założeniu, że wartość każdej współrzędnej jest obciążona błędem nie większym niż 0,5 piksla i są największymi co do modułu wartościami błędów. Wykresy 9.11 i 9.12 przedstawiają porównanie minimalnych i maksymalnych wartości odchylenia od wartości rzeczywistej. Wykresy 9.13 i 9.14 przedstawiają porównanie szerokości przedziałów wartości współrzędnych. Dane z wykresów można interpretować następująco:

1. Pierwszy punkt pomiarowy jest obciążony bardzo dużym błędem. Ponieważ jest to najbardziej odległy punkt w całej serii (współrzędne 180, 320), można przyjąć że w tak dużej odległości błąd wyznaczenia współrzędnych znacznika w obrazie jest znacznie większy niż 0,5 piksla. Podobnie większe niż teoretyczna wartości błędów są dla punktu (140, 320) czyli drugiego punktu pomiarowego.
2. Pozostałe punkty leżą w strefie, w której dokładność jest lepsza niż oczekiwana.

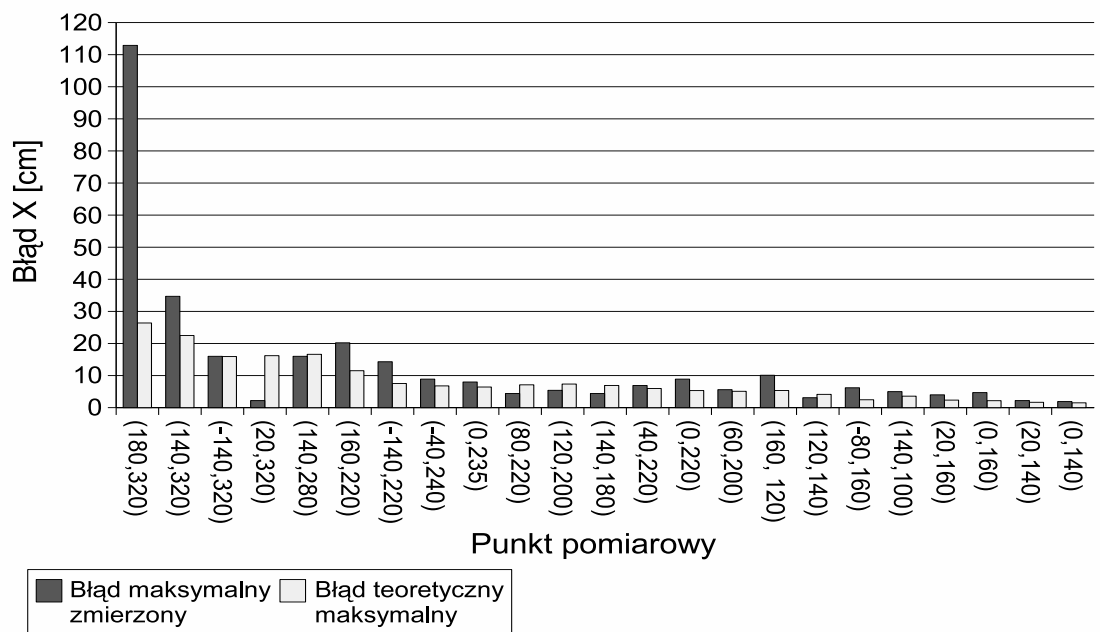
Dla niektórych punktów widać znaczne przesunięcie środka przedziału wartości, co powoduje że np. dla punktu 8 (-140, 220) prawie wszystkie pomiarowe błędy są tego samego znaku. Można to interpretować albo jako błąd wyznaczenia położenia rzeczywistego kamery (dokładność tego pomiaru wynosi ok. 1-2 cm dla dużych przedziałów), albo jako wpływ skończonej dokładności korekcji obrazu. Dokładność wyznaczenia współrzędnych siatki korygującej wynosi 1 piksel, a ten sam znak błędu dla sąsiednich węzłów może spowodować obserwowany efekt.

Przy porównaniu z wynikami poprzedniej metody z rozdziału 9.1 widoczna jest mniejsza dokładność wyników z niniejszego rozdziału. Ponadto widać, że w strefie przed znacznikiem niniejsza metoda jest najdokładniejsza, w przeciwieństwie do metody opisanej w rozdziałach 8.4 i 9.1.

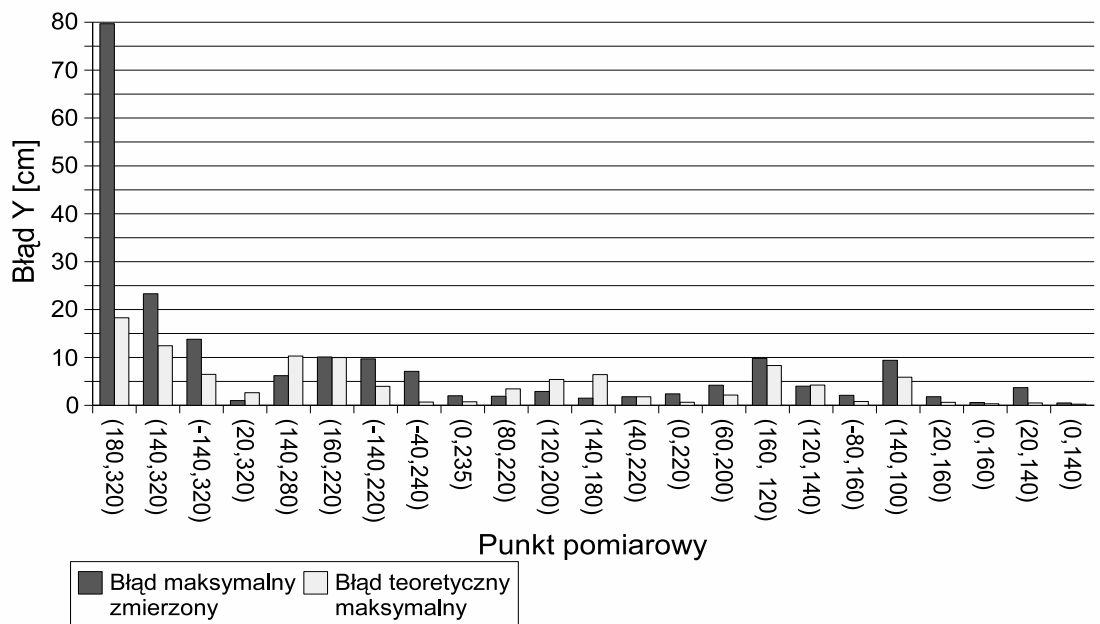
9.2.2 Własności statystyczne metody

Tabela 9.8 przedstawia podstawowe parametry statystyczne pomiarów: wartości minimalną, maksymalną i średnią oraz wariancję i odchylenie standardowe pomiarów obu współrzędnych. W kolumnach $XAvg$ i $YAvg$ znajdują się wartości średnie pomiarów. Standardowe odchylenia od wartości średniej znajdują się w kolumnach $XDev$ i $YDev$, zaś wariancje wartości współrzędnych znajdują się w kolumnach $XVar$ i $YVar$. Ostatnia kolumna, $Corr$, zawiera współczynniki korelacji między obiema współrzędnymi. Wartości uzyskane opisywaną metodą są bardziej skorelowane niż wyniki z rozdziału 9.1.2, tzn współczynnik korelacji jest większy prawie dla wszystkich przypadków.

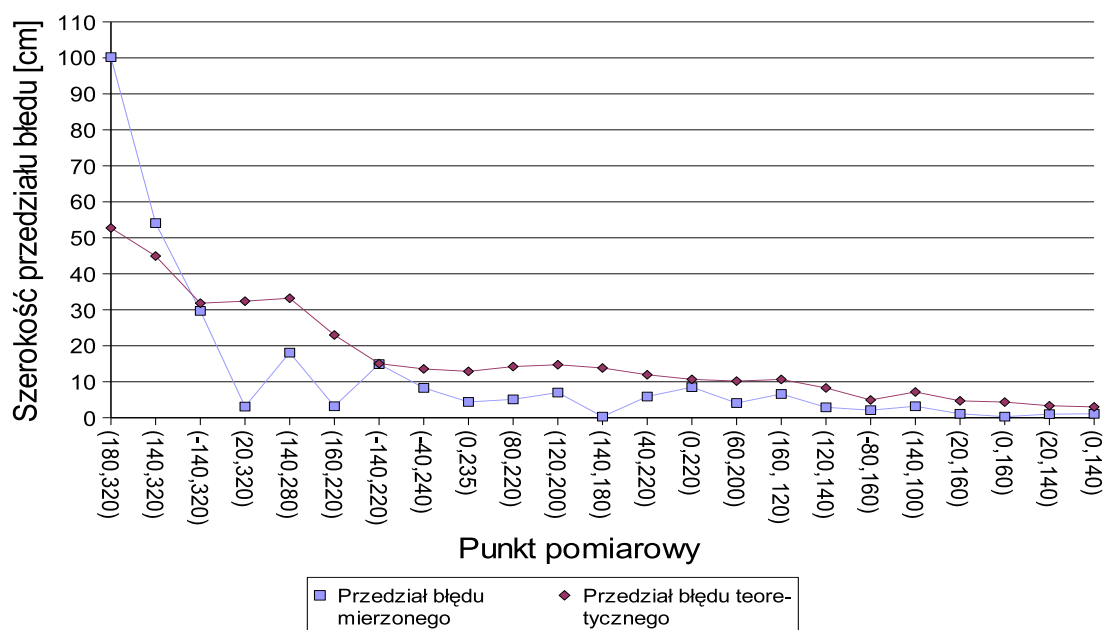
Aby zobrazować statystyczne własności metody, zamieszczono rozkład współrzędnych dla punktów o największej i najmniejszej wariancji pomiarów. Punktem o największej wariancji jest punkt najbardziej odległy od znacznika, o współrzędnych (180, 320). Jak pokazują tabela 9.9 i wykres 9.15, nawet dla najbardziej rozproszonych wyników prawie wszystkie pomiary są



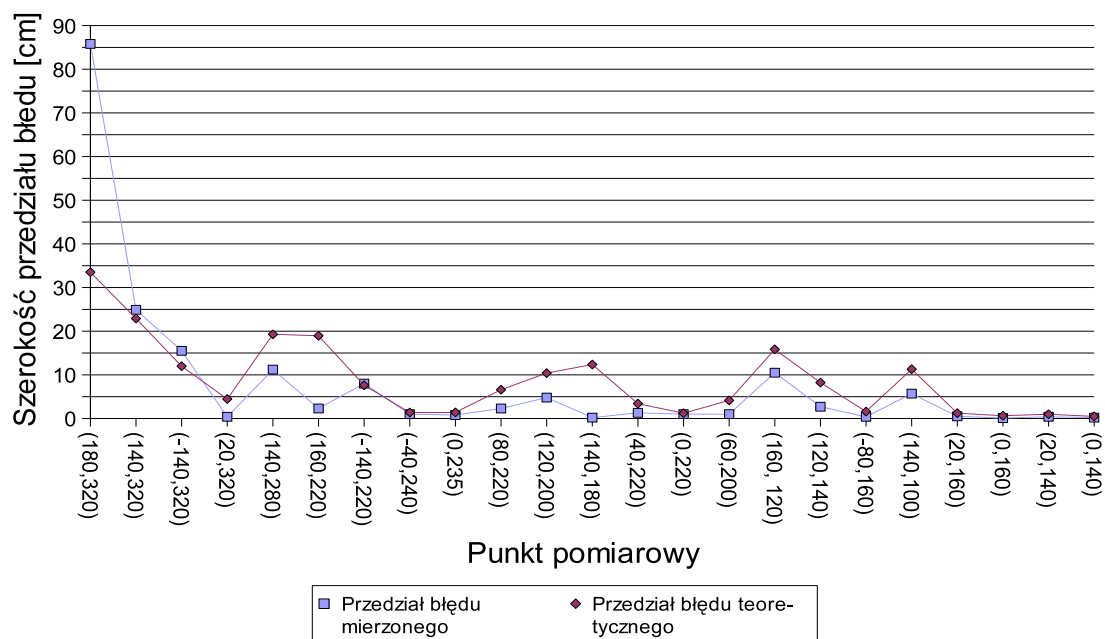
Rysunek 9.11: Porównanie minimalnych i maksymalnych błędów współrzędnej X pomierzonych i wyznaczonych teoretycznie



Rysunek 9.12: Porównanie minimalnych i maksymalnych błędów współrzędnej Y pomierzonych i wyznaczonych teoretycznie



Rysunek 9.13: Porównanie rozpiętości przedziałów błędów współrzędnej X pomierzonych i wyznaczonych teoretycznie



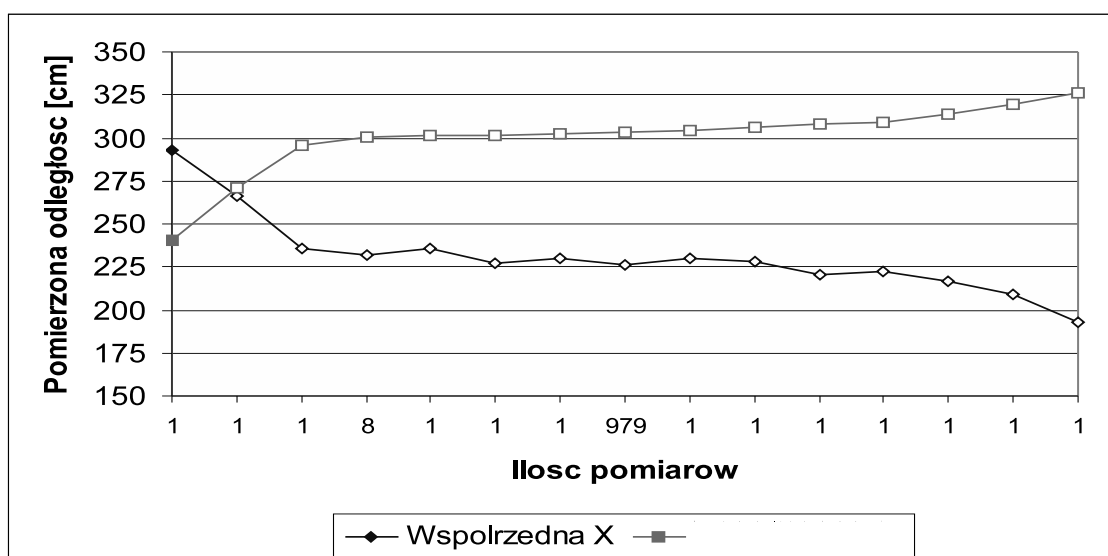
Rysunek 9.14: Porównanie rozpiętości przedziałów błędów współrzędnej Y pomierzonych i wyznaczonych teoretycznie

Y	X	XAvg	XDev	XVar	YAvg	YDev	YVar	Corr
320	180	225,9	2,9	8,2	303,6	2,5	6,1	-1,0
320	140	124,7	6,7	45,3	336,8	2,1	4,4	-0,9
320	20	21,8	0,6	0,3	320,6	0,02	0,0003	-0,2
300	-140	-149,1	3,6	13,3	302,3	1,7	3,0	1,0
280	140	147,7	3,4	11,3	277,8	2,5	6,2	-1,0
240	-40	-46,9	3,2	10,2	233,2	0,3	0,1	1,0
235	0	-5,7	1,4	2,0	233,4	0,2	0,05	-0,6
220	-140	-127,7	3,8	14,2	229,0	2,1	4,5	1,0
220	40	45,7	1,2	1,5	218,4	0,3	0,1	-0,7
220	0	3,9	0,8	0,6	221,9	0,3	0,1	0,9
100	140	144,1	0,9	0,8	92,0	1,7	2,7	-1,0
120	160	151,3	2,1	4,4	127,5	3,0	8,9	-1,0
140	0	1,1	0,4	0,2	139,6	0,1	0,008	1,0
140	20	21,5	0,1	0,004	136,6	0,02	0,0006	-1,0
140	120	121,3	0,6	0,3	137,8	0,4	0,2	-0,9
160	0	4,6	0,05	0,002	160,6	0,003	0,00001	0,1
160	20	23,3	0,4	0,1	161,6	0,2	0,02	-1,0
160	-80	-74,6	0,8	0,6	161,9	0,1	0,02	0,9
180	140	144,1	0,009	0,00009	178,7	0,006	0,00004	-1,0
200	60	56,7	0,4	0,1	196,3	0,1	0,02	-0,9
200	120	122,3	1,7	3,0	200,6	1,2	1,4	-1,0
220	80	82,2	0,9	0,8	219,2	0,5	0,3	-1,0
220	160	178,3	0,1	0,007	211,8	0,7	0,5	-0,1

Tablica 9.8: Parametry statystyczne pomiarów dla metody z pomiarem dwóch odległości

X [cm]	Y [cm]	ilość wystąpień
292,9	240,3	1
266,5	271,2	1
235,6	296,1	1
231,9	300,5	8
235,4	301,5	1
226,9	301,7	1
229,8	302,1	1
225,8	303,7	979
229,7	304,5	1
228,1	306,2	1
220,2	308,4	1
222	309,3	1
216,4	313,9	1
209,1	319,5	1
192,7	326,1	1

Tablica 9.9: Ilość odczytów różnych wartości współrzędnych w serii 1000 pomiarów dla punktu (180, 320)



Rysunek 9.15: Ilości odczytów różnych wartości współrzędnych

identyczne. Poszczególne wartości współrzędnych X i Y są skorelowane w 100%, czyli każdej zmierzonej wartości współrzędnej X odpowiada dokładnie jedna zmierzona wartość Y.

Dla punktu o współrzędnych (0, 140) wariancja jest najmniejsza, podobnie jak odległość punktu pomiarowego od znacznika. Pomierzone współrzędne X i Y są skorelowane, ale nie w 100%. Ilości wystąpień współrzędnych z uwzględnieniem par przedstawia tabela 9.10. Wartość statystyczną pomiarów dla tak małej odległości należy traktować z dużą ostrożnością, ponieważ dla współrzędnej Y zmierzono tylko 3 wartości, położone w przedziale o szerokości 2 mm, zaś dla współrzędnej X – 6 wartości w przedziale 8 mm. Warto zauważyć, że dokładność wyznaczenia położenia kamery za pomocą taśmy mierniczej wynosi ok 1-2 cm, zatem dla małej odległości między kamerą i znacznikiem wszystkie pomiary są praktycznie jednakowe.

X [cm]	Y [cm]	ilość wystąpień
0,8	139,5	654
0,9	139,5	1
0,9	139,6	1
1,1	139,5	58
1,7	139,6	1
1,7	139,7	284
1,9	139,6	1

Tablica 9.10: Ilość odczytów różnych wartości współrzędnych w serii 1000 pomiarów dla punktu (0, 140)

9.3 Nawigacja oparta o kamerę jako sensor

Celem eksperymentu było porównanie nawigacji bazującej wyłącznie na odometrii z nawigacją bazującą na odometrii wspomaganą pomiarami pochodzącymi z sensora-kamery opracowanego w niniejszej pracy. Eksperyment przeprowadzono używając robota Pioneer firmy RWI

jako robota wyposażonego w odometrię o słabej dokładności oraz robota NAVIGATOR skonstruowanego na Wydziale Mechatroniki Politechniki Warszawskiej. Robot NAVIGATOR jest wyposażony w różne sensory (niektóre można montować jako moduły), w tym głowicę obrotową z kamerą. Głowica ma możliwość obrotu o 90° w prawo i w lewo od kierunku jazdy robota, co umożliwia objęcie wizją obszaru kąтового ponad 180° . Na czas eksperymentu robot NAVIGATOR wraz z głowicą stanowił element zamocowany na pokrywie robota Pioneer. Konstrukcję przedstawiono na rysunku 9.16.



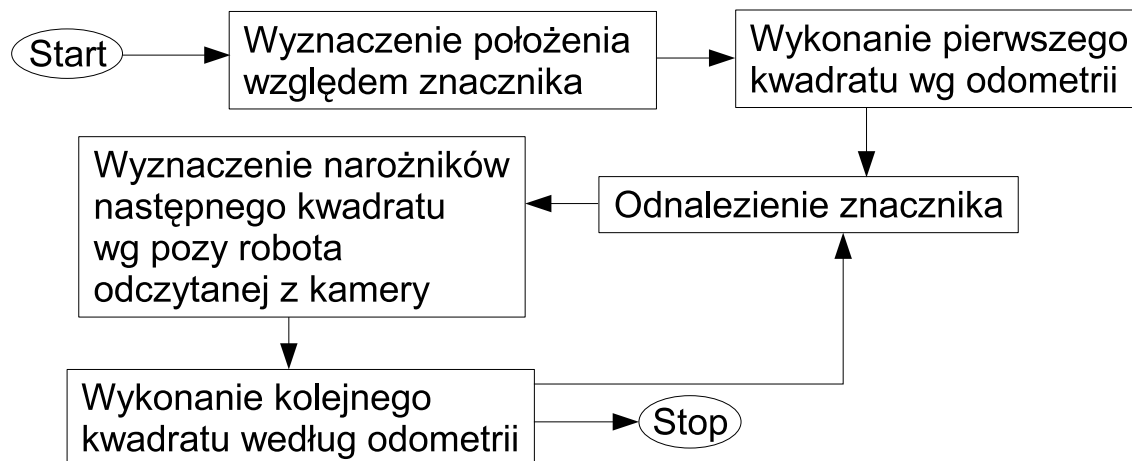
Rysunek 9.16: Roboty Pioneer i NAVIGATOR w zestawie używanym do eksperymentów

Podczas eksperymentu jazdą robota sterował moduł SAPHIRA, będący standardowym oprogramowaniem dla robotów Pioneer. SAPHIRA umożliwia jazdę robotem z określoną prędkością i obrót robota, oraz dostarcza informacji o bieżących współrzędnych i orientacji robota w układzie współrzędnych związanych z odometrią robota. Ze względu na niedokładności odometrii wraz z ilością wykonanych skrętów wzrasta błąd określenia położenia.

Przeprowadzony eksperyment polegał na wykonaniu testu UMBMark, opisanego w rozdziale 2.2.1. Wyniki tego testu przyjęto za standardową miarę ilościową błędów systematycznych odometrii. Następnie wykonano jeszcze raz test UMBMark, z dodatkowym określaniem położenia robota za pomocą kamery po przejechaniu każdego kwadratu. Algorytm sterowania przedstawia rysunek 9.17.

Podczas nawigacji z zastosowaniem kamery, pierwszym etapem było ustalenie położenia robota według widocznego znacznika. Następnie robot planował trasę jazdy po kwadracie w układzie współrzędnych związanym ze znacznikiem. Współrzędne te zostały zapamiętane i według nich robot wykonywał kolejne kwadraty. Przed wykonaniem kwadratu współrzędne były przeliczane na układ współrzędnych związany z odometrią robota. Jazda po kwadracie była oparta na pomiarach odometrycznych. Po przejechaniu każdego kwadratu następowało

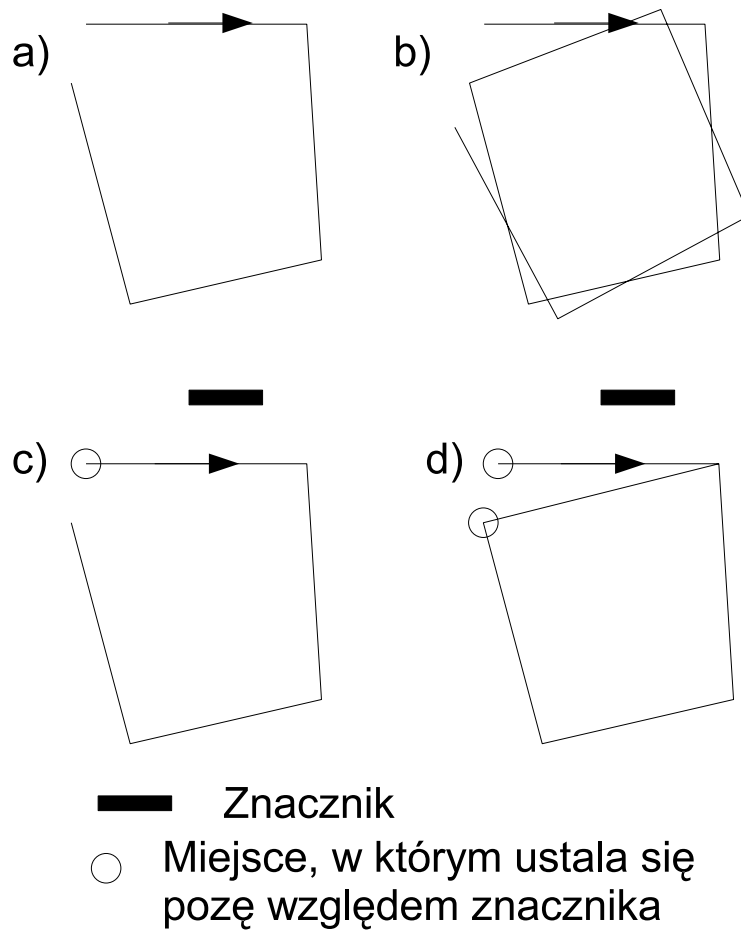
wyznaczenie pozy robota w układzie współrzędnych związanym ze znacznikiem i na tej podstawie oraz znajomości pozy robota według odometrii następowalo przeliczenie współrzędnych kwadratu na nowe współrzędne w układzie współrzędnych odometrii. W ten sposób końcowy bład pozy był błędem przejechania jednego kwadratu na podstawie odometrii.



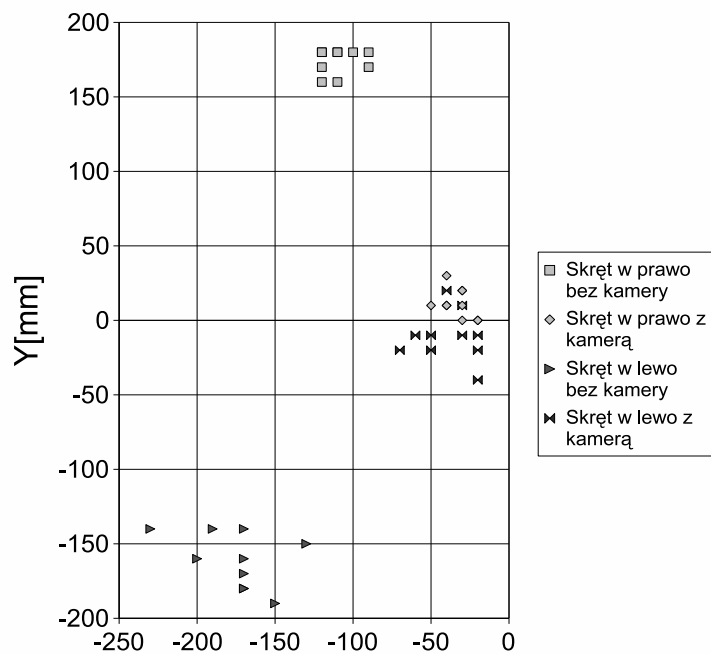
Rysunek 9.17: Algorytm jazdy w teście UMBMark z korekcją pozy za pomocą kamery.

Rysunek 9.18 przedstawia schematycznie trajektorię robota podczas jazdy wyłącznie na podstawie odometrii i ze wspomaganie kamerą. Można łatwo zauważyć, że bład spowodowany niedokładnością odometrii był zerowany przez określenie pozy robota na podstawie znacznika.

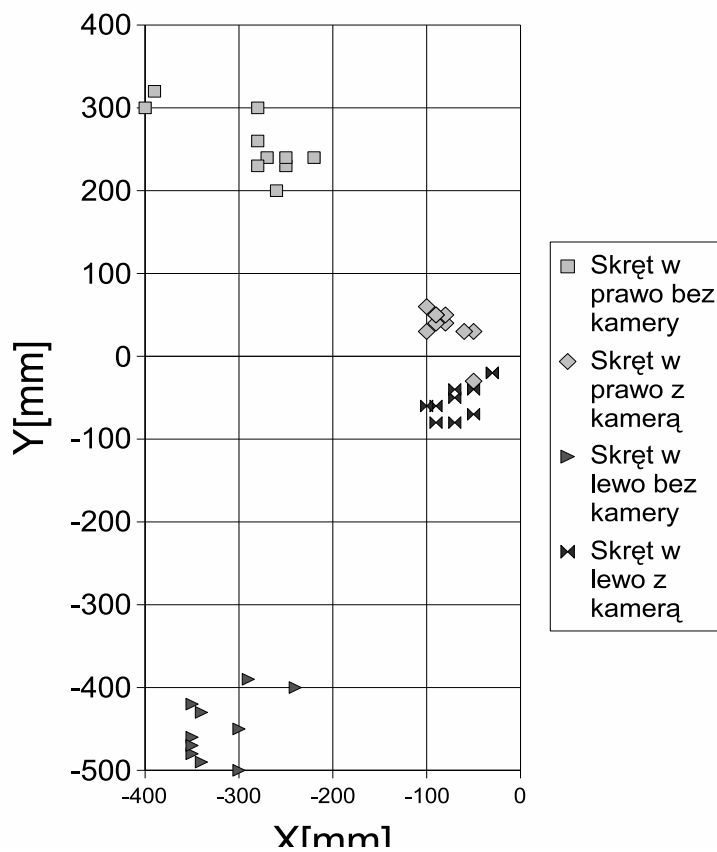
Ze względu na ograniczenia przestrzenne wykonano pomiary dla kwadratów o boku 1 i 2 metrów. Wyniki przedstawiono na wykresach 9.19 dla kwadratów o boku 1 m i 9.20 dla kwadratów o boku 2 metrów. W obu wariantach bład podczas jazdy wyłącznie na podstawie odometrii jest kilkakrotnie większy niż bład przy okresowym określaniu położenia za pomocą kamery i znacznika.



Rysunek 9.18: Trajektorie robota: a) po przejechaniu pierwszego kwadratu na podstawie odometrii; b) po przejechaniu dwóch kwadratów na podstawie odometrii; c) po przejechaniu pierwszego kwadratu na podstawie znacznika i odometrii; d) jazda po drugim kwadracie według znacznika i odometrii



Rysunek 9.19: Wyniki testu dla kwadratów o boku 1 metra



Rysunek 9.20: Wyniki testu dla kwadratów o boku 2 metrów

Rozdział 10

Podsumowanie i wnioski

10.1 Dowód tezy

Obecnie lokalizację robota mobilnego przeprowadza się nie na podstawie jednego pomiaru, lecz serii pomiarów wykonywanych najczęściej podczas jazdy. Pojedynczy pomiar dostarcza częściowych danych. Dokładne wyznaczenie położenia robota na podstawie jednego pomiaru z zastosowaniem kamery było możliwe przez stereowizję, czyli układ kamer sprzężonych ze sobą, lub przez triangulację, czyli pomiar azymutu do trzech obiektów znajdujących się na jednym zdjęciu. Też pracą było sprawdzenie możliwości lokalizacji robota przy użyciu jednej kamery na podstawie jednego obrazu jednego znacznika. **W niniejszej pracy udowodniono, że jest to możliwe.** Dowód teoretyczny (matematyczny) został potwierdzony dowodem praktycznym, czyli miarodajną próbą nawigacji robota z zastosowaniem opracowanej metody lokalizacji.

Do lokalizacji nadają się znaczniki płaskie o dwóch wyraźnych pionowych krawędziach zewnętrznych. Kryterium to pozwala na zastosowanie szerokiej gamy przedmiotów znajdujących się w typowym otoczeniu robota jako znaczników naturalnych – mogą to być drzwi, meble, kalendarze i plakaty i inne obiekty o kształcie prostokątnym. Kształty znaczników pasujące do metody lokalizacji przedstawiono w rozdziale 6.3.

Kształt i barwa obiektów są cechami, na podstawie których przeprowadzono identyfikację. W rozdziale 7 przedstawiono technikę detekcji i dokładnego pomiaru sztucznych znaczników o jednolitej barwie dających się dość dobrze wpisać w prostokąt. Przedstawiono metodę detekcji obszarów barwnych o jednolitej barwie i obszarów o zbliżonej jasności w obrazie szarym. Przedstawiono też metodę detekcji obiektów o określonych kształtach, opartą na momentach centralnych i niezmiennikach H_u , do której zastosowano szybki algorytm wyznaczający te wartości dla wszystkich obszarów w obrazie podczas jednego skanowania obrazu. Przedstawiona metoda pozwala na identyfikację w przypadku gdy w otoczeniu robota znajdują się odpowiednie znaczniki, w zakresie odległości umożliwiającą dobrą lokalizację.

Ze względu na wyznaczanie odległości na podstawie widoczności pionowych krawędzi wybrane **znaczniki muszą posiadać 2 wyraźne krawędzie pionowe oddalone od siebie o znaną odległość.** Takich elementów w otoczeniu robota jest wiele, znacznikami naturalnymi mogą być drzwi, niektóre elementy wyposażenia - np szafa, także okna i znajdujące się na ścianach plakaty lub kalendarze. Takie obiekty mają kształt prostokątny. W rozdziale 6.3 przedstawiono też inne możliwe kształty znaczników.

W rozdziałach 8.4 i 8.5 opisano **metodę lokalizacji na podstawie sztucznego znacznika.** Opisane w rozdziale 8 metody wyznaczania odległości, azymutu i położenia są dokładne, dopóki odległość między kamerą a znacznikiem nie przekroczy ok. 10 wysokości znacznika, zaś

współrzędna Y - ok. 7 wysokości znacznika. Przebadane znaczniki miały większą wysokość niż szerokość, miały one rozmiary: 210x297 mm (A4), 50x40cm, 70x50cm. Znacznik o rozmiarze 210x297 mm pozwalał na lokalizację przy współrzędnych Y mniejszych niż ok. 2 metry, i współrzędnych X mniejszych niż 1,5 metra.

Położenie i orientację robota wyznacza się dwoma metodami na podstawie współrzędnych 4 narożników prostokąta. Ponieważ jedna z metod, opisana w rozdziale 8.4, wyznacza położenie na podstawie niesymetrycznych wartości, można stosować dwa warianty tej metody, "lewy" i "prawy", które są względem siebie symetryczne. W rozdziale 8.4 opisano jeden z wariantów tej metody. Druga metoda, opisana w rozdziale 8.5, bazuje na symetrycznej informacji i nie ma wariantów. W rozdziałach 8.4.1 i 8.5.1 przedstawiono strefy, w których poszczególne współrzędne mogą być określone z dokładnością 2%, 5% i 10%. Strefy te uzupełniają się umożliwiając dokładną lokalizację w całym obszarze przed znacznikiem, pod warunkiem że kąt między osią kamery a płaszczyzną znacznika nie jest zbyt ostry, tzn. jest powyżej 15°. W miejscach gdzie strefy obu metod zachodzą na siebie, stosowano metodę o lepszej dokładności w danym miejscu. Podczas eksperymentów wyznaczano trzy pozy robota, dwie metodą z wyznaczeniem odległości jednej krawędzi (oddzielnie dla każdej krawędzi) oraz jedną pozę metodą z wyznaczeniem odległości dwu krawędzi. Dla każdej wyznaczonej pozy robota wyznaczano wrażliwość jej wyznaczenia na błędy jako maksymalną różnicę między wartościami wyznaczonymi i pierwszymi pochodnymi względem współrzędnych narożników. Za wartości najbliższe rzeczywistym uznawano dane z metody najmniej wrażliwej na błędy w danym punkcie.

Uzyskanie możliwie najdokładniejszego pomiaru wymagało opracowania metody eliminacji zniekształceń spowodowanych optyką kamery oraz metody jak najdokładniejszego wyznaczenia położenia znacznika w obrazie z kamery.

W rozdziale 4 opisano autorską metodę eliminacji zniekształceń obrazu, która może być stosowana również gdy znaczną część zniekształceń stanowią zniekształcenia inne niż regularne (np pierścieniowe) lub proste (np asymetryczne położenia osi optycznej w obrazie). Podstawą kalibracji i późniejszej korekcji obrazu jest mapa zniekształceń sporządzona na podstawie jednego widoku siatki. Opracowano proces kalibracji kamery, wraz z programem umożliwiającym kalibrację każdej kamery. Sporządzony program potrafi wykryć przebieg linii siatki w obrazie uzyskanym z kamery, wyznaczyć punkty przecięcia linii poziomych i pionowych, wyznaczyć równania linii do korekcji. Wyznaczone współrzędne węzłów mogą być zapisane w postaci obrazu z naniesionymi odpowiednimi punktami, i również mogą być z obrazu odtworzone. Zapis mapy w postaci obrazu pozwala na przyspieszenie przygotowania programu do korekcji obrazu, ponieważ odtworzenie punktów przecięcia z odpowiednio przygotowanego obrazu jest o kilka rzędów szybsze niż każdorazowe wykrycie linii przez analizę obrazu uzyskanego podczas kalibracji. Do utworzonej siatki program buduje tablice przekształcenia korygującego obraz, co pozwala na wykonanie korekcji w bardzo krótkim czasie. Dzisiaj stosowane komputery wykonują korekcję obrazu o rozmiarach 640x480 pikseli w czasie znacznie krótszym niż 0,01 sekundy. Dokładność przekształcenia korygującego wynosi 1 piksel, tzn. współrzędne w uzyskanym obrazie bez zniekształceń różnią się od współrzędnych w obrazie idealnym bez zniekształceń o nie więcej niż jeden piksel.

W rozdziale 7.3 przedstawiono sposób wyznaczenia brzegów obszaru wpisanego w prostokąt, oraz dokładnego pomiaru rozmiarów obszaru. Pomiar polega na wyznaczeniu prostych w bliskim otoczeniu obszaru, identyfikacji wśród nich brzegów obszaru oraz dokładnym wyznaczeniu punktów przecięcia linii na podstawie równań linii. Pomijając błąd wprowadzony przez korekcję zniekształceń, współrzędne narożników są wyznaczane z dokładnością do 0,1 piksla.

10.2 Porównanie z podobnymi metodami lokalizacji

W pracach dotyczących robotyki przedstawionych w rozdziale 2 położenie robota wyznacza się stosując jeden sensor i wiele znaczników (metoda triangulacji) lub stosując wiele sensorów i wyznaczając głębię widocznej sceny (stereowizja). Metoda którą zaproponowano można uznać za rozszerzony wariant triangulacji z czterema znacznikami w określonej geometrii. W porównaniu z klasyczną triangulacją metoda jest znacznie mniej wymagająca: pozwala na dokładne określenie położenia na podstawie jednego znacznika, który można uważać za cztery wirtualne znaczniki czyli narożniki znacznika położone blisko siebie. W pracy [65] analiza dokładności lokalizacji na podstawie blisko położonych znaczników mieszczących się w jednym obrazie pokazuje, że mogą wystąpić niekorzystne konfiguracje powodujące bardzo duży błąd lokalizacji. Z kolei dobrą dokładność przy triangulacji zawsze uzyskuje się jeżeli znaczniki są znacznie oddalone od siebie kątowno, co wymaga porównania kilku zdjęć lub zastosowania skanera laserowego umożliwiającego prawie jednoczesny pomiar w obszarze o szerokich granicach kątowych. Metoda lokalizacji z jednym prostokątnym znacznikiem jest pozbawiona wymienionych niedogodności triangulacji, tzn jeden prostokątny znacznik z czterema dokładnie wyznaczonymi narożnikami pozwala na dokładną lokalizację (do pewnej odległości granicznej).

Druga z stosowanych metod, stereowizja, wymaga stosowania kilku sensorów (najczęściej kamer) jednocześnie. Porównanie widoku tych samych obiektów z różnych kamer oddalonych od siebie o znaną odległość pozwala na wyznaczenie odległości do widocznych obiektów. Choć rezultaty stereowizji są zadowalające, jest to metoda pracochłonna, ponieważ wymaga nie tylko kalibracji indywidualnej każdej z kamer, lecz także wyznaczenia wspólnych parametrów opisujących układ kamer. Podobnie przy wykrywaniu obiektów nie wystarczy detekcja znaczników, lecz konieczne jest ich dopasowanie w obrazach. Przy udanym dopasowaniu można określić odległość do dopasowanych elementów obrazu, tworząc trójwymiarowy obraz. Zaproponowana metoda jest prostsza niż stereowizja, wystarczy prosta kalibracja jednej kamery i wykrywanie znaczników w pojedynczym obrazie bez dopasowania widoku znacznika do innego obrazu.

Wydaje się, że elementem nowatorskim jest zastosowanie pojedynczego znacznika o odpowiednim kształcie do lokalizacji robota, co zostało opisane w pracy w rozdziale 8 i częściowo w pracy [47]. Opisana w rozdziale 4 metoda kalibracji oraz opisana w rozdziale 7 metoda detekcji znaczników mogą również zawierać elementy nowatorskie.

Literatura

- [1] R. Alur, A. Das, J. Esposito, R. Fierro, Y. Hur, G. Grudic, V. Kumar, I. Lee, J. P. Ostrowski, G. Pappas, J. Southall, J. Spletzer, and C. Taylor. A framework and architecture for multirobot coordination. In *ISER00, Seventh International Symposium on Experimental Robotics*, pages 289–300, December 2000.
- [2] C. Andersen, S. James, and J. Crowley. Appearance based processes for visual navigation. In *Proc. of the 5th International Symposium on Intelligent Robotic Systems SIRS'97*, pages 227–236. Royal Institute of Technology, Jul 1997.
- [3] Claus S. Andersen, Stephed D. Jones, and James L. Crowley. Appearance based processes for visual navigation. In *Proceedings of the 5th International Symposium on Intelligent Robot Systems, SIRS'97*, July 1997.
- [4] J.M. Armingol, L. Moreno, A. de la Escalera, and M.A. Salichs. Landmark perception planning for mobile robot localization. In *IEEE International Conference on Robotics and Automation*, 1998.
- [5] Romuald Aufrere, Jay Gowdy, Christoph Mertz, Chuck Thorpe, Chieh-Chih Wang, and Teruko Yata. Perception for collision avoidance and autonomous driving. *Mechatronics*, 13(10):1149–1161, December 2003.
- [6] Jon Azpiazu, Tim Smithers, and Iñaki Rañó. An adaptive colour based face detector system for a social mobile robot. In *Proceedings of the 2nd European Conference on Mobile Robots ECMR05*, 2005.
- [7] A. Bal, H. Palus, and P. Wołczyk. Selected properties of perceptual colour spaces. In *Proceedings of 2nd Conference on Computer Recognition Systems KOSYR'01*. Miłków, 2001.
- [8] David Ball, Gordon Wyeth, and Stephen Nuske. A global vision system for a robot soccer team.
- [9] Thorsten Bandlow, Alexa Hauck, Tobias Einsele, and Georg Färber. Recognising objects by their silhouette. Technical report, Lab. for Process Control and Real-Time Systems, Technische Universität München, Germany, 1990.
- [10] B. Barshan and H. F. Durrant-Whyte. *Inertial Navigation Systems for Mobile Robots*, volume 11, pages 328–342. June 1995.
- [11] Robert Bączyk. Metody korekcji zniekształceń obrzów w systemie lokalizacji robota mobilnego. In *VII Krajowa konferencja robotyki*. Instytut Cybernetyki Technicznej Politechniki Wrocławskiej, 2001.

- [12] Robert Bączyk and Piotr Skrzypczyński. A framework for vision-based positioning in a distributed robotic system. In *Proceedings of the 1st European Conference on Mobile Robots ECMR03*, pages 153–158, 2003.
- [13] M. Betke and K. Gurvits. Mobile robot localization using landmarks. In *Proceedings of the IEEE International Conference on Robotics and Automation*, volume 2, pages 135–142, 1994.
- [14] M. Betke and K. Gurvits. *Mobile robot localization using landmarks*, volume 13, pages 251–263. 1997.
- [15] Giovanni Bianco, Alexander Zelinsky, and Miriam Lehrer. Visual landmark learning, 2000.
- [16] J. Borenstein., H.R. Everett., L. Feng, and D. Wehe. *Mobile Robot Positioning: Sensors and Techniques.*, volume 14, pages 231–249. April 1997.
- [17] Johann Borenstein. *Internal Correction of Dead-reckoning Errors With the Compliant Linkage Vehicle*, volume 12, pages 257–273. 1995.
- [18] Johann Borenstein and Liqiang Feng. Umbmark – a method for measuring, comparing, and correcting dead-reckoning errors in mobile robots. Technical report, University of Michigan, December 1994.
- [19] Johann Borenstein and Liqiang Feng. Correction of systematic odometry errors in mobile robots. In *Proceedings of the 1995 International Conference on Intelligent Robots and Systems (IROS '95)*, pages 569–574, August 1995.
- [20] Johann Borenstein and Liqiang Feng. Umbmark: A benchmark test for measuring dead-reckoning errors in mobile robots. In *SPIE Conference on Mobile Robots*, 1995.
- [21] Olivier Chomat and James Crowley. Recognition motion using local appearance. Technical report, Lab GRAVIR-IMAG Saint-Martin, France, 1998.
- [22] C. Cohen and F. Koss. A comprehensive study of three-object triangulation. In *SPIE Mobile Robots VII*, 1992.
- [23] Aveek K. Das, Rafael Fierro, R. Vijay Kumar, B. Southall, John R. Spletzer, and Camillo J. Taylor. Real-time vision-based control of a nonholonomic mobile robot. In *Proceedings of the IEEE International Conference on Robotics and Automation 2001 (ICRA 2001)*, volume 2, pages 1714–1719. IEEE, 2001.
- [24] V. de Verdiere and J. Crowley. Local appearance space for recognition of navigation landmarks. In *Journal of Robotics and Autonomous Systems*. 1999.
- [25] V. de Verdière and J. Crowley. A prediction–verification strategy for object recognition using local appearance, 1999.
- [26] Guilherme N. DeSouza and Avinash C. Kak. Vision for mobile robot navigation: A survey. In *IEEE Transactions on pattern analysis and machine intelligence*, volume 24, pages 237–267. IEEE, Feb 2002.

- [27] Frédéric Devernay and Olivier Faugeras. Automatic calibration and removal of distortion from scenes of structured environments.
- [28] Tom Duckett, Stephen Marsland, and Jonathan Shapiro. Simultaneous localization and mapping - a new algorithm for a compass-equipped mobile robot. In *Proceedings of Seventeenth International Joint Conference on Artificial Intelligence IJCAI-2001 Workshop on Reasoning with Uncertainty in Robotics*, pages 27–34, August 2001.
- [29] R.O. Duda and P.E. Hart. Use of the hough transform to detect lines and curves in pictures. *Communications of the ACM*, 15(1):11–15, January 1972.
- [30] R.O. Duda and P.E. Hart. Use of the hough transform to detect lines and curves in pictures. In *CMetImAly77*, pages 204–208, 1977.
- [31] Rafael Fierro, Aweek Das, John Spletzer, Joel Eposito, Vijay Kumar, James P. Ostrowski, George Pappas, and Camillo J. Taylor. A framework and architecture for multi-robot coordination. In *The International Journal of Robotics Research*, volume 21, pages 977–995. SAGE Publications, 2002.
- [32] Jan Flusser and Tomáš Suk. Pattern recognition by affine moment invariants. In *Pattern Recognition*, volume 26, pages 167–174, 1993.
- [33] Jan Flusser and Tomáš Suk. Affine moment invariants: A new tool for character recognition. In *Pattern Recognition Letters*, volume 15, pages 433–436. DBPL, 1994.
- [34] Jan Flusser and Barbara Zitová. Combined invariants to linear filtering and rotation. In *International Journal of Pattern Recognition and Artificial Intelligence*, volume 13, pages 1123–1135. World Scientific Publishing Company, August 1999.
- [35] Jan Flusser and Barbara Zitová. Landmark recognition using invariant features. In *Pattern Recognition Letters*, number 20, pages 541–547. Elsevier Science B.V., 1999.
- [36] Jan Flusser, Barbara Zitová, and Tomáš Suk. Invariant-based registration of rotated and blurred images.
- [37] Udo Frese and Tom Duckett. A multigrid approach for accelerating relaxation-based SLAM. In *Proc. IJCAI Workshop on Reasoning with Uncertainty in Robotics (RUR 2003)*, pages 39–46, Acapulco, Mexico, 2003.
- [38] Rafael Gonzalez. *Digital Image Processing*. Addison-Wesley Publishing Company, 1987.
- [39] Dimitry O. Gorodnichy and William W. Armstrong. Single camera stereo for mobile robot world exploration. In *Vision Interface 99*, pages 528–535, May 1999.
- [40] Nicolas Guil, Julián Cozár, and Emilio Zapata. Planar 3d object detection by using the generalizad hough transform. In *10th International Conference on Image analysis and Processing (ICIAP 1999)*, pages 358–363. IEEE Computer Society, 1999.
- [41] Dirk Hähnel, Sebastian Thrun, Ben Wegbreit, and Wolfram Burgard. An efficient fastslam algorithm for generating maps of large-scale cyclic environments from raw laser range measurements. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, volume 1/2003, pages 206–211, 2003.

- [42] Dirk Hähnel, Sebastian Thrun, Ben Wegbreit, and Wolfram Burgard. *Towards Lazy Data Association in SLAM*, volume 15/2005, page 421. Springer Berlin / Heidelberg, 2005.
- [43] Ireneusz Hallmann. Wykrywanie obiektów o charakterystycznych barwach. In *Automation 2003 - nowości i perspektywy*, 2003.
- [44] Ireneusz Hallmann. Detection of rectangular landmarks. In *Advanced Soft in Computing. Proceedings of the 4th International Conference on Computer Recognition Systems CORES'05*, pages 395–402. Springer, 2005.
- [45] Ireneusz Hallmann and Adam Borkowski. Korekcja zniekształceń nieliniowych obrazu. In *Automation 2002 - nowości i perspektywy*, pages 429–433, 2002.
- [46] Ireneusz Hallmann and Adam Borkowski. Zautomatyzowana kalibracja kamery szerokokątnej. In *Automation 2005 - nowości i perspektywy*, pages 383–391, 2005.
- [47] Ireneusz Hallmann and Barbara Siemiątkowska. Artificial landmark navigation system. In *Proceedings of the 9th Symposium on Intelligent Robotic Systems SIRS'2001*, pages 219–228, 2001.
- [48] P.V.C. Hough. Machine analysis of bubble chamber pictures. In *International Conference on High Energy Accelerators and Instrumentation*. CERN, 1959.
- [49] P.V.C. Hough. Method and means for recognizing complex patterns, 1962.
- [50] Kohsia S. Huang and Mohan M. Trivedi. Distributed video arrays for tracking, human identification and activity analysis. In *Proceedings of the 4th IEEE International Conference on Multimedia and Expo*, volume 2, pages 9–12, Jul 2003.
- [51] Strona internetowa. <http://www.aethon.com>.
- [52] Gijeong Jang, Sungho Kim, and Inso Kweon. Robust self-localization of mobile robot using artificial and natural landmarks. In *Proceedings 2003 IEEE International Symposium on Computational Intelligence in Robotics and Automation*, pages 412–417, 2003.
- [53] Jaber Juntu, Jan Sijbers, Dirk Van Dyck, and Jan Gielen. Bias field correction for mri images. In *Advanced Soft in Computing. Proceedings of the 4th International Conference on Computer Recognition Systems CORES'05*, pages 543–551. Springer, 2005.
- [54] Tomasz Kapuściński and Marian Wysocki. Hand gesture recognition for man-machine interaction. In *Second Workshop on Robot Motion and Control*, October 2001.
- [55] Gregor Klančar, Matej Kristan, and Rihard Karba. Wide-angle camera distortions and non-uniform illumination in mobile robot tracking. In *Robotics and Autonomous Systems*, volume 46, pages 125–133. Elsevier Science B.V., 2004.
- [56] Bogdan Kwolek. The usage of hidden markov models in a vision system of a mobile robot. In *Second Workshop on Robot Motion and Control*, October 2001.
- [57] Kyoung-Mi Lee and W. Nick Street. Dynamic learning of shape for automatic object recognition.

- [58] B. C. Li. *Applications of moment invariants to neurocomputing for pattern recognition*. PhD thesis, The Pennsylvania State University, 1990.
- [59] James Little, Jiping Lu, and Don Murray. Selecting stable image features for robot localization using stereo. In *Proceeding of the 1999 IEEE International Conference on Robotics and Automation*, 1998.
- [60] Salvatore Livatino and Claus B. Madsen. Autonomous robot navigation with automatic learning of visual landmarks. In *Proceedings of the 7th International Symposium on Intelligent Robot Systems*, July 1999.
- [61] Jorge Lobo, Carlos Quieroz, and Jorge Dias. Vertical world feature detection and mapping using stereo vision and accelerometers. In *Proceedings of the 9 International Symposium on Intelligent Robot Systems, SIRS'97*, pages 229–238, 2001.
- [62] David Lowe. Object recognition from local scale-invariant features. In *Proceedings of the International Conference on Computer Vision*, September 1999.
- [63] David G. Lowe. Towards a computational model for object recognition in IT cortex. In *Biologically Motivated Computer Vision*, pages 20–31, 2000.
- [64] Lili Ma, YangQuan Chen, and Kevin L Moore. A family of simplified geometric distortion models for camera calibration, 2003.
- [65] Claus B. Madsen and Claus B. Andersen. Optimal landmark selection for triangulation of robot position, 1998.
- [66] D. Marr and E. Hildreth. Theory of edge detection. *Proc. of the Royal Society*, B-207:187–217, 1980.
- [67] Emmanuele Menegatti, G. Gatto, Enrico Pagello, Hiroshi Ishiguro, and Takashi Minato. Distributed vision system for robot localisation in indoor environment. In *Proceedings of the 2nd European Conference on Mobile Robots ECMR05*, 2005.
- [68] Emmanuele Menegatti, Mauro Zoccarato, Enrico Pagello, and Hiroshi Ishiguro. Hierarchical image-based localisation for mobile robots with monte-carlo localisation. In *Proceedings of the 1st European Conference on Mobile Robots ECMR03*, pages 13–20, 2003.
- [69] Wojciech Mokrzycki. Stereoskopowe systemy postrzegania głębi. In *Machine Graphics and vision vol 1*, volume 1, pages 342–392. IPI PAN, 1992.
- [70] Eugenia Montiel, Alberto S. Augado, and Mark S. Nixon. On resolving ambiguities in arbitrary-shape extraction by the hough transform. In *Proceedings of the 11th British Machine Vision Conference BMVC2000*, 2000.
- [71] Don Murray and Cullen Jennings. Stereo vision based mapping and navigation for mobile robots. In *In Proc. IEEE International Conference on Robotics and Automation*, 1997.
- [72] Don Murray and Jim Little. Using real-time stereo vision for mobile robot navigation. In *Autonomous Robots*, volume 8, pages 161–171. Springer Netherlands, 2000.

- [73] Janeš Pers and Stanislav Kavačič. Nonparametric, model-based radial lens distortion correction using tilted camera assumption. In *Proceedings of the Computer Vision Winter Workshop 2002*, pages 286–295, 2002.
- [74] Filip Ponulak and Piotr Skrzypczyński. Experiments with a multi-agent perception system for mobile robots. In *Automation 2002 - nowości i perspektywy*, pages 252–260, 2002.
- [75] Yong Rui, Anoop Gupta, and JJ Cadiz. Viewing meetings captured by an omni-directional camera, 2000.
- [76] Stephen Se, Jim Little, and David Lowe. Global localization using distinctive visual features. In *Proceedings of the 2002 IEEE/SRJ International Conference on Intelligent Robots and Systems*, 2002.
- [77] Stephen Se, Jim Little, and David Lowe. *Mobile Robot Localization And Mapping with Uncertainty Using Scale-Invariant Visual Landmarks*, pages 735–758. 2002.
- [78] Stephen Se, David Lowe, and Jim Little. Local and global localization for mobile robots using visual landmarks. In *Proceedings of the 2001 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 414–420, 2001.
- [79] Stephen Se, David Lowe, and Jim Little. Vision-based mobile robot localization and mapping using scale-invariant features. In *Proceedings of IEEE International Conference on Robotics and Automation, ICRA 2001*, pages 2051–2058, May 2001.
- [80] Barbara Siemiątkowska. Wykorzystanie transformaty Hough’a do lokalizacji robota mobilnego. *Pomiar Automatyka Kontrola*, (4/2004), 2004.
- [81] Barbara Siemiątkowska and Rafał Chojecki. Mobile robot navigation based on omnidirectional sensor. In *Proceedings of the 1st European Conference on Mobile Robots ECMR03*, pages 101–106, Sep 2003.
- [82] Cyrill Stachniss, Giorgio Grisetti, and Wolfram Burgard. Recovering particle diversity in a rao-blackwellized particle filter for slam after actively closing loops. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pages 667–672, 2005.
- [83] Cyrill Stachniss, Dirk Häehnel, and Wolfram Burgard. Exploration with active loop-closing for FastSLAM. In *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2004.
- [84] Gideon Stein. Lens distortion calibration using point correspondences, October 1996.
- [85] Rahul Swaminathan and Shree K. Nayar. Non-metric calibration of wide-angle lenses.
- [86] Rahul Swaminathan and Shree K. Nayar. Non-metric calibration of wide-angle lenses and polycameras.
- [87] S. Thrun, M. Montemerlo, D. Koller, B. Wegbreit, J. Nieto, and E. Nebot. Fastslam: An efficient solution to the simultaneous localization and mapping problem with unknown data association, 2004.

- [88] Robert Valkenburg. Classification of camera calibration techniques. In *Image and Vision Computing NZ Conference IVCNZ95*, pages 43–48, 1995.
- [89] Felix von Hundelshausen, Michael Schreiber, and Raúl Rojas. A constructive feature detection approach for robotic vision. In *RobuCup*, pages 72–83, 2004.
- [90] Christopher D. Watkins, Alberto Sadun, and Stephen Marenka. *Nowoczesne metody przetwarzania obrazu*. Wydawnictwa Naukowo-Techniczne, 1995.
- [91] KikJin Yoon and InSo Kweon. Landmark desing and real-time landmark tracking for mobile robot localization. In *Proceedings of SPIE*, volume 4573, pages 219–226, 2002.
- [92] Zhengyou Zhang. A flexible technique for camera calibration. Technical report, Microsoft Research Microsoft Corporation, 1998.
- [93] Primo Zingaretti and Luca Regini. A framework for appearance-based robot navigation. In *Proceedings of the 1st European Conference on Mobile Robots ECMR03*, pages 147–152, Sep 2003.
- [94] Barbara Zitová, Jaroslav Katusky, Gabriele Peters, and Jan Flusser. Robust detection of significant points in multiframe images. In *Pattern Recognition Letters*, volume 20, pages 199–206. Elsevier Science B.V., 1999.

Errata do rozprawy doktorskiej pt
Lokalizacja robota mobilnego
względem automatycznie wybieranych obiektów

Ireneusz Hallmann

30 kwietnia 2008

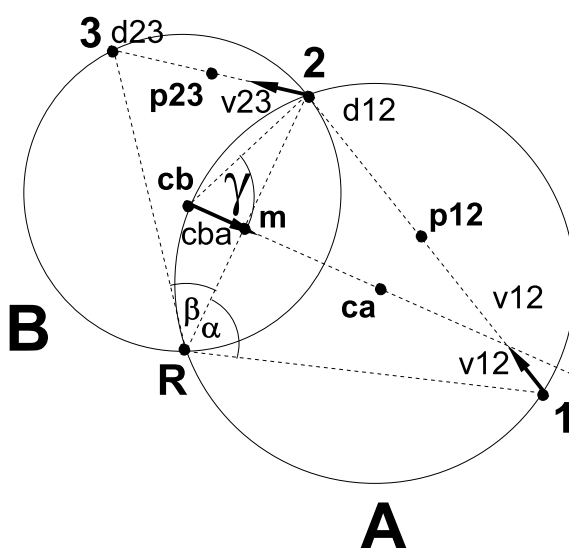
Część I

Poprawki do treści i rysunków

Rozdział 2.3, Str. 14 – prawidłowa postać równania (2.4):

$$\begin{cases} d12^2 = z_1^2 + z_2^2 - 2z_1z_2 \cos(\alpha) \\ d23^2 = z_2^2 + z_3^2 - 2z_2z_3 \cos(\beta) \\ d13^2 = z_1^2 + z_3^2 - 2z_1z_3 \cos(\alpha + \beta) \end{cases} \quad (2.4)$$

Rozdział 2.3, Str. 16 – prawidłowy rysunek 2.3 z wszystkimi symbolami:



Rysunek 2.3: Trzy znaczniki pozwalają na jednoznaczne wyznaczenie położenia robota – metoda geometrycznego przecinania okręgów [22]

Rozdział 2.4, Str. 19 – tekst i równania na początku strony 19 powinny mieć postać:

Jak widać z równań (2.18) i (2.19), zakłada się liniowość wszystkich transformacji. Gęstość prawdopodobieństwa dla wielowymiarowego wektora stanu jest następujący:

$$p(x) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp -\frac{1}{2}(x - \mu)^T \Sigma^{-1}(x - \mu) \quad (2.20)$$

Transformacje liniowe z początkowymi dystrybucjami gaussowskimi dają w wyniku również dystrybucje Gaussa, co symbolicznie przedstawia równanie (2.21). Równanie (2.22) przedstawia złożenie dwu jednowymiarowych dystrybucji Gaussa w dystrybucję dwuwymiarową:

$$\left. \begin{array}{l} X \sim N(\mu, \sigma) \\ Y = aX + b \end{array} \right\} \Rightarrow Y \sim N(\mathbf{A}\mu + \mathbf{B}, \mathbf{A}\Sigma\mathbf{A}^T), \quad (2.21)$$

$$\left. \begin{array}{l} X_1 \sim N(\mu_1, \Sigma_1^2) \\ X_2 \sim N(\mu_2, \Sigma_2^2) \end{array} \right\} \Rightarrow p(X_1) * p(X_2) \sim N\left(\frac{\Sigma_2}{\Sigma_1 + \Sigma_2}\mu_1 + \frac{\Sigma_1}{\Sigma_1 + \Sigma_2}\mu_2, \frac{1}{\Sigma_1^{-1} + \Sigma_2^{-1}}\right). \quad (2.22)$$

Rozdział 2.7, Str. 28 – fragment pierwszego akapitu powinien mieć postać:

System taki jest przedmiotem badań w kilku miejscach m.in. na Wydziale Mechatroniki Politechniki Warszawskiej i na Instytucie Automatyki i Inżynierii Informatycznej Politechniki Poznańskiej. W publikacjach [74] i [12] opisano wieloagentowy system lokalizacji robota mobilnego, którego elementem lokalizującym jest podwieszona kamera.

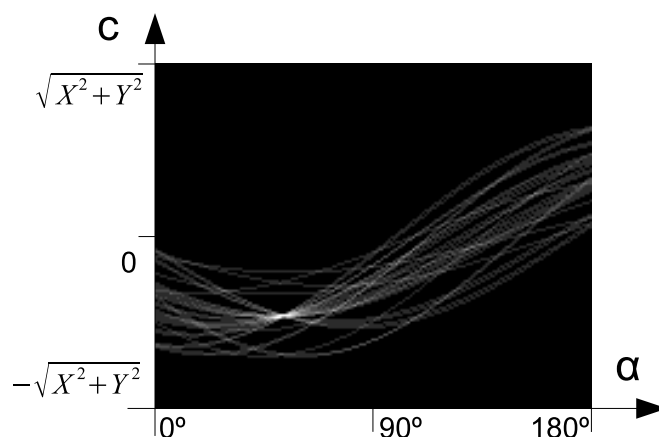
Rozdział 3.2, Str. 33 – prawidłowa postać równania (3.5):

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left(-\frac{(x^2 + y^2)}{2\sigma_x^2\sigma_y^2}\right), \quad (3.5)$$

Rozdział 3.4.2, Str. 39 – prawidłowa postać równania (3.17):

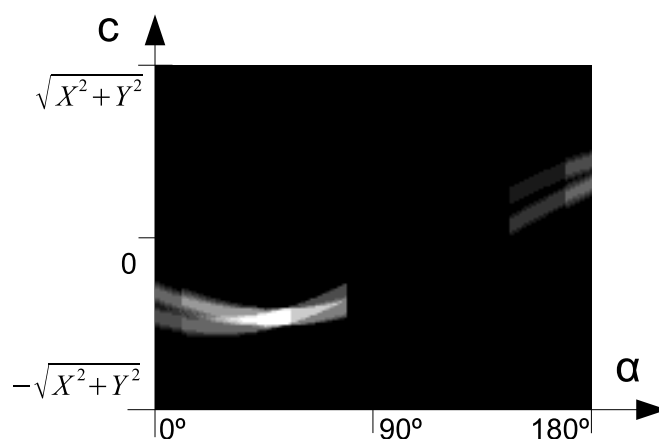
$$\begin{bmatrix} a_{00} & a_{01} & a_{02} & b_0 \\ a_{10} & a_{11} & a_{12} & b_1 \end{bmatrix} \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \\ y_1 & y_2 & y_3 & y_4 \\ x_1y_1 & x_2y_2 & x_3y_3 & x_4y_4 \\ 1 & 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} x'_1 & x'_2 & x'_3 & x'_4 \\ y'_1 & y'_2 & y'_3 & y'_4 \end{bmatrix} \quad (3.17)$$

Rozdział 3.5, Str. 41 – rysunek (3.15) z osiami współrzędnych:



Rysunek 3.15: Obraz transformaty Hougha obrazu z rysunku 3.14

Rozdział 3.5, Str. 42 – rysunek (3.16) z osiami współrzędnych:



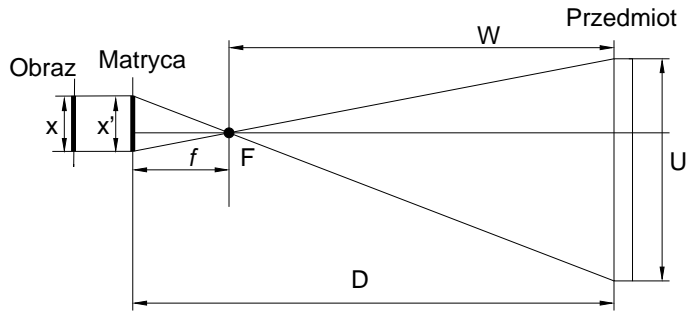
Rysunek 3.16: Obraz rozszerzonej transformaty Hougha obrazu z rysunku 3.14

Rozdział 4.1, Str. 43 – prawidłowa postać wzoru (4.1) z symbolami ujednoliconymi w obrębie rozdziału 4:

Znając **rozmiar fizyczny obrazu** U , ogniskową f oraz rozmiar obrazowy x przedmiotu, odległość D między obiektywem a przedmiotem wyznacza się wg wzoru:

$$D = f + W = f + f \frac{U}{x'} = f + f \frac{U}{s x} \quad (4.1)$$

Rozdział 4.1, Str. 44 – rysunek 4.2 z symbolami ujednoliconymi w obrębie rozdziału 4:



Rysunek 4.2: Parametry kamery - ogniskowa (f), ognisko (F) i związek pomiędzy rozmiarem na obrazie oraz rozmiarem rzeczywistym

Rozdział 4.2, Str. 45 – wstęp do rozdziału 4.2 powinien mieć postać:

Krótki przegląd i klasyfikację strategii kalibracji można znaleźć w [88]. Poniżej przedstawiono definicje i klasyfikację podejść do kalibracji zaczerpnięte z tego artykułu uzupełnione przykładami konkretnych metod.

Metody kalibracji kamery różnią się następującymi czynnikami: *złożonością modelu, klasyfikacją parametrów modelu, miarą dopasowania modelu i techniką rozwiązania numerycznego*. Model systemu można opisać równaniem

$$F(\mu) = 0, \quad (4.2)$$

gdzie μ jest wektorem parametrów łączącym wszystkie punkty obrazu, punkty świata rzeczywistego oraz parametry kamery. Z wektorem parametrów μ można skojarzyć wektorową zmienną losową M . Niech m będzie wektorem pomiarów M , zaś ϵ – błędem tego pomiaru (tzn $\nu = m - \epsilon$). Przy założeniu że M ma rozkład normalny z wartością średnią μ i macierzą kowariancji C^{-1} (tzn. $M \sim N(\mu, C^{-1})$), łatwo można pokazać że estymata maksymalnego podobieństwa $\hat{\mu}$ jest dana rozwiązaniem równania:

$$\min_{\mu} \epsilon^T C \epsilon \text{ należy do } F(m - \epsilon) = 0. \quad (4.3)$$

W ten sposób problem kalibracji można przedstawić, z pewnym uogólnieniem, jako optymalizację nieliniową.

Rozdział 5.1, Str. 59 – początek rozdziału powinien mieć postać:

Stosowane do identyfikacji obiektów momenty geometryczne są wartościami wyznaczanymi dla funkcji dwu zmiennych. Dla funkcji ciągłej dwóch zmiennych moment geometryczny rzędu j, k definiuje się wg wzoru (5.3):

$$M_{jk} = \int_{x=-\infty}^{+\infty} \int_{y=-\infty}^{+\infty} \psi_{jk}(x, y) f(x, y) dx dy; j, k = 0, 1, 2, \dots + \infty \quad (5.1)$$

Momenty funkcji dyskretnej dwu zmiennych wyznacza się według wzoru (5.4):

$$M_{jk} = \sum_{x=-\infty}^{+\infty} \sum_{y=-\infty}^{+\infty} \psi_{jk}(x, y) P(x, y) \quad (5.2)$$

W praktyce momenty wyznacza się dla obszarów o skończonych rozmiarach, jakimi są obrazy. Dodatkowo, zazwyczaj przed identyfikacją wyodrębnia się obszar który należy identyfikować, w którym wartość funkcji jasności $f(x, y)$ wynosi 1 dla punktów należących do obszaru i 0 dla punktów nie należących do obszaru. Oznaczając rozmiary obszaru przez X i Y , otrzymuje się:

$$M_{j,k} = \sum_{x=0}^X \sum_{y=0}^Y x^j y^k. \quad (5.5)$$

Bardziej praktyczne zastosowanie znajdują dyskretne momenty centralne, wyznaczone wg wzoru (5.6) i znormalizowane momenty centralne, wyznaczone wg wzoru (5.7):

$$\mu_{j,k} = \sum_{x=0}^X \sum_{y=0}^Y (x - \bar{x})^j (y - \bar{y})^k, \quad (5.6)$$

$$\eta_{j,k} = \frac{\mu_{j,k}}{\mu_{0,0}^{\frac{1+j+k}{2}}}. \quad (5.7)$$

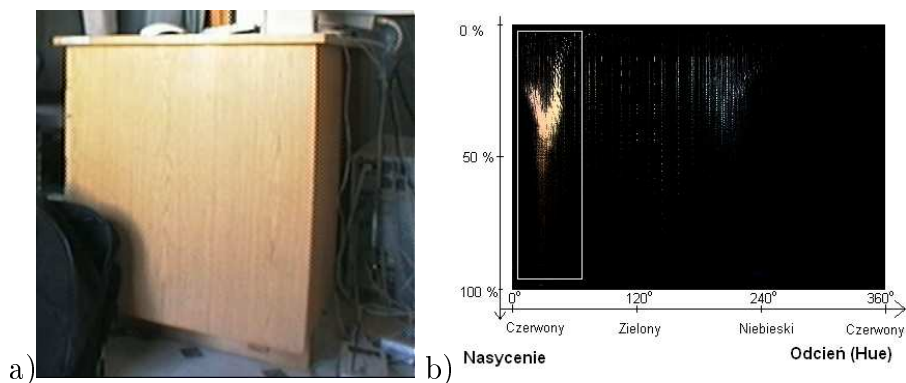
Rozdział 5.1.2, Str. 61 – prawidłowa postać wzorów (5.14), (5.15) i (5.16):

$$H5 = (\eta_{3,0} - 3\eta_{1,2})(\eta_{3,0} + \eta_{1,2})[(\eta_{3,0} + \eta_{1,2})^2 - 3(\eta_{2,1} + \eta_{0,3})^2] + (3\eta_{2,1} - \eta_{0,3})(\eta_{2,1} + \eta_{0,3})[3(\eta_{3,0} + \eta_{1,2})^2 - (\eta_{2,1} + \eta_{0,3})^2], \quad (5.14)$$

$$H6 = (\eta_{2,0} + \eta_{0,2})[(\eta_{3,0} + \eta_{1,2})^2 - 3(\eta_{2,1} + \eta_{0,3})^2] + 4\eta_{1,1}(\eta_{3,0} + \eta_{1,2})(\eta_{2,1} + \eta_{0,3}), \quad (5.15)$$

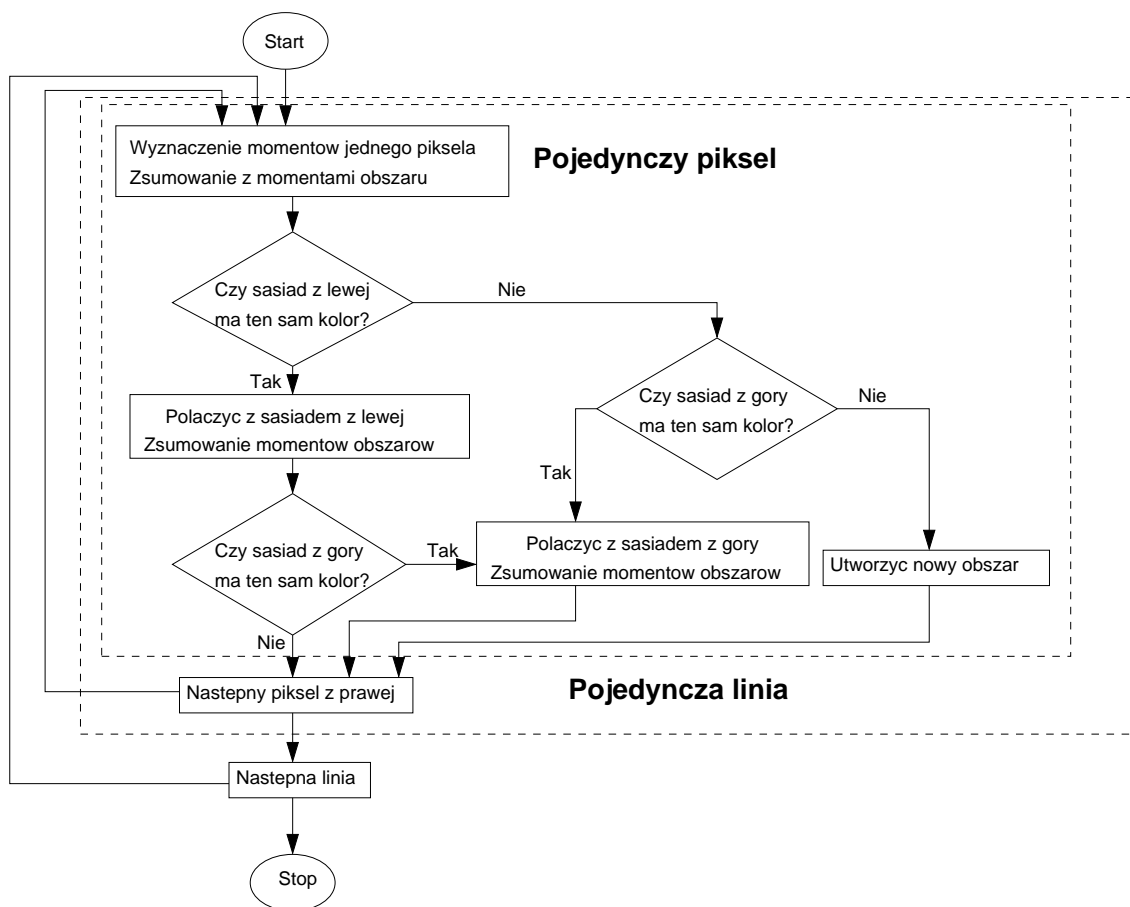
$$H7 = (3\eta_{2,1} - \eta_{0,3})(\eta_{2,1} + \eta_{0,3})[(\eta_{3,0} + \eta_{1,2})^2 - 3(\eta_{2,1} + \eta_{0,3})^2] - \eta_{3,0} - 3\eta_{1,2})(\eta_{3,0} + \eta_{1,2})[3(\eta_{3,0} + \eta_{1,2})^2 - (\eta_{2,1} + \eta_{0,3})^2]. \quad (5.16)$$

Rozdział 5.1.2, Str. 73 – rysunek 7.1b) z osiami współrzędnych:



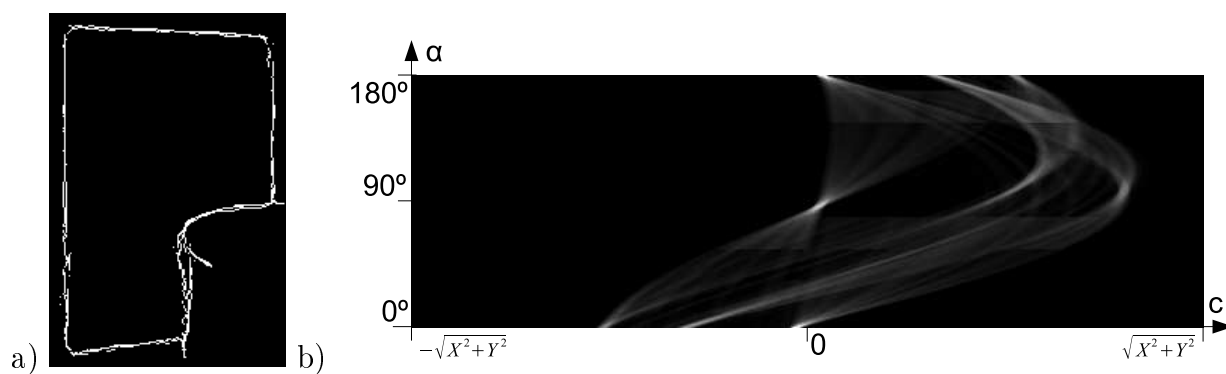
Rysunek 7.1: Histogram barw: Obiekt nierównomiernie oświetlony i histogram obrazu w modelu HSV

Rozdział 7.2, Str. 88 – rysunek 7.17 z zaznaczonymi miejscami obliczania momentów:

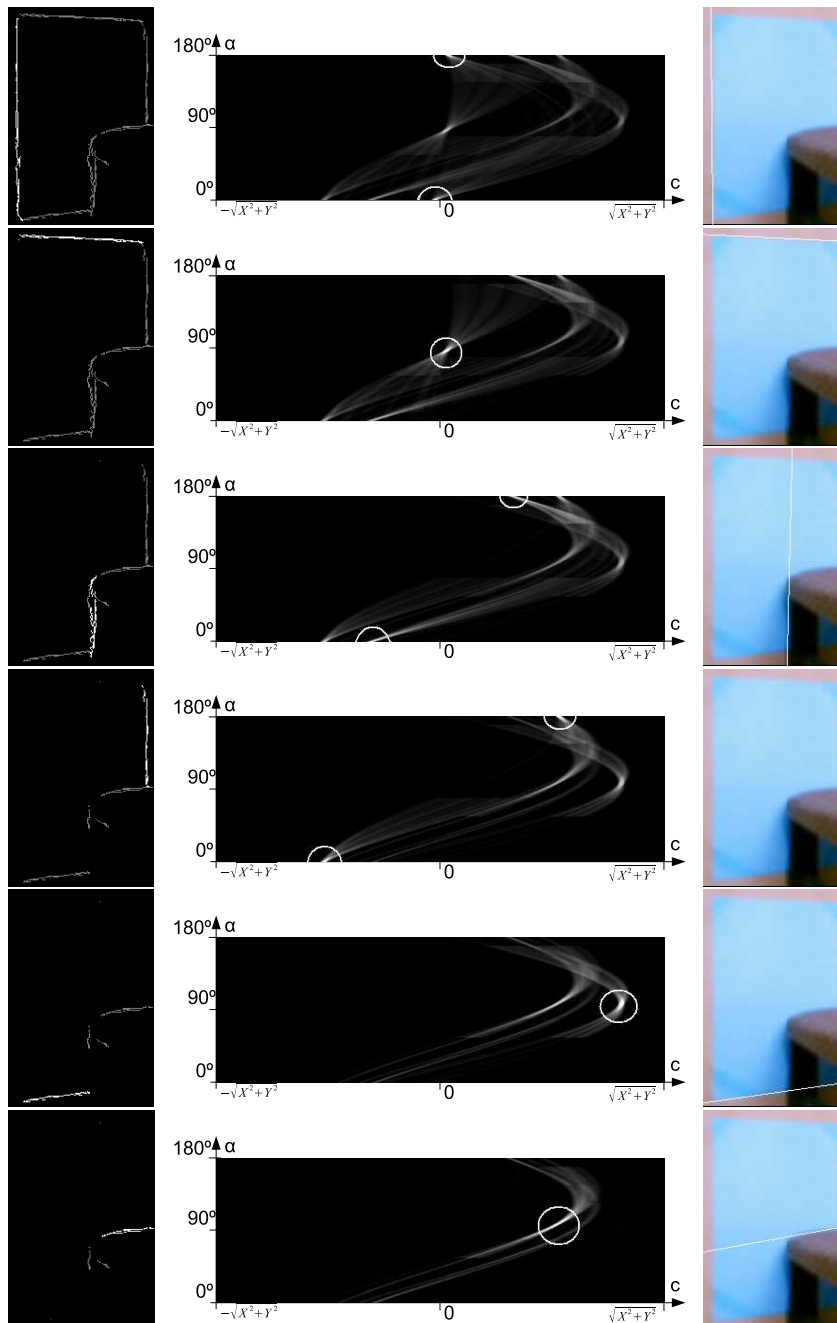


Rysunek 7.17: Schemat algorytmu wyznaczającego momenty geometryczne wszystkich obszarów w obrazie podczas jednego skanowania obrazu

Rozdział 7.3.3, Str. 93 – rysunek 7.23 z osiami współrzędnych:

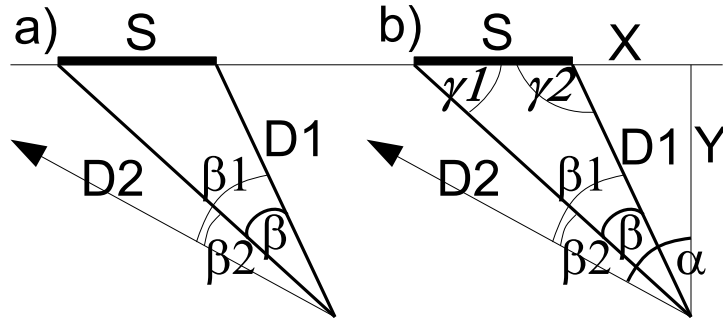


Rysunek 7.23: Transformata Hougha fragmentu obrazu z zakłóconym prostokątem: a) obraz wejściowy; b) obraz transformaty Hougha tego obrazu



Rysunek 7.24: Linie wykrywane w kolejnych cyklach algorytmu detekcji linii. W kolumnach od lewej: wejściowe piksele krawędzi z wyróżnioną najdłuższą linią; wejściowy obraz transformaty Hougha z wyróżnioną najdłuższą linią; wykryta linia na tle obrazu

Rozdział 8, Str. 103 – rysunek 8.5 powinien zawierać więcej szczegółów:



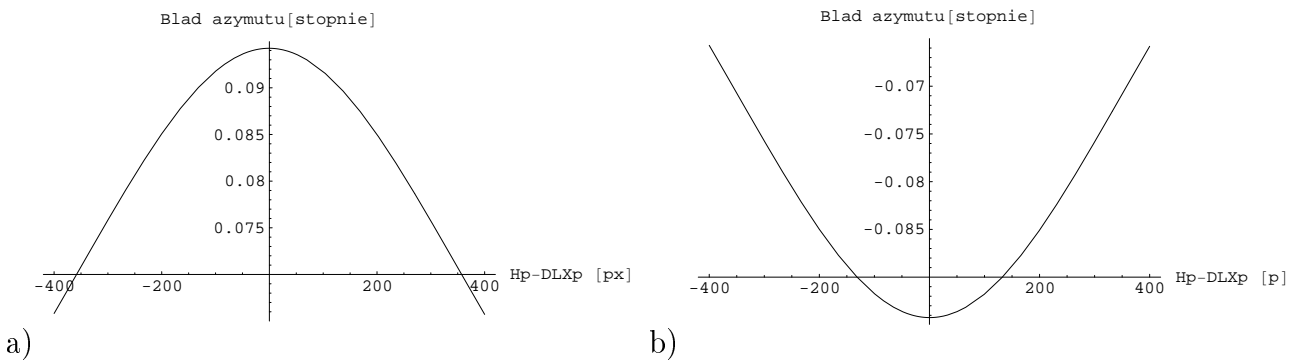
Rysunek 8.5: Schemat etapów wyznaczania pozycji robota metodą opartą o pomiar dwu odległości: a) wielkości mierzone bezpośrednio (D1, D2, β) i znane (S); b) pozycja robota (X i Y)

Rozdział 8.1, Str. 105 – wzory (8.6) i (8.7) powinny mieć postać:

$$\frac{\partial \beta}{\partial DX_{cm}} = \frac{Dist_{cm}(H_p - DLX_p)DX_p}{(H_p - DLX_p)^2 DX_{cm}^2 + Dist_{cm}^2 DX_p^2} = \frac{32(H_p - DLX_p)}{369664 + (H_p - DLX_p)^2} \quad (8.6)$$

$$\frac{\partial \beta}{\partial Dist_{cm}} = \frac{DX_{cm}(H_p - DLX_p)DX_p}{(H_p - DLX_p)^2 DX_{cm}^2 + Dist_{cm}^2 DX_p^2} = -\frac{38(H_p - DLX_p)}{369664 + (H_p - DLX_p)^2} \quad (8.7)$$

Rozdział 8.1, Str. 106 – wykres 8.8 z poprawioną skalą na osi poziomej:



Rysunek 8.8: Zależność błędu wyznaczenia azymutu od błędu wejściowych współrzędnych wynoszącego: a) +1 piksel, b) -1 piksel

Rozdział 8.2, Str. 108 – wzór (8.12) powinien mieć postać:

$$D = \frac{D_r}{\cos \beta} = D_r \sqrt{1 + \tan^2 \beta} = Dist_{cm} \frac{DY_p * V}{V_p * DY_{cm}} \sqrt{1 + \frac{(H_p - DLX_p)^2 * DX_{cm}^2}{DX_p^2 Dist_{cm}^2}} = \frac{DY_p * V}{DX_p * V_p * DY_{cm}} \sqrt{DX_p^2 * Dist_{cm}^2 + (H_p - DLX_p)^2 * DX_{cm}^2} \quad (8.12)$$

Rozdział 8.2, Str. 109 – akapit zaczynający się na dole strony 109 powinien mieć postać:

Odległość D i dokładność jej wyznaczenia zależą od kąta β , dlatego poniższe wykresy przedstawiają zależność zmiany D od V_p i β . Rysunek 8.11 przedstawia zależność wyznaczonej odległości D od rozmiaru obrazowego V_p i odległości obrazowej od osi optycznej $H_p - DLX_p$. Przekroje tego wykresu wzdłuż osi wysokości obrazowej V_p i odległości poziomej od osi optycznej $H_p - DLX_p$ są przedstawione na rysunkach 8.12 i 8.13. Rysunek 8.12 przedstawia zależność D od obrazowej odległości poziomej obiektu od osi optycznej $H_p - DLX_p$ dla wybranych wartości wysokości obrazowej V_p : 100, 200, 300, 400 i 500 pikseli. Zależność ta jest rosnąca szybciej niż liniowo. Wpływ azymutu na zmierzoną odległość nie jest bardzo duży. Odległość jest tylko elementem umożliwiającym wyznaczenie położenia kamery i pominięcie azymutu w modelu powodowałoby znaczne błędy pomiarów. Rysunek 8.13 przedstawia zależność odległości D od wysokości obrazowej obiektu V_p dla wybranych wartości odległości od osi optycznej $H_p - DLX_p$: 0, 100, 200, 300 i 400 pikseli. Zależność jest w przybliżeniu hiperboliczna. Z wykresu wynika, że dla zastosowanej kamery i znacznika o wysokości 70 cm zakres odległości w których znacznik jest dobrze widoczny i można tę odległość mierzyć wynosi 5 metrów. Skala wykresów kończy się na wartości wysokości obrazowej $V_p = 100px$, ponieważ mniejsze rozmiary obrazowe znacznika powodują gorszą dokładność pomiarów i mogą utrudniać rozpoznanie znacznika.

Rozdział 8.2, Str. 111 – akapit zaczynający się na dole strony 111 powinien mieć postać:

W bardziej złożonym, ale mającym znaczenie praktyczne przypadku, interesująca jest dokładność wyznaczenia odległości zależna od położenia kamery w lokalnym i globalnym układzie współrzędnych i przy różnych kierunkach patrzenia kamery. W celu wyznaczenia tych zależności wprowadzono funkcje pomocnicze:

- rzeczywista wysokość obrazowa $Vl_p = Vl_p(\beta_n, D_n)$,
- rzeczywista odległość obrazowa obiektu od osi optycznej $Hl_p = Hl_p(\beta_n, D_n)$.

Funkcje te oznaczają oczekiwane położenie i wysokość obiektu w obrazie przy określonych współrzędnych i orientacji kamery.

Wartości tych funkcji są wyznaczone jako liczby rzeczywiste, z pominięciem skończonego rozmiaru piksela. Wartość Hl_p zależy w lokalnym układzie współrzędnych tylko od kąta β i parametrów kamery, zaś Vl_p od odległości D , kąta β i parametrów kamery:

$$Hl_p = \frac{Dist_{cm}DX_p \tan \beta_n}{DX_{cm}} \quad (8.20)$$

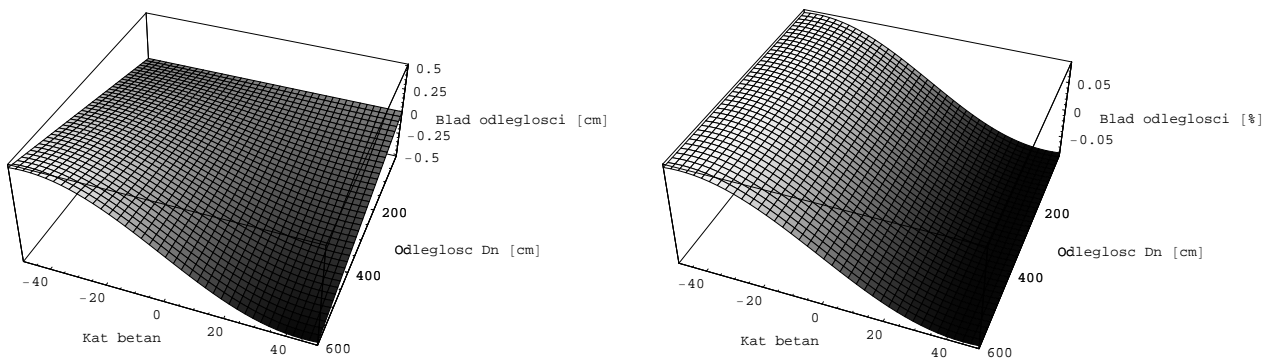
$$Vl_p = \frac{Dist_{cm}DY_p V}{DY_{cm} D_n \cos \beta_n} \quad (8.21)$$

Rzeczywiste wartości odległości i rozmiarów są liczbami całkowitymi jeżeli liczy się piksele, lub rzeczywistymi jeżeli wyznaczono je na podstawie równań prostych. Błąd maksymalny zależy od przyjętego maksymalnego błędu położenia i rozmiaru. Wyznaczenie tego błędu opisano w rozdziale 7.3 – zwykle jest on mniejszy niż 1 piksel. Błąd pomiaru dla określonych $D = D_n$, $\beta = \beta_n$ wyznacza się następująco:

- Wyznaczenie Hl_p i Vl_p na podstawie D_n i β_n ,

- Zniekształcenie Hl_p i/lub Vl_p o określoną wartość, wynikającą z przyjętego maksymalnego błędu współrzędnych,
- Wyznaczenie D i β dla zniekształconych Hl_p i Vl_p i porównanie z wartościami początkowymi.

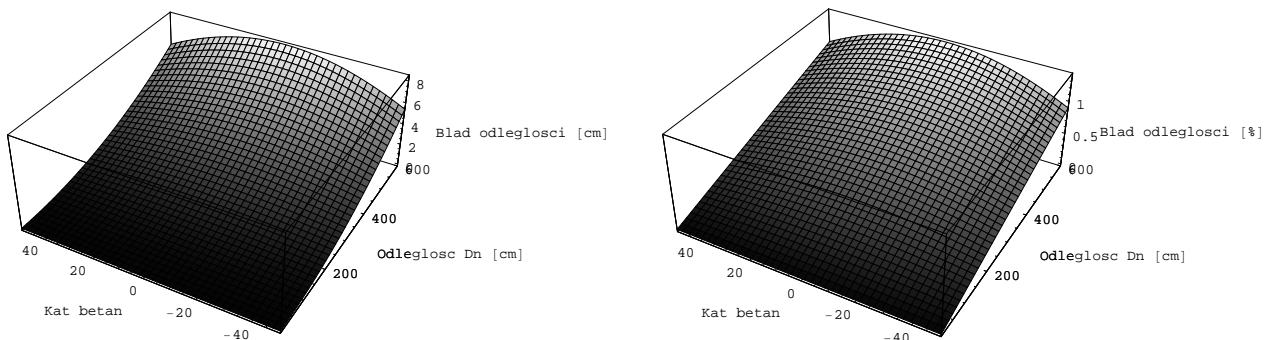
Rozdział 8.2, Str. 114 – poprawione oznaczenia na rysunkach 8.16 i 8.17:



a)

b)

Rysunek 8.16: Błąd wyznaczenia odległości w układzie lokalnym przy przesunięciu Hl_p o 1 piksel a) w cm b) procentowy

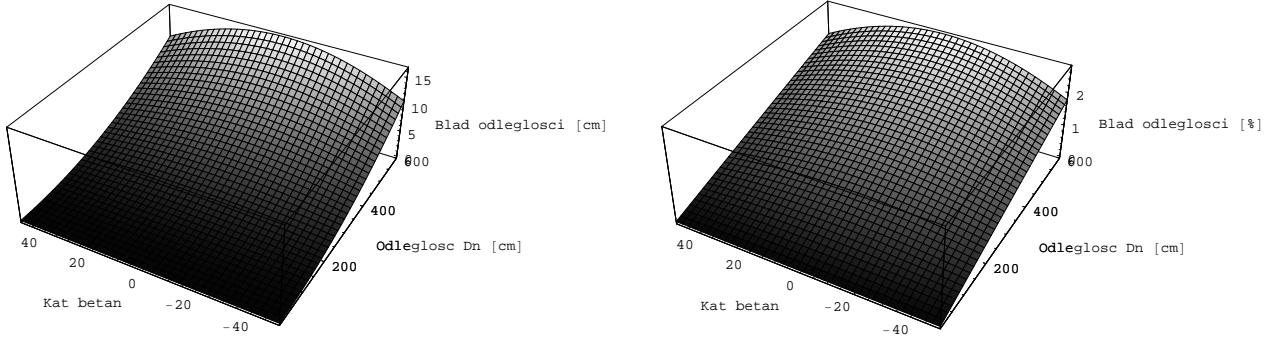


a)

b)

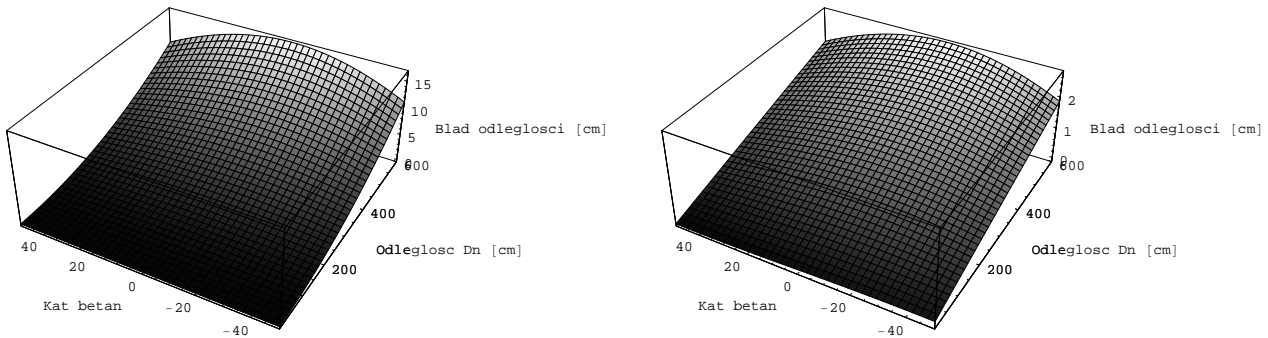
Rysunek 8.17: Błąd wyznaczenia odległości w układzie lokalnym przy przesunięciu Vl_p o 1 piksel a) w cm b) procentowy

Rozdział 8.2, Str. 115 – poprawione oznaczenia na rysunkach 8.18 i 8.19:



a) b)

Rysunek 8.18: Błąd wyznaczenia odległości w układzie lokalnym przy przesunięciu Vl_p o 2 piksele a) w cm b) procentowy



a) b)

Rysunek 8.19: Błąd wyznaczenia odległości w układzie lokalnym przy przesunięciu Hl_p o 1 piksel a) w cm b) procentowy i Vl_p o 2 piksele

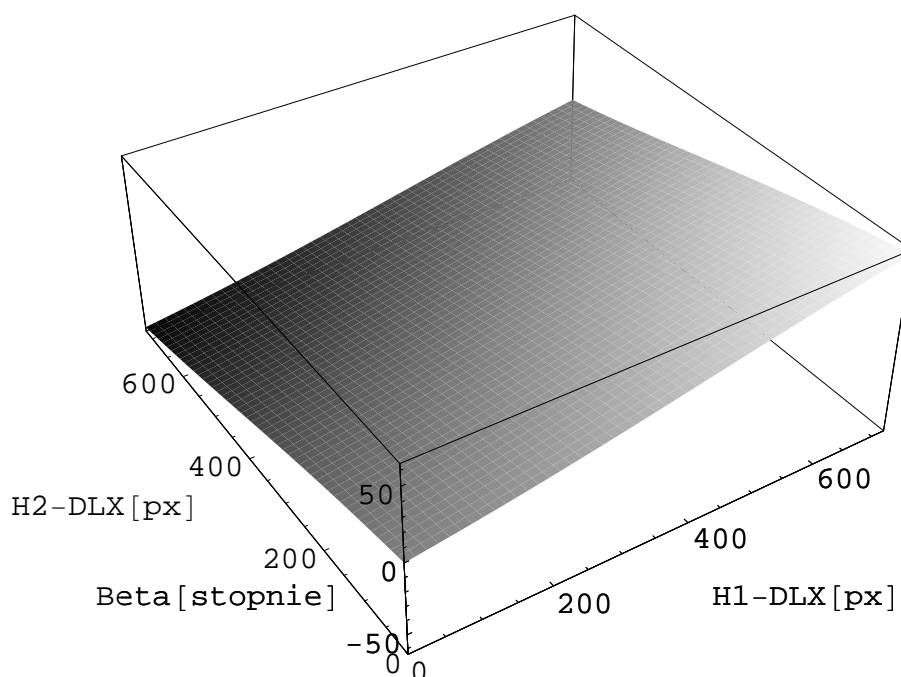
Rozdział 8.3, Str. 118 – poprawiony tekst podrozdziału oraz oznaczenia na rysunkach 8.18, 8.19 i 8.24:

Oczywistym ograniczeniem dokładności pomiaru jest skończona rozdzielczość obrazu w pikselach. Dlatego analiza dokładności obu metod powinna wyznaczyć zmianę kąta β przy zmianie o jeden piksel wartości $H1$ i $H2$ dla obu metod.

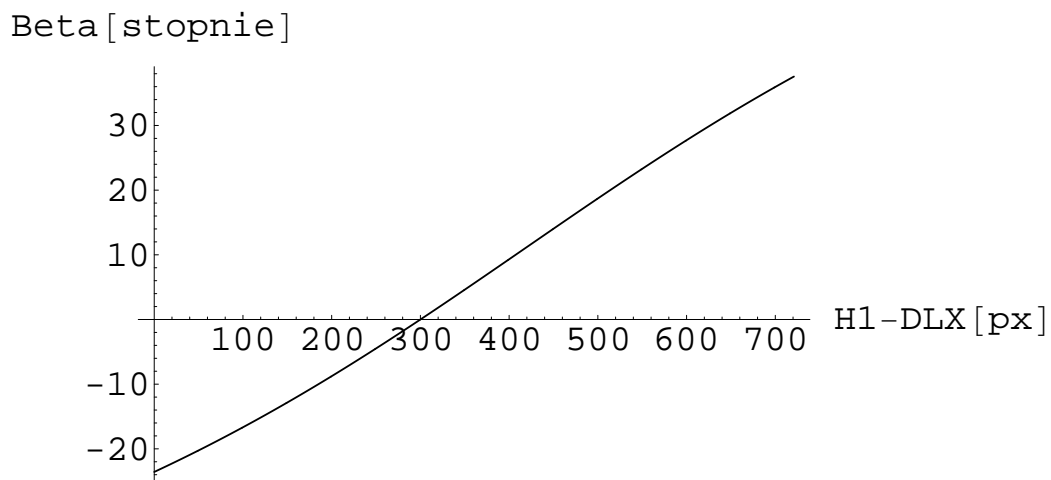
Wartości $Dist_{cm} = 16,0\text{cm}$, $X_{cm} = 19,0\text{cm}$, $X_p = 722$, $DLX_p = 418$ zostały wyznaczone podczas kalibracji kamery (rozdział 4). Dla tych parametrów kamery kąt β wyznacza się według równania (8.30):

$$\beta = \frac{180}{\pi}(\arctan 0,00164931(H1_p - 418) - \arctan 0,00164931(H2_p - 418)) \quad (8.30)$$

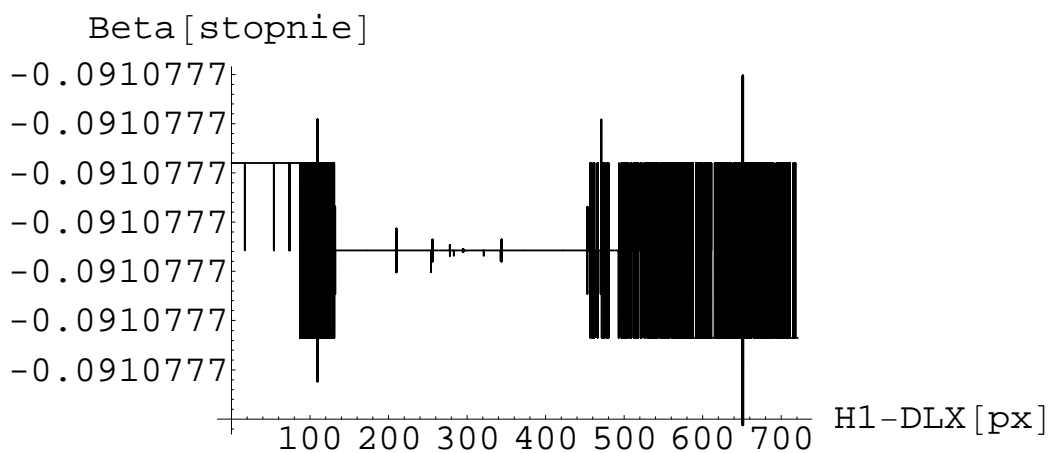
Wykres zamieszczony na rysunku 8.22 przedstawia zależność kąta β od wartości B1 i B2 wyrażonych w pikslach. Wykres 8.23 przedstawia przekrój przez wykres dla B2=300 piksli. Jak widać, wykres przypomina funkcję tangens. W zakresie kątów mierzonych w rzeczywistości ($<30^\circ$) i gdy mierzony obiekt znajduje się w pobliżu osi optycznej wykres jest prawie liniowy. Oznacza to, że kąt można wyznaczyć z określoną rozdzielczością prawie niezależnie od jego wartości. Różnicę funkcji $\beta(H1_p + 1) - \beta(H1_p)$ przedstawiono na wykresie 8.24. Z ostatniego wykresu wynika, że przy zmianie współrzędnej o jeden piksel dla małych kątów zmiana kąta jest praktycznie stała i dla używanej kamery wynosi $0,091^\circ/px$. Najlepszą rozdzielczość kątową i dokładność wyznaczenia kąta uzyskuje się, gdy obiekt znajduje się w pobliżu osi optycznej kamery. W zakresie współrzędnych poziomych 200-500 czyli w pobliżu osi optycznej kamery na wykresie z rysunku 8.24 błąd wynosi 0, co jest spowodowane skończoną dokładnością i rozdzielczością obliczeń numerycznych.



Rysunek 8.22: Zależność kąta β od poziomych współrzędnych obrazowych obiektu



Rysunek 8.23: Wartości kąta β przy współrzędnej obrazowej $X2_p = 300$



Rysunek 8.24: Zmiana kąta β przy zmianie współrzędnej obrazowej $X1_p$ o 1 piksel dla $X2_p = 300$

Część II

Zestawienie symboli

- A – macierz współczynników przekształcenia afinicznego
- a – współrzędna pozioma x piksela w prostokącie po przekształceniu dwuliniowym korygującym zniekształcenia geometryczne
- \underline{A} – macierz współczynników przekształcenia biliniowego
- $A1$ – odległość od kamery do kierunku do lewej krawędzi znacznika rzutowanego na płaszczyznę siatki kalibrującej, w odległości $Dist_{cm}$, wyrażona w cm.
- $A2$ – odległość od kamery do kierunku do prawej krawędzi znacznika rzutowanego na płaszczyznę siatki kalibrującej, w odległości $Dist_{cm}$, wyrażona w cm.
- AX – czyli szerokość obrazowa znacznika rzutowana na płaszczyznę siatki kalibrującej, wyrażona w cm.
- $AMI[n]$ (rozdz. 7) – n -ty niezmiennik AMI
- A_t – macierz (o rozmiarach $n \times n$) zmiany stanu z chwili $t - 1$ do chwili t
- B – wektor przesunięcia przekształcenia afinicznego
- b (rozdz. 2.5) – odległość między ogniskami kamer
- b (rozdz. 4.3) – współrzędna pionowa y piksela w prostokącie po przekształceniu dwuliniowym korygującym zniekształcenia geometryczne
- \underline{B} – wektor przesunięcia przekształcenia biliniowego
- $B1$ – odległość pozioma kierunku do lewej krawędzi znacznika rzutowanej na płaszczyznę siatki kalibrującej od lewego brzegu widocznej części płaszczyzny siatki kalibrującej (w odległości $Dist_{cm}$).
- $B2$ – odległość pozioma kierunku do prawej krawędzi znacznika rzutowanej na płaszczyznę siatki kalibrującej od lewego brzegu widocznej części płaszczyzny siatki kalibrującej (w odległości $Dist_{cm}$).
- $\bar{bel}(x_t)$ – przewidywana gęstość prawdopodobieństwa wartości wektora stanu x w chwili t , opisująca przewidywane wartości na podstawie wartości wyznaczonych dla chwili $t - 1$
- $bel(x_t)$ – gęstość prawdopodobieństwa wartości wektora x wyznaczana na podstawie pomiarów w chwili t
- B_t – macierz (o rozmiarach $n \times m$) opisująca zmianę sterowania z chwili $t - 1$ do chwili t
- $c2m$ – odległość między cb a m
- ca, cb – środki okręgów A i B
- C_e – rozdzielczość licznika w impulsach na obrót
- c_m – współczynnik przeliczenia impulsów licznika na przebytą odległość
- C_t – macierz (o rozmiarach $k \times n$) opisująca przejście z wektora stanu x_t na wektor obserwacji z_t
- D – odległość między kamerą a obiektem wyrażona w cm
- $D1$ – odległość od kamery do lewej krawędzi znacznika (widocznej po prawej stronie obrazu)
- $D1_r$ – rzeczywista odległość kamery (roboty) od lewej krawędzi znacznika
- $D2$ – odległość od kamery do prawej krawędzi znacznika (widocznej po lewej stronie obrazu)
- $D2_r$ – rzeczywista odległość kamery (roboty) od lewej krawędzi znacznika
- d_1, d_2, d_3, d_4 – narożniki czworokąta w obrazie przed korekcją przekształceniem dwuliniowym, także ogólne współrzędne tych narożników

- d_{12} – odległość między znacznikami 1 i 2
- d_{13} – odległość między znacznikami 1 i 3
- d_{23} – odległość między znacznikami 2 i 3
- dHp_1 – błąd wyznaczenia współrzędnej poziomej lewej krawędzi znacznika w pikselach
- dHp_2 – błąd wyznaczenia współrzędnej poziomej prawej krawędzi znacznika w pikselach
- $Dist_{cm}$ – odległość od kamery do płaszczyzny siatki kalibrującej, wyrażona w cm
- DLX – odległość między osią optyczną a brzegiem obrazu w kierunku poziomym, czyli współrzędna pozioma rzutu osi optycznej w obrazie, wyrażona w pikselach (DLXp). Posiada odpowiednik metryczny DLXcm będący odległością rzutu osi na płaszczyznę siatki kalibrującej do brzegu widzianego obszaru podczas kalibracji kamery.
- D_n (rozd. 2.2) – nominalna średnica kół
- D_n (rozd. 8.2) – rzeczywista odległość do obiektu wyznaczona na podstawie znajomości X , Y i α
- $DR1$ – odległość od kamery do płaszczyzny prostopadłej do osi kamery i przechodzącej przez lewą krawędź znacznika
- $DR2$ – odległość od kamery do płaszczyzny prostopadłej do osi kamery i przechodzącej przez prawą krawędź znacznika
- D_r – odległość między kamerą a obiektem rzutowana na oś optyczną wyrażona w cm
- dVp_1 – błąd wyznaczenia wysokości lewej krawędzi w pikselach
- dVp_2 – błąd wyznaczenia wysokości lewej krawędzi w pikselach
- d_x – różnica współrzędnych x środka obrazu i rzutu osi optycznej
- DX_{cm} – szerokość pozioma pola widzenia podczas kalibracji
- DX_p – szerokość obrazu w pikselach
- d_y – różnica współrzędnych x środka obrazu i rzutu osi optycznej
- DY_{cm} – wysokość pola widzenia podczas kalibracji
- DY_p – wysokość obrazu w pikselach
- e_1, e_2 – punkty pomocnicze przy przekształceniu dwuliniowym
- F – równanie łączące parametry kamery, współrzędne rzeczywistych punktów i odpowiadające im współrzędne obrazowe
- f – ogniskowa soczewki kamery
- F_l, F_p – ognisko lewej i prawej kamery
- F_x, F_y – równanie F dla współrzędnych obrazowych x i y
- g' – udział składowej zielonej w ogólnej jasności w modelu barw RGB
- H – odległość od obiektu do osi optycznej wyrażona w cm
- $H1 \div H7$ (rozd. 5) – niezmienniki Hu od 1 do 7
- $H1_{cm}$ – współrzędna lewej krawędzi znacznika rzutowana na płaszczyznę siatki kalibrującej, czyli odległość od brzegu widocznej siatki kalibrującej do kierunku do lewej krawędzi w odległości $Dist_{cm}$, wyrażona w cm
- $H1_p$ – współrzędna obrazowa lewej krawędzi znacznika
- $H2_{cm}$ – współrzędna prawej krawędzi znacznika rzutowana na płaszczyznę siatki kalibrującej, wyrażona w cm
- $H2_p$ – współrzędna obrazowa prawej krawędzi znacznika
- H_{cm} (rozd. 8.1) – odległość pozioma na płaszczyźnie siatki kalibrującej od brzegu widocznego obszaru do kierunku do lewego brzegu siatki
- H_{cm}, H_m (rozd. 8.2) – odległość od rzutu osi optycznej do kierunku do obiektu rzutowanego na płaszczyznę w odległości $Dist_{cm}$ siatki kalibrującej wyrażona w cm
- Hl_p – oczekiwana współrzędna obrazowa obiektu wyrażona w pikselach

- H_p – współrzędna obrazowa obiektu wyrażona w pikselach
- $Hu[n]$ (rozd. 7) – n-ty niezmiennik Hu
- I – czas zliczania impulsów liczników odometrycznych
- I_n (rozd. 5) – n-ty niezmiennik AMI
- L – jasność piksela wyznaczana jako średnia składowych R,G,B
- Lx_i, Ly_i – współrzędne x i y znacznika i , np Lx_0
- Lo_i – azymut znacznika i względem przodu robota
- la – odległość między ca a odcinkiem łączącym znaczniki 1 i 2
- lb – odległość między cb a odcinkiem łączącym znaczniki 2 i 3
- m – środek odcinka $z2$ łączącego robota ze znacznikiem 2
- $M_{jk}, M_{j,k}$ – moment geometryczny rzędy j,k
- N_R, N_L – liczby impulsów zliczone przez liczniki prawy i lewy
- n – współczynnik biegu – przeliczenie obrotów silnika na obroty kół
- P – obserwowany punkt w przestrzeni
- $p12$ – środek odcinka $d12$
- $p23$ – środek odcinka $d23$
- P_l – obraz punktu P w lewej kamerze
- P_r – obraz punktu P w prawej kamerze
- ra, rb – promienie okręgów A i B
- $r_{c.g.cw}$ – błąd odometrii przy skręcaniu w prawo
- $r_{c.g.ccw}$ – błąd odometrii przy skręcaniu w lewo
- Ro – orientacja robota
- Rx i Ry – współrzędne robota
- r' – udział składowej czerwonej w ogólnej jasności w modelu barw RGB
- S – szerokość znacznika
- s, s_x, s_y – współczynnik przeliczenia współrzędnych metrycznych na obrazowe
- U – współrzędna punktu w układzie kamery, odpowiadająca współrzędnej obrazowej x
- u_t – wektor (o rozmiarze m) sterowania w chwili t
- V (rozd. 4) – współrzędna punktu w układzie kamery, odpowiadająca współrzędnej obrazowej y
- V (rozd. 8.2) – wysokość rzeczywista obiektu wyrażona w cm
- $v12, v23, cba$ – wektory jednostkowe o kierunku odcinków $d12, d23, c2m$
- V_{cm}, V_m – wysokość obiektu rzutowanego na płaszczyznę w odległości $Dist_{cm}$ siatki kalibrującej wyrażona w cm
- Vl_p – oczekiwana wysokość obrazowa obiektu wyrażona w pikselach
- V_p – wysokość obrazowa obiektu wyrażona w pikselach
- W – współrzędna punktu w układzie kamery, odpowiadająca odległości od kamery rzutowanej na oś optyczną
- X (rozd. 2.5) – współrzędna rzeczywista X w globalnym układzie współrzędnych
- X (rozd. 8) – odległość od prawej krawędzi znacznika do kamery rzutowana na płaszczyznę znacznika (współrzędna X kamery w układzie współrzędnych znacznika)
- x (rozd. 4) – współrzędna obrazowa x
- \bar{x} – współrzędna x środka ciężkości
- x_1, x_2, x_3, x_4 – współrzędne x narożników czworokąta w obrazie przed korekcją przekształceniem dwuliniowym
- X_l – współrzędna rzeczywista X w układzie współrzędnych lewej kamery
- x_l – współrzędna obrazowa x punktu p w obrazie lewej kamery

- X_r (rozd. 2.5) – współrzędna rzeczywista X w układzie współrzędnych prawej kamery
- X_r (rozd. 8) – rzeczywista współrzędna X kamery (robota) w układzie współrzędnych znacznika
- x_r – współrzędna obrazowa x punktu p w obrazie prawej kamery
- x_t – wektor (o rozmiarze n) stanu w chwili t
- x' – rozmiar przedmiotu przutowany na matrycę kamery, w jednostkach metrycznych
- Y (rozd. 2.5) – współrzędna rzeczywista Y w globalnym układzie współrzędnych
- Y (rozd. 8) – odległość od kamery do płaszczyzny znacznika (współrzędna Y kamery w układzie współrzędnych znacznika)
- y (rozd. 4) – współrzędna obrazowa y
- \bar{y} – współrzędna y środka ciężkości
- y_1, y_2, y_3, y_4 – współrzędne y narożników czworokąta w obrazie przed korekcją przekształceniem dwuliniowym
- Y_l – współrzędna rzeczywista Y w układzie współrzędnych lewej kamery
- y_l – współrzędna obrazowa y punktu p w obrazie lewej kamery
- Y_r (rozd. 2.5) – współrzędna rzeczywista Y w układzie współrzędnych prawej kamery
- Y_r (rozd. 8) – rzeczywista współrzędna Y kamery (robota) w układzie współrzędnych znacznika
- y_r – współrzędna obrazowa y punktu p w obrazie prawej kamery
- Z – współrzędna rzeczywista Z w globalnym układzie współrzędnych
- Z_l – współrzędna rzeczywista Z w układzie współrzędnych lewej kamery
- Z_r – współrzędna rzeczywista Z w układzie współrzędnych prawej kamery
- z_t – wektor (o rozmiarze k) obserwacji w chwili t
- α (rozd. 2.3) – kąt między azymutami znaczników 1 i 2
- α (rozd. 8) – kąt między osią kamery a osią OX układu znacznika
- α_r – rzeczywista orientacja kamery (robota) w układzie współrzędnych znacznika
- β (rozd. 2.3) – kąt między azymutami znaczników 2 i 3
- β (rozd. 8.1, 8.2) – kąt między osią optyczną kamery a kierunkiem do obiektu (azymut do obiektu)
- β (rozd. 8.3, 8.4, 8.5) – kąt widzenia znacznika (szerokość katowa)
- β_1 – kąt między osią kamery a kierunkiem od kamery do lewej krawędzi znacznika
- β_{1r} – rzeczywisty kąt między osią kamery a kierunkiem od kamery do lewej krawędzi znacznika.
- β_2 – kąt między osią kamery a kierunkiem od kamery do prawej krawędzi znacznika
- β_{2r} – rzeczywisty kąt między osią kamery a kierunkiem od kamery do prawej krawędzi znacznika.
- β_n – rzeczywisty kąt między osią kamery i kierunkiem do obiektu wyznaczony na podstawie znajomości X, Y i α
- γ_1 – kąt między kierunkiem od kamery do prawej krawędzi znacznika a płaszczyzną znacznika
- γ_2 – kąt między kierunkiem od kamery do lewej krawędzi znacznika a płaszczyzną znacznika
- δ_t – zmienna losowa reprezentująca zakłócenia pomiaru w filtrze Kalmana
- ϵ_t – zmienna losowa reprezentująca zakłócenia przetwarzania w filtrze Kalmana
- η_{jk} – znormalizowany moment centralny geometryczny rzędu j, k
- μ_{jk} – moment centralny geometryczny rzędu j, k
- Σ – kąt o jaki są obrócone kamery w układzie zbieżnoosiowym